Mathematical Statistics
Stockholm University

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

Ekaterina Fetisova, Anders Moberg and Gudrun Brattström

Research Report 2017:13

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 **2017:13**,
http://www.math.su.se

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

Ekaterina Fetisova* Anders Moberg† Gudrun Brattström‡

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) variables. In Part I, several latent factor models were suggested for evaluation of temperature data from climate model simulations, forced by a varying number of forcings, against climate proxy data from the last millennium. Here, in Part II, focusing on climatological characteristics of forcings, we deepen the discussion by suggesting two alternative latent variable models that can be used for evaluation of temperature simulations forced by *five specific* forcings of natural and anthropogenic origin. The first statistical model is formulated in line with *confirmatory factor analysis* (CFA), accompanied by a more detailed discussion about the interpretation of latent temperature responses and their mutual relationships. Introducing further *causal links* between some latent variables, the CFA model is extended to a *structural equation model* (SEM), which allows us to reflect more complicated climatological relationsnhips with respect to all SEM's variables. Each statistical model is developed for use with data from a single region, which can be of any size. Associated with different hypotheses, the CFA and SEM models can, as a beginning, be fitted to observable simulated data only, which allows us to investigate the underlying latent structure associated with the simulated climate system. Then, the best-fitting model can be fitted to the data with real climate proxy data included, to test the consistency between the latent simulated temperature responses and their real-world counterparts embedded in observations. The performance of both these statistical models and some models suggested in Part I 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 Physical Geography, Stockholm university, Sweden; anders.moberg@natgeo.su.se
‡Department of Mathematics, Stockholm University; gudrun@math.su.se

1 Introduction

Climate models are powerful tools for improving our understanding of how the climate system works, for making predictions of the future climate and for assessing potential impacts of climatic changes ([11]). Using a mathematical representation of the real climate system, climate models are defined as systems of complex differential equations based on physical, biological and chemical principles. In the virtual world of climate models, climatologists can perform experiments that are not feasible in the real world climate system; for example, to neglect or simplify all but one process, in order to identify the role of this particular process clearly, e.g. the influence of changes in solar irradiance on the radiative properties of the atmosphere, or to test hypotheses related to this process. In an analogous fashion, the overall effect of several processes, acting jointly, can be investigated.

Climatologically, in order to evaluate and compare the magnitude of the effects of the processes in question on the climate, it is often convenient to analyse their impact on the radiative balance of the Earth ([12]). The net change in the Earth's radiative balance at the tropopause (incoming energy flux minus outgoing energy flux expressed in W/m^2) caused by a change in a climate driver is called *a climate (radiative) forcing* (see glossary p.1460 in [18] for a definition and e.g. [23] for an overview discussion about the concept *radiative forcing*).

External natural drivers of climate change, such as changes in solar radiation or in the orbital position of the Earth, will result in radiative forcing of climate. Volcanic eruptions, ejecting small particles and various chemical compounds into the atmosphere and thereby affecting climate (during a few years), is another example of a natural external agent that induces climate forcing. The ongoing release of carbon dioxide to the atmosphere, primarily by burning fossil fuels is also an example of external forcing, but being of anthropogenic origin. As concluded in [28], "it is unequivocal that anthropogenic increases in the well-mixed greenhouse gases have substantially enhanced the greenhouse effect, and the resulting forcing continues to increase". Other examples of human influence on climate are changes in land-use and the emissions of aerosols through various industrial and burning processes, which are also associated with radiative forcing of climate.

Causes of the internal climate variability are various processes internal to the climate system itself. Ocean and atmosphere circulation and their variations and mutual interactions are examples of processes that are clearly internal to the climate system. In some situations, in particular in modelling experiments, climate scientists can regard internal causes for climate change as forcings. For example, natural variations in atmospheric greenhouse gas concentrations or aerosols can be seen as drivers of climate change, although they are rather occurring due to various biogeochemical processes within the climate system.

The range of types of climate models is very wide. Here, our focus is on the most sophisticated state-of-the art climate models referred to as Global Climate Models (GCMs) or Earth System Models (ESMs). As computing capabilities have evolved during the past years, the complexity of GCMs and ESMs has substantially increased: for instance, the number of components of the Earth system that can be included and coupled in GCMs and ESMs have increased, or the previous equilibrium simulations can now be replaced by transient changes, e.g. in the atmospheric greenhouse gases and aerosol loading (see e.g. [4]; [24]; [6]). However, despite great advances achieved during the past decades, some simplifications are unavoidable, e.g. due to the time scales involved and/or incomplete knowledge about some processes. As a consequence, the complexity of even the most sophisticated climate models is still far from the complexity of the real climate system. Further, it should be kept in mind that even a careful design cannot guarantee that each component of climate modelling, e.g. parameterisation of subgrid-scale processes, has been employed in its optimal form. All these together may affect the accuracy of model simulations.

Another issue that may affect the accuracy of model simulations is uncertainties in forcing reconstructions. As emphasised by e.g. [15], uncertainties can be large for such anthropogenic forcings as aerosol forcing and land use forcing, especially associated with the conversion of forest to agricultural land. Further, our knowledge about various feedback processes that may either amplify or damp the direct effect of a given forcing is not complete.

All the above-mentioned issues together point naturally to the importance of undertaking evaluation of climate model simulations. Clearly, the choice of evaluation approaches depends on the scientific objectives of the study for which a particular climate model has been designed. In the context of the present work, our attention is confined to two particular approaches. The first one stems from the statistical framework developed by [37] (henceforth referred to as SUN12), while the second, known as the optimal fingerprinting framework, is employed in the so-called Detection and Attribution (D&A) studies ([39], [25], [13], [14]).

A key feature common to both frameworks is that each of them deals with latent (i.e. unobservable) variables. More precisely, focusing on the near-surface temperature as a climatic variable of interest, both assume that temperature responses to forcings are not directly observable either in a simulated climate system or in the real one. Further, both frameworks incorporate simulated and observational data, where the latter consists of instrumental data when it is available and, otherwise, of temperature reconstructions derived from climate proxy data. Importantly, both frameworks are suitable for applications to the data covering the relatively recent past of about one millennium, albeit each of them can be generalised to any period in the geological past as soon as simulations and proxy data on any continuous climatic variable are available.

The differences between the frameworks lie in the statistical methods used there. SUN12 developed two test statistics - a correlation and distance-based test statistics - allowing us to determine (1) the significance of correlation between a forced climate model simulation and observational data and (2) whether a given forced climate model simulation demonstrates a significantly better agreement with observational data than an unforced (control) simulation. Ultimately, applying these two test statistics can help us to address the question as to whether the magnitude of a latent simulated temperature response to a given forcing is correctly represented by the climate model under consideration compared to its real-world counterpart embedded in observations. The same question (among others) has been addressed in a number of D&A studies but by means of linear regression models, where not only response variables but also explanatory variables are allowed to be contaminated with noise. Such regression models are referred to by statisticians as measurement error (ME) models.

Using the ideas and definitions of these two frameworks, we, [9] (henceforth referred to as Part I) formulated several latent factor models of varying complexity that can be used for evaluation of climate model simulations forced by different numbers of (reconstructed) forcings. We also focused on the link between our factor models and ME models used in D&A studies. Our theoretical discussion in Sec. 5 in Part I demonstrated that our factor models are capable of addressing questions posed in D&A studies, which justifies their use in D&A studies as an alternative approach to ME models. Furthermore, we elucidated additional advantages of reasoning in the spirit of factor analysis. However, despite those advantages, we also pointed out that factor analysis may be too restrictive for describing complicated underlying climatological relationships. Therefore, in the present work, our intention is to investigate theoretically possible extensions of our factor models in order to allow the statistical modelling of climatological relationships which cannot be described within factor analysis.

The main motive behind extensions is that in factor analysis the relationships among latent factors themselves are modelled exclusively in terms of correlations. However, assuming that two latent factors are correlated (or associated) says nothing about the underlying reasons for this association. Indeed, an association between two variables, say A and B, may arise because

(1) A influences (or causes) B, which graphically can be expressed as $\mathtt{A}\rightarrow \mathtt{B}$,

(2) B influences A, $\mathtt{A}\leftarrow \mathtt{B}$,

(3) A and B influence each other reciprocally, $\mathtt{A} \overset{\longleftarrow}{\longrightarrow} \mathtt{B}$

(4) A and B depend on some third variable(-s) (spurious correlation).

Statistical models allowing causal links between latent common factors (and between latent and observed variables as well) are known as structural equation models (SEM) and are widely used in various research fields, for example in sociology ([31],[33]), psychology ([1],[7]), and economics ([16]). In the present work, we argue that their application within climatological science is also relevant.

As a matter of fact, the notion of causality is not new to climate research. As examples, we can refer to [21] and [36], where the causal structure between atmospheric CO_2, i.e. the forcing itself, and global temperature has been studied by applying the methods based on Granger causality and the concept of information flow, respectively. The latter concept was also used by [22] to investigate the cause-effect relation between the two climate modes, El Niño and the Indian Ocean Dipole. But our questions to be addressed and the methods we use for achieving our goals are different compared to the above-mentioned works. Our main aim in Part I and here, in Part II, is to suggest statistical methods that can be used for evaluating temperature data from climate model simulations against observed/reconstructed temperatures for the last millennium in terms of latent (unobservable) temperature responses to climate forcings. These statistical methods should be capable of taking into account uncertainties in observable data, both simulated and observational, and of reflecting our substantive knowledge of the properties of the real-world system and of the climate model under consideration.

In our opinion, structural equation modelling with latent variables is an appropriate approach for achieving our goals. Admittedly, the SEM approach, combining the properties of factor analysis and path analysis, is a more sophisticated statistical technique than factor analysis, but on the other hand it will give us more flexibility in analysing and evaluating climate model simulations in case associated factor models fail to lead to clear and/or reliable conclusions.

In what follows, focusing on the properties of five specific real-world forcings, we will first present the basic conceptual ideas of possible causal links between *true* temperature responses to these real-world forcings (see Sec. 2). Based on this discussion, two schemes of modelling the relationships between latent temperature responses to forcings are suggested. The first one ignores any causal links, while the second allows their presence. In Sec. 3, each scheme is used for formulating an associated statistical model incorporating both simulated temperatures and observational data. The first model is a (mixed [1]) factor model. Although factor models have been discussed in Part I, presenting a factor model here will illustrate the consequences of assuming the negligibility of causal relationships for the interpretation of latent factors, which was not discussed in Part I. The second model is a structural equation model. We also discuss a possible mixture of these two statistical models. In Sec. 4, an overview of the main features of the statistical models presented is given.

[1] Recall from Part I, a mixed factor model combines features of an oblique factor model where all latent factors are modelled as mutually correlated, and of an orthogonal factor model, where all latent factors are mutually uncorrelated.

2 A structure for describing relations between true temperature variations and contributions from different climate forcings

As a first step, let us define the true unobservable temperature τ as follows:

$$\tau = \beta \cdot \xi^T_{\mathsf{ALL}} + \eta_{\mathsf{internal}}, \tag{2.1}$$

where ξ^T_{ALL} represents the true temperature response to all possible forcings (the superscript stands for *True*, not a transpose), η_{internal} represents the internal random variability of the real-world climate system, *including* any random variability due to the presence of the forcings, and the coefficient β represents the expected change in τ for a one-unit change in ξ^T_{ALL}. Eq. (2.1) reflects the assumption that only forcings are capable of influencing the temperature systematically, while the internal factors contribute to the temperature variability randomly, without generating trends. Notice that all components in (2.1) are given in the form of mean-centered time-series.

Following the assumptions made by e.g. [34], we assume that the forced and unforced components, i.e. ξ^T_{ALL} and η_{internal}, respectively, are mutually independent. For the purpose of our discussion, let us represent Eq. (2.1) graphically by means of a path diagram, which is an important component of structural equation modelling (see Figure 1).

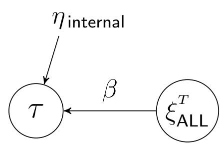

Figure 1. Path diagram associated with Eq. (2.1).

To understand a path diagram, we need to explain its symbols:

- A straight, one-headed arrow represents a causal relationship between two variables, meaning that a change in the variable at the tail of the arrow will result in a change in the variable at the head of the arrow (with all other variables in the diagram held constant). The former type of variables are referred to as *exogenous* (Greek: "of external origin") or *independent* variables because their causes lie outside the path diagram. Variables that receive causal inputs in the diagram are referred to as *endogenous* ("of internal origin") or *dependent* variables because their values are influenced by variables that lie within the path diagram.

- A curved two-headed arrow between two variables indicates that these variables may be correlated without any assumed direct relationship.

- Two straight single-headed arrows connecting two variables signifies reciprocal causation.

- Latent variables are designated by placing them in circles, observed variables by placing them in squares, while disturbance/error terms are represented as latent variables, albeit without placing them in circles.

