



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

N. Bounceur¹, M. Crucifix¹, and R. D. Wilkinson²

¹Université catholique de Louvain, Earth and Life Institute, Georges Lemaître Centre for Earth and Climate Research, Louvain-la-Neuve, Belgium

²University of Nottingham, School of Mathematical Sciences, Nottingham, UK

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

Received: 26 June 2014 – Published in Earth Syst. Dynam. Discuss.: 23 July 2014

Revised: 1 April 2015 – Accepted: 6 April 2015 – Published: 4 May 2015

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 experiments. The outputs considered 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 in LOVECLIM equally to GDD in the Northern Hemisphere, and the effect of obliquity on the response of Southern Hemisphere temperature dominates precession effects. Precession dominates precipitation changes in subtropical areas. Compared to standard approaches based on a small number of simulations, the methodology presented here allows us to identify more systematically regions susceptible 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 1000 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, 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 effect of the choice of length scales and the type of emulator, and estimate uncertainties associated with specific computational aspects, to conclude that the principal component emulator is a good option for this kind of application.

1 Introduction

The seasonal and spatial distribution of incoming solar radiation (insolation) at the top of the atmosphere is determined by three astronomical parameters: Earth's eccentricity, the longitude of the perihelion, and Earth's obliquity. The variation in these parameters causes sufficient changes in insolation to significantly affect climate, such as 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 the approach followed, for example, by the Paleoclimate 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. 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 associated with the responses of vegetation and sea ice.

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”. The fact that the response to obliquity depends on precession may be interpreted as a second-order effect.

One difficulty appears when one points to the possibility that second-order effects may be significant during certain epochs, or at certain critical locations only. A local analysis, focused on specific periods, such as the study by Crucifix and Loutre (2002), may thus by chance over- or underemphasise second-order effects. The approach of Tuenter et al. (2003) is more global in the sense that it does not refer to a specific epoch, but the coverage of astronomical forcing configurations may be too coarse to detect second-order effects reliably enough.

To see how the problem may be formulated statistically, let us use vector \mathbf{x} to express the astronomical forcing. Then, let us denote the output of a model run at configuration \mathbf{x} by $f(\mathbf{x})$. The vector \mathbf{x} varied in the past, and we can estimate its distribution by reference to existing astronomical solutions. If we assume that the climate model represents reality and that the climate is quasi-stationary with respect to the astronomical forcing, then $f(\mathbf{x})$ reflects the past evolution of climate. We can in particular enquire about the variance in $f(\mathbf{x})$ caused by the distribution of \mathbf{x} , decompose this variance into contributions from individual factors (precession and obliquity in this example), and then check for second-order effects. This is the principle of global sensitivity analysis (Homma and Saltelli, 1996; Saltelli et al., 2004), for which there exist methods specifically tailored for complex numerical simulators.

We therefore propose to 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 precipitation and temperature. We also attempt to detect regions where fast climate responses may occur in response to the slow changes in astronomical forcing. The objectives of this work are twofold. The first is climatic: we would like to determine the respective roles of astronomical 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, adding to the discussion in de Noblet et al. (1996), Claussen et al. (1999), Crucifix and Loutre (2002), Kageyama et al. (2004), and Meissner et al. (2003). 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 aerosol 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 factors), the methodological approach that will be followed here is broadly similar to that of Lee et al. (2011) and Oakley and O'Hagan (2004): (a) choose the experimental design to efficiently fill the input space (here, the space of astronomical forcings); (b) run LOVECLIM in these experiments; (c) train and validate an emulator, that is, a stochastic statistical model used to predict the function $f(\mathbf{x})$ 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 recent years as a tool to explore input spaces with the aim of *calibrating* the model on observations and estimating climate sensitivity (Rougier et al., 2009; Holden et al., 2010; Schmittner et al., 2011). We 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 palaeoclimatology, 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 $\lambda = 90^\circ$, the September equinox to $\lambda = 180^\circ$, 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 insolation¹, 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, 99 % of the time, eccentricity $e < 0.05$ and the inner 99th percentile of ε is 22.3 to 24.3° , these boundaries differing by less than 0.1° if the Berger (1978b) or Laskar et al. (2004) solution is used as the reference.

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 an *input factor matrix* denoted \mathbf{X} , which has three columns corresponding to the inputs $i_{1,2,3}$, 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, for example, the monograph by 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 is to

maximise the information that can be inferred from the experiment set. If we use a Gaussian process (GP) emulator (as here), this criterion 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. Finding an adequate design is thus an optimisation problem, and generally not a straightforward one, because the GP hyper-parameters are not known a priori, and 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 properties (where we maximise the minimum Euclidean distance between any pair of design points) or orthogonality (in this case, maximising $\mathbf{X}'\mathbf{X}$). Hybrid designs combine several such properties and can be theoretically justified as good solutions for Gaussian process emulation (Joseph and Hung, 2008). Out-of-the-box packages can be used to produce Latin hypercube designs (e.g. the *lhs* package; Carnell, 2012), but they are unsuited to our needs as we face an additional constraint: we want to avoid wasting computational effort sampling eccentricities > 0.05 , which translates into a constraint on the quantity $i_1^2 + i_2^2$. (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 that of (Vernon et al., 2010): it satisfies the constraint on eccentricity while aiming at near orthogonality and max–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 (Bounceur, 2015).

The resulting design is shown in Fig. 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 ensemble uses the pre-industrial conditions provided by default in the LOVECLIM model package. The second uses the final state of member 2 of the first ensemble. This is a so-called warm orbit (high obliquity, high eccentricity, and $\varpi \simeq 90^\circ$) that produces extensive vegetation cover in the Northern Hemisphere and in the Sahara region. Every experiment is run for 2000 years, and 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 multiple steady-state solutions. Brovkin et al. (1998), for example, found two stable equilibria for certain orbital configurations when using the same vegetation model as here, 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 (unimodal) Gaussian process as used here. Our initial intention was to build two distinct emulators: one using the warm orbit as

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

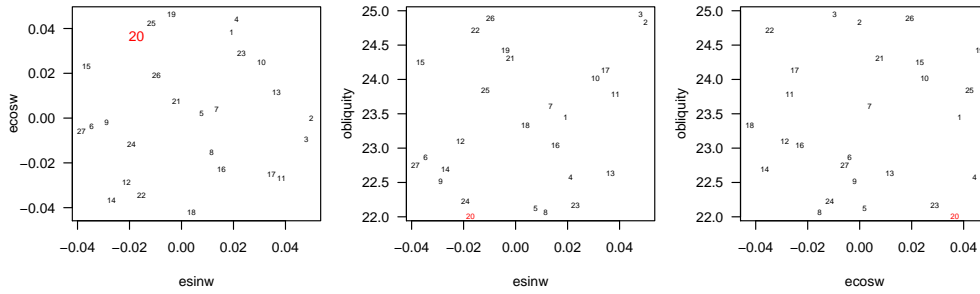


Figure 1. Two-dimensional projections of the experimental design. The experiment marked in red (experiment 20) was discarded from the analysis.

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 this was worth the cost so as to avoid the risk of missing the existence of multiple steady states. The third ensemble uses the same initial conditions as the first 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

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 $f(\mathbf{x})$ to be the simulator response at input vector \mathbf{x} , which contains the values of the three astronomical forcing terms. Let $\rho(\mathbf{x})$ 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). The *total variance* in the output associated with the variations of factors may then be expressed as

$$\mathbf{V} := \text{Var}(f(\mathbf{x})), \quad (1)$$

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

Suppose we are now interested in the variance in the simulator output caused by the variation of a subset of the input factors only. Let \mathbf{x}_i denote a subset of the components in \mathbf{x} , and let \mathbf{x}_{-i} be the remaining components. For example, if i is

obliquity, $-i$ will be the indices associated with eccentricity and precession.

