Earth System Dynamics
Discussions

Open Access

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

N. Bounceur[1], M. Crucifix[1], and R. D. Wilkinson[2]

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

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

Abstract

A global sensitivity analysis is used to describe the response of the Earth Climate Model of Intermediate Complexity LOVECLIM to components of the astronomical forcing (longitude of perihelion, obliquity, and eccentricity) assuming interglacial boundary conditions. Compared to previous studies, the sensitivity is global in the sense that it considers the full range of astronomical forcing that occurred during the Quaternary. We provide a geographical description of the variance due to the different components and their combinations and identify non-linear responses.

The methodology relies on the estimation of sensitivity measures, which due to the computational cost of LOVECLIM cannot be obtained directly. Instead, we use a fast surrogate of the climate model, called an emulator, in place of the simulator. A space filling design (a maximin Latin hypercube constrained to span the range of astronomical forcings characterising the Pleistocene) is used to determine a set of experiments to run, which are then used to train a reduced-rank Gaussian process emulator. The simulator outputs considered are the principal modes of the annual mean temperature, precipitation, and the growing degree days, extracted using a principal component analysis. The experiments are run on two distinct land surface schemes to address the effect of vegetation response on climate. Sensitivity to initial conditions is also explicitly assessed.

Precession and obliquity are found to contribute equally to growing degree days (GDD) in the Northern Hemisphere, and the effects of obliquity on the response of Southern Hemisphere temperature dominate precession effects. Further, compared to the original land-surface scheme with fixed vegetation, the LOVECLIM interactive vegetation induces non-linear responses in the Sahel-Sahara and Arctic sea-ice area. Finally, we find that there is no synergy between obliquity and precession.

1   Introduction

Any experiment with a climate model requires the specification of a number of input factors or parameters, be it parameters involved in parameterisations, or factors defining forcing conditions. It is good practice to assess carefully the sensitivity of
5 model outputs to these inputs. The objective may be either to quantify uncertainties associated with uncertain inputs (e.g., parameters), or to learn about climate's sensitivity to changes in external factors.

One classical approach for estimating a model's sensitivity consists in defining a reference state, and then changing one or two inputs (parameters or forcings) with
10 respect to the reference state. In particular, Stein and Alpert (1993) show how to identify second-order effects, that is, the extra contribution obtained by varying two factors at once, compared to the sum of variations obtained by varying the factors independently. These second-order effects are sometimes called *synergies* or *synergistic effects* in the climate literature (Ganopolski et al., 1998; Berger, 1999; Wohlfahrt et al., 2004;
15 Henrot et al., 2010).

The approach followed by Stein and Alpert (1993) and others is *local*, in the sense that it quantifies the model sensitivity between two well-defined states. By contrast, the purpose of a *global sensitivity analysis* is to explore systematically all the physically relevant space of inputs. Specifically, global sensitivity analysis consists in
20 estimating the relationships between the variances of input factors – either separate or combined – and the variances of outputs, and summarising this information by means of appropriately defined sensitivity measures (Saltelli et al., 2004).

One crucial aspect of global sensitivity analysis is the choice of experiments to be made. This is particularly important if the model is computationally expensive, as is
25 common in climate science. To illustrate this point consider a global sensitivity analysis of a model with d inputs. A factorial design is a natural choice, with experiments performed along a regular grid in the input space. However, the size of a factorial design grows exponentially with d. Even with $d = 3$, and assuming a resolution of ten

Title Page

Abstract   Introduction

Conclusions   References

Tables   Figures

|◄   ►|

◄   ►

Back   Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

values along each input axis, 1000 experiments would need to be run, which is already prohibitive for many climate simulators.

The theory of experimental design is the response to this problem. Its aim is to determine which experiments are expected to yield the best estimators of sensitivity measures. The theory was landmarked by the early studies of Sobol' (1976) and, in particular, by Sacks et al. (1989), Morris and Mitchell (1995) and Morris et al. (2008), who addressed the specificities of computer experiments in contrast to laboratory experiments.

Complementary to this approach, a number of authors, notably Sacks et al. (1989), Kennedy and O'Hagan (2000) and Oakley and O'Hagan (2004), developed a "meta-modelling" strategy. The "meta-model", often referred to as an "emulator" in this context, refers to a statistical model of the computer model (the "simulator"). The purpose of meta-modelling is to build a computationally cheap approximation to the simulator that can be used as a surrogate in any subsequent analysis. The approach is feasible because it is usually possible to formalise additional assumptions on the response structure of the simulator: typically the *smooth character* of the model's response to variations in inputs, or possible correlations between the different components of model outputs. This approach is effective for exploring the model response within the full input space. It accounts explicitly for prior probability distributions of input factors, and can be used to provide estimators of second-order effects associated with the combined variations of input factors.

The purpose of the present article is to develop an emulator-based global sensitivity analysis of a climate model to three input factors representing the astronomical forcing. This objective introduces a number of specific challenges not covered to our knowledge in the literature on global sensitivity analysis. First, we want to distinguish linear from non-linear effects. For example, the desertification of the Sahara is seen as a typical case of non-linear response to precession (Claussen et al., 1999). Can we identify this with an emulator-based approach? Second, we want to produce geographical maps, that is, deal with multivariate outputs. Thirdly, we have to deal with possible dependency

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

on initial conditions and with the possible co-existence of several stable solutions, and assess the extent to which they may induce artifactual conclusions.

2 Methodology

2.1 Emulator

Consider first the simpler case of simulators which have one-dimensional outputs. The Bayesian meta-model introduced by Sacks et al. (1989) and Kennedy and O'Hagan (2000) is used here because it is adequate for distinguishing local from global linear effects. It is also the basis of the emulator-based, global sensitivity analysis theory developed by Oakley and O'Hagan (2004).

Let $f(\mathbf{x})$ denote the simulator output when run at input vector $\mathbf{x}$. Until the simulator is evaluated at a given input, we are uncertain about its value. Thus, the Bayesian approach is to model f as an unknown stochastic quantity. Specifically, we assume

$$f(\cdot) \sim \text{GP}\left(\tilde{m}(\cdot), \tilde{V}(\cdot, \cdot)\right)$$

where this notation denotes that f is a Gaussian process (GP), with distribution fully specified by its mean and covariance functions, $\tilde{m}$ and $\tilde{V}$ (see, e.g. Rasmussen and Williams, 2005). We assume these take the form

$$\tilde{m}(\mathbf{x}) = \mathbf{h}(\mathbf{x})'\boldsymbol{\beta} \tag{1}$$

$$\tilde{V}(\mathbf{x}, \mathbf{x}^\star) = \sigma^2[c(\mathbf{x}, \mathbf{x}^\star)] \tag{2}$$

where $c(\mathbf{x}, \mathbf{x}^\star)$ is covariance function, σ^2 its variance, $\mathbf{h}(\mathbf{x})$ is a feature vector of q a priori known regression functions and $\boldsymbol{\beta}$ is the vector of corresponding regression coefficients. Note that the $()'$ is used to denote the transpose operation.