Applying the above description of the symbols to the path diagram in Figure 1 enables us to interpret Eq. (2.1) from the perspective of structural equation modelling, that is, ξ^T_{ALL} and η_{internal} can be viewed not only as components of τ but also as its causes. That is, we may say that τ is an endogenous (or dependent) variable, whose variability is accounted for by two exogenous (or independent) variables, ξ^T_{ALL} and η_{internal}. The assumption of independence between the latter two is reflected in the path diagram by the absence of a curved two-headed arrow between them. Finally, the path diagram, in contrast to Eq. (2.1), highlights that (1) all variables involved are latent, as none of them is placed in a square, and (2) η_{internal} is modelled as a disturbance term, i.e. a term influencing τ randomly.

Next, let us take a closer look at the structure of ξ^T_{ALL}, that is, at its components that may contribute to the variability of ξ^T_{ALL} either systematically or randomly. Hence, just as τ, ξ^T_{ALL} is to be viewed as a latent endogenous variable, receiving causal inputs from its components. For Eq. (2.1), it entails that β is set to 1.

By definition, ξ^T_{ALL} comprises the true temperature responses to all possible external forcings and to all kinds of interactions between them. To list all forcings, acting in the real-world climate system, is an unrealistic [2] and, fortunately, unnecessary task within our analysis. Since we are aiming at evaluating climate model simulations forced by selected forcings either individually or jointly, it is justified to confine our attention to these selected forcings. Letting ξ^T_{comb} represent the overall true temperature response to the forcing combination of interest, we may first decompose ξ^T_{ALL} as follows:

$$\xi^T_{\mathsf{ALL}} = \xi^T_{\mathsf{comb}} + \tilde{\zeta}^T_{\mathsf{ALL}}, \tag{2.2}$$

where $\tilde{\zeta}^T_{\mathsf{ALL}}$ represents the residual forced variability due to other climate forcings not included in the combination. Statistically, excluding forcings from the simulated climate system entails the assumption that the systematical influence of the corresponding real-world forcings on τ is negligible (which might be true, depending on what forcings are excluded). In other words, just as internal factors, excluded forcings are assumed to contribute to the temperature variability randomly and independently from the forcings included in the combination. This corresponds to viewing $\tilde{\zeta}^T_{\mathsf{ALL}}$ in Eq. (2.2) as a disturbance term, independent from ξ^T_{comb}.

[2] Some of the forcings might be unknown to us due to our incomplete knowledge of the real-world climate system, for example regarding many processes related to forcings from aerosols ([3])

The next step is to discuss the structure of ξ^T_{comb}, viewed in (2.2) as a latent endogenous variable. To this end, let the following five forcings be in focus:

1. Changes in the solar irradiance (Sol),
2. Changes in the orbital position of the Earth (Orb),
3. Changes in the amount of stratospheric aerosols of volcanic origin (Volc),
4. Changes in vegetation and land cover caused by natural and anthropogenic factors (Land), and
5. Changes in the concentrations of greenhouse gases in the atmosphere (Ghg) also of both natural and anthropogenic origin.

The reason behind this choice is that these five forcings are regarded as main drivers of the climate change during the last millennium ([19]). Thus, state-of-the-art Earth System Model (ESM) simulations driven by these (or some of these) forcings both individually *and* jointly are already available ([29]) and further simulations are planned ([19]), thereby making the issue of their evaluation relevant.

Following the notations of Part I, let the individual temperature responses to each of the specified forcings be denoted ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, ξ^T_{Land}, and ξ^T_{Ghg}, respectively. The last two temperature responses deserve special attention because each of them represents the overall (joint) temperature response both to natural and anthropogenic changes in vegetation and in the concentrations of Ghg:s, respectively. Put differently, they are two-component temperature responses, decomposed as follows:

$$\xi^T_{\mathsf{Land}} = \xi^T_{\mathsf{Land\,(natural)}} + \xi^T_{\mathsf{Land\ (anthr)}} \tag{2.3}$$

$$\xi^T_{\mathsf{Ghg}} = \xi^T_{\mathsf{Ghg\,(natural)}} + \xi^T_{\mathsf{Ghg\ (anthr)}} \tag{2.4}$$

Undoubtedly, in the real-world climate, the range of possible causes of natural changes in Land and Ghg, which give rise to $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$, may be very wide. More precisely, these changes can occur not only due to forcings, but also due to internal factors. However, under the assumption of the independence between the forced and unforced components of τ and between the two components of ξ^T_{ALL} in (2.2), internal factors and the forcings, not included in the combination of interest, are not regarded as possible causes of natural changes in the Land and Ghg forcings capable of influencing the temperature systematically. Consequently, taking these independence assumptions into account, natural changes in vegetation and in the levels of Ghg:s can be explained only by the forcings that are part of the combination of interest. In our study, they are the solar, orbital and volcanic forcings.

In Figure 2, giving a graphical overview of these relationships (among others to be discussed further), the relations between (ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}), ($\xi^T_{\mathsf{Land\,(natural)}}$, $\xi^T_{\mathsf{Ghg\,(natural)}}$)

and the associated forcings are highlighted by blue arrows. Following these arrows, we may say that the first three temperature responses can be viewed as indirect causes (i.e. through the Land and Ghg forcings) of the last two.

At this point, it is important to stress that, just as in Part I, we wish to analyse temperature responses to the forcings, not the forcings themselves. In other words, we are interested in the relationships depicted in Figure 2 where the forcings are excluded. A direct consequence of excluding the forcings is that ξ^T_{Sol}, ξ^T_{Orb} and ξ^T_{Volc} become direct causes of $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$, which is not true from the physical viewpoint: a temperature response cannot physically be a direct cause of another temperature response. Nevertheless, from the pure statistical perspective, this issue is not as relevant as from the physical one. Without interpreting cause-effect relations between temperature responses literally, viewing ξ^T_{Sol}, ξ^T_{Orb} and ξ^T_{Volc} as direct causes of $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$ would allow us to apply another statistical method of analysing pairwise associations between latent variables representing these temperature responses. Indeed, instead of relating them to each other through correlations, which, in fact, is done in the 'optimal fingerprinting' approach used in many D&A studies and in our factor models from Part I, cause-effect relations justifies the use of regression models, where the 'causes' play the role of explanatory (independent, exogenous) variables, while the 'effects' are response variables, i.e. dependent (endogenous) variables.

Replacing correlations by regressions offers the advantage of statistical modelling to some extent the presence of feedbacks in the climate system, meaning in the context of the present work that natural changes in vegetation and in the levels of Ghg:s are processes that are physically dependent on the solar, orbital and volcanic forcings. As already mentioned in the introduction, this replacement is motivated when statistical models, where the relationships between latent variables are modelled in terms of correlations, failed to provide clear and reliable conclusions. Note that increasing the degree of complexity of a statistical model by introducing causal links does not guarantee that the resulting model will lead to acceptable results. But bearing in mind that the climatological relations can be complicated, the development of more complicated statistical models is highly motivated.

Although cause-effect relations between temperature responses are not to be taken literally, it does not mean that the direction of influence between them can be determined arbitrarily. Each link should be justified from the climatological point of view, which inevitably requires the involvement of forcings, although they are not represented in our statistical models explicitly. For example, according to Figure 2, the Land- and Ghg-forcings (whether natural or anthropogenic) cannot impact the temperature by inducing changes in the solar, orbital, and volcanic forcings. Hence, regression models with ($\xi^T_{\mathsf{Land\,(natural)}}$, $\xi^T_{\mathsf{Ghg\,(natural)}}$) as causes of ($\xi^T_{\mathsf{Sol}}$, ξ^T_{Orb}, ξ^T_{Volc}), would be senseless.

Based on the discussion above, we define $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$ as *'causally'* *dependent* temperature responses, each of which depends on ξ^T_{Sol}, ξ^T_{Orb} and ξ^T_{Volc}. Consequently, the last three temperature responses are defined as *'causally' independent* with respect to all temperature responses involved, including themselves. Statistically, causal independence implies that the variables in question can be related to each other only through correlations (or equivalently, covariances).

It should also be noted that the natural changes in vegetation can also have an impact on the level of greenhouse gases in the atmosphere through the carbon cycle and vice versa, thereby establishing a reciprocal relationship between $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$. In other words, these two temperature responses can be 'causally' dependent on each other as well. In Figure 2, for the sake of neatness, this possible reciprocal relationship is highlighted by one two-headed blue arrow relating the Land and Ghg forcings.

The relationships associated with human activity are highlighted in Figure 2 by brown arrows. Notice that in and of itself human activity is not a forcing, but its presence in Figure 2 is definitely needed.

We regard human activity as a process physically independent of the natural forcings (we do not discuss here any possible influence of the changed climate on the actions of humanity). Therefore, anthropogenic changes in Land and Ghg are also regarded as forcings physically independent from the natural ones. This makes it reasonable to classify $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$ as 'causally' independent temperature responses with respect to the temperature responses to the natural forcings. However, with respect to each other, they can be defined either as (1) 'causally' independent, or as (2) 'causally' dependent due to possible reciprocal or unidirectional causal relationships between them. Compare with the temperature responses to the natural forcings, which are defined exclusively as 'causally' independent with respect to all temperature responses, including themselves.

Finally, according to Figure 2, there is one more 'causally' independent variable that may be viewed as a 'direct cause' of $\xi^T_{\mathsf{Land\,(natural)}}$ and $\xi^T_{\mathsf{Ghg\,(natural)}}$, namely the temperature response to all possible interactions between the (physically independent) natural and anthropogenic forcings. In Figure 2, this is denoted $\xi^T_{\mathsf{interact}}$. Admittedly, it would be more appropriate to separate the interactions between the natural forcings from the interaction between anthropogenic ones. But, keeping in mind the main aim of our analysis, requiring ultimately involving climate model simulations in the discussion, we have to take the issue of the availability of simulated data into account. Just as in Part I, we assume here that climate model simulations driven by all possible combinations of forcings are not available. Thus, $\xi^T_{\mathsf{interact}}$ cannot be split into several terms representing temperature responses to interactions between various combinations of the given forcings.

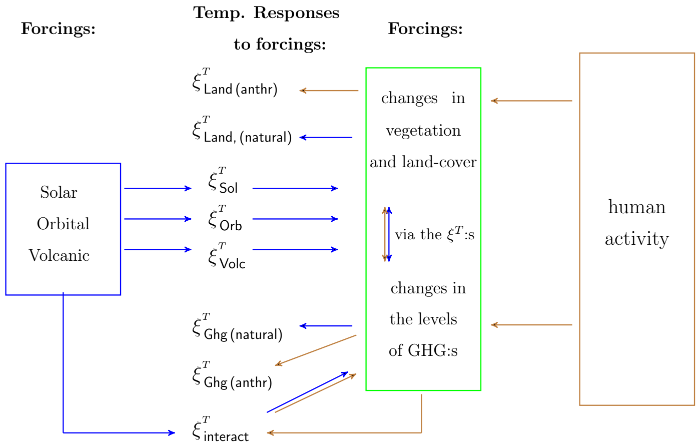

Figure 2. *Schematical (and simplified for the purposes of our analysis) description of the influences of the real-world natural and anthropogenic forcings on the temperature represented here by its responses to the five selected forcings of natural and anthropogenic character. Natural influences are highlighted by blue arrows, anthropogenic influences by brown ones.*

Definitely, the structure suggested in Figure 2 is not a simple structure, which immediately gives rise to questions as to (1) how the strength of the real-world relationships between the individual temperature responses can be statistically assessed, and (2) whether the same relationships hold within the simulated climate system under consideration. In the present paper, we suggest two possible ways of reasoning, which we call Scheme 1 and Scheme 2. In what follows, we present the basic ideas and assumptions associated with each of the schemes, which will constitute a basis for formulating statistical models incorporating single-forcing and multi-forcing climate model simulations of interest.

2.1 Scheme 1: *only 'causally' independent temperature responses*

Scheme 1 arises when all causal inputs to ξ^T_{Land} and ξ^T_{Ghg} from ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc} and ξ^T_{interact} are ignored. That is, the natural components, $\xi^T_{\text{Land (natur)}}$ and $\xi^T_{\text{Ghg (natur)}}$, are not related to τ in a systematic way. Instead, they are thought to be a part of the random internal temperature variability represented by η_{internal} (see Eq. (2.1)). Hence, Scheme 1 is associated with the assumption that the effect of natural changes in vegetation and in the levels of Ghg:s on the temperature is negligible.

Consequently, ξ^T_{Land} and ξ^T_{Ghg} are no longer overall temperature responses, but are one-component responses containing only $\xi^T_{\text{Land (anthr)}}$ and $\xi^T_{\text{Ghg (anthr)}}$, respectively. Notice that under Scheme 1, $\xi^T_{\text{Land (anthr)}}$ and $\xi^T_{\text{Ghg (anthr)}}$ cannot be modelled as 'causally' dependent on each other.