We define the following quantities (e.g. Saltelli et al., 2004): the *main effect*, $\eta(\mathbf{x}_i)$, is the expected output conditional on the value of \mathbf{x}_i , i.e. $\eta(\mathbf{x}_i) = \mathbb{E}[f(\mathbf{x})|\mathbf{x}_i]$; the *main effect variance* is the variance in the main effect with respect to \mathbf{x}_i : $\mathbf{V}_i = \text{Var}[\eta(\mathbf{x}_i)]$.

On the other hand, the output variance associated with factors \mathbf{x}_i varying while the factors \mathbf{x}_{-i} are fixed is denoted $\text{Var}(f(\mathbf{x})|\mathbf{x}_i)$. This is a function of the value of the fixed factors \mathbf{x}_{-i} . If we further average this variance over possible values of \mathbf{x}_{-i} , that is, we take the expectation of this quantity with respect to \mathbf{x}_{-i} , we obtain the total effect variance associated with factor \mathbf{x}_i , denoted \mathbf{T}_i :

$$\mathbf{T}_i := \mathbb{E}[\text{Var}(f(\mathbf{x})|\mathbf{x}_{-i})].$$

$$\text{Similarly, } \mathbf{T}_{-i} = \mathbb{E}[\text{Var}(f(\mathbf{x})|\mathbf{x}_i)]. \quad (2)$$

The law of total variance implies $\mathbf{V}_i := \mathbf{V} - \mathbf{T}_{-i}$, 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 \mathbf{x}_i . The standardised quantity \mathbf{V}_i/\mathbf{V} is commonly referred to as the main effect index, and \mathbf{T}_i/\mathbf{V} as the total effect 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 (2004) for one-dimensional scalar simulator outputs to multi-dimensional vector outputs. Let \mathcal{X} , \mathcal{X}_i , and \mathcal{X}_{-i} be the domains of the input factors \mathbf{x} , \mathbf{x}_i , and \mathbf{x}_{-i} , respectively. The *total variance* can be explicitly written as

$$\begin{aligned} \mathbf{V} &= \int_{\mathcal{X}} f(\mathbf{x})f(\mathbf{x})'\rho(\mathbf{x})d\mathbf{x} \\ &\quad - \iint_{\mathcal{X}_{-i} \times \mathcal{X}_i} f(\mathbf{x})f(\mathbf{x}^*)'\rho(\mathbf{x})\rho(\mathbf{x}^*)d\mathbf{x}d\mathbf{x}^*. \end{aligned} \quad (3)$$

The main effect is computed as follows:

$$\eta(\mathbf{x}_i) = \int_{\mathcal{X}_{-i}} f(\mathbf{x})\rho(\mathbf{x}_{-i}|\mathbf{x}_i)d\mathbf{x}_{-i}, \quad (4)$$

where $\rho(\mathbf{x}_{-i}|\mathbf{x}_i)$ is the conditional density of \mathbf{x}_{-i} given \mathbf{x}_i . Because our knowledge of $\mathbf{f}(\mathbf{x})$ is limited to the ensemble of model runs, we are uncertain about the value of $\boldsymbol{\eta}(\mathbf{x}_i)$ for all values of \mathbf{x}_i . The approach described in Oakley and O’Hagan (2004) is to build an emulator of $\mathbf{f}(\mathbf{x})$ that can predict its value for any input configuration \mathbf{x} . They use a Gaussian process (GP) model for the emulator, with mean function $\mathbf{m}(\mathbf{x})$ and covariance function between $\mathbf{f}(\mathbf{x})$ and $\mathbf{f}(\mathbf{x}^*)$, which we will denote $\boldsymbol{\Sigma}(\mathbf{x}, \mathbf{x}^*)$. Because the emulator is a Gaussian process, the main effect $\boldsymbol{\eta}(\mathbf{x}_i)$, which is a linear transformation of $\mathbf{f}(\mathbf{x})$, is also a Gaussian process, and has mean and variance functions

$$\mathbf{M}_i(\mathbf{x}_i) = \int_{\mathcal{X}_{-i}} \mathbf{m}(\mathbf{x})\rho(\mathbf{x}_{-i}|\mathbf{x}_i) d\mathbf{x}_{-i}, \quad (5)$$

$$\mathbf{S}_i(\mathbf{x}_i, \mathbf{x}_i^*) = \iint_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} \boldsymbol{\Sigma}(\mathbf{x}, \mathbf{x}^*)\rho(\mathbf{x}_{-i}|\mathbf{x}_i)\rho(\mathbf{x}_{-i}^*|\mathbf{x}_i^*) d\mathbf{x}_{-i}d\mathbf{x}_{-i}^*. \quad (6)$$

It follows that

$$\mathbb{E}_f(\mathbf{V}_i) = \int_{\mathcal{X}_i} [\mathbf{M}_i(\mathbf{x}_i)\mathbf{M}_i'(\mathbf{x}_i) + \mathbf{S}_i(\mathbf{x}_i, \mathbf{x}_i)]\rho_i(\mathbf{x}_i) d\mathbf{x}_i - \mathbf{C}, \quad (7)$$

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

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

$$\mathbf{M}_{\text{tot}} = \int_{\mathcal{X}} \mathbf{m}(\mathbf{x})\rho(\mathbf{x}) d\mathbf{x}, \quad (9)$$

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

The expectation of the total variance \mathbf{V} of the simulator output is

$$\mathbb{E}_f(\mathbf{V}) = [\mathbf{M}_{\text{tot}}\mathbf{M}_{\text{tot}}' + \mathbf{S}_{\text{tot}}] - \mathbf{C}, \text{ with } \mathbf{S}_{\text{tot}} = \int \boldsymbol{\Sigma}(\mathbf{x}, \mathbf{x})\rho(\mathbf{x})d\mathbf{x}. \quad (11)$$

From now on, it will be implicit that the \mathbf{V} , \mathbf{V}_i , and \mathbf{T}_i are estimated with the emulator and the symbol \mathbb{E}_f will be dropped. Specifically, we refer to $\mathbf{T}_{\{\text{e}\varpi\}}$ as the total effect variance associated with precession, and \mathbf{T}_ε as that associated with obliquity.

Finally, we refer to the quantity $\mathbf{V} - \mathbf{T}_{\{\text{e}\varpi\}} - \mathbf{T}_\varepsilon = \mathbf{T}_\varepsilon - \mathbf{T}_\varepsilon$ as the *synergy* term. More generally, $\mathbf{T}_i - \mathbf{V}_i$ is a measure of how much the factor i is 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 individually. Unlike the global sensitivity indices discussed here, synergy terms are classically estimated on the basis of a reference experiment, and comparing changes one at a time with combined changes in the different factors, following a method called factor separation analysis (Stein and Alpert, 1993; Alpert and Sholokhman, 2011; see, for example, Crucifix and Loutre, 2002; Ganopolski et al., 1998; Claussen, 2009; Braconnot et al., 1999; Berger, 1999; Henrot et al., 2010; Wohlfahrt et al., 2004).

2.4 Emulator

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, which is done by Lee et al. (2011). There are, however, several other possibilities for multivariate emulation (see, for example, Rougier, 2008). The main challenge is defining a covariance function for generating the covariance matrix $\boldsymbol{\Sigma}(\mathbf{x}, \mathbf{x}^*)$ in order to produce 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 advantages of this choice are commented on in Sect. 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 Eq. (18). To our knowledge, this has not been given elsewhere. The reader mainly interested in the emulator performance and climatological analysis may immediately jump to Sect. 3.