The interpretation of this model is that the simulator response is the sum of a global response function expressed as the linear combination of regressions $\tilde{m}(\mathbf{x}) = h(\mathbf{x})'\boldsymbol{\beta}$,

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

and a smooth local stochastic component (a zero-mean Gaussian process) that absorbs deviations from this global response.

Let $\mathcal{D} = \{x_i, y_i\}_{i=1}^{n}$ denote the set of simulator runs obtained in the designed experiment, where x_i is the input vector associated with the ith component of the experimental design, and y_i is the corresponding output. Let $y = (y_1, \ldots, y_n)'$ be the vector of all n outputs, and let $\mathbf{H}$ be the design matrix which has row i equal to the regressors $h(x_i)'$. The assumed GP model then states that the prior distribution of y is Gaussian, with $y \sim N(\mathbf{H}\beta, \sigma^2 A)$, where $\mathbf{A}$ is the Gram matrix with $A_{i,j} = c(x_i, x_j)$. This appears deceptively simple: although the distribution is restricted to being Gaussian, GPs form an extremely rich class of models, incorporating functional behaviour ranging from Brownian motion, to cubic-splines.

This prior distribution can then be updated using the output of the simulator experiments, $\mathcal{D}$. Assuming the vague prior $(\beta, \sigma^2) \propto \sigma^{-2}$ proposed by Berger et al. (2001) and used by, e.g., Oakley and O'Hagan (2002) and Bastos and O'Hagan (2009), the posterior distribution of the simulator output follows a Student−t distribution with $n - q$ degrees of freedom, with mean and variance

$$m(x) = h(x)'\beta + T(x^\star)'\mathbf{A}^{-1}(y - \mathbf{H}\hat{\beta}) \tag{3}$$

$$V(x, x^\star) = \hat{\sigma}^2[c(x, x^\star) - T(x)\mathbf{A}^{-1}T(x)' + P(x)(\mathbf{H}'\mathbf{A}^{-1}\mathbf{H})^{-1}P(x^\star)'], \tag{4}$$

respectively, with

$$\hat{\sigma}^2 = \frac{1}{n - q - 2}(y - \mathbf{H}\hat{\beta})'\mathbf{A}^{-1}(y - \mathbf{H}\hat{\beta}), \quad \text{and} \quad \hat{\beta} = (\mathbf{H}'\mathbf{A}^{-1}\mathbf{H})^{-1}\mathbf{H}'\mathbf{A}^{-1}y, \tag{5}$$

and $T(x)_j = c(x, x_j)$ and $P(x) = h(x)' - T(x)\mathbf{A}^{-1}\mathbf{H}$. Note that in the following we conveniently approximate the Student t distribution by a normal distribution, which although in principle is true only as $n \to \infty$, is accurate enough in practice for values of $n - q$ larger than 10.

Within this framework, the choices of the regression functions $h(x)$ and the Gaussian process correlation function c are application dependent. This is where the user may

inject knowledge of the expected response of the climate model, while reserving the option of considering other choices if the emulator performs poorly.

Given that our purpose is to distinguish linear global trends from local non-linear deviations, it is natural to set $h(x)' = (1, x')$ (linear regressors), so that the local stochastic component of the Gaussian process absorbs non-linearities.

One possible choice for c is the squared exponential decay with nugget, discussed in length in Andrianakis and Challenor (2012):

$$c(x_i, x_j) = \exp[-(x_i'\Lambda^{-2}x_j)] + v\|_{i=j},\tag{6}$$

where Λ is a scaling matrix (commonly called the length-scales), chosen to be diagonal with components λ_i. A popular alternative is the Matérn covariance function Berger et al. (2001), though it was not seen here to yield substantial improvement. The nugget term, $v\|_{i=j}$, was originally introduced to account for measurement errors in geospatial data analysis (Cressie, 1993). In emulators, it may also be introduced and justified, either as a regularisation *ansatz* to avoid poor matrix conditioning (Pepelychev, 2010), as a way to account for non-explicitly specified inputs (in the present case: initial conditions, sampling time and length), or as a way to account for the mis-specification in the correlation function (Gramacy and Lee, 2012).

There is no universal recommendation for choosing λ_i and v, and an analytical treatment of priors and posteriors is intractable. A standard approach is to choose values that lead to high likelihoods and that the user is prepared to defend as acceptable given what is already known about the simulator. The logarithm of the likelihood is

$$\log L(v, \Lambda) \propto -\frac{1}{2}(\log(|A||H'A^{-1}H|) + (n-q)\log(\hat{\sigma}^2)).\tag{7}$$

Andrianakis and Challenor (2012) further recommend the use of a *penalised likelihood*

$$\log L_p(v, \Lambda) = \log L(v, \Lambda) - 2\frac{\bar{M}(v, \Lambda)}{\epsilon\bar{M}(\infty)},\tag{8}$$

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

with $\bar{M}(\nu, \lambda) = \frac{\nu^2}{n}(\boldsymbol{y} - \mathbf{H}\hat{\boldsymbol{\beta}})'\mathbf{A}^{-2}(\boldsymbol{y} - \mathbf{H}\hat{\boldsymbol{\beta}})$ and $\bar{M}(\infty) = \frac{1}{n}(\boldsymbol{y} - \mathbf{H}\hat{\boldsymbol{\beta}})'(\boldsymbol{y} - \mathbf{H}\hat{\boldsymbol{\beta}})$. In effect this rejects high nugget values, and allows us to avoid emulators with error larger than those obtained in ordinary least-squares regression. We use $\epsilon = 1$. An alternative approach for ameliorating the problem with mis-specified covariance functions, is to integrate out the uncertainty in hyper-parameters (Agarwal and Gelfand, 2005). This is not possible analytically, and a Monte Carlo approach has to be taken instead. This approach is more computationally expensive, and not explored here.

2.2 Multivariate generalisation

We now focus on spatial resolution, and assume $f : \mathbb{R}^d \to \mathbb{R}^p$, with $m \gg 1$. Typically, the output will be a spatial field resolved onto a grid on the Earth's surface, and so m will be large. We hence need a multivariate emulator capable of mapping from multidimensional input to high dimensional output. The original version of Rougier (2008)[1], as well as Hankin (2012) provide useful summaries of available techniques.

For the present application we consider the PC emulator introduced by Higdon et al. (2008) and Wilkinson (2010) – who both consider the additional problem of calibrating on actual observations. The formalism used here is that of Wilkinson (2010).

Let $\mathbf{Y}$ denote the matrix in which each column represents the output of one experiment, i.e., $\mathbf{Y} = [\boldsymbol{y}(\boldsymbol{x}_1), \ldots, \boldsymbol{y}(\boldsymbol{x}_n)]$. The singular value decomposition (SVD) of the centered matrix $\mathbf{Y}^\star = \mathbf{Y} - \bar{\mathbf{Y}}$ (where $\bar{\mathbf{Y}}$ is the vector of row averages of $\mathbf{Y}$), is

$$\mathbf{Y}^\star = \mathbf{UDV}',$$

where $\mathbf{D}$ is diagonal, and $\mathbf{U}$ and $\mathbf{V}$ are square and orthonormal. The columns of $\mathbf{U}$ represent the basis vectors, $\{\boldsymbol{u}_k\}$, and the projection coefficients are given by $\mathbf{VD}$. For the jth experiment, the coefficient for the kth basis vector is $a_k(\boldsymbol{x}_j) = \mathbf{V}_{j,k}\mathbf{D}_{kk}$.