To summarise, the temperature responses of interest under Scheme 1 are: ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, $\xi^T_{\text{Land (anthr)}}$, $\xi^T_{\text{Ghg (anthr)}}$, and ξ^T_{interact}. Since each of them is 'causally' independent with respect to the others, the structure of ξ^T_{comb} and, thus, of ξ^T_{ALL}, can be expressed by one equation, namely:

$$\xi^T_{\mathsf{ALL}} = \underbrace{\beta_1 \cdot \xi^T_{\mathsf{Sol}} + \beta_2 \cdot \xi^T_{\mathsf{Orb}} + \beta_3 \cdot \xi^T_{\mathsf{Volc}} + \beta_4 \cdot \xi^T_{\mathsf{Land\,(anthr)}} + \beta_5 \cdot \xi^T_{\mathsf{Ghg\,(anthr)}} + \beta_6 \cdot \xi^T_{\mathsf{interact}}}_{=\xi^T_{\mathsf{comb}}} + \tilde{\zeta}^T_{\mathsf{ALL}}. \tag{2.5}$$

where ξ^T_{comb} and $\tilde{\zeta}^T_{\mathsf{ALL}}$ are defined in (2.2), and each coefficient β_i is a partial coefficient, meaning that it represents the expected change in τ for a one-unit change in the corresponding ξ^T, when the remaining ξ^T:s are held at constant values.

Keeping in mind that $\tilde{\zeta}^T_{\mathsf{ALL}}$ is assumed to be independent of ξ^T_{comb}, inserting (2.5) into the expression for τ in (2.1) yields:

$$\begin{aligned} \tau &= \xi^T_{\mathsf{comb}} + \zeta^T_{\mathsf{ALL}} = \\ &= \beta_1 \cdot \xi^T_{\mathsf{Sol}} + \beta_2 \cdot \xi^T_{\mathsf{Orb}} + \beta_3 \cdot \xi^T_{\mathsf{Volc}} + \beta_4 \cdot \xi^T_{\mathsf{Land\,(anthr)}} + \beta_5 \cdot \xi^T_{\mathsf{Ghg\,(anthr)}} + \beta_6 \cdot \xi^T_{\mathsf{interact}} + \tilde{\nu} \end{aligned} \tag{2.6}$$

where $\tilde{\nu} = \tilde{\zeta}^T_{\mathsf{ALL}} + \eta^T_{\mathsf{internal}}$ is independent of ξ^T_{comb}, and of each individual ξ^T.

Next, the relation between the individual temperature responses in (2.6) needs to be discussed. 'Causal' independence entails that the variables in question are related to each other through correlations. As motivated earlier in this section, $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$ might be correlated to each other, but not to ξ^T_{Sol}, ξ^T_{Orb}, and ξ^T_{Volc}. Concerning the last three temperature responses, we argue that they are rather mutually uncorrelated than correlated. This is because the forcings causing them are acting on different time scales and with different character of their temporal evolutions. It is thus reasonable to expect that their temperature responses will not demonstrate a more or less similar shape, i.e. a temporal pattern. On the other hand, we found it difficult to hypothesise zero-correlations between $\xi^T_{\mathsf{interact}}$ and the 'causally' independent temperature responses. Thus $\xi^T_{\mathsf{interact}}$ is allowed to be correlated with $\xi^T_{\mathsf{Land\,(anthr)}}$, $\xi^T_{\mathsf{Ghg\,(anthr)}}$, ξ^T_{Sol}, ξ^T_{Orb}, and ξ^T_{Volc}. All these assumptions about correlations are reflected in a path diagram plotted in Figure 3.

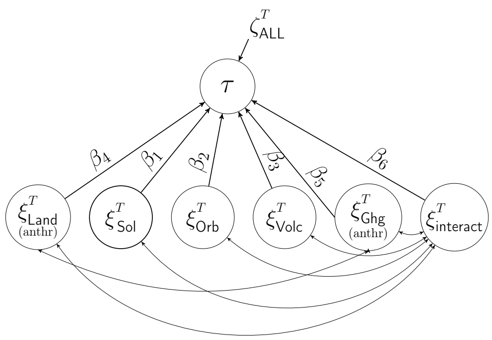

Figure 3. Path diagram for Eq. (2.6) associated with Scheme 1.

2.2 Scheme 2: *both 'causally' dependent and 'causally' independent temperature responses are involved*

Scheme 2 arises when causal inputs to ξ^T_{Land} and ξ^T_{Ghg} from ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc} and/or $\xi^T_{\mathsf{interact}}$ are allowed. This in turn permits us to relax the assumption that the effect of natural changes in vegetation and in the levels of Ghg:s on the temperature is negligible. Consequently, ξ^T_{Land} and ξ^T_{Ghg} under Scheme 2 represent the overall two-component temperature responses. Recall also from the earlier discussion that they are allowed to be 'causally' dependent not only on ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, $\xi^T_{\mathsf{interact}}$ but also on each other either reciprocally or unidirectionally. Statistically, causal dependence implies that the relations between such variables are modelled by means of (linear) regression models.

Clearly, to express the above relationships, one equation for τ is not sufficient: a multiequation model is needed. Indeed, expressing ξ^T_{Land} and ξ^T_{Ghg} as a linear function of ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, $\xi^T_{\mathsf{interact}}$ and of each other leads to the following nonrecursive, i.e. with reciprocal loops (see Appendix A1), system of equations:

$$\begin{aligned}\tau = \beta_1 \cdot \xi^T_{\mathsf{Sol}} + \beta_2 \cdot \xi^T_{\mathsf{Orb}} + \beta_3 \cdot \xi^T_{\mathsf{Volc}} + \beta_4 \cdot \xi^T_{\mathsf{Land}} + \\ + \beta_5 \cdot \xi^T_{\mathsf{Ghg}} + \beta_6 \cdot \xi^T_{\mathsf{interact}} + \zeta^T_{\mathsf{ALL}} \end{aligned} \qquad (2.7)$$

$$\xi^T_{\mathsf{Land}} = a_1 \cdot \xi^T_{\mathsf{Sol}} + a_2 \cdot \xi^T_{\mathsf{Orb}} + a_3 \cdot \xi^T_{\mathsf{Volc}} + a_4 \cdot \xi^T_{\mathsf{interact}} + c_1 \cdot \xi^T_{\mathsf{Ghg}} + \xi^T_{\mathsf{Land\,(anthr)}} \qquad (2.8)$$

$$\xi^T_{\mathsf{Ghg}} = b_1 \cdot \xi^T_{\mathsf{Sol}} + b_2 \cdot \xi^T_{\mathsf{Orb}} + b_3 \cdot \xi^T_{\mathsf{Volc}} + b_4 \cdot \xi^T_{\mathsf{interact}} + c_2 \cdot \xi^T_{\mathsf{Land}} + \xi^T_{\mathsf{Ghg\,(anthr)}}. \qquad (2.9)$$

where $\zeta^T_{\mathsf{ALL}} = \tilde{\zeta}^T_{\mathsf{ALL}} + \eta^T_{\mathsf{internal}}$. Notice that, although the same notations are used, ζ^T_{ALL} in (2.7) differs from ζ^T_{ALL} in (2.6) because under Scheme 1 the natural components are modelled as a part of ζ^T_{ALL}, whereas under Scheme 2 they are not.

Another important remark about Eq. (2.8)-(2.9) is that $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$ are considered as disturbance terms (or equivalently, errors in equations), i.e. terms contributing to the temperature variability randomly. Although disturbance terms are by definition 'causally' independent variables, which $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$ are, treating these temperature responses as disturbance terms obviously prevents us from analysing statistically possible systematic effects of the anthropogenic changes in vegetation and in the levels of Ghg:s on the temperature.

This is a direct implication of treating ξ^T_{Land} and ξ^T_{Ghg} as joint temperature responses, whose simulated counterparts are also assumed to be joint. The latter originates from our assumption that under Scheme 2 climate model simulations driven by the Land(natur)-, Land(anthr)-, Ghg(natur)- and Ghg(anthr)-forcings separately are not available. Instead, there are climate model simulations driven by the sum of natural and anthropogenic Land and Ghg, respectively. Given this limitation, it is not possible to model the four corresponding temperature responses as latent factors, and thus it is not possible to estimate coefficients associated with these latent factors. Instead, the contribution of anthropogenic changes in vegetation and in the levels of Ghg:s to the variability of the temperature can be assessed by judging the significance of the variance of the corresponding disturbance terms.

Regarding the structure of ξ^T_{Land} and ξ^T_{Ghg}, it should be pointed out that Scheme 2 represents a general situation subsuming other situations as special cases. We do not exclude that depending on the availability of climate model simulations, one of the two-component temperature responses can be modelled as one-component, while the other remains two-component. Such a situation would require the mixing of Scheme 1 and Scheme 2. Later, in Sec. 3.2.1, we consider one special case when only ξ^T_{Ghg} is modelled as a two-component temperature response, and we shall see how it changes the structure of the structural equation model associated with Scheme 2.

In the terminology of structural equation modelling, regression equations in (2.7)-(2.9) are called *structural equations*, where the term "structural" stands for the assumption that the regression coefficients (in this context also called structural) are not just descriptive measures of association but rather that they reveal an invariant causal relation. A graphical representation of Eq. (2.7)- (2.9) is given in Figure 4. The figure also reflects the fact that the assumptions concerning the correlatedness between the latent exogenous variables remain the same as under Scheme 1 (compare to Figure 3) except that the correlations between $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{interact}}$ and between $\xi^T_{\mathsf{Ghg\,(anthr)}}$ and $\xi^T_{\mathsf{interact}}$ are set to zero. This is done in order to meet the basic assumption of our statistical models that latent factors are uncorrelated with disturbance terms.

Other conceivable paths that climatologists may wish to add to Figure 4 are the paths from τ to ξ^T_{Land} and to ξ^T_{Ghg}. From the climatological perspective, this would allow us to reflect the idea that the changing climate itself can be a cause of subsequent changes in the Land and Ghg forcings. Notice that adding τ from (2.7) into Eq. (2.9)-(2.8) entails the addition of the ζ^T_{ALL} in these equations. Since ζ^T_{ALL}, defined in (2.6), comprises the residual forced variability and the internal variability due to the internal factors, freeing the paths from τ to ξ^T_{Land} and to ξ^T_{Ghg} corresponds to allowing even the excluded forcings and the internal factors be possible contributors of natural changes in the Land and Ghg forcings. Note that we are still assuming that the forced and unforced components of τ are independent.

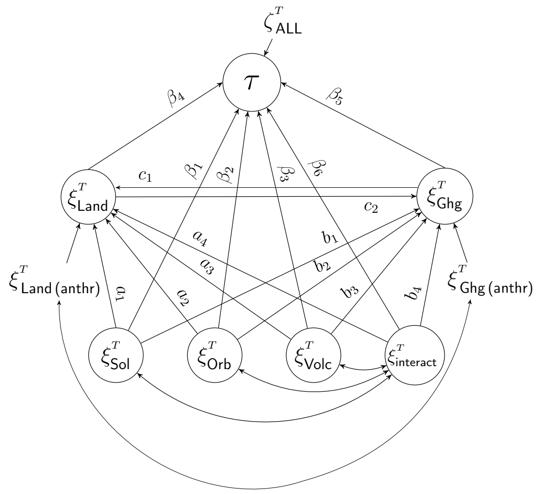

Figure 4. Path diagram of cause-effect relationships between the true temperature responses to the five selected forcings, represented in Eq. (2.7)-(2.9).

Up to now, we have discussed possible ways of relating only the *true* latent forcing effects to each other. The next step is to involve the simulated climate system to enable addressing the question of interest, i.e. the evaluation of climate model simulations against climate proxy and instrumental records of the near-surface temperature for the last millennium. Just as in Part I, this can be done by applying the concept of common factors, that is, factors common for the real-world latent temperature responses and their simulated counterparts. In the next section, we will demonstrate this process and describe the statistical models associated with each structure.

As mentioned earlier in the Introduction, the first scheme is associated with a factor model, while the second, involving causal links between latent variables, requires a *structural equation model* (SEM). A general description of a factor model was given in Appendix A in Part I, while a general definition of a structural equation model can be found in Appendix A here. We conclude this section by pointing out that a general factor model is a special case of a general SEM, which implies that the issues of estimation, hypothesis testing, identifiability, and model evaluation for SEM parallel those associated with factor analysis.

3 Statistical models involving both true and simulated temperature responses: moving from factor models to structural equation models

Let x_{comb} denote a time series of simulated temperatures generated by a climate model driven by a combination of reconstructed forcings, sampled over the same spatial and temporal domain that is represented by the true temperature τ. Analogously to τ, the *mean-centered* x_{comb} can also be decomposed into the forced and unforced components:

$$x_{\mathsf{comb}} = \xi^S_{\mathsf{comb}} + \tilde{\delta}_{\mathsf{comb}} \,, \tag{3.1}$$

where

ξ^S_{comb} - the fixed Simulated overall temperature response to reconstructed forcings in question,

$\tilde{\delta}_{\mathsf{comb}}$ - the simulated internal random temperature variability, *including* any random variability due to the presence of the forcings.

Note that if a combination of forcings is represented by only one forcing, i.e. comb $\equiv$ single forcing, the definition from (3.1) is applicable even to simulated temperatures generated by single-forcing climate models.

3.1 Statistical model under Scheme 1: a factor model

Although we have already demonstrated in Part I the process of formulating factor models, let us, for the convienience of the readers, repeat the main steps of this process.