2.4.2 Principal component decomposition

Let \mathbf{Y} denote the matrix in which each column represents the output of one experiment, i.e. $\mathbf{Y} = [\mathbf{f}(\mathbf{x}^{(1)}), \dots, \mathbf{f}(\mathbf{x}^{(n)})]$, where $\mathbf{x}^{(j)}$ is the input of experiment j . For example, if we want to emulate annual mean surface temperature, \mathbf{f} is a vector of $p = 2048$ components (the number of grid points). Denote the p vector of row averages of \mathbf{Y} by $\bar{\mathbf{Y}}$. We now define the centred matrix $\mathbf{Y}^* = \mathbf{Y} - \bar{\mathbf{Y}}\mathbf{1}'_p$, with $\mathbf{1}_p$ a vector of length p with all components equal to 1, and consider the singular value decomposition (SVD) $\mathbf{Y}^* = \mathbf{U}\mathbf{D}\mathbf{V}'$, where \mathbf{D} is a diagonal matrix and \mathbf{U} and \mathbf{V} are square orthonormal matrices. The columns of \mathbf{U} represent the basis vectors, $\{\mathbf{u}_k\}$, and the projection coefficients are given by $\mathbf{V}\mathbf{D}$. For the j th experiment, the coefficient for the k th basis vector is $a_k(\mathbf{x}^{(j)}) = \mathbf{V}_{j,k}\mathbf{D}_{kk}$. Wilkinson (2010) keeps the first ℓ eigenvectors only (ordered by decreasing eigenvalues). This is sometimes known

as “hard-thresholding regularisation” (Silverman, 1996), and results in a reduced-order model: $\mathbf{f}(\mathbf{x}) \approx \sum_{k=1}^{\ell} a_k(\mathbf{x}) \mathbf{u}_k$.

Thus, for any input \mathbf{x} , the emulator output prediction will lie in the space spanned by the $\{\mathbf{u}_1 \dots \mathbf{u}_\ell\}$. The error associated with hard-thresholding is accounted for in the covariance of the estimator (see Sect. 2.5).

2.4.3 Emulation of PC scores

We then need to predict the $a_k(\mathbf{x})$ based on the experiment output $a_k(\mathbf{x}_i)$. This is done by considering ℓ Gaussian process (GP) models (Rasmussen and Williams, 2005). We suppose the prior mean of each GP is $\mathbf{h}(\mathbf{x})' \boldsymbol{\beta}_k$, where $\mathbf{h}(\mathbf{x})$ is a vector of q known regression functions and $\boldsymbol{\beta}_k$ 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, $\mathbf{h}(\mathbf{x})$ is simply defined as $(1, x_1, x_2, x_3)'$ (linear regressors), so that $\mathbf{h}(\mathbf{x})' \boldsymbol{\beta}_k = 1 + \beta_{k,1}x_1 + \beta_{k,2}x_2 + \beta_{k,3}x_3$.

We then define the following:

- \mathbf{H} , the matrix which has row i equal to the regressors $\mathbf{h}(\mathbf{x}_i)$;
- $\mathbf{y}_k = (a_k(\mathbf{x}_1), \dots, a_k(\mathbf{x}_n))'$, the vector of PC scores we wish to emulate;
- $c_k(\mathbf{x}, \mathbf{x}^*) = \text{cor}(a_k(\mathbf{x}), a_k(\mathbf{x}^*))$, the correlation function for $a_k(\cdot)$. This is typically a monotonically decreasing function of the distance between the two points. We let \mathbf{A}_k be the $n \times n$ Gram matrix, with $\mathbf{A}_k[i, j] = c_k(\mathbf{x}_i, \mathbf{x}_j)$.

If each $a_k(\cdot)$ is modelled as a Gaussian process, then we have a Gaussian prior predictive distribution for \mathbf{y} conditional on $\boldsymbol{\beta}_k$ and σ_k , with $\mathbf{y}_k \sim N(\mathbf{H}\boldsymbol{\beta}_k, \sigma_k^2 \mathbf{A}_k)$.

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.

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 that is a Student t distribution with $n - q$ degrees of freedom, with mean and variance functions

$$m_k(\mathbf{x}) = \mathbf{h}(\mathbf{x})' \hat{\boldsymbol{\beta}}_k + \mathbf{t}_k(\mathbf{x})' \mathbf{e}_k, \quad (12)$$

$$\Sigma_k(\mathbf{x}, \mathbf{x}^*) = \hat{\sigma}_k^2 [c_k(\mathbf{x}, \mathbf{x}^*) - \mathbf{t}_k(\mathbf{x})' \mathbf{A}_k^{-1} \mathbf{t}_k(\mathbf{x}^*) + \mathbf{p}_k(\mathbf{x})' (\mathbf{H}' \mathbf{A}_k^{-1} \mathbf{H})^{-1} \mathbf{p}_k(\mathbf{x}^*)], \quad (13)$$

where

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

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

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

$$c_k(\mathbf{x}_i, \mathbf{x}_j) = \exp[-(\mathbf{x}_i' \mathbf{A}_k^{-2} \mathbf{x}_j)] + \nu_k \mathbb{1}_{i=j}, \quad (14)$$

where \mathbf{A}_k is a scaling matrix chosen to be diagonal and ν_k is a “nugget”, which we will discuss shortly. The diagonal elements of \mathbf{A}_k are commonly called the length scales. The parameters \mathbf{A}_k and ν_k are then found by optimising a penalised likelihood $L_k^{(\text{pen})}$ of the Gaussian process associated with the principal component k , as a function of \mathbf{A}_k and ν_k . The expression of L_k is given in (Andrianakis and Challenor, 2012):

$$\log L_k(\nu_k, \mathbf{A}_k) = -\frac{1}{2} (\log(|\mathbf{A}_k|) + \mathbf{H}' \mathbf{A}_k^{-1} \mathbf{H}) + (n - q) \log(\hat{\sigma}_k^2), \quad (15)$$

$$\log L_k^{(\text{pen})}(\nu_k, \mathbf{A}_k) = \log L_k(\nu_k, \mathbf{A}_k) - 2\nu_k^2 \frac{(\mathbf{y}_k - \mathbf{H} \hat{\boldsymbol{\beta}}_k)' \mathbf{A}_k^{-2} (\mathbf{y}_k - \mathbf{H} \hat{\boldsymbol{\beta}}_k)}{(\mathbf{y}_k - \mathbf{H} \hat{\boldsymbol{\beta}}_k)' (\mathbf{y}_k - \mathbf{H} \hat{\boldsymbol{\beta}}_k)}. \quad (16)$$

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

The nugget term, $\nu_k \mathbb{1}_{i=j}$, was originally introduced to account for measurement errors in geospatial data analysis (Cressie, 1993). In emulators of deterministic systems, the nugget may also be justified as a regularisation ansatz to avoid poor matrix conditioning (Pepelychev, 2010) or as a way to account for 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 “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 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.

Finally, two options will be considered to estimate the different \mathbf{A}_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. $\Lambda_k = \Lambda$, and $\nu_k = \nu$ and optimise $\sum_{k=1}^{\ell} L_k$. 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 Eqs. (12) and (13) are conditional on the values of these parameters.

2.5 Recombination of PC scores

The emulator posterior distribution for predictions of LOVECLIM’s outputs $f(\mathbf{x})$ follows a distribution with mean and covariances functions given by

$$\mathbf{m}(\mathbf{x}) = \sum_{k=1}^{\ell} m_k(\mathbf{x})\mathbf{u}_k, \quad (17)$$

$$\Sigma(\mathbf{x}, \mathbf{x}^*) = \underbrace{\sum_{k=1}^{\ell} \Sigma_k(\mathbf{x}, \mathbf{x}^*)\mathbf{u}_k\mathbf{u}'_k}_{\Sigma^{(gp)}} + \underbrace{\sum_{k=\ell+1}^n \frac{\mathbf{D}_{kk}^2}{n}\mathbf{u}_k\mathbf{u}'_k}_{\Sigma^{(pc)}}. \quad (18)$$

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.

Variance indices may now be obtained by plugging m_k and *boldsymbol* Σ_k (Eqs. 12 and 13) into Eqs. (17) and (18) to obtain \mathbf{m} and Σ and then using these expressions in Eq. (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.

2.6 Short discussion of possible advantages over the independent emulator 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 be briefly summarised as follows:

- The PC emulation is based on the calibration of ℓ Gaussian process models per output field (temperature, precipitation, GDD). We use $\ell = 10$ (justified below). We 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 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.
- As the PC emulation approach requires fewer emulators, more time can be spent individually validating each of them.