---

[1]Still available as http://www.mucm.ac.uk/Pages/Downloads/Technical%20Reports/07-12.pdf, last access: 11 July 2014.

Wilkinson (2010) keeps the first n' eigenvectors ordered by decreasing eigenvalues only. This is sometimes known as "hard-thresholding regularization" (Silverman, 1996), and results in a reduced-order model:

$$y(x) \approx \sum_{k=1}^{n'} a_k(x)u_k.$$

By using the SVD, we guarantee that the loss of information caused by this procedure is minimized, as the SVD gives the best low rank approximation to the full matrix as measured by the Frobenius norm.

If we model

$$a_k(x) \sim \text{GP}\left(m_k(\cdot), V_k(\cdot, \cdot)\right),$$

then $y(x)$ is a multi-output Gaussian process with mean and covariance function

$$m(x) = \sum_{k=1}^{n'} m_k(x)u_k \tag{9}$$

$$V(x, x^\star) = \underbrace{\sum_{k=1}^{n'} V_k(x, x^\star)u_k u'_k}_{V^{(gp)}} + \underbrace{\sum_{k=n'+1}^{n} \frac{D_{kk}^2}{n} u_k u'_k}_{V^{(pc)}}, \tag{10}$$

respectively.

Two options are considered for estimating Λ and v: (1) optimise, for every principal component independently, the penalised maximum likelihood Eq. (8), and (2) use the same v and Λ for all principal components.

Although option (2) may not be attractive because it does *not* maximise the likelihood of the emulator taken as a whole, it is shown in Sect. 2.5.4 that it presents strong computational benefits in the context of global sensitivity analysis. Hence, its performance for reconstructing spatial fields needs to be critically assessed with respect to option (1).

2.3 Input factors

The astronomical forcing is determined by three parameters relating to the Earth's orbit: eccentricity e, the heliocentric longitude of the perihelion ϖ, and obliquity ε.

The time-evolution of astronomical parameters over the Pleistocene is well known (Berger, 1978b; Laskar et al., 2004). Relevant to our purpose, eccentricity $e < 0.05$ over 99 % of the time, and the inner 99th percentile of obliquity is 22.3–24.3°, with differences of less than 0.1°, regardless of whether the Berger (1978b) or the Laskar et al. (2004) solution is used as the reference.

In palæoclimatology it is common to refer to the time of the year using the true solar longitude (λ), that is, the heliocentric angle between the vernal equinox and the position occupied by the Earth at any point during the year. For example, the June solstice corresponds to $\lambda = 90°$, the September equinox to $\lambda = 180°$ etc. For the purpose of computing insolation at a given time of year, we need the true solar longitude at perihelion, that is, the true solar longitude corresponding to the shortest Earth-Sun distance. This quantity is denoted ϖ. 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 well approximated as a linear combination of $i_1 = e\sin\varpi$, $i_2 = e\cos\varpi$ and $i_3 = \varepsilon$ (Loutre, 1993). The quantity i_1 is often referred to as the climatic precession parameter. As i_1, i_2 and i_3 are not correlated, the three inputs can be viewed as a canonical form of the astronomical forcing parameters. In particular, their signature on the season-latitude distribution of incoming solar radiation is characteristic: i_1 and i_2 control the Earth–Sun distance at any true solar longitude with very little effect on annual mean insolation[2], and i_3 controls the seasonal contrast and the annual mean insolation.

---

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

2.4 Experiment design

The climate model is the ocean-atmosphere-vegetation model of intermediate complexity "LOVECLIM" (Goosse et al., 2010). We concentrate here on the choice of inputs for the experiments to be run. This information is encoded in the design matrix $\mathbf{X}$, which has three columns (corresponding to the inputs $i_{1\ldots3}$), and as many rows as ensemble members.

Finding the experimental design that minimises the posterior variance of a Gaussian process emulator was first formalised by Sacks et al. (1989). This however, is only one aspect of the problem: one also wants minimum bias in the emulator prediction. Thus, finding a good design is largely a heuristic process based on good practice rules and a posteriori evaluation of algorithmic choices (e.g., Santner et al., 2003).

Theoretical considerations and experience both suggest that Latin-hypercube designs (McKay et al., 1979; Morris and Mitchell, 1995) are a good choice for computer experiment design. Urban and Fricker (2010) compare Latin-hypercube and factorial designs of the same size in the context of emulation. They find a small but significant improvement in emulator performance when using a Latin hypercube design, both for emulator prediction and for the estimation of first-order sensitivity indices.

This said, there are many possible Latin-hypercube designs for a given set of factors, and they are not all equally satisfactory. One generally fruitful approach consists in combining a Latin hypercube design with a maxi-min property. The maxi-min property consists in maximising the minimum Euclidean distance between two design points, and is thus a space-filling criteria. Johnson et al. (1990) showed that maximising the minimum distance is asymptotically equivalent to maximising the determinant of the covariance matrix $\mathbf{A}$ associated with the design points (assuming some generally met conditions on the covariance function). Joseph and Hung (2008) then propose an algorithm to combine maxi-min with the property of orthogonality, i.e., minimise pairwise correlation between design points, and show how this approach effectively leads to maximising the determinant of the information matrix.

Unfortunately these algorithms cannot be straightforwardly applied here, because we need to exclude the region $e > 0.05$ from the input space defined by $e\cos\varpi$, $e\sin\varpi$. The good news, however, is that the number of input factors (3) is sufficiently small to adopt a search of the optimal design by Monte Carlo generation of independent trials. The following hybrid-design generation algorithm provided satisfactory results.

1. The three factors are first standardised so that they cover the ranges $[-1, 1]$. Let $\mathbf{x} = (x_1, x_2, x_3)$ be a point of the input space, where the x_i are the three factors.

2. Set $d_{m0} = 0$, and $\det_{init} = 0$.

3. Sample a Latin hypercube design of $N = 27$ points in the 3-dimensional cube $[-1, 1]^3$, as follows:

   (a) Divide the interval $[-1, 1]$ into N equal-width intervals and number the middle of each interval. Let $x(i)$ be the midpoint of the ith interval.

   (b) Generate three random permutations of $i = 1...N$, denoted $\mathbf{n}_j = \{n_{i,j}\}_{i=1}^{N}$ for $j = 1, 2, 3$.

   (c) Form the design matrix $\mathbf{X}$, with $X_{i,j} = x\left(n_{i,j}\right)$

4. For every point of the design, check if the constraint on e is verified. 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 steps 4. and 5. until the dimension of the design $\mathbf{X}$ equals N

7. Calculate the minimum distance d_m between any two points of the design $\mathbf{X}$. If $d_m > d_{m0}$:

   (a) Set $d_{m0} = d_m$.