The first step is to express x_{comb} and τ as (linear) functions of common factors, which are the true temperature responses to the forcings under consideration. Under Scheme 1, they are: ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, $\xi^T_{\mathsf{Land\,(anthr)}}$, $\xi^T_{\mathsf{Ghg\,(anthr)}}$, and $\xi^T_{\mathsf{interact}}$. As a matter of fact, τ is already represented as a linear function of these temperature responses in Eq. (2.6). Nevertheless, we repeat the same equation, but with the coefficients, used in our statistical models. To write ξ^S_{comb} as a linear function of the common factors, the latter are to be extracted from ξ^S_{comb}, which yields:

$$\begin{aligned}
\tau &= \{(2.6)\} = \xi^T_{\mathsf{comb}} + \tilde{\nu} = \\
&= \text{Strue} \cdot \xi^T_{\mathsf{Sol}} + \text{Otrue} \cdot \xi^T_{\mathsf{Orb}} + \text{Vtrue} \cdot \xi^T_{\mathsf{Volc}} + \text{Ltrue} \cdot \xi^T_{\mathsf{Land\,(anthr)}} + \\
&\quad + \text{Gtrue} \cdot \xi^T_{\mathsf{Ghg\,(anthr)}} + \text{Itrue} \cdot \xi^T_{\mathsf{interact}} + +\tilde{\nu} \,, \qquad (3.2)\\
x_{\mathsf{comb}} &= \xi^S_{\mathsf{comb}} + \tilde{\delta}_{\mathsf{comb}} = \\
&= \text{Ssim} \cdot \xi^T_{\mathsf{Sol}} + \text{Osim} \cdot \xi^T_{\mathsf{Orb}} + \text{Vsim} \cdot \xi^T_{\mathsf{Volc}} + \text{Lsim} \cdot \xi^T_{\mathsf{Land\,(anthr)}} + \\
&\quad + \text{Gsim} \cdot \xi^T_{\mathsf{Ghg\,(anthr)}} + \text{Isim} \cdot \xi^T_{\mathsf{interact}} + \underbrace{\tilde{\zeta}^S_{\mathsf{comb}} + \tilde{\delta}_{\mathsf{comb}}}_{=\delta_{\mathsf{comb}}} \qquad (3.3)
\end{aligned}$$

where

1. $\tilde{\zeta}^S_{\mathsf{comb}}$ represent the residual part of ξ^S_{comb}, which remains after extracting the common factors from ξ^S_{comb}. This residual term is assumed to be independent of all common factors, $\tilde{\delta}_{\mathsf{comb}}$, and of $\tilde{\nu}$. Hence, $\tilde{\nu}$ and $\delta_{\mathsf{comb}} = \tilde{\zeta}^S_{\mathsf{comb}} + \tilde{\delta}_{\mathsf{comb}}$ are mutually independent. and independent of each common factor.

2. The coefficients (Ssim, Osim, ..., Itrue) are standardised partial coefficients (or

factor loadings). They are standardised because the variances of all common factors are standardised to have a unit variance. That is, we are talking about changes measured in standard deviation units. Standardised coefficients are particularly useful when comparisons are to be made across different variables. It makes it easier to judge the relative importance of latent variables.

Analogously, we decompose the single-forcing simulated temperatures, assumed to be available (for example, just as in [29]):

$$
\begin{aligned}
x_{\mathsf{Sol}} &= \xi^{S}_{\mathsf{Sol}} + \tilde{\delta}_{\mathsf{Sol}} = \mathrm{Ssim} \cdot \xi^{T}_{\mathsf{Sol}} + \underbrace{(\tilde{\zeta}^{S}_{\mathsf{Sol}} + \tilde{\delta}_{\mathsf{Sol}})}_{=\delta_{\mathsf{Sol}}}, \\
x_{\mathsf{Orb}} &= \xi^{S}_{\mathsf{Orb}} + \tilde{\delta}_{\mathsf{Orb}} = \mathrm{Osim} \cdot \xi^{T}_{\mathsf{Orb}} + \underbrace{(\tilde{\zeta}^{S}_{\mathsf{Orb}} + \tilde{\delta}_{\mathsf{Orb}})}_{=\delta_{\mathsf{Orb}}}, \\
x_{\mathsf{Volc}} &= \xi^{S}_{\mathsf{Volc}} + \tilde{\delta}_{\mathsf{Volc}} = \mathrm{Vsim} \cdot \xi^{T}_{\mathsf{Volc}} + \underbrace{(\tilde{\zeta}^{S}_{\mathsf{Volc}} + \tilde{\delta}_{\mathsf{Volc}})}_{=\delta_{\mathsf{Volc}}}, \\
x_{\mathsf{Land}} &= \xi^{S}_{\mathsf{Land}} + \tilde{\delta}_{\mathsf{Land}} = \{\text{under Scheme 1}\} = \\
&= \xi^{S}_{\mathsf{Land\,(anthr)}} + \tilde{\delta}_{\mathsf{Land}} = \mathrm{Lsim} \cdot \xi^{T}_{\mathsf{Land\,(anthr)}} + \underbrace{\tilde{\zeta}^{S}_{\mathsf{Land}} + \tilde{\delta}_{\mathsf{Land}}}_{=\delta_{\mathsf{Land}}}, \\
x_{\mathsf{Ghg}} &= \xi^{S}_{\mathsf{Ghg}} + \tilde{\delta}_{\mathsf{Ghg}} = \{\text{under Scheme 1}\} = \\
&= \xi^{S}_{\mathsf{Ghg\,(anthr)}} + \tilde{\delta}_{\mathsf{Ghg}} = \mathrm{Gsim} \cdot \xi^{T}_{\mathsf{Ghg\,(anthr)}} + \underbrace{\tilde{\zeta}^{S}_{\mathsf{Ghg}} + \tilde{\delta}_{\mathsf{Ghg}}}_{=\delta_{\mathsf{Ghg}}}.
\end{aligned}
\tag{3.4}
$$

Further, on comparing (3.4) with (3.3) one notes that $\xi^{T}_{\mathsf{single\ forcing}}$ is expected to have equal (direct) influence (or contribution, which might be a more suitable notion in the climatic perspective) on the associated single-forcing simulation and on the multi-forcing simulation x_{comb}. To exemplify, the (direct) influence of ξ^{T}_{Sol} on x_{Sol} and x_{comb} is represented by Ssim. This can be justified only under the condition that the same reconstruction and implementation of a single forcing in question has been employed to generate $x_{\mathsf{single\ forcing}}$ and x_{comb} and, of course, that the same climate model is used in both cases.

The second step is to replace the unobservable τ by observational data, v, consisting of instrumental data when available and/or temperature reconstructions from proxies (see also Eq. (4.1.3) in Part I, Sec. 4.1). Replacing τ in (3.2) by v leads to

$$
\begin{aligned}
v = \mathrm{Strue} \cdot \xi^{T}_{\mathsf{Sol}} + \mathrm{Otrue} \cdot \xi^{T}_{\mathsf{Orb}} + \mathrm{Vtrue} \cdot \xi^{T}_{\mathsf{Volc}} + \mathrm{Ltrue} \cdot \xi^{T}_{\mathsf{Land\,(anthr)}} + \\
+ \mathrm{Gtrue} \cdot \xi^{T}_{\mathsf{Ghg\,(anthr)}} + \mathrm{Itrue} \cdot \xi^{T}_{\mathsf{interact}} + \nu
\end{aligned}
\tag{3.5}
$$

where ν is the sum of $\tilde{\nu}$ from (3.2) and the residual non-climatic variation. Just as in Part I, the latter is assumed to be uncorrelated with τ, and it is also, in the context of this article, assumed to have constant variance, implying that the variance of ν, σ^2_{ν}, is constant.

The third step is to combine the equations for observed and simulated temperatures in a factor model. Combining (3.3), (3.2) and (3.4) leads to a 7-indicator 6-factor model, abbr. FA(7,6), presented in Table 2. As follows from this table, a priori knowledge of the specific-factor variances, associated with simulations, is required, otherwise the model is underidentified. A possible estimator of σ^2_δ, based on the availability of ensembles, can be found in Appendix B in Part I.

Availability of ensembles allows us also to analyse ensemble-mean sequences instead of single members of ensembles. As known (e.g. [5]), averaging over replicates of the same type of forced model leads to a time series with an enhanced forced climate signal and a reduced effect of the internal variability of the corresponding forced climate model. The use of mean-sequences requires replacing the specific-factor variances $\sigma^{2\,*}_{\delta_{\mathsf{f}i}}$ by $\sigma^{2\,*}_{\delta_{\mathsf{f}i}}/k_{\mathsf{f}i}$, where $k_{\mathsf{f}i}$ is the number of replicates in the associated ensemble.

However, as discussed in Part I, a disadvantage of using the suggested independent estimator of σ^2_δ is that this estimator estimates the variance of $\tilde{\delta}$, not the variance of δ. The latter, according to (3.3) and (3.4), is the sum of the residual term $\tilde{\zeta}^S$ and $\tilde{\delta}$. If the variance of $\tilde{\zeta}^S$ is not negligible, the use of this estimator might lead to the biasedness of some parameter estimates. Despite this, the factor model in Table 2 is to be evaluated under the assumption of the negligibility of the variance of $\tilde{\zeta}^S$, because freeing up the δ-factor variances would lead to underidentifiability.

In Part I, we suggested to use replicates of each single-forcing climate model as additional indicators in order to investigate whether this assumption is appropriate for single-forcing simulations (see for example model (4.1.10) in Part I, Sec. 4.1). A similar procedure can be applied even to the factor model presented in Table 2, or to its final version. As a further comment on this factor model, let us note that although all indicators in the model are assumed to be constructed by averaging over replicates. we do not use the bar notation to designate the mean sequences.

Table 2. Parameters of the 7-indicator 6-factor model, abbr. FA(7,6).

Indicator			**Common**	**factors**			**Specific-**
	factor 1	factor 2	factor 3	factor 4	factor 5	factor 6	**-factor**
	$\boldsymbol{\xi}^T_{\mathsf{Sol}}$	$\boldsymbol{\xi}^T_{\mathsf{Orb}}$	$\boldsymbol{\xi}^T_{\mathsf{Volc}}$	$\boldsymbol{\xi}^T_{\mathsf{Land\,(anthr)}}$	$\boldsymbol{\xi}^T_{\mathsf{Ghg\,(anthr)}}$	$\boldsymbol{\xi}^T_{\mathsf{interact}}$	**variances**
1. x_{Sol}	Ssim	0	0	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{Sol}}}$
2. x_{Orb}	0	Osim	0	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{Orb}}}$
3. x_{Volc}	0	0	Vsim	0	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{Volc}}}$
4. x_{Land}	0	0	0	Lsim	0	0	$\sigma^{2\,*}_{\delta_{\mathsf{Land}}}$
5. x_{Ghg}	0	0	0	0	Gsim	0	$\sigma^{2\,*}_{\delta_{\mathsf{Ghg}}}$
6. x_{comb}	Ssim	Osim	Vsim	Lsim	Gsim	Isim	$\sigma^{2\,*}_{\delta_{\mathsf{comb}}}$
7. v	Strue	Otrue	Vtrue	Ltrue	Gtrue	Itrue	σ^2_ν

Correlations among Common Factors

1	0	0	0	0	ϕ_{SI}
	1	0	0	0	ϕ_{OI}
		1	0	0	ϕ_{VI}
			1	ϕ_{LG}	ϕ_{LI}
				1	ϕ_{GI}
					1

* the parameter assumed to be known a priori, i.e. estimated independently.

As pointed out by [26], free parameters, i.e. parameters to be estimated, are not associated with hypotheses because nothing is specified by freeing the parameter, meaning that no restriction(-s), imposed on the implied variance-covariance matrix of the indicators[3], is associated with this parameter. The estimated value of the parameter may turn out to be negative, positive, or zero! Nevertheless, the sign and strength of parameter estimates are important aspects for judging how reasonable numerical results are. If estimates cannot be linked to (in our case climatological) properties of latent factors, then the model can hardly be accepted as a good approximation of the underlying latent structure. By taking into consideration such aspects like

- the time period, time unit, seasons,
- region and its size,
- our knowledge about the real-world forcings,
- the properties of the reconstructions of forcings used to generate climate model simulations,
- results from previous studies,

researchers can arrive at different conceptions about expected magnitudes of the estimates of factor loadings. For example, it seems to be reasonable to expect that the influence of the anthropogenic land use forcing in Antarctica during the last millennium prior to the industrialisation period is negligible. So it would be difficult to accept a numerical result leading to the opposite conclusion.

When discussing the expected signs of the factor loadings, such properties of the forcings like positiveness/negativeness can be added to the above-mentioned aspects. For example, consider orbital forcing. In the summer of the northern hemisphere, this forcing is associated with a negative trend in incoming solar radiation throughout the millennium, while the corresponding trend during the summer of the southern hemisphere is positive. That would motivate letting Osim be negative if we study summer temperatures in Europe but positive if we study summer temperatures in Australia.

What is important to keep in mind when determining the expected sign is that the solution remains unique even if the observed sign is changed to an opposite one

[3]Recall from Part I that the basic idea of confirmatory factor analysis is that the population variance-covariance matrix of the indicators, $\mathbf{\Sigma}$, can be represented as a function of the model parameters $\boldsymbol{\theta}$. The resulting matrix, denoted $\mathbf{\Sigma}(\boldsymbol{\theta})$, is called the implied (or model's reproduced) variance-covariance matrix of the indicators.

in accordance with substantive justifications. In general, a sign change corresponds merely to changing the sign of the factor, which, however, might require a sign change of other parameters associated with this factor. In our factor model, other parameters are correlations.