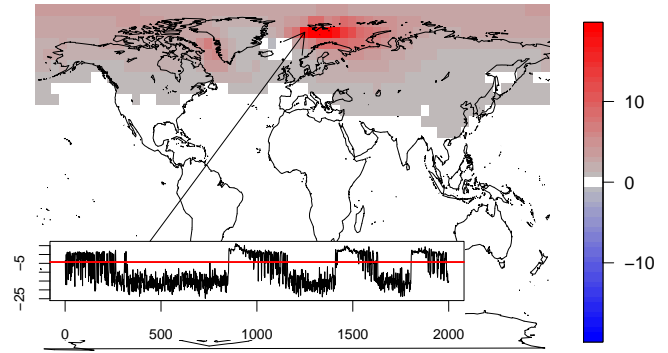


Figure 2. Slow oscillations developing in experiment 20 ($e = 0.040$, $\varpi = 334.6^\circ$, $\varepsilon = 22^\circ$). The surface annual temperature over one of the North Atlantic grid points is shown (inset, in $^\circ\text{C}$) along with the geographic distribution of the difference between the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

- 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, moreover, 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 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 (see Supplement). Experiment 27 shows a higher-amplitude variability pattern, but clearly oscillates around one mean value and, as we will shortly see, this mean is correctly captured by the emulator. It is therefore kept without further discussion for all subsequent analyses. Experiment 20 used the lowest configuration of obliquity (22°). This is lower than any 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 (Fig. 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 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 unclear. We choose to ignore the experiment for the time being (the following diagnostics ignore experiment 20), but we discuss the possible consequences of this 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 output fields: annual precipitation, growing degree days (GDD), and annual mean temperature. GDD is defined here as the annual sum of daily temperatures (in degrees Celsius) exceeding 0°C . It is used as a calendar-independent indicator of summer intensity and length in extra-tropical regions, 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 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 four (annual temperature) to eight principal components (Fig. 3). As discussed in the methodology section, two options are considered for the estimation of the length scales and nugget variance Λ and ν . We first attempt to use different correlation parameters obtained by maximising the penalised likelihood for each principal component emulator, independently of the others. It is then observed that the Gaussian process likelihood decreases with the index of the PC (Fig. 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 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 $\ell = 10$. The alternative approach consists in using the same correlation parameters for all 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 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 credible intervals. A well-calibrated emulator would, on average, correctly predict 66, 95, and 99% of the points, respectively, in each category. Some remarks are in order:

1. Based on this diagnostic only, using constant rather than PC-specific correlation parameters does not significantly affect the overall performance. This is explicitly shown for GDD, but is also true of the other fields.
2. However, all fields exhibit an excessive number of 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 are an excessive number of predictions within the 66th central credible interval.

3.3 Astronomically forced variance vs. other effects

The total variance \mathbf{V} resulting from the astronomical forcing can be estimated from Eq. (11). Here we compare the estimated total variance with other effects that may broadly be described as sources of uncertainties on this quantity (Fig. 6). 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

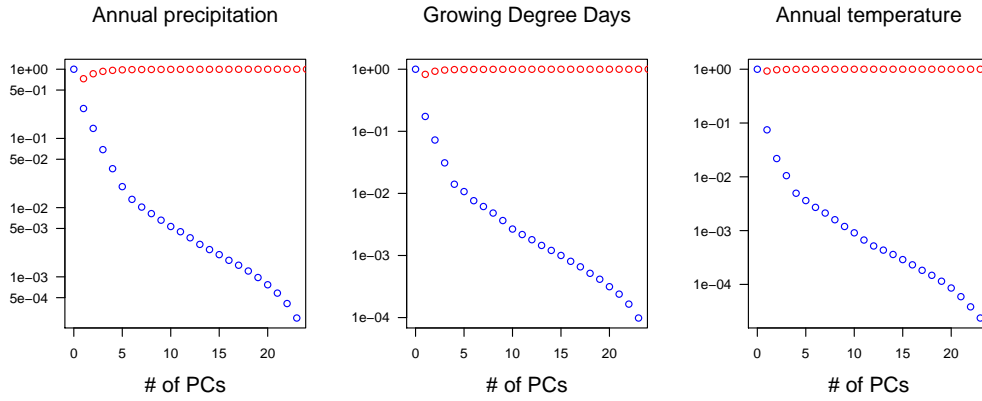


Figure 3. PCA decomposition of the logarithm of annual precipitation, GDD, and annual temperature fields based on the experiment design. Shown are the cumulated normalised variances resolved by the principal components (red) and the remaining variance (blue), which is modelled as white noise in Eq. (18), as a function of ℓ . Quantities are grid-cell averages.

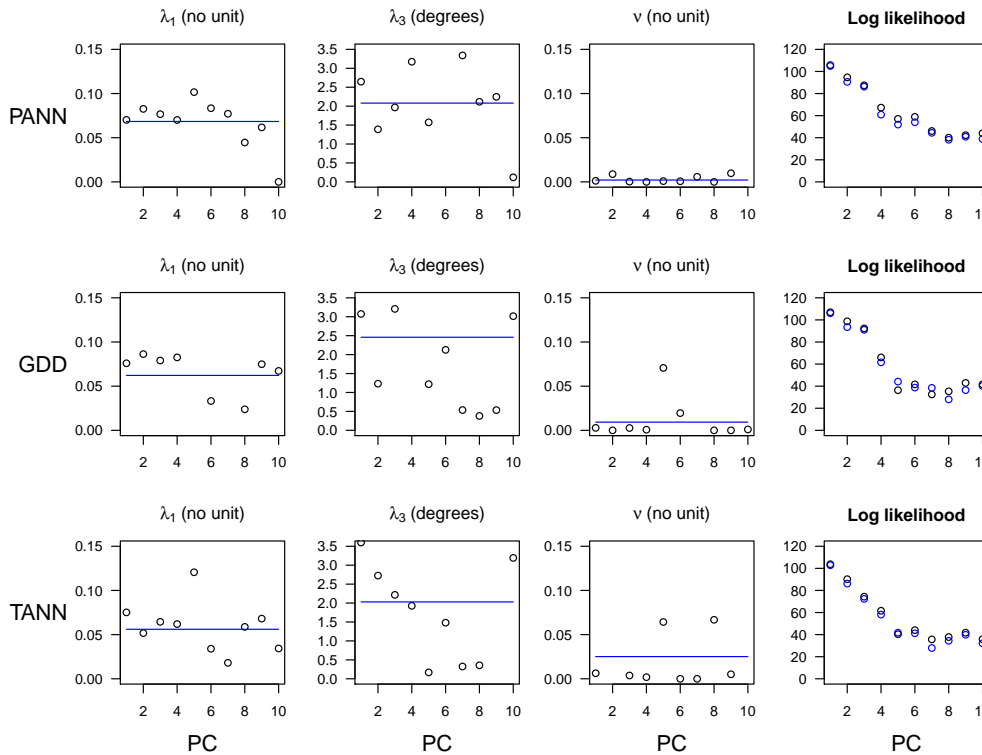


Figure 4. Gaussian process parameters maximising the penalised log likelihood for three variables: GDD, log(annual precipitation), and annual temperature, either (black) optimised for each PC independently or (blue) optimised based on the product of the likelihoods of the first 10 PCs, assuming that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

Carlo integrals with the analytical solution in the particular case of a uniform distributions for x ($\rho(x) \propto 1$), 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. 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 Supplement further shows that the patterns of the global sensitivity indices are similar regard-

