Referee N°1

We thank once more the referee for his careful review and constructive criticism. Most of the responses will be clearly clear/visible in the diff-document, and we provide complementary information point by point below.

Major Comments

1 First point

page 8

3 sets of experiments with the same experiment design We understand the point of view of the referee and appreciate that they may room for further improvements of the experiment design. We, however, defend our choice as follows: The following is the corresponding text in the paper

"Our initial intention was to build two distinct emulators, one using the warm orbit as initial conditions, and one using the cold orbit. In practice, the two emulators ended up being redundant (sect. 3.1), and thus a third of our computational budget was thus spent on replicating essentially identical runs. However, this could not have been easily foreseen, and we feel was worth the cost to avoid the risk of missing the existence of multiple steady-states. The third ensemble uses the same initial conditions as the first, but uses the original ECBILT surface scheme with fixed vegetation (Opsteegh et al., 1998) and not the VECODE model. We built an independent emulator in this case, but we note that 'multi-level emulation' approaches, which associate similar but different simulators Cumming and Goldstein (2009), could potentially have increased our predictive accuracy by allowing us to make better use of the limited computational resource. We leave this option as a possible subject for further investigation."

Second and third points page 8 and 9

Section 2.3.2 is moved as suggested. This is followed by the definition of the sensitivity measures as well as the clarification of different interpretations by adding the corresponding equations explicitly.

page 10

The estimation of sensitivity measures by the aid of a Gaussian process emulator is expressed after the two points above.

Fourth point page 10

This part is detailed in a way to make the document self contained. Explicit equations and the corresponding interpretations are added in a way to show where the emulator intervene in calculations of the sensitivity measures.

Fifth point page 21

This comment is taken into account and the text is set to:

"The oscillation itself could be of physical relevance for past climate variability, but the limits of the phase space region in which the oscillation occurs are also likely to depend on many other uncertain parameters of the model. One possible action would be to use a sequential design strategy to delineate the region of occurrence of the phenomenon and develop an emulator aimed at characterising this oscillation. In particular, history matching theory provides adequate concepts and methods to this end Williamson et al. (2013). Given the likely sensitivity of the oscillation on model parameters, the significance of this enterprise for palaeoclimate interpretation is unsure. We rather choose to ignore the experiment for the time being (the following diagnostics ignore experiment 20), but discuss the possible consequences of this choice in the final discussion. In statistical terms, we provisionnally condition the analysis on the hypothesis that these oscillations do not occur in the phase space."

Sixth point page 27

A text is added in order to clarify the fact the results presented in the section and paper are conditioned on the use of modeling strategy and more specifically conditioned on the use of LOVECLIM model. That the conclusions do not represent reality and is a guidance to interpret and gain understanding of reality.

Minor Comments

1-4 points 1 to 4 are taken into account

5 these papers are now cited.

6-15 these points are taken into account

16 SVD

Let us denote by $\mathbf{Y_{full-space}}$ the $N \times n$ output matrix which contains simulation results at the training design $\mathbf{X}$ of N distinct experiments and s values of the output (multivariate). For example, in the case of LOVECLIM mean surface temperature, the output per simulation has 2048 values (grid points). The decomposition is realized on the centered output $\mathbf{Y}^{'} = \mathbf{Y_{full-space}} - \mu\mathbf{1}$, where the row mean μ is subtracted from the output training data and $\mathbf{Y}^{'}$ is an $N \times n = 27 * 2048$ matrix.

The singular value decomposition of $\mathbf{Y}^{'}$ is $\mathbf{UDV}^T$ where $\mathbf{U}, \mathbf{V}$ are the left and right singular vectors of $\mathbf{Y}^{'}$respectively;

$\mathbf{U}$ is a $27 * 2048$,

$\mathbf{V}$ is a $2048 * 27$, and
$\mathbf{D}$ is a $27 * 27$.
$\mathbf{D}$ is diagonal and contains the non-negative singular values of $\mathbf{Y}^{'}$ in decreasing order : N values, a square matrix of 27 by 27 elements.

The principal components are the projection coefficients given by $\mathbf{VD}$ (2048*27x27*27)such that the coefficient of the k^{th} eigenvector u_k(2048*1) in $\mathbf{U}$ for a given j^{th} experiment is $\mathbf{V}_{jk}\mathbf{\Gamma}_{kk}$.(2048*1 x 1*1)

17 This is modified in pages 14 section 2.4.2, bottom of the paragraph. "*and the error associated with this assumption is accounted for in the covariance of the estimator (see section 2.5).*" See also In section 2.5: , equation 18 left-hand-side.

18 This is modified in the text and we no longer say it is standard.

19-23 in subsection 2.4.3, these points are modified/corrected and thus taken into account

24 in page 14 subsection 2.4.3 a sentence is added to clarify this point : ?1 line 6

25 pages 14-15

In the present paper, ν_k is not considered an indicator of internal variability. It is inappropriate to interpret it as so, as we use long climate averages (500 years). Moreover, we anticipate that internal variability will be of the same order as to the error associated with the truncation of principal components.

The following is the text in the paper set to *"(...) In climate model applications, the nugget may also be justified as a way to account for 'internal variability'. Indeed, the chaotic dynamics of the simulator are such that a particular climate average over a given time window, can be viewed as a stochastic quantity, even though the simulator is deterministic. In the climate modelling parlance, the effect is referred to as un- certainty associated with the internal simulator variability. For example, in Araya-Melo et al.(2015), we found that our estimate of the nugget variance is consistent with the assumption that this term models the uncertainty associated due to simulator variability. This is also the inter- pretation adopted by Williamson et al. (2014) (both studies use the climate model HadCM3). In the the present application, we use rather long climate averages (500 years) and we anticipate that internal variability will be of the same order as to the error associated with the truncation of principal components. It may thus not be appropriate to interpret ν_k as an indicator of internal variability, and we therefore chose not to do so."*

26-28 All these points are taken into account and are changed in the text

29 Formula of C is explicitly given in the revised version.

30 "Output variable" is changed to "output field (temperature, precipitation, GDD)" in page 16, first bullet point

31 This is taken into account in section 2.6, page 16. The bullet point 2.6 is split into two points and the third bullet point modified in purpose.

32 The term is suppressed

33 The point is suppressed

34-35-37-38 These points are taken into account and already corrected/changed in the text

36 We changed "wrong predictions" into "prediction outside the 99th credible interval".

Referee N°2

General comments

1 First bullet point

It is correct that parallelizing the code is favorable for optimization computing time, but here we downweighted the argument on computer speed and further emphasized on the conceptual advantages, this is explained in page 19.

2 see specific comments

Specific Comments11

1-2 These are corrected/changed inn the text

3 equations 15-19 , are moved into appendix. C expression is explicitly specified in page 11, equation 9.

4 Argument is taken into account in section 2.6

the third bullet point is set to:

"In some regions there may be only small variation in the simulated output as the input parameters change. If independent emulators are used for each grid cell, estimating the hyper-parameters for these cells can be difficult without applying some sort of parameter regularisation, and besides, the computational effort of building the emulator is unnec- essary (as the output is constant). The global principal component emulator is therefore preferable in these situations, as these constant regions are automatically accounted for in the principal component variance decomposition."

5 "diagnostic" is corrected

6 S_{tot} is referenced now and pointed to equation (10) following the new notations.

7 It is correct that a full probabilistic processing of hyperparameters (with priors and posterios), such as in Farah, 2011, is in principle needed to provide a full Bayesian quantification of the uncertainty on variance indices. This said, we noticed little sensitivity of sensitivity indices estimates on the length scales, and in no case did we notice sensitivitity that would impact the general scientific conclusions. On the other hand, most of the existing literature focusing on emulators formulated through Gaussian process (as here) considers, just as we did, semi-Bayesian estimate of the main effects with fixed hyperparameters. This is, in particular, the approach of Oakley and O'Hagan (2004). The penalised likelihood maximisation adopted here follows quite exactly Andrianikis and Challenor (2012). Is summary, we comply with mainstream literature and verified that reasonably small changes in length-scales and nuggets have no scientific incidence. This was enough to convince us that the semi-Bayesian approach adopted here is adequate. Providing posterior bounds on hyperparameters would require a computational cost that does not seem justified here.

Other changes

We added references to the important works of Tuenter and other collaborators of the Utrecht group. The manuscript underwent an additional number of further editorial corrections (typos, English, etc.)

Manuscript prepared for Earth Syst. Dynam. Discuss.
with version 2014/05/30 6.91 Copernicus papers of the LaTeX class copernicus.cls.
Date: 1 April 2015

Global sensitivity analysis of the climate-vegetation system to astronomical forcing: an emulator-based approach

Nabila Bounceur[1], Michel Crucifix[1], and Richard Wilkinson[2]

[1]Université catholique de Louvain , Earth and Life Institute, Georges Lemaître Centre for Earth and Climate Research, Louvain-la-Neuve, Belgium
[2]School of Mathematics, University of Nottingham, United Kingdom

Correspondence to: M. Crucifix (michel.crucifix@uclouvain.be)

Abstract

A global sensitivity analysis is performed to describe the effects of astronomical forcing on the climate-vegetation system simulated by the model of intermediate complexity LOVECLIM in interglacial conditions. The methodology relies on the estimation of sensitivity measures, using a Gaussian process emulator as a fast surrogate of the climate model, calibrated on a set of ~~well chosen~~ well-chosen experiments. The outputs considered ~~here~~ are the annual mean temperature and precipitation, and the Growing Degree Days (GDD). The experiments were run on two distinct land surface schemes to estimate the importance of vegetation feedbacks on climate variance. This analysis provides a spatial description of the variance due to the factors and their combinations, in the form of “fingerprints" obtained from the covariance indices. The results are broadly consistent with the current understanding of Earth’s climate response to the astronomical forcing. In particular, precession and obliquity are found to contribute ~~equally to (GDD) in~~ LOVECLIM equally to GDD in the northern hemisphere, and the ~~effects~~ effect of obliquity on the response of southern hemisphere temperature dominate precession effects. Precession dominates precipitation changes in subtropical areas. Compared to standard approaches based on a small number of simulations~~for well-defined past epochs~~, the methodology presented here allows us to identify more systematically regions susceptible ~~of~~ to experiencing rapid climate change in response to the smooth astronomical forcing change. In particular, we find that using interactive vegetation significantly enhances the expected rates of climate change, specifically in the Sahel (up to 50% precipitation change in 1,000 years) and in the Canadian Arctic region (up to 3° in 1000 ~~,~~years). None of the tested astronomical configurations were found to induce multiple steady states, but~~we observe~~, at low obliquity, we observed the development of an oscillatory pattern that has already been reported in LOVECLIM. Although the mathematics of the analysis are fairly straightforward, the emulation approach still requires considerable care in its implementation. We discuss the ~~effects~~ effect of the choice of ~~length scales~~length-scales, the type of emulator, and estimate uncertainties associated with specific computational aspects, to conclude that the ~~PC~~ principal component emulator is a ~~reasonable~~ good option for this kind of application.

1 Introduction

The seasonal and spatial ~~distributions~~ distribution of incoming solar radiation (insolation) at the top of the atmosphere ~~are~~ is determined by three astronomical parameters: the Earth's eccentricity, the longitude of the perihelion, and Earth's obliquity. The variation of these parameters causes sufficient changes in insolation to significantly affect climate, in particular, the distribution of surface temperature, vegetation cover, monsoon rainfall, Arctic sea-ice etc. These changes can be simulated and studied by means of experiments with global climate models. One classical approach consists in identifying two epochs in the past for which sufficient data are available, running the climate model (with the implicit assumption that simulated climate is quasi-stationary with respect to the astronomical forcing), and then comparing the two resulting simulated climates. This is, for example, the approach followed by the Palaeoclimate Modelling Intercomparison project (Braconnot et al., 2007).

The difference between the two epochs may then be further decomposed into contributions of several factors. Suppose we want to compare the beginning and the end of the last interglacial. These two periods, distant by 11,000 years, are characterised by significant climatic differences (e.g. Sanchez-Goñi et al., 1999). The difference in insolation forcing is caused by a change in the position of the perihelion on Earth's orbit (this is the precession process) and a decrease in obliquity. ~~One may then ask what are the individual effects of precession and obliquity, and whether these effects combine linearly, or, to the contrary, whether second-order effects are significant. In the climate literature the second order terms are often termed "synergy". In this particular context the input factors are physical forcing parameters (Crucifix and Loutre, 2002; Henrot et al., 201~~ ~~but the word synergy is also commonly used when input factors describe the possibility to "activate" interactive model components such as dynamic vegetation or oceanic circulation compared to a situation where vegetation or ocean temperatures are fixed (Ganopolski et al., 1998; Brac~~ ~~Returning to the example of the last interglacial~~

Based on a series of transient experiments, Crucifix and Loutre (2002) suggested that the sum of individual effects of precession and obliquity during the Eemian are less than their combined effects. In the model experiments discussed in that article, this is caused by feedbacks associ-

ated with the responses of vegetation and sea-ice. ~~However,~~ Tuenter et al. (2003) investigated the difference between individual and combined effects more generically, without reference to specific epochs of the Quaternary. They performed a set of seven experiments with different combinations of precession and obliquity spanning the range of values reached during the Quaternary. They found that "the amplitude of the [North African] precipitation response to obliquity depends on precession, while the precipitation response to precession is independent of obliquity". One problem, though, is that such second-order effects may ~~well be important~~ be significant during certain epochs only, ~~where non-linear transitions between different states occur (e. g., Green Sahara vs White Sahara). One may therefore argue that more robust conclusions will be obtained if~~ or at certain critical locations. Consequently, a more robust and comprehensive approach to estimate individual and combined contributions of astronomical factors to climate change would be to scan the whole domain of possible forcing configurations~~is systematically scanned. This is the purpose of the~~ . As we show next this objective can be reached by reference to the concept of global sensitivity analysis (Homma and Saltelli, 1996; Saltelli et al., 2004)~~. To illustrate this concept, let us call~~ , which may be introduced as follows. Call $\boldsymbol{x}$ a particular forcing configuration — in the present context, $\boldsymbol{x}$ contains the astronomical configuration. Denote the output of a model run at configuration $\boldsymbol{x}$ by $\boldsymbol{f}(\boldsymbol{x})$. The vector $\boldsymbol{x}$ varied in the past and we can estimate its distribution by reference to existing astronomical solutions. If we assume that the climate model ~~is correct~~ represents reality and that the climate is quasi-stationary with respect to the astronomical forcing, then $\boldsymbol{f}(\boldsymbol{x})$ reflects the past evolution of climate. We can in particular enquire about the variance of $\boldsymbol{f}(\boldsymbol{x})$ caused by the distribution of $\boldsymbol{x}$, and decompose this variance into contributions from individual factors (precession and obliquity in this example).

Here, we carry out a global sensitivity analysis of the climate model of intermediate complexity LOVECLIM (Goosse et al., 2010) to the astronomical forcing. We provide geographic distributions of the contributions of obliquity and precession on ~~climate on~~ precipitation and temperature~~, and estimate synergy effects~~. We also attempt to detect regions where fast climate responses may occur in response to the slow changes of astronomical forcing. The objectives of this work are two-fold. The first ~~objective is climatic.~~ is climatic: We would like to determine the respective roles of astronomical ~~components~~ factors on interglacial climate

change, and contribute to the discussion on the mechanisms of glacial inception and interglacial duration with a focus on climate-vegetation interactions, ~~in the line of the works of de Noblet et al. (1996); Claussen et al. (1999); Crucifix and Loutre (2002); Kageyama et al. (2004).M~~ ~~The~~ adding to the discussion in de Noblet et al. (1996); Claussen et al. (1999); Crucifix and Loutre (20
Our second objective is methodological. While global sensitivity analysis is well-established in statistics, it has only recently been applied to climate problems. Lee et al. (2011), for example, performs a global sensitivity analysis of a global atmospheric model associated with a complex aeorosol model, in order to decompose model output uncertainty into contributions from eight uncertain parameters (see also Lee et al. (2013); Carslaw et al. (2013)). Although our scientific objective is different (we want to decompose forced climate variances induced by variances in forcing ~~components~~factors), the methodological approach is broadly similar to that of Lee et al. (2011) and follows from Oakley and O'Hagan (2004) who proposed probabilistic sensitivity analysis as a generalization of global sensitivity analysis to complex simulators: (a) ~~sample a space-filling experiment design ,~~ choose the experimental design to efficiently fill the input space (here, the space of astronomical forcings) (b) run LOVECLIM at these ~~design points~~experiments, (c) ~~calibrate~~ train and validate an emulator, that is, a stochastic statistical model used to predict the function $\boldsymbol{f}(\boldsymbol{x})$ at any input point, based on the output of the experiments actually run~~,~~; and (d) use the emulator to estimate sensitivity indices. Emulation has been increasingly used in climate science in ~~the~~ recent years as a tool to explore input spaces with the ~~objective~~ aim of *calibrating* the model on observations and estimating climate sensitivity (Rougier et al., 2006, 2009; Holden et al., 2010; Schmittner et al., 2011). We ~~therefore want to~~ explore the potential of this methodology for our specific application, with a discussion of its possible advantages, challenges, and drawbacks compared to more classical approaches.