Regarding correlations among the latent factors, caution is needed when too high estimates are observed, say over 0.8 in absolute value. This is because (1) a high correlation means that two temperature responses are almost proportional, which is difficult to interpret physically, and (2) it can in effect indicate problems with identifiability rather than two temperature responses being correlated.

Under the assumptions that the specific factors are uncorrelated and their variances can be estimated a priori, the FA(7,6)-model in Table 2 is (over-)identified with 11 degrees of freedom. Nevertheless, setting only Isim to zero makes the associated correlation coefficients underidentified, i.e. each of them can take on any real value without changing the variance-covariance matrix of the observed variables. So when one wishes to test whether the interaction effect is negligible or not, it is necessary to eliminate all correlation coefficients associated with $\xi^T_{\mathsf{interact}}$ from the vector of the model parameters. This increases the degrees of freedom to 18.

The hypothesis of main interest within our analysis, i.e. the hypothesis of consistency between the latent simulated and true temperature responses, is tested by imposing the following six equality constraints: Ssim=Strue, Osim=Otrue, Vsim=Otrue, Lsim=Ltrue, Gsim=Gtrue, and Isim=Itrue. This gives us six additional degrees of freedom: one degree of freedom for each equality constraint. It is also possible to introduce only some subset of these equality constraints, which, however, reduces the degrees of freedom accordingly.

We do not discuss in detail all possible models nested within the least restricted FA(7,6)-model because the way of reasoning is similar to that associated with the FA(5,4)-model, given in Part I (see Table 4). In addition, the FA(7,6)-model is analysed practically in Part III (see [10]) so more details can be found there.

Prior to moving on to the discussion about a structural equation model, arising under Scheme 2, we summarise the FA(7,6)-model graphically by means of a path diagram (see Figure 5), which might contribute to a better understanding of differences and similarities between these two statistical models.

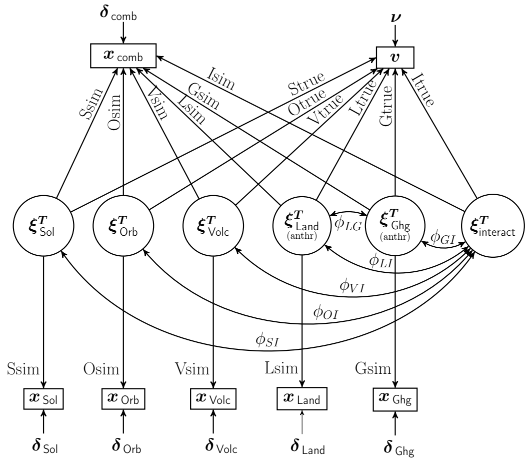

Figure 5. Path diagram describing the relationships among the latent common temperature responses under the assumption of their mutual *causal* independence.

3.2 Statistical model under Scheme 2: a structural equation model (SEM)

Under Scheme 2, the general underlying structure of ξ^T_{comb} and ξ^S_{comb} includes the following common factors: ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc}, the two-component factors ξ^T_{Land} and ξ^T_{Ghg}, and, finally, $\xi^T_{\mathsf{interact}}$. Rewriting Eq. (2.8) and (2.9) with the coefficients, used in our statistical models, we get

$$\begin{aligned}
\xi^T_{\mathsf{Land}} &= \xi^T_{\mathsf{Land\,(natural)}} + \xi^T_{\mathsf{Land\,(anthr)}} = \\
&= \mathrm{SL}\cdot\xi^T_{\mathsf{Sol}} + \mathrm{OL}\cdot\xi^T_{\mathsf{Orb}} + \mathrm{VL}\cdot\xi^T_{\mathsf{Volc}} + \mathrm{IL}\cdot\xi^T_{\mathsf{interact}} + \mathrm{GL}\cdot\xi^T_{\mathsf{Ghg}} + \xi^T_{\mathsf{Land\,(anthr)}} \\
\xi^T_{\mathsf{Ghg}} &= \xi^T_{\mathsf{Ghg\,(natural)}} + \xi^T_{\mathsf{Ghg\,(anthr)}} = &(3.6)\\
&= \mathrm{SG}\cdot\xi^T_{\mathsf{Sol}} + \mathrm{OG}\cdot\xi^T_{\mathsf{Orb}} + \mathrm{VG}\cdot\xi^T_{\mathsf{Volc}} + \mathrm{IG}\cdot\xi^T_{\mathsf{interact}} + \mathrm{LG}\cdot\xi^T_{\mathsf{Land}} + \xi^T_{\mathsf{Ghg\,(anthr)}} &(3.7)
\end{aligned}$$

In the equations above, the common factors ξ^T_{Sol}, ξ^T_{Orb}, ξ^T_{Volc} and $\xi^T_{\mathsf{interact}}$ are standardised to have a unit variance. The common factors $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$, because of being modelled as disturbance terms [4,5], have unstandardised variances, and coefficients fixed to 1,00. Based on significance of the estimates of these variances, conclusions about the contribution of the anthropogenic changes in Ghg:s to the temperature variability can be drawn.

As for ξ^T_{Land} and ξ^T_{Ghg}, their variances cannot be standardised either, which is due to

[4] The reason for this was discussed earlier in connection with Eq. (2.8)-(2.9).

[5] It should be remarked that modelling $\xi^T_{\mathsf{Land\,(anthr)}}$ and $\xi^T_{\mathsf{Ghg\,(anthr)}}$ as disturbance terms should not be interpreted as that their effects on the temperature are necessarily smaller than the effects of the natural forcings. Recall from Part I that the effect of the internal climate processes on the temperature variability is also modelled as a disturbance term not only in our analysis but also in measurement models used in many D&A studies. Nevertheless, this does not preclude that the contribution of internal factors to the temperature variability might be strong.

the fact that the variances of latent endogenous variables in structural equation models are not model parameters (but can be calculated afterwards, see (A1.2) and (A3.7) in Appendix). Therefore, following the theory of structural equation modelling, measurement scales of these latent factors ought be established by fixing the coefficients for ξ^T_{Land} and ξ^T_{Ghg} to 1,00 in relation to one of their indicators, which are: v, x_{comb}, x_{Land} or x_{Ghg}. To be able to analyse only simulations without involving the climate record v, it was decided to fix the coefficients in relation to x_{comb}. This immediately implies that the coefficients are to be fixed to 1,00 in relation to the single-forcing indicators as well, because the same reconstruction and implementation of a given reconstructed forcing is employed both in the multi-forcing climate model and in the associated single-forcing climate model.

Based on the discussion above, the following expressions for v, x_{comb}, x_{Land}, and x_{Ghg} are obtained:

$$
\begin{aligned}
v = \mathrm{Strue} \cdot \xi^T_{\mathsf{Sol}} + \mathrm{Otrue} \cdot \xi^T_{\mathsf{Orb}} + \mathrm{Vtrue} \cdot \xi^T_{\mathsf{Volc}} + \mathrm{Ltrue} \cdot \xi^T_{\mathsf{Land}} + \\
+ \mathrm{Gtrue} \cdot \xi^T_{\mathsf{Ghg}} + \mathrm{Itrue} \cdot \xi^T_{\mathsf{interact}} + \nu,
\end{aligned}
\tag{3.8}
$$

(recall from Sec. 3.1 that within the present work ν is assumed to have a constant variance, denoted σ^2_ν),

$$
\begin{aligned}
x_{\mathsf{comb}} = \xi^S_{\mathsf{comb}} + \tilde{\delta}_{\mathsf{comb}} = \\
= \mathrm{Ssim} \cdot \xi^T_{\mathsf{Sol}} + \mathrm{Osim} \cdot \xi^T_{\mathsf{Orb}} + \mathrm{Vsim} \cdot \xi^T_{\mathsf{Volc}} + \xi^T_{\mathsf{Land}} + \\
+ \xi^T_{\mathsf{Ghg}} + \mathrm{Isim} \cdot \xi^T_{\mathsf{interact}} + \underbrace{\tilde{\zeta}^S + \tilde{\delta}_{\mathsf{comb}}}_{=\delta_{\mathsf{comb}}}.
\end{aligned}
\tag{3.9}
$$

$$
x_{\mathsf{Land}} = \xi^S_{\mathsf{Land}} + \tilde{\delta}_{\mathsf{Land}} = \xi^T_{\mathsf{Land}} + \underbrace{(\tilde{\zeta}^S_{\mathsf{Land}} + \tilde{\delta}_{\mathsf{Land}})}_{=\delta_{\mathsf{Land}}}
\tag{3.10}
$$

$$
x_{\mathsf{Ghg}} = \xi^S_{\mathsf{Ghg}} + \tilde{\delta}_{\mathsf{Ghg}} = \xi^T_{\mathsf{Ghg}} + \underbrace{(\tilde{\zeta}^S_{\mathsf{Ghg}} + \tilde{\delta}_{\mathsf{Ghg}})}_{=\delta_{\mathsf{Ghg}}}.
\tag{3.11}
$$

The expressions for x_{Sol}, x_{Orb}, and x_{Volc} remain the same as under Scheme 1.

Prior to combining Eq. (3.6)-(3.11) into a structural equation model, it is important to stress that v and x_{comb} are *common* indicators for the latent exogenous and latent endogenous variables. However, according to the definition of a general structural equation model (see Appendix A), the sets of indicators for latent exogenous and latent endogenous variables should be disjoint [6]. One way to extricate ourselves from this difficulty is to regard v and x_{comb} as latent variables. This is achieved by introducing two 'new' variables, say x^+_{comb} and v^+, such that $x^+_{\mathsf{comb}} = x_{\mathsf{comb}}$ and $v^+ = v$, meaning that

[6] Note that this requirement is completely satisfied when only single-forcing simulations are analysed. However, the disadvantages of such an analysis is that inferences about interactions between the forcings and about the ability of involved climate models to simulate the observed climate change are not possible.

the disturbance variances of x^{+}_{comb} and v^{+} are set to zero. Hence, x^{+}_{comb} and x_{comb} are the same variable. Analogously, v^{+} and v are the same variable. The resulting structural equation model is graphically presented in Figure 6, while the associated equations are summarised in Appendix B.

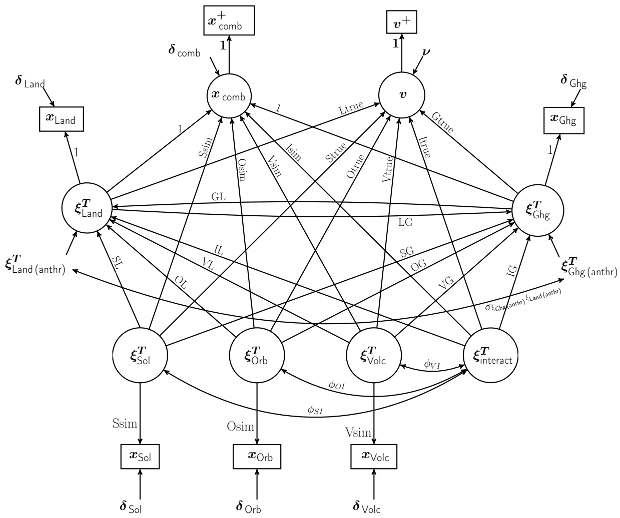

Figure 6. Path diagram for a nonrecursive Structural Equation Model arising under Scheme 2.

We use this model as a point of departure for constructing different SEM models, meaning that the model can be modified in many ways by deleting some of the depicted paths or by adding new ones. For example, as discussed earlier, x_{comb} and/or v can influence back the endogenous ξ_{Land} and/or ξ_{Ghg}, which amounts to expressing the idea that the changing climate itself can be a cause of subsequent changes. In other words, it permits us to reflect the idea that not only the forcings under consideration but also the excluded forcings and/or the internal factors may contribute to natural changes in vegetation and in the levels of Ghg:s.

From the perspective of structural equation modelling, freeing paths from observed variables to latent ones entails the movement from the general standard representation of a SEM model to its alternative representation. More details about these two representations and their connection to each other can be found in Appendix A.

If an initial SEM model demonstrates a reasonable fit [7], model simplifications might be of more interest than model expansions. On the other hand, if the initial model has a bad fit, introducing additional relationships might improve the fit. Useful means in providing clues to specific model modification are standardised (normalised) residuals [8],

[7] See Appendix A in Part I for descriptions of the χ^2 test and heuristic goodness-of-fit indices that can be used as measures of the overall model fit to the data

[8] Residuals are elements in the residual matrix, defined as the difference between $\boldsymbol{S}$ and the estimated reproduced variance-covariance matrix $\boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$, $\boldsymbol{S} - \boldsymbol{\Sigma}(\widehat{\boldsymbol{\theta}})$. A standardised residual is a fitted residual divided by its asymptotic standard error. Normalised residuals that exceed 1.96 or 2.58 in absolute

and the modification indices [9]. Importantly, introducing additional paths should be done judiciously because we do not want to free too many parameters because it leads to a decrease in the degrees of freedom.