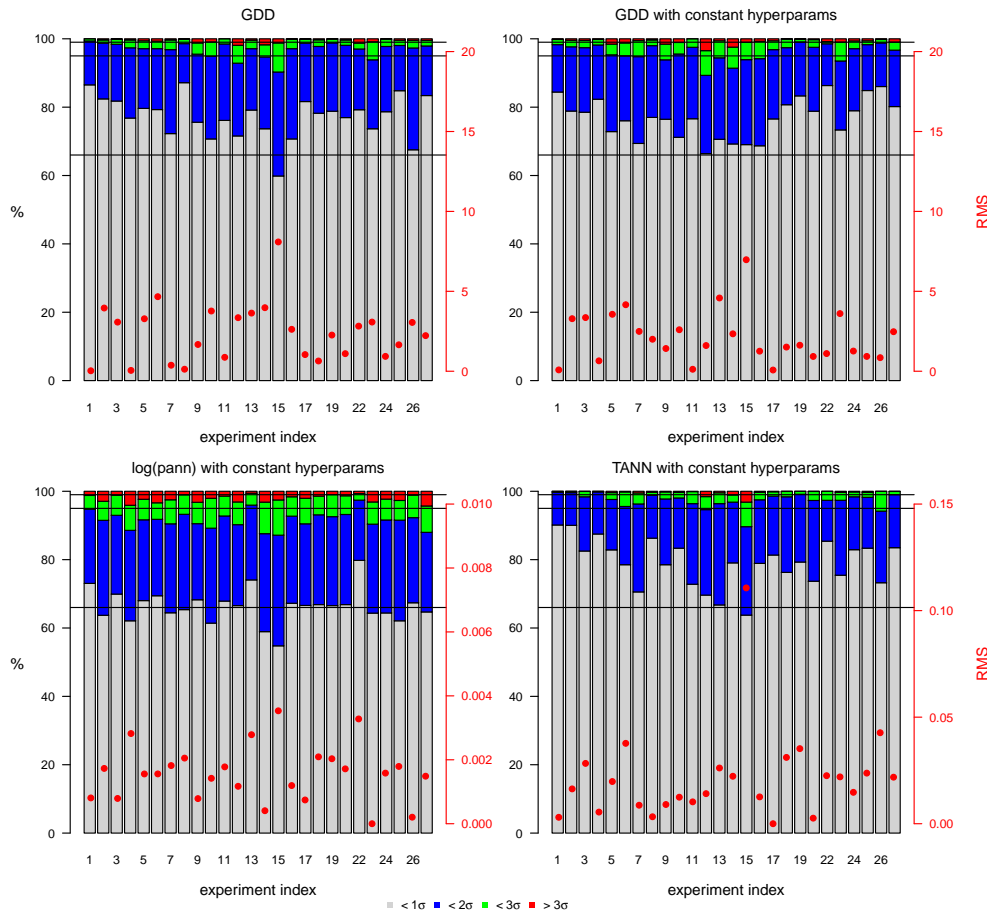


Figure 5. Evaluation of the PC emulators. The bars give the fraction of grid points for which the emulation correctly predicts the value of the experiment left out of the training set, within the 66th, 95th, and 99th inner quantiles of the distribution. The horizontal lines indicate the theoretical position 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>.

less which emulator is used, so that this choice is of no consequence for the scientific interpretation.

- The term “GP var” explicitly refers to S_{tot} defined in Eq. (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.
- The absolute value of the synergy term, measured as the difference between the total variance and the sum of the mean sensitivity indices, 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 \mathbf{V} , $\mathbf{T}_{\{e\varpi\}}$, and \mathbf{T}_{ε} . They pro-

vide essentially the same information as could be obtained using emulators independently calibrated over all grid boxes, as follows (see Fig. 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, otherwise expected, may be a consequence 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.

- 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 \mathbf{V} and \mathbf{T}_i are available, linear fingerprints may be obtained by performing an SVD of these covariance matrices (Fig. 8). Specifically, the eigenvectors $\mathbf{T}_{\{e\varpi\}}$ and \mathbf{T}_e are called the “fingerprints” of precession and obliquity, respectively. The first fingerprint of obliquity explains more than 90 % of the variance in 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 that it provides information on the in-phase or anti-phase relationships between climate variables, as follows:

- Obliquity produces in-phase effects on monsoon-related precipitation both in the Northern and in the Southern Hemisphere (compare Africa and northern Australia, 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 at 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 the Supplement). 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 of the longitude of the perihelion for three points along the African monsoon flow. We assume obliquity and eccentricity typical of the Holocene (Fig. 9), and indicatively denote longitudes of perihelion corresponding to the time of the beginning of the Holocene (11 000 years ago), as well as 6000 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 9000 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 Bounceur, 2015). This exercise may, 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 thousands years (Fig. 10). In order to enhance the palaeoclimatological interest of this discussion, we compare the experiments with interactive vegetation to 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 ky^{-1} ; a variable responding linearly and exclusively to obliquity with 1 standard deviation would show a maximum rate of change of 0.35 ky^{-1} . On this basis, comparing Fig. 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 1000 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 North American sector of the Arctic, including both northern Canada and sea-ice-covered regions: rates of temperature changes are of the order of 3°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 also suggests a role for sea-ice cover,

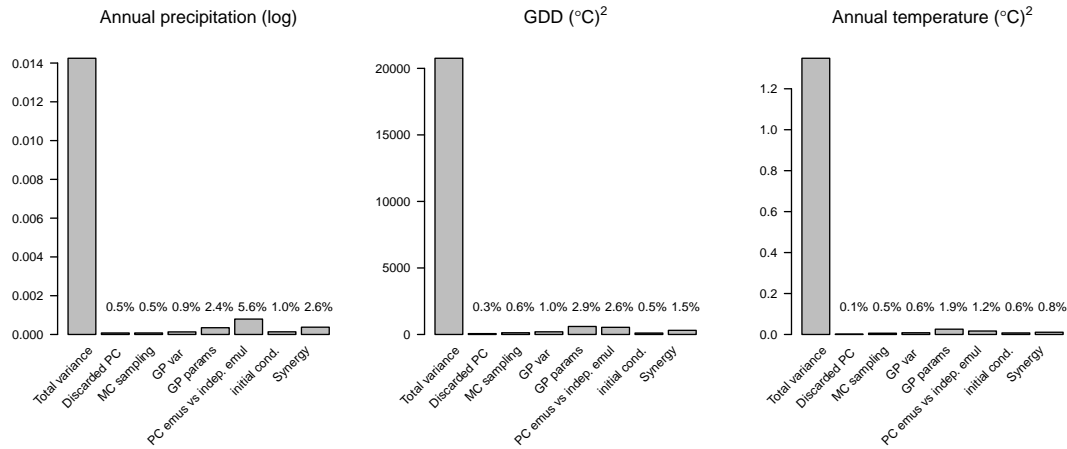


Figure 6. Total variances associated with inputs considered grid-point-wise, and then averaged over all grid points. These variances are then compared to different elements of variances associated with the PC emulator, namely variance in the discarded principal components, error estimated from Monte 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 (experiment 20 excluded), and synergy term defined as $\mathbf{V} - \mathbf{T}_{\{e\varpi\}} - \mathbf{T}_{\varepsilon}$.