2 Methodology

2.1 Choice of input factors

In palæoclimatology, it is common to refer to the time of the year using the true solar longitude (λ), that is, the geocentric angle between the vernal equinox and the position occupied by the Earth at any point during the year. For example, the June solstice corresponds to $\lambda = 90°$, the September equinox to $\lambda = 180°$ etc. For the purpose of computing insolation at a given time of year, we need the true solar longitude at perihelion, that is, the true solar longitude corresponding to the shortest Earth-Sun distance. This quantity is denoted ϖ. The shape of the Earth's orbit is elliptic and characterised by eccentricity e. Finally, the angle between the ecliptic and the equator is called the obliquity and denoted ε. It may then be shown that the secular evolution of the top-of-the-atmosphere incoming solar radiation at any latitude and any true solar longitude is reasonably well approximated by a linear combination of $i_1 = e \sin \varpi$, $i_2 = e \cos \varpi$ and $i_3 = \varepsilon$ (Loutre, 1993). The quantity i_1 is often referred to as the climatic precession parameter. As i_1, i_2 and i_3 are not correlated, the three inputs can be viewed as a canonical form of the astronomical forcing parameters. In particular, their signature on the season-latitude distribution of incoming solar radiation is characteristic: i_1 and i_2 control the Earth-Sun distance at any true solar longitude with very little effect on annual mean insolation[1], and i_3 controls the seasonal contrast as well as the annual mean insolation.

The time-evolution of the astronomical parameters over the Pleistocene is well known (Berger, 1978b; Laskar et al., 2004). Note that ~~eccentricity $e < 0.05$~~, 99 % of the time, eccentricity $e < 0.05$ and the inner 99^{th} percentile of ε is 22.3° to 24.3°, these boundaries differing by less than 0.1° whether the Berger (1978b) or the Laskar et al. (2004) solution is used as the reference.

[1] ~~a~~ A small effect on the global ~~,~~annual mean insolation emerges as a result of variations in eccentricity.

2.2 Experiment design

As specified above, the climate model we use is the ocean-atmosphere-vegetation model of intermediate complexity "LOVECLIM" (Goosse et al., 2010). The choice of inputs for the experiments are encoded in a ~~*design* *input factor* *matrix*~~ denoted $\mathbf{X}$, ~~of which the three columns correspond~~ which has three columns corresponding to the inputs ~~$i_{1...3}$~~$i_{1,2,3}$, and which has as many rows as ensemble members (27 experiments in our case). There is a rich literature on experiment design sampling techniques, and the ~~monography~~ monograph by (e.g. Santner et al., 2003) is specifically dedicated to computer experiments. As a general rule, given a fixed number of experiments to be run, one objective ~~to achieve~~ is to maximise the ~~amount of information to~~ information that can be inferred from the experiment set. If we use ~~an Gaussian process~~ a Gaussian process (GP) emulator (as here), this criteria corresponds to minimising the posterior variance in the GP predictions (Sacks et al., 1989). Another objective is to minimise the bias in the quantities to be estimated. ~~Sampling a good~~ Finding an adequate design is thus an optimisation problem, and generally not a straightforward one~~because, on the one hand, the optimum solution may depend on parameters that~~ , because the GP hyper-parameters are not known a priori, and ~~, on the other hand, efficient algorithms for finding optimal designs may be difficult to find~~so the design must trade-off the need to learn these parameters with the need to minimise the prediction variance. In practice, an effective approach consists in using latin-hypercube designs (McKay et al., 1979; Morris and Mitchell, 1995; Urban and Fricker, 2010) with specific properties, such as maxi-min ~~latin-hypercubes~~ preporties (where we maximize the minimum Euclidean ~~distances between any~~ distance between any pair of design points) or orthogonality (in this case, maximizing ~~$\mathbf{X}^T\mathbf{X}$~~$\mathbf{X}'\mathbf{X}$). Hybrid designs ~~associate~~ use several such properties and ~~may~~ can be theoretically justified as good solutions for Gaussian process emulation (Joseph and Hung, 2008). ~~Out of the box packages are available~~ Out-of-the-box packages can be used to produce latin hypercube designs ~~(e.g. the *lhs* package (Carnell, 2012)) but we face here~~ (e.g., the *lhs* package, Carnell, 2012) , but are unsuited to our needs as we face an additional constraint~~: We~~ . Namely, we want to avoid wasting ~~computing time~~ computational effort sampling eccentricities > 0.05, which translates into a constraint on the quantity $i_1^2 + i_2^2$.

~~The algorithm given in the Appendix Asatisfies this constraint , while attempting to satisfy (near-)orthogonality~~ Vernon et al. (2010) and Draguljić et al. (2012) provide examples of designs for non-rectangular input spaces. The algorithm used here (Appendix A) is similar to Vernon et al. (201 it satisfies the constraint on eccentricity while aiming at near-orthogonality and maxi-min properties. The choice of $n = 27$ experiments arises from the general recommendation of Loeppky et al. (2009) to perform 10 experiments per input dimension. We used 27 members (and not 30), in order to compare with a factorial design of 3^3 members (N. Bounceur, thesis in preparation, not discussed here).

The resulting design is shown in Figure 1. It is executed three times. Two ensembles use the standard version of LOVECLIM with the VECODE vegetation model (Brovkin et al., 1997), but with two distinct sets of initial conditions. The first ~~one is~~ ensemble uses the pre-industrial conditions provided by default in the LOVECLIM model package. ~~For the second ensemble, initial conditions are~~ The second uses the final state of ~~the~~ member #2 of the first ensemble. This ~~particular member~~ is a so-called warm orbit (high obliquity, high eccentricity and $\varpi \simeq 90°$) that produces extensive vegetation cover in the northern hemisphere and in the Sahara region. Every experiment is run 2,000 years and the last the data used for the following analysis are obtained by averaging the last 500 years.

The purpose of using two distinct initial conditions is to detect potential co-existence of ~~distinct steady state solutions.~~ multiple steady-state solutions. Brovkin et al. (1998) , for example, found two stable equilibria for certain orbital configurations when using the same vegetation model, but a different atmosphere-ocean system. If such multiple states were to co-exist, then the simulator output would need to be emulated using a multi-modal process, rather than a (uni-modal) Gaussian process as used here. Our initial intention was to build two distinct emulators, one using the warm orbit as initial conditions, and one using the cold orbit. In practice, the two emulators ended up being redundant (sect. 3.1), and thus a third of our computational budget was thus spent on replicating essentially identical runs. However, this could not have been easily foreseen, and we feel was worth the cost to avoid the risk of missing the existence of multiple steady-states. The third ensemble uses the same initial conditions as the first~~set, but uses~~ , but considers the original “ECBILT" surface scheme with fixed vegetation (Opsteegh

et al., 1998) and not the VECODE model. We chose to build a third independent emulator in this case, but we note that "multi-level emulation" approaches, which associate similar but different simulators (Cumming and Goldstein, 2009) , could potentially have increased our predictive accuracy by allowing us to make better use of the limited computational resource. We leave this option as a possible subject for further investigation.

2.3 Global variance measures

2.3.1 ~~Definition~~

~~To investigate the effect of each component of the astronomical forcing on the simulator outputs, it is necessary to partition the~~ In the present application the inputs (the astronomical forcing) are known, and we want to estimate the simulator output variance induced by separate and/or combined variations of the different forcing factors. Recall that we defined $\boldsymbol{f}(\boldsymbol{x})$ to be the simulator response at input vector $\boldsymbol{x}$~~into components. Let $\boldsymbol{x}_i$ denote a subset of the components in $\boldsymbol{x}$ (with $\mathcal{X}_i$ the space of possible values of $\boldsymbol{x}_i$), and let $\boldsymbol{x}_{-i}$ be the remaining components (taking values in $\mathcal{X}_{-i}$)~~, which contains the values of the three astronomical forcing terms. Let $\rho(\boldsymbol{x})$ be the time-wise occupation density of the input space during the Pleistocene, which can be estimated from standard astronomical solutions (Berger, 1978b; Laskar et al., 2004). ~~Using this probability distribution for~~ The *total variance* of the output associated with the variations of factors may then be expressed as

$$V := \mathbb{V}\mathrm{ar}(\boldsymbol{f}(\boldsymbol{x})), \tag{1}$$

where the variance operator means that we are sampling (varying) $\boldsymbol{x}$ following the distribution $\rho(\boldsymbol{x})$. As $\boldsymbol{f}(\boldsymbol{x})$ is a vector, V is a matrix with elements giving covariances between any output pair.

Suppose we are now interested in the variance of the simulator output caused by the variation of a subset of the input factors only. Let $\boldsymbol{x}_i$ denote a subset of the components in $\boldsymbol{x}$, and let $\boldsymbol{x}_{\sim i}$ be the remaining components. For example, if i is obliquity, $\sim i$ will be the indices associated with eccentricity and precesion.

We define the following quantities: The *main effect*, ~~we can ask what is the average state of the system given a value for a particular component of the astronomical forcing. That is , we want to find the main effect~~ $\boldsymbol{\eta}(\boldsymbol{x}_i)$, is the expected output conditional on the value of $\boldsymbol{x}_i$, ~~defined as~~

$$\boldsymbol{\eta}(\boldsymbol{x}_i) := \int_{\mathcal{X}_{\sim i}} \boldsymbol{f}(\boldsymbol{x})\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)\mathrm{d}\boldsymbol{x}_{\sim i},$$

~~where $\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)$ is the probability density function of forcing terms~~ i.e., $\boldsymbol{\eta}(\boldsymbol{x}_i) = \mathbb{E}[\boldsymbol{f}(\boldsymbol{x})|\boldsymbol{x}_i]$. The *main effect variance* is the variance of the main effect with respect to $\boldsymbol{x}_i$: $V_i = \mathbb{V}\mathrm{ar}[\boldsymbol{\eta}(\boldsymbol{x}_i)]$.

On the other hand, the output variance associated with factors $\boldsymbol{x}_i$ varying while the factors $\boldsymbol{x}_{\sim i}$ ~~given~~ are fixed, is denoted $\mathbb{V}\mathrm{ar}(\boldsymbol{f}(\boldsymbol{x})|\boldsymbol{x}_{\sim i})$. This is a function of the value of ~~$\boldsymbol{x}_{\sim i}$.~~ the fixed factors $\boldsymbol{x}_{\sim i}$. If we further average this variance over possible values of $\boldsymbol{x}_{\sim i}$, that is, we take the expectation of this quantity with respect to to $\boldsymbol{x}_{\sim i}$, we obtain the total variance associated with factor $\boldsymbol{x}_i$, denoted T_i:

$$T_i := \mathbb{E}[\mathbb{V}\mathrm{ar}(\boldsymbol{f}(\boldsymbol{x})|\boldsymbol{x}_{\sim i})]. \text{ Similarly, } T_{\sim i} = \mathbb{E}[\mathbb{V}\mathrm{ar}(\boldsymbol{f}(\boldsymbol{x})|\boldsymbol{x}_i)]. \tag{2}$$

~~In order to relate output variances to input variances, we are interested in quantities such as $\boldsymbol{V}_i := \mathbb{V}\mathrm{ar}_i(\boldsymbol{\eta}(\boldsymbol{x}_i))$ where the variance is with respect to $\boldsymbol{x}_i$ (see interpretation in section ??). However, because~~ The law of total variance implies $V_i := V - T_{\sim i}$, so that we can see that the main effect variance may be interpreted as the expected loss of output variance resulting from fixing (knowing) the value of $\boldsymbol{x}_i$. The standardised quantity V_i/V is commonly referred to as the main effect index, and T_i/V, the *total* sensitivity index (cf. Homma and Saltelli (1996) and Saltelli et al. (2004)).

To estimate these sensitivity indices, we extend the framework established by Oakley and O'Hagan (one dimensional scalar simulator outputs to multi-dimensional vector outputs. Let $\mathcal{X}$, $\mathcal{X}_i$, and $\mathcal{X}_{\sim i}$ be the domains of the input factors $\boldsymbol{x}$, $\boldsymbol{x}_i$ and $\boldsymbol{x}_{\sim i}$, respectively. The *total* variance can be

explicitly written as

$$V = \int_{\mathcal{X}} \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})'\rho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \iint_{\mathcal{X}\times\mathcal{X}} \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x}^\star)'\,\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star. \tag{3}$$

The main effect is computed as follows:

$$\boldsymbol{\eta}(\boldsymbol{x}_i) = \int_{\mathcal{X}_{\sim i}} \boldsymbol{f}(\boldsymbol{x})\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)\mathrm{d}\boldsymbol{x}_{\sim i}, \tag{4}$$

where $\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)$ is the conditional density of $\boldsymbol{x}_{\sim i}$ given $\boldsymbol{x}_i$. Because our knowledge of $\boldsymbol{f}(\boldsymbol{x})$ is limited to the ensemble of model runs, we are uncertain about the value of $\boldsymbol{\eta}(\boldsymbol{x}_i)$ for ~~every value~~ all values of $\boldsymbol{x}_i$. ~~Suppose, however, that we already have an emulator which,~~ The approach described in Oakley and O'Hagan (2004) is to build an emulator of $\boldsymbol{f}(\boldsymbol{x})$ that can predict its value for any input ~~value~~ configuration $\boldsymbol{x}$~~, provides an estimate of the model output $\boldsymbol{f}$ in the form of a mean~~. They use a Gaussian process (GP) model for the emulator, with mean function $\boldsymbol{m}(\boldsymbol{x})$ ~~, and a covariance between two inputs $\boldsymbol{\Sigma}(\boldsymbol{x},\boldsymbol{x}^\star)$. Then~~and covariance function between $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{f}(\boldsymbol{x}^*)$ which we will denote by $\boldsymbol{\Sigma}(\boldsymbol{x},\boldsymbol{x}^*)$. Because the emulator is a Gaussian process, the main effect ~~is $\boldsymbol{\eta}(\boldsymbol{x}_i)$, which is a linear transformation of $\boldsymbol{f}(\boldsymbol{x})$, is~~ also a Gaussian process, and has mean and variance ~~as follows:~~ functions

$$\boldsymbol{M}_i(\boldsymbol{x}_i) = \int_{\mathcal{X}_{\sim i}} \boldsymbol{m}(\boldsymbol{x})\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)\,\mathrm{d}\boldsymbol{x}_{\sim i}, \text{ and} \tag{5}$$

$$S_i(\boldsymbol{x}_i,\boldsymbol{x}_i^\star) = \iint_{\mathcal{X}_{\sim i}\times\mathcal{X}_{\sim i}} \boldsymbol{\Sigma}(\boldsymbol{x},\boldsymbol{x}^\star)\rho(\boldsymbol{x}_{\sim i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}_{\sim i}^\star|\boldsymbol{x}_i^\star)\mathrm{d}\boldsymbol{x}_{\sim i}\mathrm{d}\boldsymbol{x}_{\sim i}^\star. \tag{6}$$