   (b) Set $\tilde{\mathbf{X}} = \mathbf{X}$.

(c) If $\det(\tilde{\mathbf{X}}'\tilde{\mathbf{X}}) > \det_{init}$

  i. Set $\det_{init} = \det(\hat{\mathbf{X}}'\hat{\mathbf{X}})$.

  ii. Set $\tilde{\mathbf{X}} = \hat{\mathbf{X}}$.

8. Repeat steps 3. to 7. 1000 times. Keep the design $\tilde{\mathbf{X}}$.

The choice of $n = 27$ experiments came from the general recommendation of Loeppky et al. (2009) to perform 10 experiments per input dimension. We used 27 members (and not 30), in order to compare with a factorial design of 3^3 members (N. Bounceur, thesis in preparation, not discussed here).

The resulting design is shown in 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 one is the pre-industrial conditions provided by default in the model package. For the second ensemble, initial conditions are the final state of the member #2 of the first ensemble. This particular member is a so-called warm orbit (high obliquity, high eccentricity and $\varpi \simeq 90°$) that produces extensive vegetation cover in the Northern Hemisphere. The purpose of using two distinct initial conditions is to detect the potential co-existence of distinct steady state solutions. 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. Experiments 20 and 27 are discussed separately in the Supplement. At this stage, we retain the experiment design with standard initial conditions for analysis. The third ensemble uses the same standard initial conditions, but uses the original ECBILT surface scheme with fixed vegetation (Opsteegh et al., 1998) and not the VECODE model. Each experiment is run 3000 years. The last 500 years are averaged and used in the following analysis.

2.5 Global variance measures

2.5.1 Definition

Recall that we are modelling the simulator $f(x)$ as a Gaussian process, and that after training the model we have

$$f(\cdot)|\mathcal{D} \sim \text{GP}\,(m(\cdot), V(\cdot, \cdot)).$$

To investigate the effect of each component of the astronomical forcing on the simulator outputs, it is necessary to partition the input vector x into components. Let x_p denote a subset of the components in x (with x_p taking values in $\mathcal{X}_p$), and let $x_{\bar{p}}$ be the remaining components. Let $\rho(x)$ be the time-wise occupation density of the input space during the Pleistocene, which can be estimated from standard astronomical solutions (Berger, 1978b; Laskar et al., 2004). Using this probability distribution for x, we can ask what is the average state of the system given a value for a particular component of the astronomical forcing. That is, we want to find the main effect of x_p, defined as

$$\eta(x_p) := \mathbb{E}(f(x)|x_p) = \int_{\mathcal{X}_{\bar{p}}} f(x)\rho(x_{\bar{p}}|x_p)\mathrm{d}x_{\bar{p}}$$

where $\rho(x_{\bar{p}}|x_p)$ is the distribution of forcing terms $x_{\bar{p}}$ given the value of x_p. In the following to keep the notation as light as possible, we sometimes abuse notation and write $\mathrm{d}\rho(x_{\bar{p}}|x_p) = \rho(x_{\bar{p}}|x_p)\mathrm{d}x_{\bar{p}}$ when necessary and if the meaning is clear.

In order to relate output variances to input variances, we are interested in quantities such as $\mathbb{V}\mathrm{ar}(\eta(x_p))$ where the variance is with respect to x_p (see interpretation in Sect. 2.5.2). However, because our knowledge of $f(x)$ is limited to the ensemble of model runs, we are uncertain about the value of $\eta(x_p)$ for every value of x_p. Note that $\eta(x_p)$ is a linear transformation of a Gaussian process, and is thus itself a Gaussian process with mean and covariance function found by applying the same transformation

to the original mean and covariance. That is,

$$\eta(\boldsymbol{x}_p) \sim \text{GP}\left(m_p(\boldsymbol{x}_p), V_p(\boldsymbol{x}_p, \boldsymbol{x}_p^*)\right)$$

where

$$m_p(\boldsymbol{x}_p) := \mathbb{E}_f(\eta(\boldsymbol{x}_p)) = \int_{\mathcal{X}_{\bar{p}}} m(\boldsymbol{x})\rho(\boldsymbol{x}_{\bar{p}}|\boldsymbol{x}_p)\,\mathrm{d}\boldsymbol{x}_{\bar{p}} \tag{11}$$

$$V_{pp}(\boldsymbol{x}_p, \boldsymbol{x}_p^*) := \mathbb{V}\mathrm{ar}_f(\eta(\boldsymbol{x}_p)) = \iint_{\mathcal{X}_{\bar{p}} \times \mathcal{X}_{\bar{p}}} V(\boldsymbol{x}, \boldsymbol{x}^*)\rho(\boldsymbol{x}_{\bar{p}}|\boldsymbol{x}_p)\rho(\boldsymbol{x}_{\bar{p}}^*|\boldsymbol{x}_p)\,\mathrm{d}\boldsymbol{x}_{\bar{p}}\mathrm{d}\boldsymbol{x}_{\bar{p}}^* \tag{12}$$

with $\mathbb{E}_f$ and $\mathbb{V}\mathrm{ar}_f$ denoting expectation and variance with respect to the Gaussian process model for f. The function $m_p(\boldsymbol{x}_p)$ describes the expected main effect, i.e., the mean state of the system assuming that component $\boldsymbol{x}_p$ is known. $V_{pp}(\boldsymbol{x}_p, \boldsymbol{x}_p^*)$ is the variance of our knowledge of the main effect due to using an emulator rather than the simulator itself.

It is then possible to estimate the variance of the main effects associated with the variance of the different input factors. We define two measures of sensitivity of the outputs to input $\boldsymbol{x}_p$:

$$S_p := \mathbb{E}_f \mathbb{V}\mathrm{ar}(\eta(\boldsymbol{x}_p))$$

and

$$\bar{S}_p := \mathbb{E}_f \left[\mathbb{V}\mathrm{ar}(\eta(\boldsymbol{x})) - \mathbb{V}\mathrm{ar}(\eta(\boldsymbol{x}_{\bar{p}})) \right].$$

Interpretations of these measures are given below. Oakley and O'Hagan (2004) show how to estimate these quantities using Gaussian process emulators in the case of