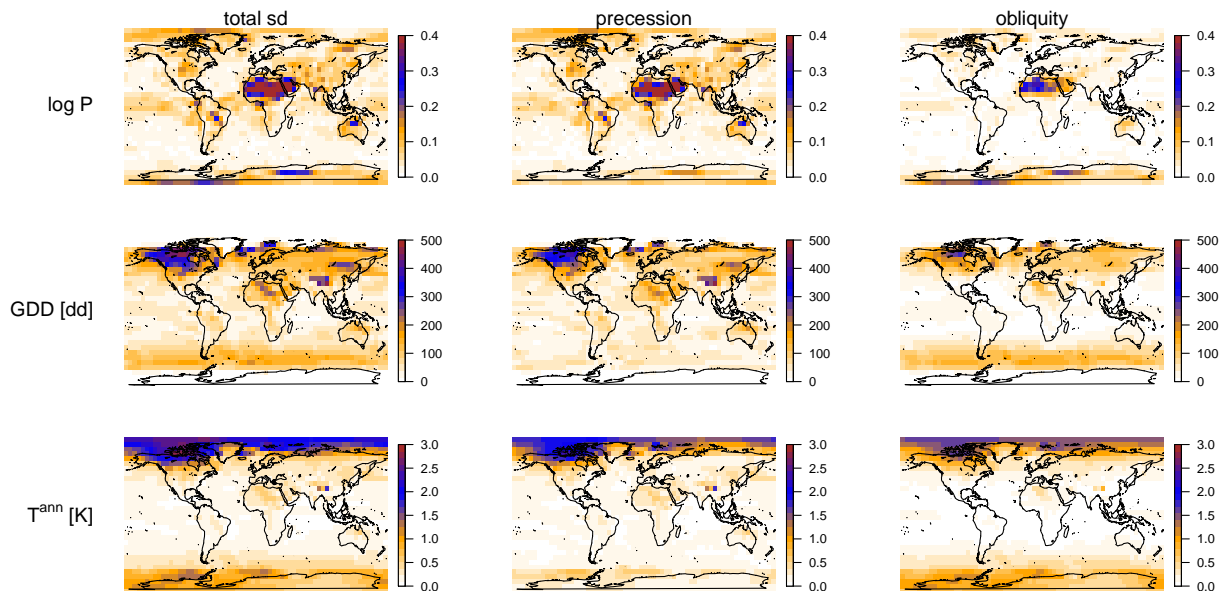


Figure 7. Geographical distribution of total variance \mathbf{V} , split into contributions from precession $\mathbf{T}_{\{\varpi\}}$ and obliquity \mathbf{T}_{ε} .

which further amplifies vegetation-induced effects. Such interactions between Arctic vegetation and sea ice have already been 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

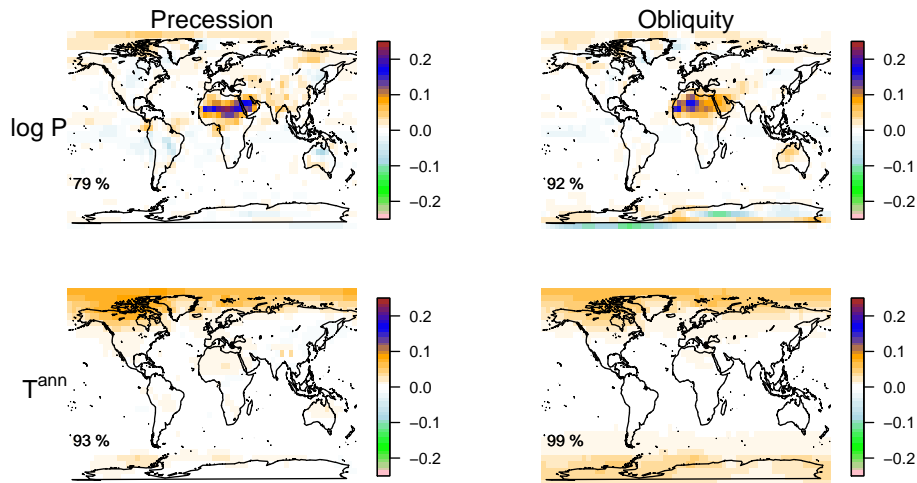


Figure 8. First eigenvectors of $T_{\{e\omega\}}$ and obliquity T_ϵ for annual precipitation (log) and annual surface temperature, along with the fraction of the explained variance by these eigenvectors. These eigenvectors are referred to here as the fingerprints.

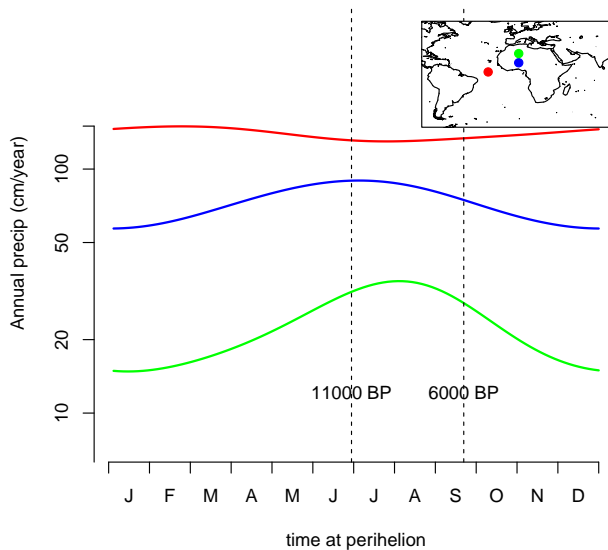


Figure 9. Emulated annual precipitation for three locations at the North Atlantic/African sector as a function of the longitude of the perihelion. For ease of interpretation, the longitude of the perihelion is denoted as the time of the year at which perihelion (closest point to the Sun) is reached.

insolation, but it may affect the annual mean climate by acting on the seasonal cycle of albedo 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 Southern Hemisphere. 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, namely

- the prominence of obliquity signals in Southern Hemisphere records, and more particularly Antarctic cores, be it CO₂ 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 northern continental records, such as Baikal’s, during isotopic stage 11 (Prokopenko et al., 2002).

2. GDD is used here as a measure of summer length and intensity. We considered it because it is used in VE-CODE as a predictor for vegetation changes. This quantity is also mathematically equivalent to the positive-degree days (PDD) index 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 the proposal by Ruddiman (2007) 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 precipitation pattern

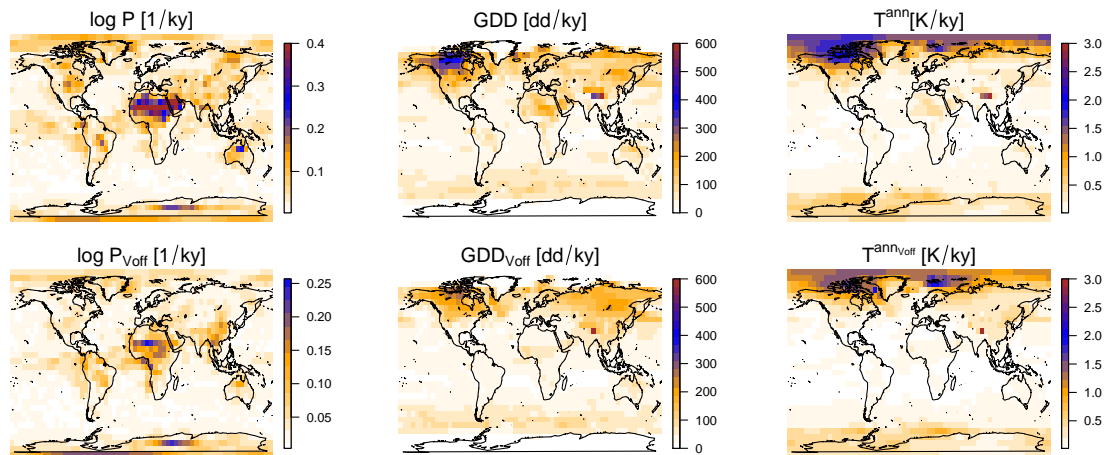


Figure 10. Maximum rate of change, in units per thousand years, estimated using the PC emulator and assuming a quasi-stationary climate and interglacial conditions.

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, for example, in Fig. 9), but not to the point of generating multiple steady states in response to a same astronomical configuration. In other words, all 27 experiments of the design converge to the same state (the particular case of experiment 20 is discussed in one of the next points). This is consistent with previous transient simulations with this model (Renssen et al., 2003), but we note that Brovkin et al. (1998) report 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, for example, 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 obliquity 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 effect variances, assessed over the whole Pleistocene, add up linearly or not. 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 have presented a global sensitivity analysis of the effects of astronomical forcing on the climate model LOVECLIM in interglacial conditions. The work relies 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 perspective, 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 to experience 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, so 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 effect of the choice of length scales and the type of emulator and estimated uncertainties associated with specific computational aspects such as the Monte Carlo estimates of integrals, and have concluded that the PC emulator is a reasonable option. We therefore recommend its use for 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 $i_3 = (\varepsilon - 23.5) / 1.5$. 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 $\det_{\text{init}} = 0$.
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 $i = 1 \dots N$, denoted $\mathbf{n}_j = \{n_{i,j}\}_{i=1}^N$ for $j = 1, 2, 3$.
 - (c) Form the matrix \mathbf{X} , with $X_{i,j} = x(n_{i,j})$
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 $d_m > d_{m0}$, then
 - (a) set $d_{m0} = d_m$,
 - (b) set $\hat{\mathbf{X}} = \mathbf{X}$.
 - (c) If $\det(\hat{\mathbf{X}}' \hat{\mathbf{X}}) > \det_{\text{init}}$, then
 - i. set $\det_{\text{init}} = \det(\hat{\mathbf{X}}' \hat{\mathbf{X}})$,
 - ii. set $\tilde{\mathbf{X}} = \hat{\mathbf{X}}$.
8. Repeat (3–7) 1000 times. Keep the design $\tilde{\mathbf{X}}$.