It follows that ~~(Oakley and O'Hagan, 2004) :~~

$$\mathbb{E}_f(V_i) = \int_{\mathcal{X}_i} \left[\boldsymbol{M}_i(\boldsymbol{x}_i)\boldsymbol{M}_i'(\boldsymbol{x}_i) + S_i(\boldsymbol{x}_i,\boldsymbol{x}_i)\right]\rho_i(\boldsymbol{x}_i)\,\mathrm{d}\boldsymbol{x}_i - \boldsymbol{C}, \mathbb{E}_f(V) = M_{\mathrm{tot}}^2 + S_{\mathrm{tot}} -, \tag{7}$$

where $\mathbb{E}_f$ denotes expectation ~~and variance~~ with respect to the emulator, and

$$\rho_i(\boldsymbol{x}_i) = \int_{\mathcal{X}_{-i}} \rho(\boldsymbol{x}_i|\boldsymbol{x}_{-i})\,\mathrm{d}\boldsymbol{x}_i, \tag{8}$$

$$\boldsymbol{M}_{\mathrm{tot}} = \int_{\mathcal{X}} \boldsymbol{m}(\boldsymbol{x})\rho(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}, \tag{9}$$

$$S_{\mathrm{tot}} = \int_{\mathcal{X}} \Sigma(\boldsymbol{x},\boldsymbol{x})\rho(\boldsymbol{x})\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{C} = \int_{\mathcal{X}} \boldsymbol{m}(\boldsymbol{x})\boldsymbol{m}(\boldsymbol{x})'\rho(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \iint_{\mathcal{X}\times\mathcal{X}} \Sigma(\boldsymbol{x},\boldsymbol{x}')\rho(\boldsymbol{x})\rho(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}'. \tag{10}$$

~~The *total* sensitivity variances $\boldsymbol{T}_i := \boldsymbol{V} - \boldsymbol{V}_{-i}$ (see again interpretation in section **??**) are estimated as $\mathbb{E}_f T_i := \mathbb{E}_f \boldsymbol{V} - \mathbb{E}_f \boldsymbol{V}_{-i}$, where the index $-i$ represents the factors complementary to the i. The main effect $\boldsymbol{\eta}$ being a vector of p components (p, the number of simulator outputs), the quantity $\mathbb{V}\mathrm{ar}(\boldsymbol{\eta})$ introduced above isa $p \times p$ matrix, and so will be~~ The expectation of the quantities V, T etc. Equations ~~are consistent with this, if $\boldsymbol{M}\boldsymbol{M}'$ is interpreted as an outer product of $\boldsymbol{M}$ with itself, and if the multivariate emulator indeed provides covariance matrices $\Sigma(\boldsymbol{x}, \boldsymbol{x}_\star)$ that are also $p \times p$, for any couple of inputs $(\boldsymbol{x}, \boldsymbol{x}_\star)$. The diagonal elements of these matrices provide sensitivities associated with each output considered independently. This is what we refer in the following as the *grid-point-wise* approach.~~ total variance V of the simulator output is:

$$\mathbb{E}_f(V) = \left[\boldsymbol{M}_{\mathrm{tot}}^2 + \boldsymbol{S}_{\mathrm{tot}}\right] - \boldsymbol{C}, \text{with} \quad S_{\mathrm{tot}} = \int \Sigma(\boldsymbol{x},\boldsymbol{x})\rho(\boldsymbol{x})\mathrm{d}\boldsymbol{x}. \tag{11}$$

(12)

2.3.1 ~~Interpretation in the case of known inputs~~

From now on, it will be implicit that the V, V_i and T_i are estimated with the emulator and the symbol $\mathbb{E}_f$ will be dropped. Specifically, we refer to $T_{\{\epsilon\varpi\}}$ as the total variance associated with precession, and T_ε, that associated with obliquity.

The measures V_i and T_i are essentially those introduced by Homma and Saltelli (1996) , further illustrated in Saltelli et al. (2004) , chap. 1, and adapted to emulator theory by Oakley and O'Ha except for the fact that these authors define scaled indices: V_i/V is known as the main effect index associated with i, and T_i/V, the corresponding total effect index. The main effect index quantifies the expected reduction in the variance of the output $f(\boldsymbol{x})$ if we were to learn $\boldsymbol{x}_i$, and the total effect index quantifies the expected reduction in the variance of the output $f(\boldsymbol{x})$ if we were to learn everything but $\boldsymbol{x}_i$.

In the present application the inputs (the astronomical forcing) are known, and we want to estimate the simulator output variance induced by separate and/or combined variations of the different components of this forcing. This particular context requires some reinterpretation of

Finally, we refer to the sensitivity measures: V_i is the output variance lost by fixing quantity $V - T_{\{e\varpi\}} - T_\varepsilon = T_e - T_\varepsilon$ as the *synergy* term. More generally, $T_i - V_i$ is a measure of how much the factor i , all other forcings varying. T_i is the output variance induced by varying i, all other forcings fixed. V_i and T_i differ if there are non-linear interactions between the different factorsis involved in interactions with any other input variable. The word "synergy" is commonly used in the climate literature to express the difference between the model response to different factors varied together, and the sum of the responses to factors considered or varied individually. Unlike the global sensitivity indices discussed here, or when the inputs are correlated. Only the first case is potentially relevant here. Concretely, we note: $T_{e\varpi} := T_{1,2}$ is the variance induced by climate precession ($e \sin \varpi$ synergy terms are classically estimated on the basis of a reference experiment, and $e \cos \varpi$ taken together), T_ε is the variance induced by obliquity, V is the variance induced by all astronomical forcing components, and $V - T_{e\varpi} - T_\varepsilon$ is the second-order term called here the *synergy* between climatic precession and obliquity.

In the following, the expectation operator $\mathbb{E}_f$ in front of Vcomparing one-at-a-time changes with combined changes in the different factors, following a method called factor separation analysis (Stein and Alpert, 1993; Alpert and Sholokhman, 2011) . See, e.g., T_i etc.will be dropped but we keep in mind that these quantities are estimated with the emulator.In particular, the quantity S_{tot} may be used as an informal measure of the amount of variance that is being introduced by using the emulator as a surrogate for the actual simulator : we expect this quantity

to remain *small* compared to the quantities of interest T_i and V. (Crucifix and Loutre, 2002; Ganopolsk

2.4 Emulator

2.4.1 Motivation

A pragmatic solution could be

2.4.1 Motivation

As the outputs of our simulator are spatially resolved climate quantities, we need to build an emulator capable of modelling multivariate outputs. A simple pragmatic solution is to train independent emulators for each grid point. This is the strategy followed , which is done by Lee et al. (2011). There are, however, several other possibilities for multivariate emulation (e.g. Rougier, 2008) . (see, e.g., Rougier, 2008) . The main challenge is defining a covariance function for generating the covariance matrix $\mathbf{\Sigma}(\boldsymbol{x}, \boldsymbol{x}^*)$, in order to produces a valid covariance function for the Gaussian process. See Fricker et al. (2013) for discussion. Here, we propose the principal component (PC) emulator (Higdon et al., 2008; Wilkinson, 2010) as a cost-effective and statistically reasonable alternative to point-wise emulation. The potential advantages of this choice are commented on in section 2.6. The derivation of sensitivity indices with PC emulation is largely based on published material, which we elaborate on here in order to introduce the notation leading to the final equation (19). To our knowledge, the latter this has not been given elsewhere. The reader mainly interested in the emulator performance and climatological analysis may immediately jump to section 3.

2.4.2 Principal components decomposition

Let $\mathbf{Y}$ denote the matrix in which each column represents the output of one experiment, i.e., $\mathbf{Y} = [\boldsymbol{y}(\boldsymbol{x}^{(1)}), \ldots, \boldsymbol{y}(\boldsymbol{x}^{(n)})]$, and $\mathbf{Y} = [\boldsymbol{f}(\boldsymbol{x}^{(1)}), \ldots, \boldsymbol{f}(\boldsymbol{x}^{(n)})]$, where $\boldsymbol{x}^{(j)}$ is the input of experiment j. For example, if we want to emulate surface annual mean temperature, y annual

mean surface temperature, $\mathbf{f}$ is a vector of $p = 2048$ components (the number of grid points). Denote ~~$\bar{\mathbf{Y}}$, the vector~~ the p-vector of row averages of $\mathbf{Y}$ ~~. The~~ by $\bar{\mathbf{Y}}$. We now define the centered matrix $\mathbf{Y}^\star = \mathbf{Y} - \bar{\mathbf{Y}}\mathbf{1}_p'$, with $\mathbf{1}_p$ a vector of length p with all components equal to 1, and consider the singular value decomposition (SVD) ~~of the centered matrix $\mathbf{Y}^\star = \mathbf{Y} - \bar{\mathbf{Y}}$ is~~ $\mathbf{Y}^\star = \mathbf{UDV}'$, where $\mathbf{D}$ is ~~diagonal~~a diagonal matrix, and $\mathbf{U}$ and $\mathbf{V}$ are square ~~and orthonormal~~ orthonormal matrices. The columns of $\mathbf{U}$ represent the basis vectors, $\{\boldsymbol{u}_k\}$, and the projection coefficients are given by $\mathbf{VD}$. For the j^{th} experiment, the coefficient for the k^{th} basis vector is $a_k(\boldsymbol{x}^{(j)}) = \mathbf{V}_{j,k}\mathbf{D}_{kk}$. Wilkinson (2010) keeps the first ~~n' eigenvectors~~ ℓ eigenvectors only (ordered by decreasing eigenvalues~~only~~). This is sometimes known as 'hard-thresholding regularization' (Silverman, 1996), and results in a reduced-order model:

$$\boldsymbol{y}\boldsymbol{f}(\boldsymbol{x}) \approx \sum_{k=1}^{n'\ell} a_k(\boldsymbol{x})\boldsymbol{u}_k.$$

Thus, ~~one assumption of our emulator is that~~ for any input $\boldsymbol{x}$, the ~~output is expected to~~ emulator output prediction will lie in the space spanned by the ~~$\{\boldsymbol{u}_1 \dots \boldsymbol{u}_{n'}\}$.~~ $\{\boldsymbol{u}_1 \dots \boldsymbol{u}_\ell\}$. The error associated with hard-thresholding is accounted for in the covariance of the estimator (see section 2.5).

2.4.3 Emulation of PC scores

We then need to predict the $a_k(\boldsymbol{x})$ based on the experiment output $a_k(\boldsymbol{x}_i)$. This is done by considering ~~n'~~ ℓ Gaussian process (GP) models (Rasmussen and Williams, 2005)~~, following the Bayesian treatment given by Oakley and O'Hagan (2002) , which has become a standard for climate model emulation~~. We suppose the prior mean of each GP is ~~$\boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}$~~$\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\beta}_k$, where $\boldsymbol{h}(\boldsymbol{x})$ is a vector of q ~~a priori~~ known regression functions and ~~$\boldsymbol{beta}$~~ $\boldsymbol{\beta}_k$ is the vector of corresponding regression coefficients. This prior mean is conditional on parameters $\boldsymbol{\beta}_k$, which will then need to be estimated. For the present application, $\boldsymbol{h}(\boldsymbol{x})$ is simply defined as $(1, x_1, x_2, x_3)'$ (linear regressors), ~~such~~ so that $\boldsymbol{h}(\boldsymbol{x})'\boldsymbol{\beta}_k = 1 + \beta_{k,1}x_1 + \beta_{k,2}x_2 + \beta_{k,3}x_3$.

~~Define, then :~~ We then define

- H: the design matrix which has row i equal to the regressors ~~$\boldsymbol{h}(\boldsymbol{x}_i)'$~~$\boldsymbol{h}(\boldsymbol{x}_i)$.

– ~~$c_k(\boldsymbol{x}, \boldsymbol{x}^\star)$~~ $\boldsymbol{y}_k = (a_k(\boldsymbol{x_1}), \ldots, a_k(\boldsymbol{x_n}))'$, the vector of PC scores we wish to emulate.

– $c_k(\boldsymbol{x}, \boldsymbol{x}^\star) = \text{cor}(a_k(x), a_k(x^*))$: the correlation function ~~defined between any two inputs,~~ for $a_k(\cdot)$. This is typically a monotone decreasing function of the distance between the two points. ~~Let~~ We let A_k be the ~~Gram matrix~~ $n \times n$ Gram matrix, with $A_k[i,j] = c_k(\boldsymbol{x}_i, \boldsymbol{x}_j)$~~, and $\boldsymbol{y}_k = (a_k(\boldsymbol{x_1}), \ldots, a_k(\boldsymbol{x_n}))'$~~.

~~Under the assumed GPmodel for the PC scores, a_k~~If each $a_k(\cdot)$ is modelled as a GP, then we have a Gaussian prior predictive distribution for $\boldsymbol{y}$ ~~, with $\boldsymbol{y}_k \sim N(H\boldsymbol{\beta}_k, \sigma^2 A_k)$~~conditional on $\boldsymbol{\beta}_k$ and σ_k, with $\boldsymbol{y}_k \sim N(H\boldsymbol{\beta}_k, \sigma_k^2 A_k)$.

The interpretation of this model is that the simulator response is the sum of a mean response function, expressed as the linear combination of regressions, and a stochastic component (a zero-mean Gaussian process) that absorbs deviations from the mean response. ~~Assuming the vague prior $(\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-2}$ proposed by Berger et al. (2001) and used by, e.g., Oakley and O'Hagan (20~~ ~~Bastos and O'Hagan (2009) , the~~

In order to estimate σ_k^2 and $\boldsymbol{\beta}_k$, we assume a conjugate non-informative prior distribution, i.e., $\pi(\boldsymbol{\beta}_k, \sigma_k^2) \propto \sigma_k^{-2}$ (Berger et al., 2001) . This allows σ_k^2 and $\boldsymbol{\beta}_k$ to be marginalised out of the analysis, resulting in a posterior distribution of the simulator output ~~follows~~ that is a Student$-t$ distribution with $n - q$ degrees of freedom, with mean and variance functions

$$m_k(\boldsymbol{x}) = \boldsymbol{h}(\boldsymbol{x})'\hat{\boldsymbol{\beta}}_k + \boldsymbol{t}_k(\boldsymbol{x})'\boldsymbol{e}_k, \quad \text{and} \tag{13}$$

$$\Sigma_k(\boldsymbol{x}, \boldsymbol{x}^\star) = \hat{\sigma}_k^2[c_k(\boldsymbol{x}, \boldsymbol{x}^\star) - \boldsymbol{t}_k(\boldsymbol{x})'A_k^{-1}\boldsymbol{t}_k(\boldsymbol{x}^\star) + \boldsymbol{p}_k(\boldsymbol{x})'(H'A_k^{-1}H)^{-1}\boldsymbol{p}_k(\boldsymbol{x}^\star)], \tag{14}$$

~~respectively, with~~ where

$$\hat{\sigma}_k^2 = \frac{1}{n-q-2}(\boldsymbol{e}_k' A_k^{-1}\boldsymbol{e}_k), \quad \hat{\boldsymbol{\beta}}_k = (H'A_k^{-1}H)^{-1}H'A_k^{-1}\boldsymbol{y}_k,$$

$\boldsymbol{t}_k(\boldsymbol{x})_j = c_k(\boldsymbol{x}, \boldsymbol{x}_j)$, $\boldsymbol{p}_k(\boldsymbol{x})' = \boldsymbol{h}(\boldsymbol{x})' - \boldsymbol{t}_k(\boldsymbol{x})'A_k^{-1}H$, and $\boldsymbol{e}_k = A_k^{-1}(\boldsymbol{y}_k - H\hat{\boldsymbol{\beta}}_k)$.