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

one-dimensional outputs. We need to define

$$\Sigma_p = \int_{\mathcal{X}_p} \left[m_p(\boldsymbol{x}_p)^2 + V_{pp}(\boldsymbol{x}_p, \boldsymbol{x}_p) \right] \mathrm{d}\rho(\boldsymbol{x}_p) = \mathbb{E}(m_p(\boldsymbol{x}_p)^2 + V_{pp}(\boldsymbol{x}_p, \boldsymbol{x}_p)), \tag{13}$$

$$\Sigma_0 = \left[m_0(\boldsymbol{x})^2 + V_{00}(\boldsymbol{x}, \boldsymbol{x}) \right], \text{ and} \tag{14}$$

$$\Sigma = \int_{\mathcal{X}} \left[m(\boldsymbol{x})^2 + V(\boldsymbol{x}, \boldsymbol{x}) \right] \mathrm{d}\rho(\boldsymbol{x}) \tag{15}$$

where the notation m_0, V_{00} implies that the space $\mathcal{X}_{\overline{p}}$ that appears in Eqs. (11) and (12) is the whole input domain $\mathcal{X}$. It can then be shown that

$$S_p = \Sigma_p - \Sigma_0$$

and

$$\bar{S}_p = \Sigma - \Sigma_{\overline{p}}.$$

2.5.2 Interpretation in the case of known inputs

The measures S_p and $\bar{S}_p$ are essentially those introduced by Homma and Saltelli (1996), further illustrated in Saltelli et al. (2004), chap. 1, and adapted to emulator theory by Oakley and O'Hagan (2004), except for the fact that these authors define scaled indices:

– $T_p = S_p/\Sigma$ is known as the main effect index associated with p, and

– $T_{\bar{p}} = \overline{S}_p/\Sigma$, the corresponding total effect index.

The main effect index quantifies the expected reduction in the variance of the output $f(\boldsymbol{x})$ if we were to learn $\boldsymbol{x}_p$, and the total effect index quantifies the expected reduction in the variance of the output $f(\boldsymbol{x})$ if we were to learn everything but $\boldsymbol{x}_p$.

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper

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

- S_p is the estimated simulator output variance lost by fixing p, all other forcings varying.

- $\overline{S}_p$ is the estimated simulator output variance induced by varying p, all other forcings fixed.

S_p and $\overline{S}_p$ differ if there are non-linear interactions between the different factors, or when the inputs are correlated. Only the first case is potentially relevant here.

Observe that Eqs. (13) and (14) contain two contributions: one from the squared mean of the emulator output, and one attached to the emulator variance. The sensitivity measure $\overline{S}_p$ may be split as follows:

$$\overline{S}_p = \overline{S}_p^m + \overline{S}_p^v. \tag{16}$$

The term $\overline{S}_p^m$ measures the estimated sensitivity of the simulator to the input variations, while $\overline{S}_p^v$ is a measure of the *uncertainty* introduced by using an emulator rather that the simulator itself. Specifically, we will refer to

- $\overline{S}_{1,2}^m$ as the variance induced by climate precession ($e\sin\varpi$ and $e\cos\varpi$ taken together),

- $\overline{S}_3^m$ as the variance induced by obliquity,

- $\overline{S}_{1,2,3}^m$ as the variance induced by all astronomical forcing components, and

- $\overline{S}_{1,2,3}^m - \overline{S}_{1,2}^m - \overline{S}_3^m$ as the second-order term representing the *synergy* between climatic precession and obliquity.

917

Finally, the computation of m_p^2 needed to compute $\overline{S}^m$ may be rewritten as follows (cf. Eq. 3):

$$m_p(\boldsymbol{x}_p)^2 = \int_{\mathcal{X}_{\bar{p}}} m(\boldsymbol{x})\mathrm{d}\rho(\boldsymbol{x}|\boldsymbol{x}_p) \int_{\mathcal{X}_{\bar{p}}} m(\boldsymbol{x}^\star)\mathrm{d}\rho(\boldsymbol{x}^\star|\boldsymbol{x}_p) = \iint_{\mathcal{X}_{\bar{p}}\times\mathcal{X}_{\bar{p}}} m(\boldsymbol{x})m(\boldsymbol{x}^\star)\mathrm{d}\rho(\boldsymbol{x}|\boldsymbol{x}_p)\mathrm{d}\rho(\boldsymbol{x}^\star|\boldsymbol{x}_p). \quad (17)$$

Concentrating on the product $m(\boldsymbol{x})m(\boldsymbol{x}^\star)$, we observe from Eq. (3) that:

$$m(\boldsymbol{x})m(\boldsymbol{x}^\star) = \underbrace{\hat{\boldsymbol{\beta}}' \boldsymbol{h}(\boldsymbol{x})\boldsymbol{h}'(\boldsymbol{x}^\star)\hat{\boldsymbol{\beta}}}_{\text{linear regression}} + \underbrace{\hat{\boldsymbol{\beta}}' \boldsymbol{h}(\boldsymbol{x})\boldsymbol{T}'(\boldsymbol{x}^\star)\boldsymbol{E} + \hat{\boldsymbol{\beta}}' \boldsymbol{h}(\boldsymbol{x}^\star)\boldsymbol{T}'(\boldsymbol{x})\boldsymbol{E} + \boldsymbol{E}'\boldsymbol{T}(\boldsymbol{x})\boldsymbol{T}'(\boldsymbol{x}^\star)\boldsymbol{E}}_{\text{higher order terms}}, \quad (18)$$

where $E = \mathbf{A}^{-1}(\boldsymbol{y} - \mathbf{H}\hat{\boldsymbol{\beta}})$. Thus, the simulator sensitivity $\overline{S}^m$ may be split into a contribution originating from the linear regression terms ($\boldsymbol{h}'\hat{\boldsymbol{\beta}}$) and an additional non-linear contribution as follows:

$$\overline{S}_p^m = \overline{S}_p^{m,\mathrm{l}} + \overline{S}_p^{m,\mathrm{nl}}, \text{ with l and nl for linear and non-linear.} \quad (19)$$

The values taken by $\overline{S}^{m,\mathrm{nl}}$ are illustrated for three specific cases in Fig. 3.

Note also that the matrices $\boldsymbol{hh}'$, $\boldsymbol{hT}'$, $\boldsymbol{TT}'$ are outer products that do not depend on the calibration data $\boldsymbol{y}$, whereas the vectors $\boldsymbol{E}$ and $\hat{\boldsymbol{\beta}}$ are independent of $\boldsymbol{x}$ but depend on the calibration data $\boldsymbol{y}$. This observation will be useful for simplifing the computations of variance measures with the PC emulator.

2.5.3 Generalisation to PC emulator

For the PC emulator, we have that

$$\boldsymbol{y}(\boldsymbol{x}) = \text{GP}\left(\sum_{k=1}^{n'} m_k(\boldsymbol{x})\boldsymbol{u}_k, \sum_{k=1}^{n'} V_k(\boldsymbol{x},\boldsymbol{x})\boldsymbol{u}_k\boldsymbol{u}_k'\right).$$

The covariance matrix is thus of dimension $m \times m$. In practice, the diagonal of this matrix, which expresses uncertainties about individual output components, will be insightful enough for our purpose. Thus, we restrict our attention to variance measures on individual output components and ignore covariances between different outputs. Specifically, the quantity $\mathbf{S}_{\bar{p}}$ is defined to be a vector, computed following Eq. (13) except that the terms entering the integrals now have the following from (dropping the p indices):

$$\mathbf{m}^2 = \sum_{k=1}^{n'} \sum_{k'=1}^{n'} m_k m_{k'} (\mathbf{u}_k \odot \mathbf{u}_{k'}) \tag{20}$$

$$\mathbf{V} = \sum_{k=1}^{n'} V_k (\mathbf{u}_k \odot \mathbf{u}_k), \tag{21}$$

where the notation $\mathbf{u}_k \odot \mathbf{u}_k$ is the component-wise multiplication of the two vectors (Hadamard product). These vectors may be efficiently computed in the particular case where all the components of the PC emulator use the same parameters $\mathbf{\Lambda}$ and v. Indeed, as observed after Eq. (18), the terms $\hat{\boldsymbol{\beta}}$ and $\mathbf{E}$ are independent of $\mathbf{x}$. Hence, integrals only need be carried over $\mathbf{h}\mathbf{T}'$, $\mathbf{T}\mathbf{T}'$, and $\mathbf{h}\mathbf{h}'$. As the latter are independent of the calibration data, they can be taken out of the summation and only need to be computed once.