Appendix B: 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

$$\mathcal{A}_i = \int_{\mathcal{X}_i} \int_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} \mathbf{h}(\mathbf{x}) \mathbf{h}(\mathbf{x}^*)' \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}_{-i}^* | \mathbf{x}_i^*) \mathrm{d}\mathbf{x}_{-i} \mathrm{d}\mathbf{x}_{-i}^* \mathrm{d}\mathbf{x}_i, \quad (\text{B1})$$

$$\mathcal{B}_i = \int_{\mathcal{X}_i} \int_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} \mathbf{t}(\mathbf{x}) \mathbf{h}(\mathbf{x}^*)' \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}_{-i}^* | \mathbf{x}_i^*) \mathrm{d}\mathbf{x}_{-i} \mathrm{d}\mathbf{x}_{-i}^* \mathrm{d}\mathbf{x}_i, \quad (\text{B2})$$

$$\mathcal{C}_i = \int_{\mathcal{X}_i} \int_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} \mathbf{t}(\mathbf{x}) \mathbf{t}(\mathbf{x}^*)' \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}_{-i}^* | \mathbf{x}_i^*) \mathrm{d}\mathbf{x}_{-i} \mathrm{d}\mathbf{x}_{-i}^* \mathrm{d}\mathbf{x}_i, \quad (\text{B3})$$

$$\mathcal{D}_i = \int_{\mathcal{X}_i} \int_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} c(\mathbf{x}, \mathbf{x}^*) \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}_{-i}^* | \mathbf{x}_i^*) \mathrm{d}\mathbf{x}_{-i} \mathrm{d}\mathbf{x}_{-i}^* \mathrm{d}\mathbf{x}_i, \quad (\text{B4})$$

$$\mathcal{E}_i = \int_{\mathcal{X}_i} \int_{\mathcal{X}_{-i} \times \mathcal{X}_{-i}} \mathbf{p}(\mathbf{x})' (\mathbf{H}' \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{p}(\mathbf{x}^*) \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}_{-i}^* | \mathbf{x}_i^*) \mathrm{d}\mathbf{x}_{-i} \mathrm{d}\mathbf{x}_{-i}^* \mathrm{d}\mathbf{x}_i, \quad (\text{B5})$$

such that

$$\mathbf{V}_i = \sum_{k=1}^{\ell} \sum_{k^*=1}^{\ell} (\boldsymbol{\beta}'_k \mathcal{A}_i \boldsymbol{\beta}_{k^*} + \boldsymbol{\beta}'_k \mathcal{B}_i \mathbf{e}_{k^*} + \mathbf{e}'_k \mathcal{B}_i \boldsymbol{\beta}_{k^*} + \mathbf{e}'_k \mathcal{C}_i \mathbf{e}_{k^*}) \mathbf{u}_k \mathbf{u}'_{k^*} + \sum_{k=1}^{\ell} \hat{\sigma}_k (\mathcal{D}_i - \text{tr}(\mathbf{A}^{-1} \mathcal{C}_i) + \mathcal{E}_i) \mathbf{u}_k \mathbf{u}'_k - \mathbf{C}, \quad (\text{B6})$$

where the formula for \mathbf{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

$$\mathcal{A}_0 = \iint_{\mathcal{X} \times \mathcal{X}} \mathbf{h}(\mathbf{x}) \mathbf{h}(\mathbf{x}^*)' \rho(\mathbf{x}) \rho(\mathbf{x}^*) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{x}^*, \quad (\text{B7})$$

$$\mathcal{B}_0 = \iint_{\mathcal{X} \times \mathcal{X}} \mathbf{t}(\mathbf{x}) \mathbf{h}(\mathbf{x}^*)' \rho(\mathbf{x}) \rho(\mathbf{x}^*) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{x}^*, \quad (\text{B8})$$

$$\mathcal{C}_0 = \iint_{\mathcal{X} \times \mathcal{X}} \mathbf{t}(\mathbf{x}) \mathbf{t}(\mathbf{x}^*)' \rho(\mathbf{x}) \rho(\mathbf{x}^*) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{x}^*, \quad (\text{B9})$$

$$\mathcal{D}_0 = \iint_{\mathcal{X} \times \mathcal{X}} c(\mathbf{x}, \mathbf{x}^*) \rho(\mathbf{x}) \rho(\mathbf{x}^*) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{x}^*, \quad (\text{B10})$$

$$\mathcal{E}_0 = \iint_{\mathcal{X} \times \mathcal{X}} \mathbf{p}(\mathbf{x})' (\mathbf{H}' \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{p}(\mathbf{x}^*) \rho(\mathbf{x}) \rho(\mathbf{x}^*) \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{x}^* \quad (\text{B11})$$

and

$$\mathbf{C} = \sum_{k=1}^{\ell} \sum_{k^*=1}^{\ell} (\boldsymbol{\beta}'_k A_3 \boldsymbol{\beta}_{k^*} + \boldsymbol{\beta}'_k \mathcal{B}_0 \mathbf{e}_{k^*} + \mathbf{e}'_k \mathcal{B}_0 \boldsymbol{\beta}_{k^*} + \mathbf{e}'_k \mathcal{C}_0 \mathbf{e}_{k^*}), \quad (\text{B12})$$

$$\mathbf{u}_k \mathbf{u}'_{k^*} + \sum_{k=1}^{\ell} \hat{\sigma}_k (\mathcal{D}_0 - \text{tr}(A^{-1} \mathcal{C}_0) + \mathcal{E}_0) \mathbf{u}_k \mathbf{u}'_{k^*}. \quad (\text{B13})$$

If the emulators for the different principal components use different Gaussian process parameters, then the integrals \mathcal{A}_i , \mathcal{B}_i , etc. need to be computed for all possible combinations k, k^* , which 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

$$\int_{\mathcal{X}_i} \int_{\mathcal{X}_i} \int_{\mathcal{X}_i} g(\mathbf{x}, \mathbf{x}^*) \rho(\mathbf{x}_{-i} | \mathbf{x}_i) \rho(\mathbf{x}^*_{-i} | \mathbf{x}^*_i) d\mathbf{x}_{-i} d\mathbf{x}^*_{-i} d\mathbf{x}_i. \quad (\text{C1})$$

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 $[0, 1]^3$), assuming a squared-exponential 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 (Gallassi et al., 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, $\mathbf{x}^{(j)}$, $j = 1 \dots q$ following the density $\rho(\mathbf{x})$, that is,
 - if we are in the case $\rho(\mathbf{x}) = 1$, then sample the uniform distribution;
 - if we are actually sampling the astronomical forcing, then generate an ensemble of possible units by computing the history of astronomical forcing (e.g. between -10^7 years and 0 at an interval of 1000 years with the astronomical solution of Berger (1978b)), and then sample q elements of this ensemble.
2. Shuffle the vectors \mathbf{x} to produce a sequence $\mathbf{x}^{*(j)}$, $j = 1 \dots q$.
3. For every j , modify the components $[i]$ of $\mathbf{x}^{*(j)}$ to make them equal to those of $\mathbf{x}^{(j)}$, i.e. $\mathbf{x}_i^{*(j)} \leftarrow \mathbf{x}_i^{(j)}$.
4. The sum $\sum_{j=1}^q g(\mathbf{x}^{(j)}, \mathbf{x}^{*(j)})$ converges to the integral (C1) as $q \rightarrow \infty$.