We primarily ~~used~~ use the squared exponential covariance function ~~for c in this work,~~ with a nugget term ~~,~~ for each c_k (as discussed at length in Andrianakis and Challenor (2012)~~:~~):

$$c_k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp[-(\boldsymbol{x}_i'\Lambda_k^{-2}\boldsymbol{x}_j)] + \nu_k \mathbb{I}_{i=j}, \tag{15}$$

where ~~A~~ Λ_k is a scaling matrix chosen to be diagonal, and ν_k a "nugget" which we discuss shortly. The diagonal elements of Λ_k are commonly called the length-scales. ~~A popular alternative is the Matérn covariance function (Stein, 1999) , though it was not seen here to yield substantial improvement.~~ The parameters Λ_k and ν_k are then found by optimising a penalized likelihood $L_k^{(\mathrm{pen})}$ of the Gaussian process associated with the principal component k, as a function of Λ_k and ν_k. The expression of L_k is given in Andrianakis and Challenor (2012) :

$$\log L_k(\nu_k, \Lambda_k) = -\tfrac{1}{2}(\log(|A_k||H' A_k^{-1} H|) + (n-q)\log(\hat{\sigma}_k^2), \tag{16}$$

$$\log L_k^{(\mathrm{pen})}(\nu_k, \Lambda_k) = \log L_k(\nu_k, \Lambda_k) - 2\nu_k^2$$

$$(\boldsymbol{y}_k - H\hat{\boldsymbol{\beta}}_k)' A_k^{-2} (\boldsymbol{y}_k - H\hat{\boldsymbol{\beta}}_k)(\boldsymbol{y}_k - H\hat{\boldsymbol{\beta}}_k)'(\boldsymbol{y}_k - H\hat{\boldsymbol{\beta}}_k). \tag{17}$$

The role of the penalty is to guarantee smaller Gaussian process variances than would be obtained by least-squares regression.

The nugget term, $\nu_k \mathbb{I}_{i=j}$, was originally introduced to account for measurement errors in geospatial data analysis (Cressie, 1993). In emulators ~~, it~~ of deterministic systems, the nugget may also be ~~introduced and justified , either~~ justified as a regularisation *ansatz* to avoid poor matrix conditioning (Pepelychev, 2010) ~~,~~ or as a way to account for ~~non-explicitly specified inputs (in the present case: initial conditions, sampling time and length), or~~ the mis-specification in the correlation function (Gramacy and Lee, 2012) . In climate model applications, the nugget may also be justified as a way to account for ~~the mis-specification in the correlation function (Gramacy and Lee, 2012) .~~ "internal variability". Indeed, the chaotic dynamics of the simulator are such that a particular climate average over a given time window, can be viewed as a stochastic quantity, even though the simulator is deterministic. In the climate modelling parlance, the effect is referred to as uncertainty associated with the internal simulator variability. For example, in Araya-Melo et al. (2015) , we found that our estimate of the nugget variance is consistent with the assumption that this term represents the uncertainty due to simulator variability. This is also the interpretation adopted by Williamson et al. (2014) (both studies use the climate model HadCM3). In the present application, we use rather long climate averages (500 years) and we anticipate that internal variability will be of the same order as the error associated with the

truncation of the principal components. It may thus not be appropriate to interpret ν_k as an indicator of internal variability, and we therefore chose not to do so.

~~There is no universal recommendation for choosing Λ_k and ν_k, and an analytical treatment of priors and posteriors is intractable. Here, we chose to maximise a Gaussian process *penalized* likelihood $L_k(\Lambda_k, \nu_k)$, the expression of which is given in Andrianakis and Challenor (2012) . More precisely~~Finally, two options ~~are considered~~ will be considered to estimate the different Λ_k and ν_k associated with the different principal components: (1) maximise L_k independently for each k, or (2) use the same parameters for all k, i.e., ~~$\Lambda_k = L$~~ $\Lambda_k = \Lambda$ and $\nu_k = \nu$ and optimise ~~$\sum_{k=1}^{n'} L_k$~~$\sum_{k=1}^{\ell} L_k$. Although option (2) does *not* maximise the likelihood of the emulator taken as a whole, it presents computational benefits in the context of global sensitivity analysis as we show next. Whichever option is used, the Λ_k and ν_k are, once chosen, considered as known. In particular, the means and variances given in equations (13) and (14) are conditional on the values of these parameters.

2.5 Recombination of PC scores

The ~~posterior~~ emulator posterior distribution for predictions of LOVECLIM's outputs $\boldsymbol{f}(\boldsymbol{x})$~~then follow~~ , then follows a distribution with mean and ~~co-variances given by :~~ covariances functions given by

$$\boldsymbol{m}(\boldsymbol{x}) = \sum_{k=1}^{n'\ell} m_k(\boldsymbol{x})\boldsymbol{u}_k \tag{18}$$

$$\Sigma(\boldsymbol{x}, \boldsymbol{x}^\star) = \underbrace{\sum_{k=1}^{n'\ell} \Sigma_k(\boldsymbol{x}, \boldsymbol{x}^\star)\boldsymbol{u}_k\boldsymbol{u}_k'}_{\Sigma^{(gp)}} + \underbrace{\sum_{k=n'+1}^{k=\ell+1} n\frac{\mathbf{D}_{kk}^2}{n}\boldsymbol{u}_k\boldsymbol{u}_k'}_{\Sigma^{(pc)}} \text{ ,~~respectively.~~}. \tag{19}$$

The covariance matrix of the emulator for LOVECLIM is thus of dimension $p \times p$ and provides information on the joint uncertainty of any two simulator outputs~~, if the product $\boldsymbol{u}_k\boldsymbol{u}_k'$ is interpreted as an outer-product~~.

~~Let us now derive the sensitivity indices . If the Gaussian process parameters ν and Λ are independent of the principal component k, then, similar to Oakley and O'Hagan (2002) , one may define~~

$$\underline{\mathcal{A}_i} \equiv \int_{\mathcal{X}_i \mathcal{X}_{\sim i} \times \mathcal{X}_{\sim i}} \iint \boldsymbol{h}(\boldsymbol{x}) \boldsymbol{h}(\boldsymbol{x}^\star)' \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i,$$

$$\underline{\mathcal{B}_i} \equiv \int_{\mathcal{X}_i \mathcal{X}_{\sim i} \times \mathcal{X}_{\sim i}} \iint \boldsymbol{t}(\boldsymbol{x}) \boldsymbol{h}(\boldsymbol{x}^\star)' \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i,$$

$$\underline{\mathcal{C}_i} \equiv \int_{\mathcal{X}_i \mathcal{X}_{\sim i} \times \mathcal{X}_{\sim i}} \iint \boldsymbol{t}(\boldsymbol{x}) \boldsymbol{t}(\boldsymbol{x}^\star)' \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i,$$

$$\underline{\mathcal{D}_i} \equiv \int_{\mathcal{X}_i \mathcal{X}_{\sim i} \times \mathcal{X}_{\sim i}} \iint c(\boldsymbol{x}, \boldsymbol{x}^\star) \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i,$$

$$\underline{\mathcal{E}_i} \equiv \int_{\mathcal{X}_i \mathcal{X}_{\sim i} \times \mathcal{X}_{\sim i}} \iint \boldsymbol{p}(\boldsymbol{x})' (H' A^{-1} H)^{-1} \boldsymbol{p}(\boldsymbol{x}^\star) \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i.$$

~~such that~~

$$V_i = \sum_{k=1}^{n'} \sum_{k^\star=1}^{n'} \left(\beta'_k \mathcal{A}_i \beta_{k^\star} + \beta'_k \mathcal{B}_i \boldsymbol{e}_{k^\star} + \boldsymbol{e}'_k \mathcal{B}_i \beta_{k^\star} + \boldsymbol{e}'_k \mathcal{C}_i \boldsymbol{e}_{k^\star} \right) \boldsymbol{u}_k \boldsymbol{u}'_{k^\star} + \sum_{k=1}^{n'} \hat{\sigma}_k (\mathcal{D}_i - tr(A_k^{-1} \mathcal{C}_i) + \mathcal{E}_i) \boldsymbol{u}_k \boldsymbol{u}'_k - \boldsymbol{C},$$

where the formula for C involves quantities $\mathcal{A}$, $\mathcal{B}$ etc. similar to the above but where $\mathcal{X}_{-i}$ is the whole domain and there is no integral over $\mathcal{X}_i$. Variance indices may now be obtained by plugging m_k and Σ_k (equations (13) and (14)) into equations (18) and (19), to obtain $\boldsymbol{m}$ and Σ, and then using these expressions in (7). Although these operations can be performed numerically, there is a computational advantage in processing the equations symbolically. Details are given in Appendices B and C.

If the emulators for the different principal components use different Gaussian process parameters, then the integrals $\mathcal{A}_i$ etc. need be to computed for all possible combinations $k, k^\star$, that is , for 10 PCs, an increase in computing cost by a factor of 45.

2.6 Short discussion of possible advantages over the independent emulator approaches approach

Now that the notation and relevant concepts have been introduced, the potential advantages and drawbacks of the PC emulator and the independent emulator approaches may briefly be summarised as follows:

- The PC emulation is based on the calibration of $n' \ell$ Gaussian process models per output variablefield (temperature, precipitation, GDD). We use $n' = 10$ $\ell = 10$ (justified below). We noted note that computational cost may be saved by using the same length-scales for all Gaussian processes, though in practice computational costs remain affordable even when using independently optimised length scaleslength-scales. Consequently, the impact of the same length-scale assumption may be assessed more easily than it would be if we used 2048 independent emulators (i.e., the number of grid points) for each output.
- The PC emulator implicitly assumes a covariance structure across the model outputs, equal to the covariance of the outputs of the experiment design. This may effectively reduce the emulator posterior variance, especially if experiment outputs are noisy (e.g.: short averages in a model with high inter-annual variability). However, this may also exacerbate the dependency of the analysis on the specific choice of experiment designAs the PC

emulation approach requires fewer emulators, more time can be spent for individually validating each of them.

- In some regions there may be only small variation in the simulated output as the input parameters change. If independent emulators are used for each grid cell, estimating the hyper-parameters for these cells can be difficult without applying some sort of parameter regularisation, and besides, the computational effort of building the emulator is unnecessary (as the output is constant). The global principal component emulator is therefore preferable in these situations, as these constant regions are automatically accounted for in the principal component variance decomposition.
- Finally, the PC emulator provides ~~co-variance~~ covariance indices between any two simulator outputs. It therefore allows us to analyse the spatial structure of the simulator response to individual and combined factors.

3 Results

3.1 Sensitivity to initial conditions

For all ensemble members but two (experiments 20 and 27), the runs with distinct initial conditions converged to the same output, modulo small variations that can be attributed to sampling variability (cf. supplementary material). Experiment 27 shows a higher-amplitude variability pattern, but clearly oscillates around one mean value and, as we will shorty see, this mean is correctly captured by the emulator. It is therefore kept without further discussion for all subsequent analyses. Experiment 20 ~~is~~ used the lowest configuration of obliquity (22°). This is ~~in fact~~ lower than any ~~actual~~ obliquity that occurred during the Pleistocene (22.07° following Laskar et al. (2004)). In this configuration, LOVECLIM develops a slow oscillation pattern that may be reminiscent of Dansgaard-Oeschger oscillations: millennial transitions between a warm and a cold North Atlantic phase, with fast warming and slow cooling (Figure 2). The phenomenon is a known feature of LOVECLIM (Goosse et al. (2002); Loutre et al. (2014)). It can be described

as the apparition of a cold North-Atlantic phase that is being visited stochastically and increasingly frequently as obliquity decreases. This cold phase is being visited shortly once during an entire (additional) experiment at obliquity of 22.5 ° (lowest 7th percentile of Pleistocene obliquities). The obliquity threshold below which such events may occur is likely to depend on the value of precession. The oscillation itself could be of physical relevance for past climate variability, but the limits of the phase space region in which the oscillation occurs are also likely to depend on many other uncertain parameters of the model. One possible action would be to use a sequential design strategy to delineate the region of occurrence of the phenomenon and develop an emulator ~~specifically~~ aimed at characterising this oscillation. In particular, history matching theory provides adequate concepts and methods to this end (Williamson et al., 2013) . Given the likely sensitivity of the oscillation on model parameters, the significance of this enterprise for palaeoclimate interpretation is ~~unsure. We rather~~ unclear. We choose to ignore the experiment for the time being (the following diagnostics ignore experiment 20), but discuss the possible consequences of this ~~choice~~ decision in the final discussion. In statistical terms, we provisionally condition the analysis on the hypothesis that these oscillations do not occur in the phase space.

3.2 Validation and choice of PC emulator

We concentrate on three ~~outputs~~output fields: annual precipitation, growing degree days (GDD), and annual mean temperature. GDD is defined here as the annual sum of daily temperatures (in Celsius) exceeding 0°C. It is used as a calendar-independent indicator of summer intensity and length in extra-tropical regions~~. Namely,~~ , and the vegetation model VECODE, used in LOVECLIM, uses GDD and annual precipitation to predict the dynamics of vegetation (Brovkin et al., 1997). For simplicity, GDD is estimated ~~here~~ from monthly means, that is, 30 ~~×~~ times the sum of monthly mean temperatures for which this temperature is above zero. In equatorial and subtropical regions this information is equivalent to annual mean temperature. Furthermore, we use the logarithm of annual precipitation rather than precipitation, as the former is closer to being Gaussian distributed than the latter.

The decomposition in principal components is effective, with 99% of the variance on average over all grid points captured by the first 4 (annual temperature) to 8 principal components (Figure 3). As discussed in the methodology section, two options are considered for the estimation of the ~~scale lengths and nugget~~ length-scales and nugget variance Λ and ν~~:~~. We first attempt to use different correlation parameters obtained by maximising the penalised likelihood for each ~~of principal component~~ principal component emulator, independently of the others. It is then observed that the Gaussian process likelihood decreases with the index of the PC (Figure 4). This is a natural result if we think of the fact that as the index of the PC increases, its spatial pattern becomes more noisy and dependent on idiosyncrasies of the analysis such as the specific experiment design, experiment length and initial conditions. They are thus less informative ~~of~~ about the model itself, and scores are more difficult to predict with a smooth Gaussian process. The likelihood stabilises around PC #10 to a minimum value that indicates that the calibrated GP is not more informative than assuming independence of outputs on inputs. We therefore use ~~$n' = 10$~~$\ell = 10$. The alternative approach consists in using the same correlation parameters for all ~~PC~~PCs, in which case they are found by maximising the product of Gaussian process likelihoods.

Our evaluation strategy is based upon the leave-one-out cross-validation approach: for each member of the experiment design, a PC emulator is trained using the remaining design runs (using the correlation parameters ~~defined~~ estimated above). The means and standard deviations of the resulting emulator are then found for the design member left out. Figure 5 shows (bars) the number of grid points correctly predicted within the central 66th, 95th and 99th ~~credibility~~ credibile intervals. A well-calibrated emulator would~~thus get~~ , on average, correctly predict 66, 95 and 99 % of the points, respectively, in each category. ~~Predictions off the 99th credibility interval may be considered as "incorrect". Some specific remarks~~Some ramarks are in order:

1. Based on this ~~diagonstic~~ diagnostic only, using constant rather than PC-specific correlation parameters does not significantly affect the overall performance. This is explicitly shown for GDD, but ~~this~~ is also true of the other fields.

2. However, all fields exhibit an excessive number of ~~incorrect predictions , compared to the ideal frequency of 1 %~~predictions outside the central 99th credible interval. Annual mean

precipitation is, in this respect, less well predicted than the others, perhaps not surprisingly, given that precipitation responds less straightforwardly to insolation changes than temperature.