Likewise, the calibration data $\mathbf{y}$ enter the posterior variance only through $\hat{\sigma}^2$, which is independent of $\mathbf{x}$. Again, the triple integrals need be computed only once for all the components of the emulator if the hyperparameters are constant across the different PCs.

2.5.4 Practical computation

Saltelli et al. (2010) reviews practical algorithms for the computation of the triple integrals needed to compute Σ_p. These integrals have the general form

$$\int_{\mathcal{X}_p} \left[\iint_{\mathcal{X}_{\bar{p}} \times \mathcal{X}_{\bar{p}}} g(\boldsymbol{x}|\boldsymbol{x}_p, \boldsymbol{x}^\star|\boldsymbol{x}_p) \mathrm{d}\rho(\boldsymbol{x}|\boldsymbol{x}_p)\mathrm{d}\rho(\boldsymbol{x}^\star|\boldsymbol{x}_p) \right] \mathrm{d}\rho(\boldsymbol{x}_p) \tag{22}$$

A Monte Carlo estimate of this expression may then be obtained with the following algorithm (Homma and Saltelli, 1996):

1. Sample m vectors, $\boldsymbol{x}_{i=1\ldots m}$ and $\boldsymbol{x}^\star_{i=1\ldots m}$ following the density $\rho(\boldsymbol{x})$,

2. Modify the components $[p]$ of $\boldsymbol{x}^\star$ to make them equal to those of $\boldsymbol{x}$, i.e., $\boldsymbol{x}_p \to \boldsymbol{x}^\star_p$.

3. The sum

$$\frac{1}{2m} \left(\sum_{i=1}^{m} g(\boldsymbol{x}_i, \boldsymbol{x}^\star_i) + \sum_{i=1}^{m} g(\boldsymbol{x}^\star_i, \boldsymbol{x}_i) \right) \tag{23}$$

converges to the integral Eq. (22) as $m \to \infty$.

The sample size m is chosen empirically to be large enough to yield acceptable estimator variance.

3 Results

3.1 Validation and choice of PC emulator

We concentrate on three outputs: annual precipitation, growing degree days (GDD), and annual mean temperature. GDD is defined here as the annual sum of daily

temperatures (in Celsius) exceeding 0 °C. It is used as a calendar-independent indicator of summer intensity and length in extra-tropical regions. Namely, the vegetation model VECODE, used in LOVECLIM, uses GDD and annual precipitation to predict the dynamics of vegetation Brovkin et al. (1997). For simplicity GDD is estimated here from monthly means, that is, 30× 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. On the other hand, we use the logarithm of annual precipitation rather than precipitation, as the first quantity is closer to being Gaussian distributed than the second one.

Note that in the following analysis, experiment 20 (the experiment with lowest obliquity, Fig. 1) is taken out of the experimental design. It is an outlier, in the sense that it is poorly predicted by the other experiments, and its inclusion degrades the predictive performance of the emulator. The case is further discussed in the Supplement.

Two options are considered for the calibration of hyperparameters Λ and v: (i) use different hyperparameters for each PC, each obtained by maximising the penalised likelihood (Eq. 8), (ii) use the same value for each of the first 10 PCs, in which case this is the product of the likelihoods of each PC that is maximised. The resulting Λ, v and log-likelihoods are shown on Fig. 6. One observes PC scores becoming decreasingly informative about the model sensitivity as their index increases. Indeed, later PCs are more design dependent, and their response is difficult to predict. This is reflected by the overall decreasing trend in the components of Λ, the occurrence of high values of v (some are out of the y-axis range) and the decrease in the likelihood. In fact, likelihood stabilises around PC #10 to a minimum value that indicates that the calibrated GP is not more informative than assuming independence of outputs and inputs.

Figure 4 shows the resolved and unresolved variances as a function of n', for the different fields. Generally, around 99 % of the mean variance is already well-captured by the first 3 PCs. However, the lower-variance PCs need to be taken into account as well: they represent variability modes such as related to tropical circulation or north-Atlantic convection sites that are locally important, even if their contribution to the mean

global variance is small. Using more PC scores consequently reduces the number of grid points that are incorrectly predicted by the PC emulator. Furthermore, the scores of the first 10 PCs can generally be correctly predicted with a suitably calibrated Gaussian process, which suggests that they are actually informative and should be retained. Hence our final choice is $n' = 10$.

The evaluation strategy relies on the leave-one-out approach: for each member of the experimental design, a PC emulator is trained using the hyperparameters defined above. Based on Eqs. (9) and (10), the means and standard deviation are predicted for each grid point of the member left out. Figure 7 shows (bars) the number of grid points correctly predicted within 1, 2 or 3 standard deviations. A well-calibrated emulator would get 66, 95 and 99 %, respectively, in each category. Points incorrectly predicted within 3 standard deviations (in red) are outliers. The result is satistfactory overall. Some specific remarks:

1. Using constant rather than PC-specific hyperparameters does not significantly affect the overall performance. This is explicitly shown for GDD, but this is also true of the other fields.

2. However, all fields exhibit an excessive number of outliers, compared to the ideal frequency of 1 %. Annual mean precipitation is, in this respect, less well predicted than the others, perhaps not surprisingly given that precipitation responds less straightforwardly to radiation changes than temperature.

3. The existence of outliers is compensated for by a larger-than-expected number of errors with less than 1 standard deviation. This is typical of heavy-tailed distributions.

4. Experiment 20 of the original ensemble was discarded based on this diagnostic. Inclusion of this experiment significantly increases the number of outliers, as is shown in the Supplement.

3.2 Total variance

3.2.1 Choice of *m* for integration

We focus first on the variance of GDD, and to be specific, consider the precession induced component $\overline{S}_{1,2}^m$. As explained in Sect. 2.5.4 this quantity is estimated from a Monte Carlo sum. To limit the computational expense we limit the reconstruction to the first two PCs, and consider different realisations of this Monte Carlo estimator, ordered by increasing Monte Carlo sample size (Fig. 2). A 10 000-member ensemble size generates a dispersion of the estimate of $\overline{S}_{1,2}^m$ of the order of 2 %. The dispersion of the integrals associated with individual PC scores (not shown) is of the order of 5–8 %, so one may consider 5 % as a rough estimate of the uncertainty on variance estimates discussed in the following.