The hypothesis of consistency is tested in a similar way as under the factor model arising under Scheme 1, that is, by testing whether the direct effects of each latent factor on x_{comb} and on v are equal or not. In terms of parameters, it requires imposing the following equality constraints: Strue=Ssim, Otrue=Osim, Vtrue=Vsim, Itrue=Isim, Ltrue=1, and Gtrue=1.

In addition to direct effects, one can also analyse indirect and total effects of latent variables. These questions are of interest even within our analysis both from the climatological and statistical perspectives. Nevertheless, we focus only on analysing direct effects. To motivate it, let us remind that the starting point of the present work is to formulate SEM models that can be use as an alternative to ME models used in many D&A studies and to our factor models. That is, SEM models should be capable of addressing the same questions as those addressed by the above-mentioned models, namely
- to investigate whether a simulated latent temperature response to a given forcing is correctly represented in the climate model under consideration, compared to its real-world counterpart embedded in observations, and
- to investigate the magnitude of the influence of real-world forcings on the observed/reconstructed temperature.

Since addressing these questions requires the comparison of just direct effects, we refrain here from discussing the issue of estimating and interpreting indirect and total effects, which may call for a separate paper.

Turning our attention back to our SEM model in Figure 6, we would like once again to stress the importance of the issue of identifiability. It should be realised that this SEM model is underidentified despite the fact that the number of the nonduplicated equations in the variance-covarince matrix of the indicators expressed as a function of unknown parameters $\boldsymbol{\theta}$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, is larger than the number of unknown parameters, $28 > 27$. The easiest way to see underidentifiability is to note that $\mathrm{cov}(x_{\mathsf{Sol}}, x_{\mathsf{Orb}})$, $\mathrm{cov}(x_{\mathsf{Sol}}, x_{\mathsf{Volc}})$, and $\mathrm{cov}(x_{\mathsf{Orb}}, x_{\mathsf{Volc}})$ are zeros, meaning that 27 parameters are in effect to be determined from 25 equations. More restrictions on the model parameters are required to achieve identifiability.

For very simple versions of our SEM model (with one or two causal links not leading to reciprocal loops), identifiability may be established algebraically by solving structural

value are considered statistically significant at the significance level of 5% and 1%, respectively. Ideally, no more than 5% of normalised residuals should be greater than 1.96. Similary, no more than 1% should be greater than 2.58.

[9] Developed by [38], these indices attempt to estimate which missing paths, if added to the current model, would result in the greatest reduction of the discrepancy between model and data. The way to use these indices is to free the fixed parameter associated with the largest reduction and reanalyse the resulting model.

covariance equations $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$ for unknown parameters in $\boldsymbol{\theta}$. However, as the model increases in complexity, the attempts of establishing the model's identifiability algebraically are very likely to be error-prone and time-consuming. Given such a situation, researchers may resort to empirical tests for identifiability. One of them is the empirical test on the information matrix defined as the matrix of second-order derivatives of the discrepancy function used to estimate the model (see Eq. A.4 in Appendix in Part I). According to [20], "if the model is identified, the information matrix is almost certainly positive definite. If the information matrix is singular, the model is underidentified". The test is automatically calculated in all statistical packages developed to estimate structural equation models, for example, LISREL, EQS, and the *R* package sem.

The inverse of the information matrix provides an estimate of the variance-covariance matrix of the asymptotic distribution of the model estimates [10]. Examining this matrix is also helpful for revealing empirical underidentifiability. If the model is nearly underidentified, it will be reflected in high covariances between two or more parameter estimates.

According to [20], identifiability can also be checked by the following two-steps test. The first step is to analyse the sample variance-covariance matrix, $\boldsymbol{S}$, as usual and to save the predicted covariance matrix based on the estimates of the model parameters, i.e. $\mathbf{\Sigma}(\widehat{\boldsymbol{\theta}})$. Next, substitute $\mathbf{\Sigma}(\widehat{\boldsymbol{\theta}})$ for $\boldsymbol{S}$ and rerun the same program. If the model is identified, the new estimates should be identical to the first ones that were generated.

The fourth possible check for identifiability is to estimate the model with different starting values for free parameters in the iterative estimation algorithm to see whether or not the algorithm converges to the same parameter estimates each time. This empirical test, however, should be used with great care. Choosing inappropriate starting values may cause the failure of convergence although the model is theoretically identified.

Finally, modification indices [11] can be used to determine whether a parameter, which is held fixed in a model, will be identified if it is set free. If a modification index for a fixed parameter is not zero and *positive*, this indicates that this parameter will be identified if it is set free.

3.2.1 Mixing Scheme 1 and Scheme 2

It is obviously impossible within the confines of this article to go through all possible models derived by modifying the SEM model depicted in Figure 6. Nevertheless, it is worthwhile to mention SEM models that combine the features of Scheme 1 and Scheme 2, where only one of the endogenous variables, ξ^T_{Land} or ξ^T_{Ghg}, is a two-component variable.

[10] Recall from Part I that under the assumption of normality of data, the Maximum Likelihood estimates are consistent and jointly asymptotically normally distributed.

[11] Developed by [38], these indices attempt to estimate which missing paths, if added to the current SEM model, would result in the greatest reduction of the discrepancy between model and data. The way to use these indices is to free the fixed parameter associated with the largest reduction and reanalyse the resulting model.

Mixing the schemes can help us take the properties of the climate model under study into account. Depending on those properties, the schemes can be mixed in different ways. To exemplify, let us assume that it was justified that ξ^T_{Land} is to be modelled as a one-component temperature response comprising only $\xi^T_{\mathsf{Land\,(anthr)}}$. This seems to be a realistic situation because the currently used implementations of land use/land cover forcings (e.g. by [30] and [17]) in many climate model simulations, e.g. such as those by [29], only represent changes in vegetation that are due to changed human agricultural and pastoral activities without including dynamic natural changes in vegetation that may occur within the climate system. This means that the type of natural vegetation is prescribed in each grid cell and held constant (at pre-specified level).

The temperature response to the Ghg forcing, on the other hand, is still modelled as a joint two-component latent factor because prescribed reconstructed greenhouse gas concentrations, used in those simulations, are likely to contain information about both natural and anthropogenic influences. The resulting SEM model is depicted in Figure 7.

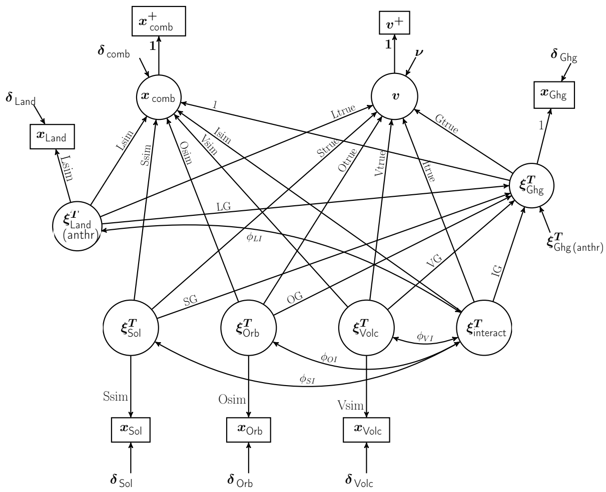

Figure 7. Path diagram for a Structural Equation Model arising when Scheme 1 and Scheme 2 are combined in such a way that ξ_{Ghg} remains a joint two-component temperature response, while ξ_{Land} is a one-component temperature response containing only $\xi_{\mathsf{Land\,(anthr)}}$.

Comparing the path diagrams in Figure 6 and 7, we can see the consequences of mixing the schemes. Since $\xi^T_{\mathsf{Land\,(anthr)}}$ is now a latent factor, i.e. not a disturbance term, it is allowed to be correlated with the interaction term, but not with $\xi^T_{\mathsf{Ghg\,(anthr)}}$, which is still modelled as a disturbance term. Further, the reciprocal relation between ξ^T_{Land} and ξ^T_{Ghg} is replaced by an unidirectional path from $\xi^T_{\mathsf{Land\,(anthr)}}$ to ξ^T_{Ghg}, which seems to be the only climatologically defensible way to relate these temperature responses to each other statistically without involving the discussion about the influence of the changing climate on human agricultural activity.

Obviously, if the influence of $\xi^T_{\mathsf{Land\,(anthr)}}$ on ξ^T_{Ghg} turns out to be statistically significant, this means that the natural component of the latter, $\xi^T_{\mathsf{Ghg\,(natur)}}$, is no longer of purely natural character. Strictly speaking, it was not such even under Scheme 2 because the influence of the joint interaction term, which can be of both natural and anthropogenic origin, may also turn out to be statistically significant.

Just as in the case of the SEM model from Figure 6, a prudent approach to analysing the SEM model from Figure 7 would be to start without involving observational data v. In case the associated SEM model is rejected, it does not make sense to proceed further with analysing the data with the observed climate record included. Further, it can be recommended to start with SEM models where the latent endogenous variables, ξ^T_{Land} and/or ξ^T_{Ghg}, are influenced by at most two 'causally' independent variables, for example, ξ^T_{Sol} and ξ^T_{Orb}.

It cannot be emphasised enough that the choice of a final or tentative model should not be made exclusively on the basis of statistical information - any modification ought to be defensible from the climatological point of view and reflect our knowledge about the real-world climate system and about the climate model under consideration. Another important aspect to highlight is that a final model obtained as a result of a purely data-driven modification process, should not be taken as a correct model. We can only say that "the model may be valid" because it does not contradict our assumptions and knowledge about the climate system.

4 Summary

In the present paper, two statistical approaches have been employed for formulating statistical models that can be used for evaluation of temperature data from forced climate model simulations against observational data for (approximately) the last millennium. The first approach is known as confirmatory factor analysis (CFA), while the second invokes structural equation modelling (SEM). Although closely related to each other (CFA is a special case of SEM), the approaches have distinct features.

One of the main differences is that CFA does not allow causal relationships between latent factors, while SEM does. As argued in Sec. 2, introducing causal links within our analysis is highly defensible from the climatological point of view because this permits us to describe to some extent some feedback mechanisms, e.g. vegetation-climate interactions. As a consequence, the associated latent temperature responses can be modelled as two-component responses, which in turn allows us to separate their variability due to natural and anthropogenic causes. Admittedly, statistical inferences about two-component variables might be not so straightforward as for single-component variables, but their use is definitely motivated for situations when separate reconstructions of natural respective anthropogenic changes in a given forcing are not available, implying that corresponding climate model simulations are not available either. Instead, there are climate model simulations driven by a reconstruction of the forcing, in which natural and anthropogenic changes are coupled together.

Of course, it is, in principle, possible to model the temperature response to a given "two-component" forcing as a single latent variable representing in that case the *overall* temperature response to this forcing. But this definitely would prevent us from getting any ideas about the magnitude of contributions of natural and anthropogenic changes in the forcing to the climate change, in particular, with respect to temperature. In addition, this also transforms the overall latent temperature response from an endogenous latent variable to an exogenous one, whose relation to other latent exogenous variables can be modelled exclusively in terms of correlations (just as in factor models, where all latent variables are exogenous). In turn, this may lead to difficulties with reflecting our knowledge about the relations and properties of forcings involved. Indeed, it might happen that one part of the overall temperature response can be motivated to be correlated with some other latent temperature responses, while the other part cannot.

Further, in factor analysis, observed variables cannot be viewed as causes of latent factors, thereby preventing us from expressing the idea that temperature changes can cause subsequent changes in some forcings, e.g. vegetation and/or the concentration of greenhouse gases in the atmosphere. In SEM models, this is possible.

At this point, we would like to emphasise that the discussion above should not be taken as an exhortation to refrain from applying factor models when evaluating climate model simulations or analysing them without involving observational data. We strongly recommend starting with considering an appropriate factor model. According to the principle of parsimony, it is always motivated to prefer a simpler model demonstrating an acceptable and adequate performance to a more complicated one. So if the factor model is not rejected then we have no reason to proceed with estimation of the associated SEM model. But if the factor model is rejected or if researchers have some doubts about the reliability of results obtained, then it becomes justified to move on to the SEM model.

Note also the general nature of our discussion above, although in the present work we have formulated a specific factor model (see Table 2 and Figure 5) with a specific corresponding SEM model (see Figure 6). They were formulated under condition that the climate model simulation to be evaluated is forced by five specific forcings, namely changes in solar radiation, changes in the orbital parameters for the Earth, changes in the amount of stratospheric aerosols of volcanic origin, *anthropogenic* changes in land use/land cover and changes in concentrations of greenhouse gases in the atmosphere of both natural and anthropogenic origin.

Confining our attention to this specific combination of forcings, our aim was to illustrate and exemplify a possible way of reasoning. Clearly, depending on the number of forcings, their climatological properties and the characteristics of climate model simulations to be evaluated, in particular, the characteristics of reconstructions of forcings, and, finally, depending on the availability of data, different factor models with different associated SEM models can be formulated.

The performance of the statistical models developed here, in Part II, is investigated and compared in a controlled numerical experiment, where the true temperature is replaced by temperature data from an appropriate climate model simulation. The results of this analysis is presented in Part III.