3. There ~~is~~ are an excessive number of predictions within the 66th central credible interval.

3.3 Astronomically-forced variance vs other effects

The total variance V resulting from the astronomical forcing can be estimated from eq. (7). Here we compare the estimated total variance V with other effects that may broadly be described as sources of uncertainties on this quantity (Figure 6). ~~Note that for~~ For the assessment specific to this subsection we considered the uniform distribution $\rho(x) = 1$ over the cube in order to be able to provide analytical integrals, and hence isolate the effects associated with Monte-Carlo sampling. The following observations can be made:

- The variance associated with the input factors largely dominates other sources of uncertainty.
- The error caused by the Monte-Carlo approximation of the integrals, estimated by comparing these Monte-Carlo integrals with the analytical solution in the particular case of ~~the uniform distributions of~~ ρa uniform distributions for $\boldsymbol{x}$ ($\rho(\boldsymbol{x}) \propto 1$), is of the order of 0.5% of the total variance.
- Slightly different variance estimates are obtained depending on whether we use the same length scales for all PC or not, or whether we use independent emulators or the PC emulator. The difference over a grid point is on average 2.5 % of the mean grid-point variance. ~~Note that different estimates will be also~~ Different estimates will also be obtained with the independent emulators over all grid-boxes depending on the length scales being used; only one length-scale was tested here. The ~~supplemental~~ supplementary material further shows that the *patterns* of the global sensitivity indices are similar whichever emulator is used, so that this choice is of no consequence for the scientific interpretation.

– The term “GP var" explicitly refers to ~~the~~ S_{tot} defined in (11) : shown here is the mean of the diagonal of this matrix, equal to the mean of the Gaussian Process variance over all grid points, averaged over the input space. Again this is a small term, which is of the same order of magnitude as the grid-box mean uncertainty associated with the choice of initial conditions.

– ~~Finally, we refer to *synergy* as the~~ The absolute value of the synergy term, measured as the difference between the total variance and the sum of the mean sensitivity indices ~~. This difference is associated with second-order effects. Perhaps surprisingly (see concluding discussion), this term is small,~~ is also of the same order as the different sources of uncertainty just discussed.

3.4 Grid-point-wise variance analysis

The variance indices over the different grid points are the diagonal elements of the matrices V, ~~$T_{\varepsilon\varpi}$~~ $T_{\{\varepsilon\varpi\}}$ and T_{ε}. They provide essentially the same information as could be obtained using emulators independently calibrated over all grid boxes. Namely (see Figure 7):

1. Precipitation is mainly controlled by precession over Western Africa and Australia. This is expected given the known control of precession on monsoon dynamics (e.g. Zhao et al., 2007). The absence of large variance patterns in South-East Asia and South America is probably to be attributed to the limitations of LOVECLIM in simulating tropical weather systems. Note also the significant influence of obliquity in the most western part of North Africa.

2. GDD exhibits distinctive responses across the hemispheres. While this quantity is controlled in the northern hemisphere by both precession and obliquity (precession dominates), southern ocean temperature is almost exclusively controlled by obliquity.

3. Annual mean temperature has the highest variance near the poles. Again, northern hemisphere temperature is equally controlled by precession and obliquity, while obliquity dominates the variance in the southern hemisphere.

3.5 Fingerprint analysis

The PC emulator, however, allows us to go one step further than the independent emulator strategy. As the full covariance matrices V and T_i are available, linear fingerprints may be obtained by performing a SVD of these covariance matrices (Figure 8). Specifically, we refer to "fingerprints" of precession and obliquity, respectively, the eigenvectors of ~~$T_{e\varpi}$~~ $T_{\{e\varpi\}}$ and T_ε. The first fingerprint of obliquity explains more than 90 % of the variance of all three variables considered here. Precession aggregates two inputs ($e\sin\varpi$ and $e\cos\varpi$). It can therefore be expected to have at least a second significant fingerprint. This is the case, but this second component represents less than 30 % of the variance. We come to that shortly. Compared to the point-wise variance analysis above, the main advantage of the fingerprint analysis is to provide information on the in-phase or anti-phase relationships between climate variables, namely

- Obliquity produces in-phase effects on monsoon-related precipitation both in the northern and in the southern hemisphere (compare Africa and, e.g., northern Australia or southern America), while precession causes opposite-phase effects. This pattern is easily explained by reference to insolation. Indeed, obliquity causes in-phase responses of summer insolation in both hemispheres, while precession causes anti-phased responses.

- Obliquity produces an equator-pole see-saw response of annual mean temperature, with, however, a weaker-amplitude response in the equator than in the extratropical regions. Again, this consistently reflects the pattern of annual mean insolation (Loutre et al., 2004).

Note also that fingerprints of precession and obliquity are not orthogonal, and thus cannot be readily recovered by principal component analysis of the model outputs. For reference, we estimated the variance decomposition of the PC scores associated with precession and obliquity (decomposition and maps of principal components available in supp. mat.). The variance analysis reveals a mixture of precession and obliquity effects on each principal component.

Coming back to precession, we expect the simulated response phase to differ from place to place. To illustrate this point, we plot the emulated precipitation as a function ~~on~~ of the longitude of the perihelion for three points along the African monsoon flow. We assume obliquity and

eccentricity typical of the Holocene (Figure 9), and indicatively denote longitudes of perihelion corresponding to the time of the beginning of the Holocene (11,000 years ago) , as well as that of 6,000 years ago, a reference period used for model intercomparison exercises (Braconnot et al., 2007). The timing of the maximum response is gradually shifted towards a late phase response as one travels northwards. This observation can be explained by considering the seasonal development of monsoon dynamics, along with the course of the zenithal sun. According to this analysis, the most favourable epoch for a "Green Sahara~~"~~" experiment with a global climate model would therefore be around 9,000 years, corresponding to the choice of early modelling experiments on this subject (Street-Perrott et al., 1990).

3.6 Detection of fast changes

Assuming that the climate system responds fast enough to changes in astronomical forcing, the time evolution of the climate system may in principle be simulated by forcing the emulator with a realistic evolution of astronomical forcing. Unfortunately, the output cannot be readily compared to observations because we are neglecting here other significant forcing elements, such as changes in land ice cover and greenhouse gas concentrations (Araya-Melo et al. (2015), as well as N. Bounceur, thesis in ~~preperation~~preparation, take this into account). This exercise may~~however~~ , however, help us to detect regions where, potentially, the climate system may respond with steep gradients to the smooth astronomical forcing changes in interglacial conditions. Specifically, we estimate the maximum rate of change of a climate variable, expressed in terms of units per ky (Figure 10). In order to enhance the palaeoclimatological interest of this discussion, we compare the experiments ~~,~~with interactive vegetation, with those with fixed land-surface properties, hereafter referred to as VOFF. As a reading guide, a climate variable responding linearly and exclusively to precession with standard deviation 1, would show a maximum rate of change of $0.82\,\mathrm{ky}^{-1}$; a variable responding ~~linearity~~ linearly and exclusively to obliquity with standard deviation 1 would show a maximum rate of change of $0.35\,\mathrm{ky}^{-1}$. On this basis, comparing Figure 10 with 7 allows us to detect two regions of potentially rapid changes:

The western Sahara: Rate of changes expressed on the log scale are of the order of 0.4 per thousand years, that is, about 50 % precipitation change in 1,000 years. This is a well-known feature explained by feedbacks between vegetation and climate (Brovkin et al., 1998), discussed specifically in LOVECLIM by Renssen et al. (2003),

The American sector of the Arctic~~,~~ including both northern Canada and sea-ice covered regions: rates of temperature changes are of the order of 3 $^{\circ}C$ per thousand years.

These *hotspots* almost disappear with the fixed-land-cover scheme, underlining the role of vegetation response. However, the fact that the North-American hotspot region extends over the Arctic ocean suggests also a role for sea-ice cover, which further amplifies vegetation-induced effects. Such interactions were between Arctic vegetation and sea-ice were already suggested to have played a role in Arctic climate change during the Holocene (Ganopolski et al., 1998).

4 Discussion

Let us first review and comment upon the results of palaeoclimate significance presented here. Naturally, they are conditional on the use of the specific simulator considered here (LOVECLIM) and must be considered critically given that LOVECLIM is an imperfect representation of reality:

1. Precession and obliquity both contribute to annual temperature. Precession generally has a greater effect in the northern hemisphere and tropical regions, and obliquity is the dominant forcing in the southern hemisphere. The fact that obliquity has proportionally more influence on southern than northern annual mean temperature in the extra-tropical latitudes is physically reasonable. Indeed, precession does not affect annual mean insolation, but it may affect the annual mean climate by acting on the seasonal cycle of albedo associated ~~associated~~ with sea-ice, snow, and vegetation feedbacks. The latter two naturally depend on the presence of continental masses, which occupy a larger fraction of the northern than the southern hemisphere. ~~In turn, this result is~~ This dichotomy between southern and

northern hemisphere responses was previously noted by Yin and Berger (2012) , based on LOVECLIM simulations of previous interglacial periods, who then referred to it as one of the elements needed to explain the occurrence of the "Mid-Brünhes event" (Yin, 2013) . It is also consistent with observations, ~~in particular~~namely:

- the prominence of obliquity signals in southern hemisphere records, and more particularly Antarctic cores, be it CO_2 concentration (Petit et al., 1999; Siegenthaler et al., 2005; Luethi et al., 2008) or deuterium excess (Vimeux et al., 2002),
- the contrasting dynamics between southern records and ~~a~~ northern continental records, such as Baïkal's, during isotopic stage 11 (Prokopenko et al., 2002).

~~It is also consistent with the proposed explanation of the mid-Brünhes event (Yin, 2013) .~~

2. GDD is used here as a measure of summer length and intensity. We considered it because it is used in VECODE as a predictor for vegetation changes. This quantity is also mathematically equivalent to the positive-degree days index (PDD) used as a predictor of net snow accumulation balance over ice sheets (e.g.: Pollard and DeConto (2005)). We see here that GDD is, in the northern hemisphere, approximately equally sensitive to precession and obliquity. Crucifix (2011), based on Berger (1978a), noted that the Milankovitch's caloric season insolation is also equally sensitive to precession and obliquity. Hence, this result is consistent with Ruddiman (2007)'s proposal to use caloric season insolation as a predictor for ice age inception.

3. We find a fairly strong obliquity effect on North African precipitation. This has been noted before, in particular by Tuenter et al. (2003) (using the same atmosphere model as LOVECLIM, but a different ocean model) and Bosmans et al. (2015) (using a higher resolution model called EC-EARTH). Both studies emphasise the global character of the climate response to obliquity and interpret the response of North African preciptiation pattern to changes in large-scale atmospheric circulation patterns.

4. Vegetation feedbacks substantially increase the climate change rate in the Arctic and in the Sahel. The response is slightly non-linear (best seen, e.g., on Figure 9), but not to the point of generating multiple steady states in response to a same astronomical configuration. Namely, all 27 experiments of the design converge to the same state (the particular case of exp. 20 is discussed in a next point). This is consistent with previous transient simulations with this model (Renssen et al., 2003), but we note that Brovkin et al. (1998) reports multi-stability of the Western-Sahara in the Early Holocene, leading to a bifurcation associated with the abrupt desertification of the Sahara during the mid-Holocene (Claussen et al., 1999). It remains unclear whether possible multiple stable states persist when more sophisticated land-surface-vegetation schemes associated with finer spatial resolution are used (see, e.g., Kleidon et al. (2007); Dekker et al. (2010)). On the other hand, model intercomparisons support the existence of a single stable state in the high latitudes (e.g. (Brovkin et al., 2003)). .

5. Our methodological approach allows us to document response phases of precipitation patterns associated with the African monsoon. In particular, we found that the northward penetration of African monsoon is at a maximum when the perihelion is reached in August. Indeed, in this configuration, average levels of spring insolation can prevent excessive warming of the Ocean surface during this season, while high positive June-July-August insolation anomalies effectively enhance the warming of the subtropical continent. This combination maximises the contrast between ocean and continental temperature during the monsoon season. Again, the result needs to be qualified with the usual cautionary remarks about the simplification of tropical dynamics in a model like LOVECLIM.

6. Instabilities of the North Atlantic circulation develop at very low obliquities. The effects of these instabilities have not been taken into account in the variance diagnostics discussed here. We note in particular that obliquities as low as those necessary to trigger the oscillations would, in the real world, also be associated with the development of ice sheets. The latter could further complicate the dynamics in the North Atlantic region. We therefore limit ourselves to observe that such oscillation dynamics, if they were to occur in

specific interglacial configurations, would dominate the astronomical sources of variance examined here.

A perhaps more surprising result of our analysis is the smallness of the synergy terms. Crucifix and Loutre (2002) outlined the significance of a synergy between precession and ~~obliqutiy~~ obliquity (Crucifix and Loutre, 2002) during the last interglacial period. They noted that effects of precession and obliquity may combine and produce non-linear effects associated with the taiga-tundra transition in the Arctic area, not incompatible with what we find here about rates of changes in the Canadian Arctic. We explain this paradox by observing that metrics provided by global sensitivity analysis are aimed at determining whether the total variances, assessed over the whole Pleistocene, ~~would~~ add up linearly or not. ~~With LOVECLIM, the answer turns out to be yes, even if~~ In LOVECLIM, total variances indeed add up roughly linearly. This does not exclude that non-linear effects may episodically dominate at critical periods.

5 Conclusions

We presented a global sensitivity analysis of the effects of astronomical forcing on the climate model LOVECLIM in interglacial conditions. The work ~~is based on three experiment plans of 27 experiments each, and we~~ rely on the methodology of PC-Gaussian process emulation to explore the input space and deliver spatially resolved variance indices . In particular, we introduce the fingerprints as the eigenvectors of the covariance indices obtained from global sensitivity analysis to provide a spatial description of the effects of the individual factors. From a palaeoclimatological prospective, the results shown here are broadly consistent with the current understanding of Earth's climate response to the astronomical forcing. Compared to standard approaches based on a small number of simulations for well-defined past epochs, the methodology presented here allows us to identify more systematically regions susceptible ~~of~~ to experiencing rapid climate change in response to the smooth astronomical forcing changes, and examine the response phase of climate change to precession. We do not have to rely on transient experiments, hence the methodology is readily applicable to more complex climate models, but

we have to rely on the assumption of quasi-stationarity of the climate response to the astronomical forcing. Although the mathematics are fairly straightforward, the emulation approach requires considerable care in its implementation. We discussed the ~~effects~~ effect of the choice of ~~length scales~~length-scales, the type of emulator, and estimated uncertainties associated to specific computational aspects such as the Monte-Carlo estimates of integrals, to conclude that the PC emulator is a reasonable option. We therefore recommend its use for ~~further~~ future applications.

Appendix A: Experiment design algorithm

The following algorithm was used to generate the experiment design.

1. The three factors are first standardised so that they cover the ranges $[-1,1]$, i.e., we use $i_1 = e\sin\varpi/0.05$, we use $i_2 = e\cos\varpi/0.05$, and ~~$i_3 = \varepsilon - 23.5)/1.5$~~$i_3 = (\varepsilon - 23.5)/1.5$. Let $\mathbf{x} = (x_1, x_2, x_3)$ be a point of the input space, where the x_i are the three factors.
2. Set $d_{m0} = 0$, and $\det_{init} = 0$.
3. Sample a Latin hypercube design of $N = 27$ points in the 3-dimensional cube $[-1,1]^3$, as follows:
   (a) Divide the interval $[-1,1]$ into N equal-width intervals and number the middle of each interval. Let $x(i)$ be the midpoint of the i^{th} interval.
   (b) Generate three random permutations of $i = 1...N$, denoted $\boldsymbol{n}_j = \{n_{i,j}\}_{i=1}^{N}$ for $j = 1,2,3$.
   (c) Form the ~~design~~ matrix $\mathbf{X}$, with $X_{i,j} = x\left(n_{i,j}\right)$
4. For every point of the design, check if the constraint $i_1^2 + i_2^2 < 1$ is verified (i.e., eccentricity smaller than 0.05). If not, omit the point and keep the design with N^* simulations
5. Sample a Latin hypercube of $N - N^*$ points in $[-1,1]^3$ and augment the design $\mathbf{X}$

6. Repeat $(4-5)$ until the dimension of the design $\mathbf{X}$ equals N
7. Calculate the minimum distance d_m between any two points of the design $\mathbf{X}$. If $d_m > d_{m0}$:
   (a) Set $d_{m0} = d_m$.
   (b) Set $\tilde{\mathbf{X}} = \mathbf{X}$.
   (c) If $\det(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}) > \det_{init}$
      i. Set $\det_{init} = \det(\hat{\mathbf{X}}'\hat{\mathbf{X}})$.
      ii. Set $\tilde{\mathbf{X}} = \hat{\mathbf{X}}$.
8. Repeat $(3-7)$ 1000 times. Keep the design $\tilde{\mathbf{X}}$.