3.2.2 Contributions to total variances

Figure 5 decomposes the variances of the different fields (grid-point averages) into their contributions from the different inputs, and compares these quantities to different elements of variances associated with the PC emulator. The following observations can be made:

1. Variances explicitly associated to inputs (precession and obliquity) largely dominate both the variances associated with using the PC emulator, and the variance associated with the choice of initial conditions. Annual precipitation is the field for which the variance associated with the discarded PCs is the largest, but even in that case it is much smaller than the variance associated with obliquity and precession.

2. Annual mean temperature is more controlled by obliquity than by precession. This result is reasonable, considering that annual mean insolation is exclusively controlled by obliquity. By contrast, precipitation is more controlled by precession

than obliquity. Again this matches the general understanding of climate dynamics, given the importance of monsoonal systems for annual precipitation, reportedly sensitive to precession. Finally, GDD is equally controlled by these two parameters, because this is a measure that is both sensitive to the peak insolation value (which is dominated by precession) and to the integrated incoming energy over the warm season (which more obliquity controlled).

3. The synergy term is extremely small. In fact it is not quantifiable given that it is smaller than uncertainties associated with the PC emulator and sampling variability.

4. In these diagrams the two components that contribute to $\overline{S}^v_{1,2,3}$ in the PC emulator, that is, the variance associated with the Gaussian process of the explicitly resolved PCs ($V^{(gp)}$, see Eq. (10)), and the variance associated with the discarded PCs (V^{pc}) are distinguished. For $n' = 10$, V^{pc} dominates. As n' is increased V^{pc} is gradually transferred to V^{gp} until the latter absorbs all the variance.

3.3 Geographical distribution of obliquity and precession effects

Figure 8 reproduces the geographical distributions of variances associated with precession and obliquity, based on Eq. (9). It is observed that:

1. Precipitation variance is dominated by some monsoon regions, namely Western Africa and Australia. The absence of large variance patterns in South-East Asia and South America is most reasonably attributed to the limitations of LOVECLIM in simulating tropical weather systems. Note also the significant influence of obliquity in the most western part of North Africa.

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

3. Annual mean temperature has the highest variance near the poles. Again, Northern Hemisphere temperature is evenly controlled by precession and obliquity, while obliquity dominates the variance in the Southern Hemisphere.

3.4 Non-linearity indices

To illustrate the potential interest of non-linearity indices we compare two sets of experiments. The first is the one which has been discussed so far, using interactive vegetation VECODE. The second set of experiments is run along the same design (with, for consistency, exp. 20 similarly omitted), but uses the original surface scheme of ECBILT Opsteegh et al. (1998) where soil properties are fixed (hereafter termed "VOFF"). An improved experimental design would admittedly have been based on VECODE with fixed-vegetation, but the current software configuration of LOVECLIM made it technically challenging.

Results are shown on Fig. 9. The immediate impression is that the VECODE scheme produces more non-linearity than the original ECBILT scheme (VOFF), in spite of total variances being broadly similar for GDD and annual temperature (not shown). Specifically:

1. VOFF does not produce the precipitation non-linearity spot in western Africa. In this region, the response of vegetation is essential to produce large response patterns. The non-linearity index is positive, typical of a sigmoid response (Fig. 3) and this effect was earlier documented in this model (Renssen et al., 2003). It is also consistent with the findings of Claussen et al. (1999).

2. In North America, the non-linearity index produced by the VECODE scheme is negative. This is typical of a saturation effect. It was also seen (but not shown here) that the total variance in this region is also slightly smaller with VECODE than with VOFF. This effect is unexpected and not easily explained: it may be a consequence of the difference in hydrological surface properties, and/or

a consequence of the difference in mean state between VON and VOFF, with possible consequences for the amplitude of the snow albedo positive feedback.

3. Runs with VECODE show significant non-linearity in the North Atlantic and over the Arctic. This implies that the model with VECODE is being brought into a regime where large sea-ice responses may occur, and play a larger amplifying role than in the VOFF experiments.

4   Conclusion and discussion

Our ability to relate input to output variances is one of the elements that characterises our understanding of a model, in general, and of a computer model (simulator), in particular.

In the current study we focused on output variances that are related to variations in the inputs caused by the slow changes in Earth's orbital parameters. The primary purpose is to compare the effects, and possible synergies between precession and obliquity. We thus ignored a number of sources of uncertainty, including the structural choice of the climate model (i.e., using LOVECLIM rather than another model and errors in LOVECLIM), and the values of the parameters involved in parameterisations.

This is thus not a full global sensitivity analysis, but it is nevertheless a significant step forward, because it allows us to approach the variety of astronomical configurations experienced during the Quaternary with no need to explicitly consider a *reference* state. The conclusions may be summarised as follows.

First, we note that the results shown in Figs. 8–9, which depicts LOVECLIM's sensitivity to astronomical forcing, are generally consistent with current knowledge on the actual climate system:

1. Precession and obliquity both contribute to annual temperature. Precession has generally more effect in the Northern Hemisphere and tropical regions, and obliquity is the dominant forcing in the Southern Hemisphere. This is physically

reasonable: precession has virtually no effect on annual mean insolation. It may thus only affect annual temperature through non-linear effects, for example associated with albedo forcing, which are more easily solicited in the Northern Hemisphere, given the presence of large continental areas. This result is also consistent with the predominance of obliquity signals in Antarctic cores, be it CO_2 concentration (Petit et al., 1999; Siegenthaler et al., 2005; Luethi et al., 2008) or Deuterium excess (Vimeux et al., 2002). This dichotomy between Northern and Southern Hemisphere sensitivities to obliquity and precession is a reasonable basis to explain the contrasting dynamics between southern records and a northern continental records, such as Baïkal's, during isotopic stage 11 (Prokopenko et al., 2002). Our results are also consistent with an explanation of the mid-Brünhes events based on this difference between northern and southern ocean responses (Yin, 2013).

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

3. The approach identifies non-linearity response spots associated with the vegetation and sea-ice. This is again generally consistent with our current understanding of climate dynamics, and in particular how slow variations in astronomical forcing may induce fairly rapid climate transitions (Claussen et al.,

1999). Non-linear responses are, by the very nature of being non-linear, sensitive to the climate background state, which itself is determined by numerous explicit and implicit modelling choices. Therefore, the location of a non-linear event in the input space – thus, here, the timing of the event, since the input space is the astronomical forcing – is generally highly sensitive to those modelling choices. This is where global sensitivity analysis shows a specific advantage compared to local approaches: as the input space is explored systematically, it does not rely on the model simulating the event with the correct timing. It rather characterises the nonlinearity of the response, and offers the possibility to locate non-linear events a posteriori (and possibly probabilistically if uncertain inputs are included in the analysis), hence yielding a much more robust analysis.