We conclude this paper by emphasising the fact that the statistical framework suggested here is the very first step in modelling more complex relationships between latent temperature responses than those associated with ME models used in many D&A studies and in factor models proposed in Part I ([9]). Therefore, we do not exclude further theoretical modifications/improvements of our framework as more experience and understanding of the problem will be gained.

Appendix A

A1

Structural Equation Model (SEM): a standard representation

A structural equation model consists of two submodels: a *latent variable model*, linking latent variables to each other, and a *measurement model*, linking latent variables to their indicators.

Submodel 1: Latent Variable Model

A structural equation for the latent variable model is given by ([2], [20]):

$$\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}, \tag{A1.1}$$

where

- $\boldsymbol{\eta}$ an $m \times 1$ vector of latent endogenous (dependent) variables, i.e. the variables that are determined within the model;
- $\boldsymbol{\xi}$ an $n \times 1$ vector of latent exogenous (independent) variables, i.e. the variables whose causes lie outside the model;
- $\boldsymbol{\zeta}$ an $m \times 1$ vector of latent errors in equations (random disturbance terms). Each ζ_i represents influences on η_i that are not included the structural equation for η_i;
- $\boldsymbol{B}$ an $m \times m$ matrix of coefficients, representing direct effects of η-variables on other η-variables. $\boldsymbol{B}$ always has zeros in the diagonal, which ensures that a variable is not an immediate cause of itself;

$\boldsymbol{\Gamma}$ an $m \times n$ matrix of coefficients, representing direct effects of ξ-variables on η-variables.

Further, model (A1.1) assumes that

- $\mathrm{E}(\boldsymbol{\eta}) = 0,\ \mathrm{E}(\boldsymbol{\xi}) = 0,\ \mathrm{E}(\boldsymbol{\zeta}) = 0,$
- $\boldsymbol{\zeta}$ is uncorrelated with $\boldsymbol{\xi}$ (otherwise, inconsistent coefficient estimators are likely).
- $\boldsymbol{I} - \boldsymbol{B}$ is nonsingular,
- $\zeta_{i\,t}$, $i = 1, 2, \ldots, m$, is homoscedastic and nonautocorrelated, meaning that the associated covariance matrix of $\boldsymbol{\zeta}$, $\boldsymbol{\Psi}$, is the same for all time points t, and that all observations on ζ_i are mutually uncorrelated.

Importantly, the structure of $\boldsymbol{\Psi}$ depends on whether a model is *recursive* or *nonrecursive.* Recursive models are systems of equations that contain no reciprocal causation, implying that the $\boldsymbol{B}$ matrix can be written as a lower triangular matrix. In this case, the errors in equations are assumed to be uncorrelated, entailing that $\boldsymbol{\Psi}$ is diagonal.

Unlike recursive models, nonrecursive models contain reciprocal causation and/or feedback loops, entailing that $\boldsymbol{B}$ is not lower triangular. Under such models, ζ-disturbances can be assumed either correlated or not.

The variance-covariance matrix of $\boldsymbol{\xi}$ is a $n \times n$ symmetrical matrix denoted $\boldsymbol{\Phi}$. That is, exogenous latent variables can be correlated, implying that $\boldsymbol{\Phi}$ in that case is not diagonal. Notice that the covariance matrix of $\boldsymbol{\eta}$ is not a free parameter matrix in the model. However, one can calculate this matrix afterwards (if required) according to the following formula:

$$\mathbf{Cov}(\boldsymbol{\eta}) = (\boldsymbol{I} - \boldsymbol{B})^{-1} \left(\boldsymbol{\Gamma}\,\boldsymbol{\Phi}\,\boldsymbol{\Gamma}' + \boldsymbol{\Psi}\right) \left[(\boldsymbol{I} - \boldsymbol{B})^{-1}\right]'. \tag{A1.2}$$

A2

Structural Equation Model (SEM): a standard representation

Submodel 2: Measurement model

As a matter of fact, vectors $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ are not observed. Instead, vectors $\boldsymbol{y}' = (y_1,\, y_2, \ldots,\, y_p)$ and $\boldsymbol{x}' = (x_1,\, x_2, \ldots,\, x_q)$ are observed, such that

$$\boldsymbol{y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{A2.1}$$
$$\boldsymbol{x} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta}, \tag{A2.2}$$

where

$\boldsymbol{y}$ a $p \times 1$ vector of observed indicators of $\boldsymbol{\eta}$;

$\boldsymbol{x}$ a $q \times 1$ vector of observed indicators of $\boldsymbol{\xi}$;

$\boldsymbol{\epsilon}$ a $p \times 1$ vector of measurement errors for $\boldsymbol{y}$ with the associated covariance matrix $\boldsymbol{\Theta}_\epsilon(p \times p)$;

$\boldsymbol{\delta}$ a $q \times 1$ vector of measurement errors for $\boldsymbol{x}$ with the associated covariance matrix $\boldsymbol{\Theta}_\delta(q \times q)$;

$\boldsymbol{\Lambda}_y$ a $p \times m$ matrix of coefficients relating $\boldsymbol{y}$ to $\boldsymbol{\eta}$;

$\boldsymbol{\Lambda}_x$ a $q \times n$ matrix of coefficients relating $\boldsymbol{x}$ to $\boldsymbol{\xi}$.

The model assumptions are:

- $\mathrm{E}(\boldsymbol{\eta}) = 0$, $\mathrm{E}(\boldsymbol{\xi}) = 0$, $\mathrm{E}(\boldsymbol{\epsilon}) = 0$, and $\mathrm{E}(\boldsymbol{\delta}) = 0$,
- $\boldsymbol{\epsilon}$ is uncorrelated with $\boldsymbol{\eta}$, $\boldsymbol{\xi}$, and $\boldsymbol{\delta}$
- $\boldsymbol{\delta}$ is uncorrelated with $\boldsymbol{\xi}$, $\boldsymbol{\eta}$, and $\boldsymbol{\epsilon}$.

To summarise, the full SEM is defined by three equations:

$$\begin{aligned} \text{Latent variable model:} \quad & \boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \\ \text{Measurement model for } \boldsymbol{y}: \quad & \boldsymbol{y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\epsilon} \\ \text{Measurement model for } \boldsymbol{x}: \quad & \boldsymbol{x} = \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta} \end{aligned} \tag{A2.3}$$

Rewriting $\boldsymbol{\eta}$ in the reduced form, that is,

$$\boldsymbol{\eta} = (\boldsymbol{I} - \boldsymbol{B})^{-1} (\boldsymbol{\Gamma}\,\boldsymbol{\xi} + \boldsymbol{\zeta})\,. \tag{A2.4}$$

and substituting (A2.4) for $\boldsymbol{\eta}$ in (A2.3) permits us to derive the expression for the covariance matrix of the observed variables as a function of the model parameters, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ ([2], p. 325):

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) & \\ \boldsymbol{\Sigma}_{xy}(\boldsymbol{\theta}) & \boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}) \end{bmatrix}, \tag{A2.5}$$

where

$\boldsymbol{\Sigma}_{yy}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_y\,\boldsymbol{A}\,(\boldsymbol{\Gamma}\,\boldsymbol{\Phi}\,\boldsymbol{\Gamma}' + \boldsymbol{\Psi})\,\boldsymbol{A}'\boldsymbol{\Lambda}_y' + \boldsymbol{\Theta}_\epsilon$

$\boldsymbol{\Sigma}_{xy}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_x\,\boldsymbol{\Phi}\,\boldsymbol{\Gamma}'\,\boldsymbol{A}'\,\boldsymbol{\Lambda}_y'$

$\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta}) = \boldsymbol{\Lambda}_x\,\boldsymbol{\Phi}\,\boldsymbol{\Lambda}_x' + \boldsymbol{\Theta}_\delta$,

where $\boldsymbol{A} = (\boldsymbol{I} - \boldsymbol{B})^{-1}$.

An important point to realise about the full SEM is that it subsumes many models as special cases. In the context of our analysis, it is relevant to mention one of the cases, namely confirmatory factor model. To see the relation between the models, set $\boldsymbol{B} = \mathbf{0}$, $\boldsymbol{\Gamma} = \mathbf{0}$, $\boldsymbol{\Theta}_\epsilon = \mathbf{0}$, $\boldsymbol{\Lambda}_y = \mathbf{0}$ and $\boldsymbol{\Psi} = \mathbf{0}$ in (A2.3) and (A2.5). This reduces the full SEM to the measurement model for $\boldsymbol{x}$ which is a general factor model associated with $\boldsymbol{\Sigma}_{xx}(\boldsymbol{\theta})$, the lower-right quadrant of (A2.5).

This straightforward connection between SEM and factor analysis immediately implies that the issues of estimation, hypothesis testing, identifiability, and model evaluation for SEM parallel those associated with factor analysis, discussed in Part 1.

To begin with, just as in factor analysis, the structure of (A2.5) depends on restrictions, imposed on the model parameters in accordance with a priori knowledge/hypotheses researchers have. Researchers can express their substantive ideas and hypotheses in form of fixed and constrained parameters (in our own analysis, the primary interest concerns constrained-equal parameters). Estimation of free parameters in SEM is performed exactly in the same way as in factor analysis, that is, under normality assumption of data, one minimises the discrepancy between $\mathbf{\Sigma}(\widehat{\boldsymbol{\theta}})$ and the sample covariance matrix of the indicators given fixed and constrained parameters (see the discrepancy function in Appendix A in Part 1). The overall fit of SEM is also assessed by the same means as the overall fit of factor model. Clearly, even the issue of identification of SEM can be addressed in the same way as it is done for factor models. As discussed in Part 1, identifiability of factor model can be established algebraically, i.e. by solving the covariance structural equations, $\mathbf{\Sigma} = \mathbf{\Sigma}(\boldsymbol{\theta})$, for the unknown free parameters. However, due to the higher complexity of SEM, the determination of its identification status algebraically can be much more tedious and thus more error-prone. In case the model of interest is very complex, researchers may resort to several rules that aid in the identification of the model, or, as advised by [20], confine themselves to determining which of the parameters can be solved for and which cannot without solving the equations explicitly.

What distinguishes SEM models from factor models are the notions of indirect and total effects. In factor analysis, it is relevant to talk only about direct effects, more precisely, direct effects of ξ-variables on their indicators, x-variables. In SEM, ξ-variables may, in addition, have direct effects on η-variables, meaning that they indirectly affect the indicators of η-variables. Direct and indirect effects together constitute the total effect. We do not proceed with discussing this topic in greater depth because (1) our main hypothesis, i.e. the hypothesis of consistency between the latent simulated and true temperature responses to forcings, concerns only direct effects of latent variables, and (2) without knowing the ability of the suggested SEM-model to address the question of interest in practice, it is quite unmotivated to discuss what additional questions can be addressed by means of this model. As mentioned in the introduction, the performance of our SEM-model and the factor model arising under Scheme 1 is investigated in Part III.

A3

An alternative representation of SEM

The representation of a general structural equation model given above is known as a standard representation. Being sufficient for capturing the relation between variables within some analyses, the standard representation might be insufficient within other analyses due to its restrictions. For example, it is not allowed that observed variables influence latent variables, in particular the endogenous ones, which in the context of the present work would prevent climatologically defensible causal links from observable temperatures (simulated and/or observed) to the latent temperature responses due to the Land and Ghg forcings. To overcome those restrictions, a two-equation model has been suggested (see [2], Ch. 9):

$$\boldsymbol{\eta}^{+} = \boldsymbol{B}^{+}\,\boldsymbol{\eta}^{+} + \boldsymbol{\zeta}^{+} \tag{A3.1}$$

$$\boldsymbol{y}^{+} = \boldsymbol{\Lambda}_y^{+}\,\boldsymbol{\eta}^{+}, \tag{A3.2}$$

where $\boldsymbol{\eta}^{+}$, $\boldsymbol{B}^{+}$, $\boldsymbol{\zeta}^{+}$, and $\boldsymbol{y}^{+}$ are related to the variables from the standard representation in the following way:

$$\boldsymbol{\eta}^{+} = \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \\ \boldsymbol{\eta} \\ \boldsymbol{\xi} \end{bmatrix}, \quad \boldsymbol{\zeta}^{+} = \begin{bmatrix} \boldsymbol{\epsilon} \\ \boldsymbol{\delta} \\ \boldsymbol{\zeta} \\ \boldsymbol{\xi} \end{bmatrix}, \quad \boldsymbol{y}^{+} = \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{x} \end{bmatrix} \tag{A3.3}$$

$$\boldsymbol{B}^{+} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}_y & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Lambda}_x \\ \mathbf{0} & \mathbf{0} & \boldsymbol{B} & \boldsymbol{\Gamma} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad \boldsymbol{\Lambda}_y^{+} = \begin{bmatrix} \boldsymbol{I}_{p+q} & \mathbf{0} \end{bmatrix}$$

where $\boldsymbol{I}_{p+q}$ is an order-$(p+q)$ identity matrix picking out the observed variables from $\boldsymbol{\eta}^{+}$. The $\boldsymbol{\Lambda}_y^{+}$ is consequently $(p+q) \times (p+q+m+n)$. Further,

- $\boldsymbol{\eta}^{+}$ and $\boldsymbol{\zeta}^{+}$ are $(p+q+m+n) \times 1$,
- $\boldsymbol{y}^{+}$ is $(p+q) \times 1$, and
- $\boldsymbol{B}^{+}$ is $(p+q+m+n) \times (p+q+m+n)$.