Appendix B: ~~Monte-Carlo estimates of integrals~~ Analytical expressions for variance indices

If the Gaussian process parameters ν and Λ are independent of the principal component k, then, similar to Oakley and O'Hagan (2002) , one may define

$$\mathcal{A}_i \approx \int_{\mathcal{X}_i} \iint_{\mathcal{X}_{-i}\times\mathcal{X}_{-i}} \boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x}_{-i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}^\star_{-i}|\boldsymbol{x}^\star_i)\mathrm{d}\boldsymbol{x}_{-i}\mathrm{d}\boldsymbol{x}^\star_{-i}\mathrm{d}\boldsymbol{x}_i, \tag{B1}$$

$$\mathcal{B}_i \approx \int_{\mathcal{X}_i} \iint_{\mathcal{X}_{-i}\times\mathcal{X}_{-i}} \boldsymbol{t}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x}_{-i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}^\star_{-i}|\boldsymbol{x}^\star_i)\mathrm{d}\boldsymbol{x}_{-i}\mathrm{d}\boldsymbol{x}^\star_{-i}\mathrm{d}\boldsymbol{x}_i, \tag{B2}$$

$$\mathcal{C}_i \approx \int_{\mathcal{X}_i} \iint_{\mathcal{X}_{-i}\times\mathcal{X}_{-i}} \boldsymbol{t}(\boldsymbol{x})\boldsymbol{t}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x}_{-i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}^\star_{-i}|\boldsymbol{x}^\star_i)\mathrm{d}\boldsymbol{x}_{-i}\mathrm{d}\boldsymbol{x}^\star_{-i}\mathrm{d}\boldsymbol{x}_i, \tag{B3}$$

$$\mathcal{D}_i \approx \int_{\mathcal{X}_i} \iint_{\mathcal{X}_{-i}\times\mathcal{X}_{-i}} c(\boldsymbol{x},\boldsymbol{x}^\star)\rho(\boldsymbol{x}_{-i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}^\star_{-i}|\boldsymbol{x}^\star_i)\mathrm{d}\boldsymbol{x}_{-i}\mathrm{d}\boldsymbol{x}^\star_{-i}\mathrm{d}\boldsymbol{x}_i, \tag{B4}$$

$$\mathcal{E}_i \approx \int_{\mathcal{X}_i} \iint_{\mathcal{X}_{-i}\times\mathcal{X}_{-i}} \boldsymbol{p}(\boldsymbol{x})'(H'A^{-1}H)^{-1}\boldsymbol{p}(\boldsymbol{x}^\star)\rho(\boldsymbol{x}_{-i}|\boldsymbol{x}_i)\rho(\boldsymbol{x}^\star_{-i}|\boldsymbol{x}^\star_i)\mathrm{d}\boldsymbol{x}_{-i}\mathrm{d}\boldsymbol{x}^\star_{-i}\mathrm{d}\boldsymbol{x}_i. \tag{B5}$$

such that

$$V_i = \sum_{k=1}^{\ell}\sum_{k^\star=1}^{\ell}\left(\boldsymbol{\beta}'_k\mathcal{A}_i\boldsymbol{\beta}_{k^\star} + \boldsymbol{\beta}'_k\mathcal{B}_i\boldsymbol{e}_{k^\star} + \boldsymbol{e}'_k\mathcal{B}_i\boldsymbol{\beta}_{k^\star} + \boldsymbol{e}'_k\mathcal{C}_i\boldsymbol{e}_{k^\star}\right)\boldsymbol{u}_k\boldsymbol{u}'_{k^\star} + \sum_{k=1}^{\ell}\hat{\sigma}_k(\mathcal{D}_i - tr(A_k^{-1}\mathcal{C}_i) + \mathcal{E}_i)\boldsymbol{u}_k\boldsymbol{u}'_k - \boldsymbol{C}, \tag{B6}$$

where the formula for C involves quantities $\mathcal{A}_0$, $\mathcal{B}_0$ etc. similar to the above but where $\mathcal{X}_{-i}$ is the whole domain and there is no integral over $\mathcal{X}_i$. That is, we define:

$$\mathcal{A}_0 \approx \iint_{\mathcal{X}\times\mathcal{X}} \boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star, \tag{B7}$$

$$\mathcal{B}_0 \approx \iint_{\mathcal{X}\times\mathcal{X}} \boldsymbol{t}(\boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star, \tag{B8}$$

$$\mathcal{C}_0 \approx \iint_{\mathcal{X}\times\mathcal{X}} \boldsymbol{t}(\boldsymbol{x})\boldsymbol{t}(\boldsymbol{x}^\star)'\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star, \tag{B9}$$

$$\mathcal{D}_0 \approx \iint_{\mathcal{X}\times\mathcal{X}} c(\boldsymbol{x},\boldsymbol{x}^\star)\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star, \tag{B10}$$

$$\mathcal{E}_0 \approx \iint_{\mathcal{X}\times\mathcal{X}} \boldsymbol{p}(\boldsymbol{x})'(H'A^{-1}H)^{-1}\boldsymbol{p}(\boldsymbol{x}^\star)\rho(\boldsymbol{x})\rho(\boldsymbol{x}^\star)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{x}^\star. \tag{B11}$$

and

$$C = \sum_{k=1}^{\ell}\sum_{k^\star=1}^{\ell} \left(\boldsymbol{\beta}_k'\mathcal{A}_0\boldsymbol{\beta}_{k^\star} + \boldsymbol{\beta}_k'\mathcal{B}_0\boldsymbol{e}_{k^\star} + \boldsymbol{e}_k'\mathcal{B}_0\boldsymbol{\beta}_{k^\star} + \boldsymbol{e}_k'\mathcal{C}_0\boldsymbol{e}_{k^\star}\right)\boldsymbol{u}_k\boldsymbol{u}_{k^\star}' + \sum_{k=1}^{\ell}\hat{\sigma}_k(\mathcal{D}_0 - tr(A_k^{-1}\mathcal{C}_0) + \mathcal{E}_0)\boldsymbol{u}_k\boldsymbol{u}_k'. \tag{B12}$$

If the emulators for the different principal components use different Gaussian process parameters, then the integrals $\mathcal{A}_i$ etc. need to be computed for all possible combinations $k, k^\star$, that is, for 10 PCs, an increase in computing cost by a factor of 45.

Appendix C: Computation of integrals

Integrals (B1)–(B5) have the general form:

$$\int_{\mathcal{X}_i} \iint_{\mathcal{X}_{\sim i} \times \mathcal{X}_i} g(\boldsymbol{x}, \boldsymbol{x}^\star) \rho(\boldsymbol{x}_{\sim i} | \boldsymbol{x}_i) \rho(\boldsymbol{x}^\star_{\sim i} | \boldsymbol{x}^\star_i) \mathrm{d}\boldsymbol{x}_{\sim i} \mathrm{d}\boldsymbol{x}^\star_{\sim i} \mathrm{d}\boldsymbol{x}_i \tag{C1}$$

The integrals may be computed analytically in the particular case of $\rho = 1$ over the cube $[0,1]^3$ for the astronomical forcing (i.e., $\rho(x) = 1$ over the cube $[0,1]^3$), ~~and use the~~ assuming squared-exponential ~~(RDW) correlation function ,~~ correlation function and linear regressors, as used here. The computations involve the *erf* and the *incomplete-gamma* functions, which can be computed with the gsl scientific library (Galassi and Gough, 2009) made available in the "gsl" R package by R. Hankin. The more complex density function ρ associated with the actual course of astronomical forcing must be accounted for by means of a Monte-Carlo algorithm, based on Homma and Saltelli (1996) :

1. Sample q vectors, $\boldsymbol{x}^{(j)}$, $j = 1 \ldots q$ following the density $\rho(\boldsymbol{x})$, that is:
   - either the uniform distribution if we suppose $\rho(\boldsymbol{x}) = 1$,
   - or either by computing the history of astronomical forcing (e.g.: between -10^7 years to 0 by interval of 1000 years) with the astronomical solution of Berger (1978b),
2. Shuffle the vectors $\boldsymbol{x}$ to produce a sequence $\boldsymbol{x}^{\star(j)}$, $j = 1 \ldots q$
3. For every j, modify the components $[i]$ of $\boldsymbol{x}^{\star(j)}$ to make them equal to those of $\boldsymbol{x}^{(j)}$ i.e., $\boldsymbol{x}_i^{\star(j)} \leftarrow \boldsymbol{x}_i^{(j)}$.
4. The sum $\sum_{j=1}^{q} g(\boldsymbol{x}^{(j)}, \boldsymbol{x}^{\star(j)})$ converges to the integral (C1) as $q \to \infty$.

By comparison with analytical integrals we found that accurate estimates are obtained in this application for $q = 10,000$. In this case, all the computations of sensitivity indices needed for

the scientific discussion take about 10 minutes on a laptop at time of writing if all PC components use the same Gaussian process parameters ν and Λ, a couple of hours if parameters are different for each component.

We carried out the analytical computations to compare independent vs PC emulators, and also to have a reference for estimating the accurracy of our Monte-Carlo simulations. The Monte-Carlo simulations accounting for the true astronomical forcing distribution are used for the palaeoclimatological discussion.

Acknowledgements. Thanks are due to Jonathan Rougier (University of Bristol) , Carlos Almeida (Université catholique de Louvain) and Ioannis Andrianakis (London School of Hygiene and Tropical Medicine) for useful interaction in the preparation of the present work. MC is research associate with the Belgian National Fund of Scientific Research (FRS-FNRS). This research is a contribution to the ITOP project, ERC-StG grant 239604. Reviewers and editors also provided particularly helpful and constructive comments. Computational resources were provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the FRS-FNRS.

References

Alpert, P. and Sholokhman, T., eds.: Factor Separation in the atmosphere: applications and future prospects, Cambrdige University Press, Cambridge, UK, 2011.

Andrianakis, I. and Challenor, P. G.: The effect of the nugget on Gaussian process emulators of computer models, Computational Statistics & Data Analysis, 56, 4215–4228, doi:10.1016/j.csda.2012.04.020, 2012.

Araya-Melo, P. A., Crucifix, M., and Bounceur, N.: Global sensitivity analysis of the Indian monsoon during the Pleistocene, Climate of the Past, 11, 45–61, doi:10.5194/cp-11-45-2015, http://www.clim-past.net/11/45/2015/, 2015.

Bastos, L. S. and O'Hagan, A.: Diagnostics for Gaussian Process Emulators, Technometrics, 51, 425–438, doi:10.1198/TECH.2009.08019, 2009.

Berger, A.: Long-term variations of caloric insolation resulting from the earth's orbital elements, Quaternary Research, 9, 139 – 167, doi:10.1016/0033-5894(78)90064-9, 1978a.

Berger, A.: The role of CO_2, sea-level and vegetation during the Milankovitch-forced glacial-interglacial cycles, in: Proceedings "Geosphere-Biosphere Interactions and Climate", Pontifical Academy of Sciences, Vatican City, 9-13 November 1998., pp. 119–146, 1999.

Berger, A. L.: Long-term variations of daily insolation and Quaternary climatic changes, Journal of the Atmospheric Sciences, 35, 2362–2367, doi:10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978b.

Berger, J. O., De Oliveira, V., and Sansó, B.: Objective Bayesian Analysis of Spatially Correlated Data, Journal of the American Statistical Association, 96, 1361–1374, doi:10.1198/016214501753382282, 2001.

Bosmans, J. H. C., Hilgen, F. J., Tuenter, E., and Lourens, L. J.: Obliquity forcing of low-latitude climate, Climate of the Past Discussions, 11, 221–241, doi:10.5194/cpd-11-221-2015, http://www.clim-past-discuss.net/11/221/2015/, 2015.

Braconnot, P., Joussaume, S., Marti, O., and de Noblet, N.: Synergistic feedbacks from ocean and vegetation on the African Monsoon response to Mid-Holocene insolation, Geophysical Research Letters, 26, 2481–2484, doi:10.1029/1999GL006047, http://dx.doi.org/10.1029/1999GL006047, 1999.

Braconnot, P., Otto-Bliesner, B. L., Harrison, S., Joussaume, S., Peterschmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, T., Hewitt, C. D., Kageyama, M., Kitoh, A., Laîné, A., Loutre, M.-F., Marti, O., Merkel, U., Ramstein, G., Valdes, P., Weber, S. L., Yu, Y., and Zhao, Y.: Results of PMIP2 coupled simulations of the Mid Holocene and Last Glacial Maximum – Part 1:

experiments and large-scale features, Climate of the Past, 3, 261–277, doi:10.5194/cp-3-261-2007, www.clim-past.net/3/261/2007/, 2007.
Brovkin, V., Ganopolski, A., and Shvirezhev, Y.: A continuous climate-vegetation classification for use in climate-biosphere studies, Ecological Modelling, 101, 251–261, doi:10.1016/S0304-3800(97)00049-5, 1997.
Brovkin, V., Claussen, M., Petoukhov, V., and Ganopolski, A.: On the stability of the atmosphere-vegetation system in the Sahara/Sahel region, J. Geophys. Res., 103, 31 613–31 624, doi:10.1029/1998JD200006, http://dx.doi.org/10.1029/1998JD200006, 1998.
Brovkin, V., Levis, S., Loutre, M. F., Crucifix, M., Claussen, M., Ganopolski, A., Kubatzki, C., and Petoukhov, V.: Stability analysis of the climate-vegetation system in the northern high latitudes, Climatic Change, 57, 119–138, 2003.
Carnell, R.: lhs: Latin Hypercube Samples, http://CRAN.R-project.org/package=lhs, r package version 0.10, 2012.
Carslaw, K. S., Lee, L. A., Reddington, C. L., Pringle, K. J., Rap, A., Forster, P. M., Mann, G. W., Spracklen, D. V., Woodhouse, M. T., Regayre, L. A., and Pierce, J. R.: Large contribution of natural aerosols to uncertainty in indirect forcing, Nature, 503, 67–71, http://dx.doi.org/10.1038/nature12674, 2013.
Claussen*, M.: Late Quaternary vegetation-climate feedbacks, Climate of the Past, 5, 203–216, doi:10.5194/cp-5-203-2009, http://www.clim-past.net/5/203/2009/, 2009.
Claussen, M., Kubatzki, C., Brovkin, V., Ganopolski, A., Hoelzmann, P., and Pachur, H.-J.: Simulation of an abrupt change in Saharan vegetation in the mid-Holocene, Geophysical Research Letters, 26, 2037–2040, doi:10.1029/1999GL900494,, 1999.
Cressie, N.: Statistics for Spatial Data, Wiley series in probability and statistics, John Wiley & Sons, Inc., Chichester, UK., 1993.
Crucifix, M.: How can a glacial inception be predicted?, The Holocene, 21, 831–842, doi:10.1177/0959683610394883, 2011.
Crucifix, M. and Loutre, F.: Transient simulations over the last interglacial period (126-115 kyr BP): feedback and forcing analysis, Climate Dynamics, 19, 417–433, doi:10.1007/s00382-002-0234-z, 2002.
Cumming, J. A. and Goldstein, M.: Small Sample Bayesian Designs for Complex High-Dimensional Models Based on Information Gained Using Fast Approximations, Technometrics, 51, 377–388, doi:10.1198/TECH.2009.08015, http://dx.doi.org/10.1198/TECH.2009.08015, 2009.

de Noblet, N., Prentice, I. C., Joussaume, S., Texier, D., Botta, A., and Haxeltine, A.: Possible role of atmosphere-biosphere interactions in triggering the last glaciation., Geophysical Research Letters, 23, 3191–3194, 1996.

Dekker, S. C., de Boer, H. J., Brovkin, V., Fraedrich, K., Wassen, M. J., and Rietkerk, M.: Biogeophysical feedbacks trigger shifts in the modelled vegetation-atmosphere system at multiple scales, Biogeosciences, 7, 1237–1245, doi:10.5194/bg-7-1237-2010, http://www.biogeosciences.net/7/1237/2010/, 2010.