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 min on a laptop at time of writing if all PC components use the same Gaussian process parameters ν and Λ , or a couple of hours if parameters are different for each component.

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

The Supplement related to this article is available online at doi:10.5194/esd-6-205-2015-supplement.

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. M. Crucifix is a 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. The reviewers and editors 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. Data, R codes, and scripts for reproducing this research are available from doi:10.6084/m9.figshare.1368671.

Edited by: M. Huber

References

- Alpert, P. and Sholokhman, T. (Eds.): Factor Separation in the atmosphere: applications and future prospects, Cambridge University Press, Cambridge, UK, 2011.
- Andrianakis, I. and Challenor, P. G.: The effect of the nugget on Gaussian process emulators of computer models, *Comput. Stat. Data Anal.*, 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, *Clim. Past*, 11, 45–61, doi:10.5194/cp-11-45-2015, 2015.
- Berger, A.: Long-term variations of caloric insolation resulting from the earth's orbital elements, *Quaternary Res.*, 9, 139–167, doi:10.1016/0033-5894(78)90064-9, 1978a.
- Berger, A.: The role of CO₂, 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, 119–146, 1999.
- Berger, A. L.: Long-term variations of daily insolation and Quaternary climatic changes, *J. Atmos. Sci.*, 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, *J. Am. Stat. Assoc.*, 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, *Clim. Past Discuss.*, 11, 221–241, doi:10.5194/cpd-11-221-2015, 2015.
- Bounceur, N.: Quaternary Climate Changes: A global sensitivity analysis based emulation approach of a climate model response to forcing, PhD thesis, Université catholique de Louvain, Louvain-la-Neuve, 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, *Geophys. Res. Lett.*, 26, 2481–2484, doi:10.1029/1999GL006047, 1999.
- Braconnot, P., Otto-Bliesner, B., Harrison, S., Joussaume, S., Peterchmitt, J.-Y., Abe-Ouchi, A., Crucifix, M., Driesschaert, E., Fichefet, Th., 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, *Clim. Past*, 3, 261–277, doi:10.5194/cp-3-261-2007, 2007.
- Brovkin, V., Ganopolski, A., and Shvirezhev, Y.: A continuous climate-vegetation classification for use in climate-biosphere studies, *Ecol. Modell.*, 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, 31613–31624, doi: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, *Clim. Change*, 57, 119–138, doi:10.1023/A:1022168609525, 2003.
- Carnell, R.: lhs: Latin Hypercube Samples, available at: <http://CRAN.R-project.org/package=lhs> (last access: 27 April 2015), 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, doi:10.1038/nature12674, 2013.
- Claussen*, M.: Late Quaternary vegetation-climate feedbacks, *Clim. Past*, 5, 203–216, doi:10.5194/cp-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, *Geophys. Res. Lett.*, 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, *Clim. Dynam.*, 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, 2009.
- 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, 2010.
- de Noblet, N., Prentice, I. C., Joussaume, S., Texier, D., Botta, A., and Haxeltine, A.: Possible role of atmosphere-biosphere inter-

- actions in triggering the last glaciation., *Geophys. Res. Lett.*, 23, 3191–3194, doi:10.1029/96GL03004, 1996.
- 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, 2012.
- Fricker, T. E., Oakley, J. E., and Urban, N. M.: Multivariate Gaussian process emulators with nonseparable covariance structures, *Technometrics*, 55, 47–56, doi:10.1080/00401706.2012.715835, 2013.
- Galassi, M., Davies, J., Theiler, J., Gough, B., Jungman, G., Booth, M., and Rossi, F., GNU Scientific Library Reference Manual, Network Theory Ltd., UK, available at: <http://www.gnu.org/s/gsl> (last access: 27 April 2015), 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, *Geophys. Res. Lett.*, 29, 1860, doi: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, *Geosci. Model Dev.*, 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, *Stat. Comput.*, 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 CO₂, continental distribution, topography and vegetation changes on the climate at the Middle Miocene: a model study, *Clim. 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, *J. Am. Stat. Assoc.*, 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, *Clim. Dynam.*, 35, 785–806, doi:10.1007/s00382-009-0630-8, 2010.
- Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, *Reliabil. Eng. Syst. 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*, 18, 171–186, 2008.
- Kageyama, M., Charbit, S., Ritz, C., Khodri, M., and Ramstein, G.: Quantifying ice-sheet feedbacks during the last glacial inception, *Geophys. Res. Lett.*, 31, L24203, doi: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, 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, *Astronom. Astrophys.*, 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, *Atmos. Chem. Phys.*, 11, 12253–12273, doi:10.5194/acp-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, *Atmos. Chem. Phys.*, 13, 8879–8914, doi:10.5194/acp-13-8879-2013, 2013.
- Loepky, 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 insolutions, PhD 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, doi:10.1016/S0012-821X(04)00108-6, 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, *Clim. Past*, 10, 1541–1565, doi:10.5194/cp-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, *Clim. Dynam.*, 21, 515–537, doi:10.1007/s00382-003-0352-2, 2003.
- Morris, M. D. and Mitchell, T. J.: Exploratory designs for computational experiments, *J. Stat. Plann. Inference*, 43, 381–402, doi:10.1016/0378-3758(94)00035-T, 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, *J. Roy. Stat. Soc. B*, 66, 751–769, doi:10.1111/j.1467-9868.2004.05304.x, 2004.
- Opsteegh, J., Haarsma, R., Smelten, F., and Kattenberg, A.: EC-BILT: 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., 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 Planet. Change*, 45, 9–21, doi:10.1016/j.gloplacha.2004.09.011, 2005.
- Prokopenko, A., Williams, D., Kuzmin, M., Karabanov, E., Khur-sevich, 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, *Geophys. Res. Lett.*, 30, 1184, doi:10.1029/2002GL016636, 2003.
- Rougier, J.: Efficient Emulators for Multivariate Deterministic Functions, *J. Comput. Graph. Stat.*, 17, 827–843, doi:10.1198/106186008X384032, 2008.
- 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, *J. Climate*, 22, 3540–3557, doi:10.1175/2008JCLI2533.1, 2009.
- Ruddiman, W. F.: The early anthropogenic hypothesis a year later: challenges and responses, *Rev. Geophys.*, 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, *Stat. Science*, 4, 409–423, doi:10.1214/ss/1177012413, 1989.
- Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M.: *Sensitivity analysis in practice*, John, W. Sussex, England, 2004.
- Sanchez-Goni, 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 Planet. Sci. Lett.*, 171, 123–137, doi:10.1016/S0012-821X(99)00141-7, 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, 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, doi:10.1126/science.1120130, 2005.
- Silverman, B. W.: Smoothed functional principal components analysis by choice of norm, *The Ann. Stat.*, 24, 1–24, doi:10.1214/aos/1033066196, 1996.
- Stein, U. and Alpert, P.: Factor Separation in Numerical Simulations, *J. Atmos. Sci.*, 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.: Milankovitch and albedo forcing of the tropical monsoons: a comparison of geological evidence and numerical simulations for 9000 yBP, *Earth Environ. Sci. Trans. Roy. Soc. Edinburgh*, 81, 407–427, doi:10.1017/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 Planet. Change*, 36, 219–235, doi:10.1016/S0921-8181(02)00196-0, 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, *Comput. Geosci.*, 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, 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 Planet. Sci. Lett.*, 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., 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, *Clim. Dynam.*, 41, 1703–1729, doi: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, *Clim. Dynam.*, 1–26, doi:10.1007/s00382-014-2378-z, in press, 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, *Clim. Dynam.*, 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, *Clim. Dynam.*, 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, *Clim. Dynam.*, 28, 533–551, doi:10.1007/s00382-006-0196-7, 2007.