The final matrix for this alternative representation is the covariance matrix for $\boldsymbol{\zeta}^{+}$ denoted $\boldsymbol{\Psi}^{+}$. Its relation to the standard parameters is

$$\boldsymbol{\Psi}^{+} = \begin{bmatrix} \boldsymbol{\Theta}_{\epsilon} & & & \\ \mathbf{0} & \boldsymbol{\Theta}_{\delta} & & \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Psi} & \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Phi} \end{bmatrix}. \tag{A3.4}$$

Substituting the reduced form of $\boldsymbol{\eta}^{+}$, given by

$$\boldsymbol{\eta}^{+} = \left(\boldsymbol{I} - \boldsymbol{B}^{+}\right)^{-1} \boldsymbol{\zeta}^{+}, \tag{A3.5}$$

into (A3.1), the reproduced covariance matrix of $\boldsymbol{\eta}^{+}$ is derived:

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta}^{+}}(\boldsymbol{\theta}) = \left(\boldsymbol{I} - \boldsymbol{B}^{+}\right)^{-1} \boldsymbol{\Psi}^{+} \left((\boldsymbol{I} - \boldsymbol{B}^{+})^{-1}\right)'. \tag{A3.6}$$

Inserting (A3.5) into (A3.2) gives the reproduced covariance matrix of the observed variables only:

$$\mathbf{\Sigma}_{\boldsymbol{y}^+}(\boldsymbol{\theta}) = \left(\mathbf{\Lambda}_y^+ \left(\boldsymbol{I} - \boldsymbol{B}^+\right)^{-1}\right) \mathbf{\Psi}^+ \left(\mathbf{\Lambda}_y^+ \left(\boldsymbol{I} - \boldsymbol{B}^+\right)^{-1}\right)' \tag{A3.7}$$

The matrices $\boldsymbol{B}^+$ from (A3.3) and $\mathbf{\Psi}^+$ from (A3.4) make explicit the implicit constraints of the standard representation. However, by changing the fixed zero elements in these matrices we can relax many of those constraints. An important point to keep in mind, when relaxing the assumptions of the standard representation, is that the resulting model should be identified.

Appendix B

Equations for the nonrecursive Structural Equation Model depicted in Figure 6: the standard representation

Combining our notations used in Figure 6 and the notations associated with the standard general representation of a structural equation model given in Appendix A1 and A2, the structural equation model depicted in Figure 4 has the following equations:

The latent variable model:

$$\underbrace{\begin{bmatrix} \xi^T_{\mathsf{Land}} \\ \xi^T_{\mathsf{Ghg}} \\ x_{\mathsf{comb}} \\ v \end{bmatrix}}_{=\boldsymbol{\eta}} = \underbrace{\begin{bmatrix} 0 & \text{GL} & 0 & 0 \\ \text{LG} & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \text{Ltrue} & \text{Gtrue} & 0 & 0 \end{bmatrix}}_{=\boldsymbol{B}} \cdot \underbrace{\begin{bmatrix} \xi^T_{\mathsf{Land}} \\ \xi^T_{\mathsf{Ghg}} \\ x_{\mathsf{comb}} \\ v \end{bmatrix}}_{=\boldsymbol{\eta}} + \underbrace{\begin{bmatrix} \text{SL} & \text{OL} & \text{VL} & \text{IL} \\ \text{SG} & \text{OG} & \text{VG} & \text{IG} \\ \text{Ssim} & \text{Osim} & \text{Vsim} & \text{Isim} \\ \text{Strue} & \text{Otrue} & \text{Vtrue} & \text{Itrue} \end{bmatrix}}_{=\mathbf{\Gamma}} \cdot \underbrace{\begin{bmatrix} \xi^T_{\mathsf{Sol}} \\ \xi^T_{\mathsf{Orb}} \\ \xi^T_{\mathsf{Volc}} \\ \xi^T_{\mathsf{interact}} \end{bmatrix}}_{=\boldsymbol{\xi}} + \underbrace{\begin{bmatrix} \xi^T_{\mathsf{Land\,(anthr)}} \\ \xi^T_{\mathsf{Ghg\,(anthr)}} \\ \delta_{\mathsf{comb}} \\ \nu \end{bmatrix}}_{=\boldsymbol{\zeta}},$$

where the variance-covariance matrices of $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are given by

$$\Phi_{\boldsymbol{\xi}} = \begin{pmatrix} 1 & 0 & 0 & \phi_{SI} \\ & 1 & 0 & \phi_{OI} \\ & & 1 & \phi_{VI} \\ & & & 1 \end{pmatrix} \quad \text{and} \quad \Psi_{\boldsymbol{\zeta}} = \begin{pmatrix} \sigma^2_{\xi_{\mathsf{Land\,(anthr)}}} & \sigma_{\xi_{\mathsf{Land\,(anthr)}}\,\xi_{\mathsf{Land\,(anthr)}}} & 0 & 0 \\ & \sigma^2_{\xi_{\mathsf{Ghg\,(anthr)}}} & 0 & 0 \\ & & \sigma^{2\,*}_{\delta_{\mathsf{comb}}} & 0 \\ & & & \sigma^2_\nu \end{pmatrix},$$

where $\sigma^{2\,*}_{\delta_{\mathsf{comb}}}$ is assumed to be known a priori.

The measurement model for $\boldsymbol{x}$-variables, i.e. the indicators of the latent exogenous variables $\boldsymbol{\xi}$:

$$\underbrace{\begin{bmatrix} x_{\mathsf{Sol}} \\ x_{\mathsf{Orb}} \\ x_{\mathsf{Volc}} \end{bmatrix}}_{=\boldsymbol{x}} = \underbrace{\begin{bmatrix} \text{Ssim} & 0 & 0 & 0 \\ 0 & \text{Osim} & 0 & 0 \\ 0 & 0 & \text{Vsim} & 0 \end{bmatrix}}_{=\mathbf{\Lambda}_x} \cdot \underbrace{\begin{bmatrix} \xi^T_{\mathsf{Sol}} \\ \xi^T_{\mathsf{Orb}} \\ \xi^T_{\mathsf{Volc}} \\ \xi^T_{\mathsf{interact}} \end{bmatrix}}_{=\boldsymbol{\xi}} + \underbrace{\begin{bmatrix} \delta_{\mathsf{Sol}} \\ \delta_{\mathsf{Orb}} \\ \delta_{\mathsf{Volc}} \end{bmatrix}}_{=\boldsymbol{\delta}},$$

where the variance-covariance matrix of $\boldsymbol{\delta}$ is given by $\mathbf{\Theta}_{\boldsymbol{\delta}} = \text{diag}\left(\sigma^{2\,*}_{\delta_{\mathsf{Sol}}},\ \sigma^{2\,*}_{\delta_{\mathsf{Orb}}},\ \sigma^{2\,*}_{\delta_{\mathsf{Volc}}}\right)$.

Each of the three variances is assumed to be known a priori.

The measurement model for $\boldsymbol{y}$-variables, i.e. the indicators of the latent endogenous variables $\boldsymbol{\eta}$:

$$\underbrace{\begin{bmatrix} x_{\mathsf{Land}} \\ x_{\mathsf{Ghg}} \\ x^{+}_{\mathsf{comb}} \\ v^{+} \end{bmatrix}}_{=\boldsymbol{y}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{=\boldsymbol{\Lambda}_y} \cdot \underbrace{\begin{bmatrix} \xi^{T}_{\mathsf{Land}} \\ \xi^{T}_{\mathsf{Ghg}} \\ x_{\mathsf{comb}} \\ v \end{bmatrix}}_{=\boldsymbol{\eta}} + \underbrace{\begin{bmatrix} \delta_{\mathsf{Land}} \\ \delta_{\mathsf{Ghg}} \\ 0 \\ 0 \end{bmatrix}}_{=\boldsymbol{\epsilon}},$$

where the variance-covariance matrix of $\boldsymbol{\epsilon}$ is given by $\boldsymbol{\Theta}_{\boldsymbol{\epsilon}} = \text{diag}\left(\sigma^{2\,*}_{\delta_{\mathsf{Land}}},\ \sigma^{2\,*}_{\delta_{\mathsf{Ghg}}},\ 0,\ 0\right)$, where $\sigma^{2\,*}_{\delta_{\mathsf{Land}}}$ and $\sigma^{2\,*}_{\delta_{\mathsf{Ghg}}}$ are regarded as known a priori.

Having elucidated the correspondence between our notations and the notations associated with the general structural equation model given in Appendix A1 and A2, it is not difficult to rewrite the equations above in accordance with the alternative representation summarised in Eq. (A3.1)-(A3.3) in Appendix A3. However, we omit here the alternative representation of the SEM model in Figure 6 due to the considerable size of the resulting matrices.

Acknowledgements

The authors would like to thank Qiong Zhang (Department of Physical Geography, Stockholm University) for helpful explanations of some aspects of climate modelling. 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] Angraini Y., et al, *The Relationships between Individualism, Nationalism, Ethnocentrism, and Authoritarianism in Flanders: A Continuous Time-Structural Equation Modeling Approach*, Multivariate Behavioral Research, 49:1, 41-53, doi: 10.1080/00273171.2013.836621, 2014

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

[3] Boucher, O. et al.: Clouds and Aerosols. In: *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, 2013.

[4] CUBASH U. et al.: *Introduction. 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.

[5] DESER, C., et al.: *Uncertainty in climate change projections: the role of internal variability.* Climate Dynamics, **38**, 527-546, doi:10.1007/s00382-010-0977-x, 2012.

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

[7] DIMITROV, D.M. and Raykov T.: *Validation of Cognitive Structure: A Structural Equation Modelig Approch*, Multivariate Behavioral research, Vol 38, 1, DOI:10.1207/S15327906MBR3801-1, 2010.

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

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

[10] 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.

[11] FLATO G.J. et al.: *Evaluation of Climate Models. 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.

[12] GOOSSE, H.: *Climate system dynamics and modelling*, Cambridge university press, 2015.

[13] HEGERL, G.C. et al.: *Understanding and Attributiong 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.

[14] 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.

[15] 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.

[16] He, C., Huang, Z, and Wang, R.: *Land use change and economic growth in urban China: A structural equation analysis*, Urban Studies, 2014, Vol. **51**(13), 2880–2898, dot:10.1177/0042098013513649, 2014.

[17] Hurtt, G. C., Chini, L. P., Frolking, S. et al.: *Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands*, Climatic Change, 1, **109**, doi:10.1007/s10584-011-0153-2, 2011.

[18] IPCC: *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.

[19] 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.

[20] Jöreskog, K.G. and Sörbom, D.: *LISREL 7: a guide to the program and applications*, SPSS INC, 1988.

[21] Kodra, A. et al.: *Exploring Granger causality between global average observed time series of carbon dioxide and temperature*, Theor. Appl. Climatol., 104, 325-335, doi:10.1007/s00704-010-0342-3, 2011.

[22] Liang, X.S.: *Unraveling the cause-effect relation between time series*, Physical Review E, **90**, 052150, doi: 10.1103/PhysRevE.90.052150, 2014.

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

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

[25] 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.

[26] Mulaik, S. A.: *Linear causal modeling with structural equations*, Chapman&Hall/CRC, 2009.

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

[28] Myhre, G. D. et al.: *Anthropogenic and Natural Radiative Forcings.* In: *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, 2013.

[29] 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, doi:10.1175/BAMS-D-14-00233.I, 2016.

[30] Pongratz, J., Reick, C., Raddatz, T., and Claussen, M.: *A reconstruction of global agricultural areas and land cover for the last millennium*, Global Biogeochem. Cycles, 22, GB3018, doi:10.1029/2007GB003153, 2008.

[31] Pudrovska, T. & Anishkin, A.: *Early-Life Socioeconomic Status and Physical Activity in Later Life: Evidence From Structural Equation Models*, Journal of Aging and Health, 25(3), 383-404, DOI: 10.1177/0898264312468601, 2012

[32] Rindskopf, D.: *Structural Equation Models: empirical identification, Heywood cases, and related problems*, 2nd edition, Sociological methods & Research, **13**(1), 109-119, 1984.

[33] Saxena A. & Khanna U.: *Advertising on Social Network Sites: A Structural Equation Modelling Approach*, Vision, 17(1) 17–25, SAGE Publications, DOI: 10.1177/0972262912469560, 2013

[34] Schurer, A. P., Hegerl G. C., Tett S. F. and Phipps S. J.: *Separating Forced from Chaotic Climate variability over the Past Millenium*, J. Clim, 26, 6954-6973, doi:10.1175/JCLI-D-12-00826.1, 2013.

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

[36] Stips, A. et al.:*On the causal structure between CO_2 and global temperature*, Scientific reports, 6:21691, doi:10.1038/srep21691, 2016.

[37] 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, doi: 10.5194/cp-8-1339-2012, 2012.

[38] Sörbom, D.: *Model Modification*, Psychometrika , **54**, 371 – 384, doi:10.1007/BF02294623, 1989.

[39] 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