Draguljić, D., Santner, T. J., and Dean, A. M.: Noncollapsing Space-Filling Designs for Bounded Nonrectangular Regions, Technometrics, 54, 169–178, doi:10.1080/00401706.2012.676951, http://dx.doi.org/10.1080/00401706.2012.676951, 2012.

Fricker, T. E., Oakley, J. E., and Urban, N. M.: Multivariate Gaussian process emulators with nonseparable covariance structures, Technometrics, 55, 47–56, 2013.

Galassi, M. and Gough, B.: GNU Scientific Library: Reference Manual, GNU manual, Network Theory, http://books.google.be/books?id=CUuNPgAACAAJ, 2009.

Ganopolski, A., Kubatzki, C., Claussen, M., Brovkin, V., and Petoukhov, V.: The influence of vegetation-atmosphere-ocean interaction on climate during the mid-Holocene, Science, 280, 1916–1919, doi:10.1126/science.280.5371.1916, 1998.

Goosse, H., Renssen, H., Selten, F. M., Haarsma, R. J., and Opsteegh, J. D.: Potential causes of abrupt climate events: A numerical study with a three-dimensional climate model, Geophysical Research Letters, 29, 1860, doi:10.1029/2002GL014993, http://dx.doi.org/10.1029/2002GL014993, 2002.

Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F., Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M.-F., Morales Maqueda, M. A., Opsteegh, T., Mathieu, P.-P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, A., and Weber, S. L.: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2, Geoscientific Model Development, 3, 603–633, doi:10.5194/gmd-3-603-2010, 2010.

Gramacy, R. and Lee, H. H.: Cases for the nugget in modeling computer experiments, Statistics and Computing, 22, 713–722, doi:10.1007/s11222-010-9224-x, 2012.

Henrot, A.-J., François, L., Favre, E., Butzin, M., Ouberdous, M., and Munhoven, G.: Effects of CO_2, continental distribution, topography and vegetation changes on the climate at the Middle Miocene: a model study, Climate of the Past, 6, 675–694, doi:10.5194/cp-6-675-2010, 2010.

Higdon, D., Gattiker, J., Williams, B., and Rightley, M.: Computer Model Calibration Using High-Dimensional Output, Journal of the American Statistical Association, 103, 570–583, doi:10.1198/016214507000000888, 2008.

Holden, P., Edwards, N., Oliver, K., Lenton, T., and Wilkinson, R.: A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1, Climate Dynamics, 35, 785–806, doi:10.1007/s00382-009-0630-8, http://dx.doi.org/10.1007/s00382-009-0630-8, 2010.

Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, Reliability Engineering & System Safety, 52, 1–17, doi:10.1016/0951-8320(96)00002-6, 1996.

Joseph, V. R. and Hung, Y.: Orthogonal-maximin latin hypercube designs, Statistica Sinica, pp. 171–186, 2008.

Kageyama, M., Charbit, S., Ritz, C., Khodri, M., and Ramstein, G.: Quantifying ice-sheet feedbacks during the last glacial inception, Geophysical Research Letters, 31, doi:10.1029/2004GL021339, http://dx.doi.org/10.1029/2004GL021339, 2004.

Kleidon, A., Fraedrich, K., and Low, C.: Multiple steady-states in the terrestrial atmosphere-biosphere system: a result of a discrete vegetation classification?, Biogeosciences, 4, 707–714, doi:10.5194/bg-4-707-2007, http://www.biogeosciences.net/4/707/2007/, 2007.

Laskar, J., Robutel, P., Joutel, F., Boudin, F., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astronomy and Astrophysics, 428, 261–285, doi:10.1051/0004-6361:20041335, 2004.

Lee, L. A., Carslaw, K. S., Pringle, K. J., Mann, G. W., and Spracklen, D. V.: Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters, Atmospheric Chemistry and Physics, 11, 12 253–12 273, doi:10.5194/acp-11-12253-2011, http://www.atmos-chem-phys.net/11/12253/2011/, 2011.

Lee, L. A., Pringle, K. J., Reddington, C. L., Mann, G. W., Stier, P., Spracklen, D. V., Pierce, J. R., and Carslaw, K. S.: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmospheric Chemistry and Physics, 13, 8879–8914, doi:10.5194/acp-13-8879-2013, http://www.atmos-chem-phys.net/13/8879/2013/, 2013.

Loeppky, J. L., Sacks, J., and Welch, W. J.: Choosing the Sample Size of a Computer Experiment: A Practical Guide, Technometrics, 51, 366–376, doi:10.1198/TECH.2009.08040, 2009.

Loutre, M. F.: Paramètres orbitaux et cycles diurne et saisonnier des insolations, Ph.D. thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 1993.

Loutre, M. F., Paillard, D., Vimeux, F., and Cortijo, E.: Does mean annual insolation have the potential to change the climate?, Earth Planet. Sc. Lett., 221, 1–14, 2004.

Loutre, M. F., Fichefet, T., Goosse, H., Huybrechts, P., Goelzer, H., and Capron, E.: Factors controlling the last interglacial climate as simulated by LOVECLIM1.3, Climate of the Past, 10, 1541–1565, doi:10.5194/cp-10-1541-2014, http://www.clim-past.net/10/1541/2014/, 2014.

Luethi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., and Stocker, T. F.: High-resolution carbon dioxide concentration record 650,000-800,000 years before present, Nature, 453, 379–382, doi:10.1038/nature06949, 2008.

McKay, M. D., Beckman, R. J., and Conover, W. J.: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, Technometrics, 21, 239–245, doi:10.2307/1268522, 1979.

Meissner, K. J., Weaver, A. J., Matthews, H. D., and Cox, P. M.: The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model, Climate Dynamics, 21, 515–537, doi:10.1007/s00382-003-0352-2, http://dx.doi.org/10.1007/s00382-003-0352-2, 2003.

Morris, M. D. and Mitchell, T. J.: Exploratory designs for computational experiments, Journal of Statistical Planning and Inference, 43, 381 – 402, doi:10.1016/0378-3758(94)00035-T, http://www.sciencedirect.com/science/article/B6V0M-3YS8YHW-12/2/da5cccb75e7de917b8d6256ba1fa6772, 1995.

Oakley, J. and O'Hagan, A.: Bayesian Inference for the Uncertainty Distribution of Computer Model Outputs, Biometrika, 89, 769–784, doi:10.1093/biomet/89.4.769, 2002.

Oakley, J. E. and O'Hagan, A.: Probabilistic sensitivity analysis of complex models: a Bayesian approach, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66, 751–769, doi:10.1111/j.1467-9868.2004.05304.x, 2004.

Opsteegh, J., Haarsma, R., Smelten, F., and Kattenberg, A.: ECBILT: a dynamic alternative to mixed boundary conditions in ocean models, Tellus, 50A, 348–367, doi:10.1034/j.1600-0870.1998.t01-1-00007.x, 1998.

Pepelychev, A.: The role of the nugget term in the Gaussian process method, in: mODa 9 Advances in Model-Oriented Design and Analysis: In: Contributions to Statistics, edited by Giovagnoli, A., Atkinson, A. C., Torsney, B., and May, C., pp. 149–156, Physica Verlag (Springer), Heidelberg, 2010.

Petit, J. R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J.-M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V. M., Legrand, M., Lipenkov, V. Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., and Stievenard, M.: Climate and atmospheric history of the past 420, 000 years from the Vostok ice core, Antarctica, Nature, 399, 429–436, doi:10.1038/20859, 1999.

Pollard, D. and DeConto, R. M.: Hysteresis in Cenozoic Antarctic ice-sheet variations, Global and Planetary Change, 45, 9 – 21, doi:10.1016/j.gloplacha.2004.09.011, long-term changes in Southern high-latitude ice sheets and climate, the Cenozoic history, 2005.

Prokopenko, A., Williams, D., Kuzmin, M., Karabanov, E., Khursevich, G., and Peck, J.: Muted climate variations in continental Siberia during the mid-Pleistocene epoch, Nature, 418, 65–68, doi:10.1038/nature00886, 2002.

Rasmussen, C. and Williams, C.: Gaussian Processes for Machine Learning, Adaptive Computation And Machine Learning, MIT Press, Cambridge MA, 2005.

Renssen, H., Brovkin, V., Fichefet, T., and Goosse, H.: Holocene climate instability during the termination of the African Humid Period, Geophysical Research Letters, 30, n/a–n/a, doi:10.1029/2002GL016636, 2003.

Rougier, J.: Efficient Emulators for Multivariate Deterministic Functions, Journal of Computational and Graphical Statistics, 17, 827–843, doi:10.1198/106186008X384032, http://www.tandfonline.com/doi/abs/10.1198/106186008X384032, 2008.

Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D. A.: A probability distribution of climate sensitivity for the HadAM3 simulator, Journal of Climate, 2006.

Rougier, J., Sexton, D. M. H., Murphy, J. M., and Stainforth, D.: Analyzing the Climate Sensitivity of the HadSM3 Climate Model Using Ensembles from Different but Related Experiments, Journal of Climate, 22, 3540–3557, doi:10.1175/2008JCLI2533.1, http://journals.ametsoc.org/doi/abs/10.1175/2008JCLI2533.1, 2009.

Ruddiman, W. F.: The early anthropogenic hypothesis a year later: challenges and responses, Reviews of Geophysics, 45, RG4001, doi:10.1029/2006RG000207, 2007.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P.: Design and Analysis of Computer Experiments, Statistical Science, 4, 409–423, doi:10.1214/ss/1177012413, 1989.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M.: Sensitivity analysis in practite, Johx, W. Sussex, England, 2004.

Sanchez-Goñi, M. F., Eynaud, F., Turon, J. L., and Shackleton, N. J.: High resolution palynological record off the Iberian margin: direct land-sea correlation for the Last Interglacial complex, Earth and Planetary Science Letters, 171, 123–137, 1999.

Santner, T., Williams, B., and Notz, W.: The Design and Analysis of Computer Experiments, Springer, New York, 2003.

Schmittner, A., Urban, N. M., Shakun, J. D., Mahowald, N. M., Clark, P. U., Bartlein, P. J., Mix, A. C., and Rosell-Melé, A.: Climate Sensitivity Estimated from Temperature Reconstructions of the Last

Glacial Maximum, Science, 334, 1385–1388, doi:10.1126/science.1203513, http://www.sciencemag.org/content/334/6061/1385.abstract, 2011.

Siegenthaler, U., Stocker, T. F., Lüthi, D., Schwander, J., Stauffer, B., Raynaud, D., Barnola, J.-M., Fisher, H., Masson-Delmotte, V., and Jouzel, J.: Stable carbon cycle-climate relationship during the Late Pleistocene, Science, 310, 1313–1317, 2005.

Silverman, B. W.: Smoothed functional principal components analysis by choice of norm, The Annals of Statistics, 24, 1–24, doi:10.1214/aos/1033066196, 1996.

Stein, M.: Interpolation of Spatial Data: Some Theory for Kriging, Springer Series in Statistics, Springer, New York, http://books.google.be/books?id=5n_XuL2Wx1EC, 1999.

Stein, U. and Alpert, P.: Factor Separation in Numerical Simulations, Journal of the Atmospheric Sciences, 50, 2107–2115, doi:10.1175/1520-0469(1993)050<2107:FSINS>2.0.CO;2, 1993.

Street-Perrott, F. A., Mitchell, J. F. B., Marchand, D. S., and Brunner, J. S.errott, F. A.: Milankovitch and albedo forcing of the tropical monsoons: a comparison of geological evidence and numerical simulations for 9000 yBP, Earth and Environmental Science Transactions of the Royal Society of Edinburgh, 81, 407–427, doi:10.1017/S0263593300020897, http://journals.cambridge.org/article_S0263593300020897, 1990.

Tuenter, E., Weber, S. L., Hilgen, F. J., and Lourens, L. J.: The response of the African summer monsoon to remote and local forcing due to precession and obliquity, Global and Planetary Change, 36, 219–235, doi:http://dx.doi.org/10.1016/S0921-8181(02)00196-0, http://www.sciencedirect.com/science/article/pii/S0921818102001960, 2003.

Urban, N. M. and Fricker, T. E.: A comparison of Latin hypercube and grid ensemble designs for the multivariate emulation of an Earth system model, Computers & Geosciences, 36, 746–755, doi:10.1016/j.cageo.2009.11.004, 2010.

Vernon, I., Goldstein, M., and Bower, R. G.: Galaxy formation: a Bayesian uncertainty analysis, Bayesian Anal., 5, 619–669, doi:10.1214/10-BA524, http://dx.doi.org/10.1214/10-BA524, 2010.

Vimeux, F., Cuffey, K. M., and Jouzel, J.: New insights into Southern Hemisphere temperature changes from Vostok ice cores using Deuterium excess correction, Earth and Planetary Science Letters, 203, 829 – 843, doi:10.1016/S0012-821X(02)00950-0, 2002.

Wilkinson, R. D.: Bayesian Calibration of Expensive Multivariate Computer Experiments, in: Large-Scale Inverse Problems and Quantification of Uncertainty, edited by Biegler, L., Biros, G., Ghattas, O., Heinkenschloss, M., Keyes, D., Mallick, B., Marzouk, Y., Tenorio, L., van Bloemen Waanders, B., and Willcox, K., pp. 195–215, John Wiley & Sons, Ltd, Chichester, UK., doi:10.1002/9780470685853.ch10, 2010.

Williamson, D., Goldstein, M., Allison, L., Blaker, A., Challenor, P., Jackson, L., and Yamazaki, K.: History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble, Climate Dynamics, 41, 1703–1729, doi:10.1007/s00382-013-1896-4, http://dx.doi.org/10.1007/s00382-013-1896-4, 2013.

Williamson, D., Blaker, A., Hampton, C., and Salter, J.: Identifying and removing structural biases in climate models with history matching, Climate Dynamics, pp. 1–26, doi:10.1007/s00382-014-2378-z, http://dx.doi.org/10.1007/s00382-014-2378-z, 2014.

Wohlfahrt, J., Harrison, S. P., and Braconnot, P.: Synergistic feedbacks between ocean and vegetation on mid- and high-latitude climates during the mid-Holocene, Climate Dynamics, 22, 223–238, doi:10.1007/s00382-003-0379-4, 2004.

Yin, Q.: Insolation-induced mid-Brunhes transition in Southern Ocean ventilation and deep-ocean temperature, Nature, 494, 222–225, doi:10.1038/nature11790, 2013.

Yin, Q. and Berger, A.: Individual contribution of insolation and CO2 to the interglacial climates of the past 800,000 years, Climate Dynamics, 38, 709–724, doi:10.1007/s00382-011-1013-5, 2012.

Zhao, Y., Braconnot, P., Harrison, S. P., Yiou, P., and Marti, O.: Simulated changes in the relationship between tropical ocean temperatures and the western African monsoon during the mid-Holocene, Climate Dynamics, 28, 533–551, doi:10.1007/s00382-006-0196-7, 2007.

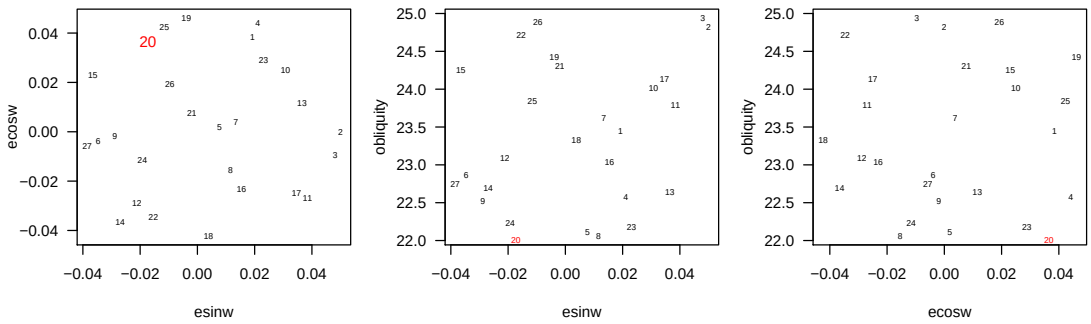

Figure 1. ~~Experiment~~ Two dimensional projections of the experimental design. The experiment marked in red (experiment 20) was discarded from the analysis.

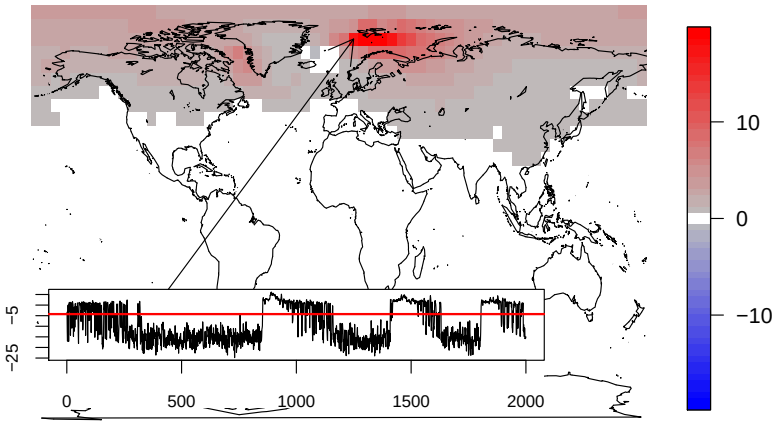

Figure 2. Slow oscillations developing in experiment 20 ($e = 0.040$, $\varpi = 334.6°$, $\varepsilon = 22°$). The surface annual temperature over one of the North Atlantic grid points is shown (inset, in ° C) along with the geographic distribution of the difference between the warm and the cold phases. The horizontal red line in the inset is the emulator prediction, calibrated on the 26 remaining experiments.

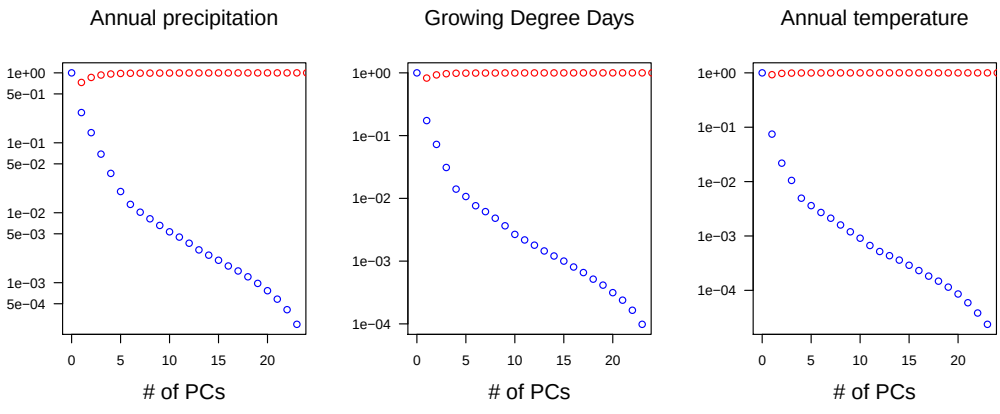

Figure 3. PCA decomposition of the logarithm of annual precipitation, GDD and annual temperature fields based on the experiment design. Shown are the cumulated normalised variances resolved by the principal components (red) and the ~~left~~ remaining variance (blue), which is modelled as white noise in (19), as a function of ~~n'~~ ℓ. Quantities are grid-cell averages.

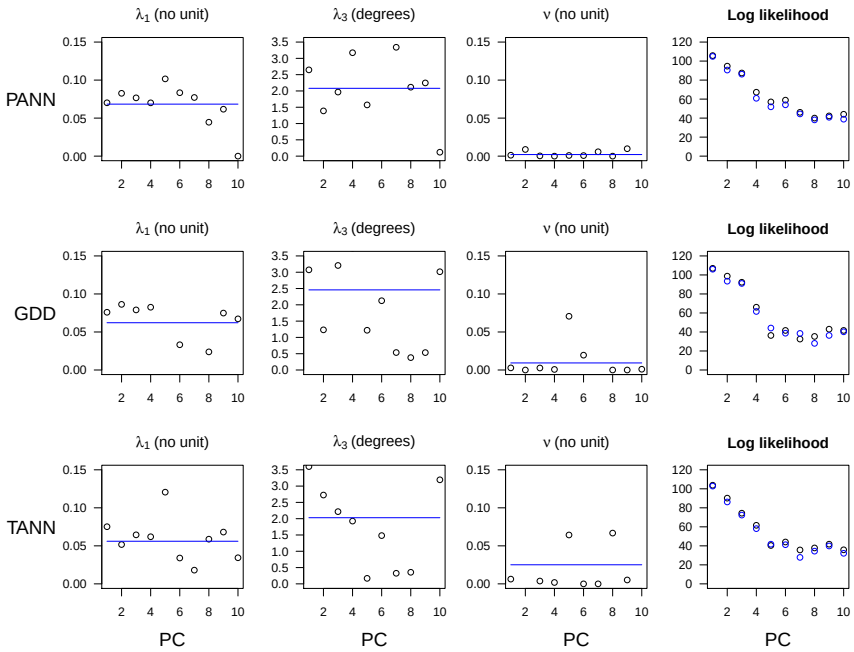

Figure 4. Gaussian process parameters maximising the penalised log-likelihood for three variables : GDD, log(Annual Precipitation) and Annual temperature, either *(black)* optimised for each PC independently or *(blue)* optimised based on the product of the likelihoods of the first 10 ~~PC~~PCs, assuming that the same correlation parameters are used on all PCs. The maximized log-likelihood associated with each PC is given for reference.

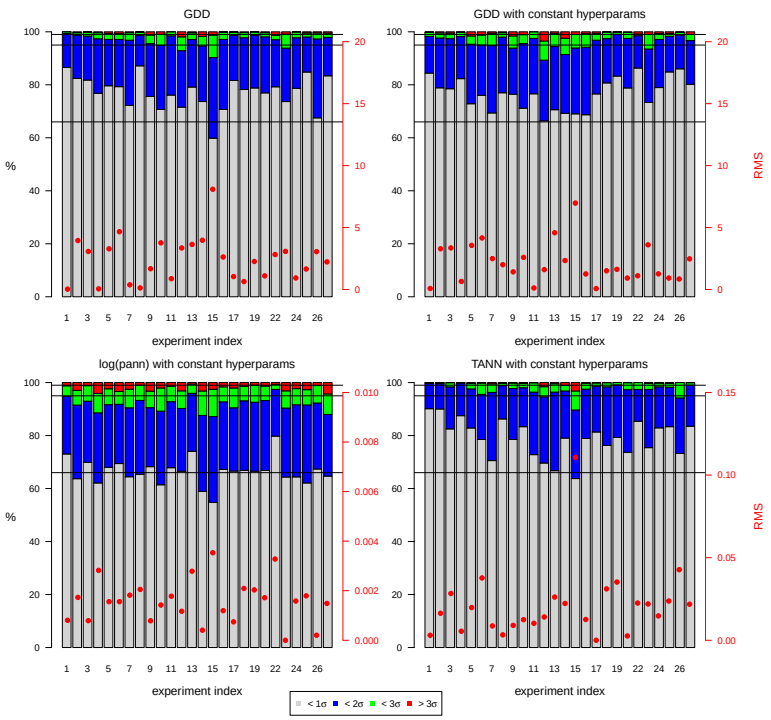

Figure 5. Evaluation of the PC emulators. The bars give the ~~fractions~~ fraction of grid points for which the emulation correctly predicts the value of the experiment left out of the training set, within the 66th, 95th and 99th inner quantiles of the distribution. The horizontal lines indicate the theoretical position of the percentiles for well-calibrated emulators. Dots provide root mean squares of the differences between predicted and actual values. The graphical layout is adapted from the recommendation of the "Modelling Uncertainty in Computer Model" project, http://mucm.aston.ac.uk/MUCM/MUCMToolkit/index.php?page=ExamMultipleOutputsPCA.html.

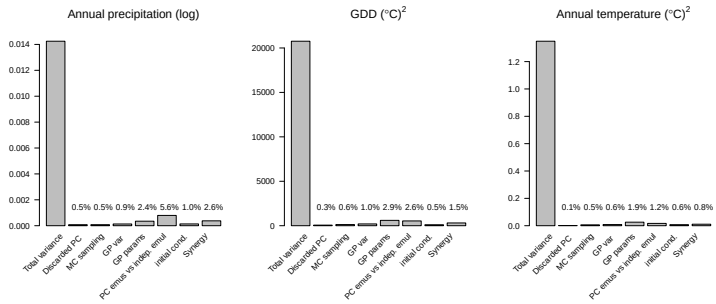

Figure 6. Total variances associated to inputs (V), averaged over all grid points. These variances are then compared to different elements of variances associated with the PC emulator, namely : variance of the discarded principal components, error estimated from Monte-Carlo sampling (estimated by comparison with analytical integrals), Gaussian Process variance (Σ_{tot}), difference in variance associated with the choice of correlation parameters (constant for all PC, or PC-dependent), or the choice of emulator (PC emulator vs independent emulators), difference due to the choice of initial conditions (exp. 20 excluded), and synergy term defined as ~~$V - T_{cw} - T_{\varepsilon}$~~ $V - T_{\{cw\}} - T_{\varepsilon}$.

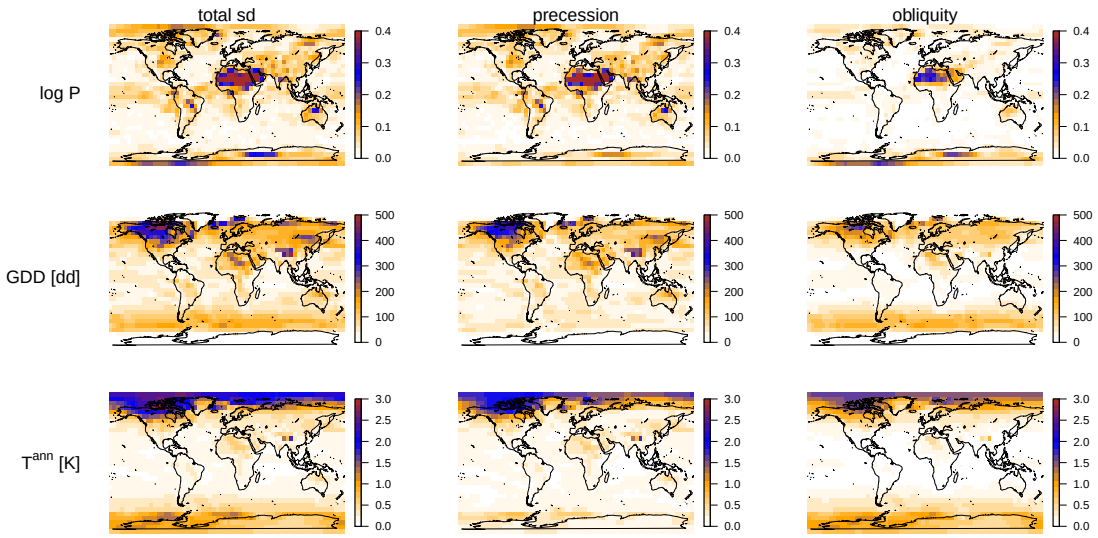

Figure 7. Geographical distribution of total variance V, split into contributions from precession ~~$T_{e\varpi}$~~ $T_{\{\varpi\}}$ and obliquity T_ε.

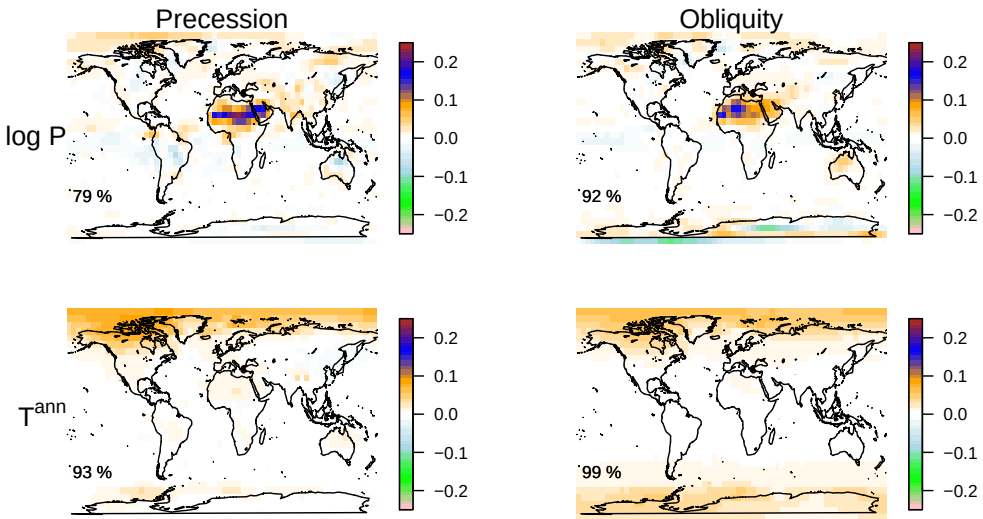

Figure 8. First eigenvectors of $T_{e\varpi}$ $T_{\{e\varpi\}}$ and obliquity T_ε for annual precipitation (log) and annual surface temperature, along with the fraction of the explained variance by these eigenvectors. These eigenvectors are referred to here as the fingerprints.

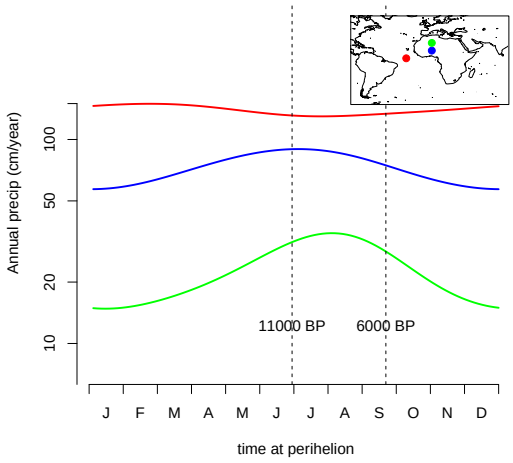

Figure 9. Emulated annual precipitation for three locations at the North Atlantic / African sector, as a function of the longitude of the perihelion. For ~~easy reading~~ ease of interpretation, the longitude of the perihelion is ~~here~~ denoted as the time of the year at which perihelion (closest point to the Sun) is reached.

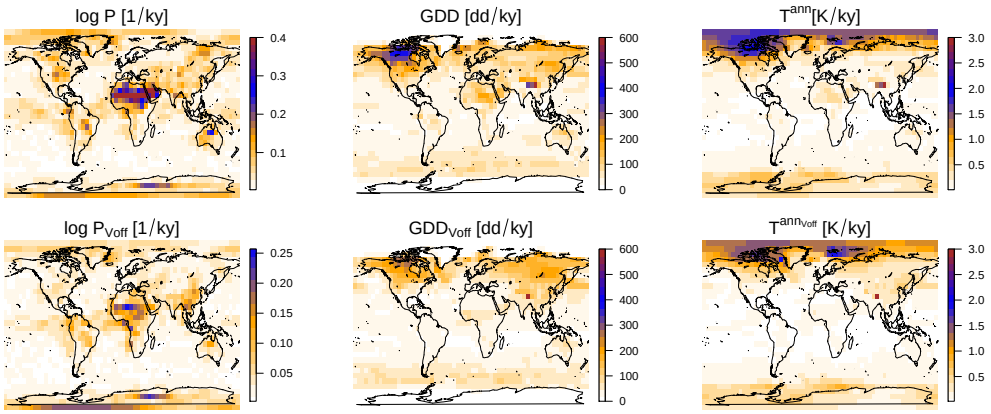

Figure 10. Maximum rate of change, in units per thousand years, estimated using the PC emulator and assuming a quasi-stationary climate and interglacial conditions.