In addition, one result contrasts with previously published conclusions:

4. The current analysis reveals no synergy between obliquity and precession. Such synergy had been previously proposed as an important contributor to climate change during interglacials (Crucifix et al., 2002; Crucifix and Loutre, 2002). In fact, we have seen climate does indeed respond non-linearly to insolation, but this non-linearity is not specific to the association of precession and obliquity changes. This is the reason why the synergy measure is here essentially null.

The spirit of sensitivity analysis in general, and of global sensitivity analysis in particular, is to aim for robustness and uncertainty quantification. This can be achieved, but at the cost of added conceptual and technical complexity compared to conventional approaches. Indeed, we used an emulator that involves a number of validation steps that are not immediately related to the climate model itself (namely, the number of principal components, the methodology for finding hyper-parameters, and the quality of the experiment design). Paradoxically, while aiming for robustness it is necessary to introduce additional assumptions and parameters in the analysis process. In fact, some of the assumptions made explicit here are implicit in the more conventional approaches, in particular the smoothness of the simulator response, or the representativeness of the

reference experiment. This may be appreciated through the case of experiment "20", which was rejected as an outlier. Admittedly, this experiment is useful because it may reveal significant dynamics occurring at low obliquity, but the emulator shows that it does not represent the simulator's response over most of the input space. A simpler experiment design, that would have only compared experiment 20 with a reference state, would have missed this point.

The Supplement related to this article is available online at doi:10.5194/esdd-5-901-2014-supplement.

References

Agarwal, D. K. and Gelfand, A. E.: Slice sampling for simulation based fitting of spatial data models, Stat. Comput., 15, 61–69, doi:10.1007/s11222-005-4790-z, 2005. 908

Andrianakis, I. and Challenor, P. G.: The effect of the nugget on Gaussian process emulators of computer models, Comput. Stat. Data An., 56, 4215–4228, doi:10.1016/j.csda.2012.04.020, 2012. 907

Bastos, L. S. and O'Hagan, A.: Diagnostics for Gaussian process emulators, Technometrics, 51, 425–438, doi:10.1198/TECH.2009.08019, 2009. 906

Berger, A. L.: 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. 927

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

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. 910, 914

Berger, A.: The role of CO_2, sea-level and vegetation during the Milankovitch-forced glacial-interglacial cycles, in: Proceedings "Geosphere-Biosphere Interactions and Climate", Pontifical Academy of Sciences, Vatican City, 9–13 November 1998, 119–126, available at: http://www.casinapioiv.va/content/dam/accademia/pdf/sv96pas.pdf, last access: 11 July 2014, 1999. 903

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. 906, 907

Brovkin, V., Ganopolski, A., and Shvirezhev, Y.: A continuous climate–vegetation classification for use in climate-biosphere studies, Ecol. Model., 101, 251–261, doi:10.1016/S0304-3800(97)00049-5, 1997. 913, 921

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. 904, 925, 927

Cressie, N.: Statistics for Spatial Data, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Chichester, UK, 1993. 907

Crucifix, M.: How can a glacial inception be predicted?, Holocene, 21, 831–842, doi:10.1177/0959683610394883, 2011. 927

Crucifix, M. and Loutre, M. F.: Transient simulations over the last interglacial period (126–115 kyr BP): feedback and forcing analysis, Clim. Dynam., 19, 419–433, doi:10.1007/s00382-002-0234-z, 2002. 928

Crucifix, M., Loutre, M. F., Tulkens, P., Fichefet, T., and Berger, A.: Climate evolution during the Holocene: a study with an Earth system model of intermediate complexity, Clim. Dynam., 19, 43–60, doi:10.1007/s00382-001-0208-6, 2002. 928

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. 903

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.,

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

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. 911

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. 907

Hankin, R. K. S.: Introducing multivator: A Multivariate Emulator, J. Stat. Softw., 46, 1–20, available at: http://www.jstatsoft.org/v46/i08, last access: 11 July 2014, 2012. 908

Henrot, A.-J., François, L., Favre, E., Butzin, M., Ouberdous, M., and Munhoven, G.: Effects of CO_2, continental distribution, topography and vegetation changes on the climate at the Middle Miocene: a model study, Clim. Past, 6, 675–694, doi:10.5194/cp-6-675-2010, 2010. 903

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. 908

Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, Reliab. Eng. Syst. Safe., 52, 1–17, doi:10.1016/0951-8320(96)00002-6, 1996. 916, 920

Johnson, M. E., Moore, L. M., and Ylvisaker, D.: Minimax and maximin distance designs, J. Stat. Plan. Infer., 26, 131–148, doi:10.1016/0378-3758(90)90122-B, 1990. 911

Joseph, V. R. and Hung, Y.: Orthogonal-maximin latin hypercube designs, Stat. Sinica, 18, 171–186, 2008. 911

Kennedy, M. C. and O'Hagan, A.: Predicting the output from a complex computer code when fast approximations are available, Biometrika, 87, 1–13, doi:10.1093/biomet/87.1.1, 2000. 904, 905

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, Astro. Astrophys., 428, 261–285, doi:10.1051/0004-6361:20041335, 2004. 910, 914

Loeppky, J. L., Sacks, J., and Welch, W. J.: Choosing the sample size of a computer experiment: a practical guide, Technometrics, 51, 366–376, doi:10.1198/TECH.2009.08040, 2009. 913

Loutre, M. F.: Paramètres orbitaux et cycles diurne et saisonnier des insolations, Ph. D. thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 1993. 910

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

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. 927

5 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. 911

Morris, M. D. and Mitchell, T. J.: Exploratory designs for computational experiments, J. Stat. Plan. Infer., 43, 381–402, doi:10.1016/0378-3758(94)00035-T, 1995. 904, 911

10 Morris, M. D., Moore, L. M., and McKay, M. D.: Using orthogonal arrays in the sensitivity analysis of computer models, Technometrics, 50, 205–215, doi:10.1198/004017008000000208, 2008. 904

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. 906

15 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. 904, 905, 915, 916

Opsteegh, J., Haarsma, R., Smelten, F., and Kattenberg, A.: ECBILT: a dynamic alternative to mixed boundary conditions in ocean models, Tellus A, 50, 348–367, doi:10.1034/j.1600-0870.1998.t01-1-00007.x, 1998. 913, 925

20 Pepelychev, A.: The role of the nugget term in the Gaussian process method, in: mODa 9 Advances in Model-Oriented Design and Analysis, Contributions to Statistics, edited by: Giovagnoli, A., Atkinson, A. C., Torsney, B., and May, C., Physica Verlag, Springer, Heidelberg, 149–156, 2010. 907

25 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. 927

30 Pollard, D. and DeConto, R. M.: Hysteresis in Cenozoic Antarctic ice-sheet variations, Long-term changes in Southern high-latitude ice sheets and climate, the Cenozoic history, Global Planet. Change, 45, 9–21, doi:10.1016/j.gloplacha.2004.09.011, 2005. 927

Prokopenko, A., Williams, D., Kuzmin, M., Karabanov, E., Khursevich, G., and Peck, J.: Muted climate variations in continental Siberia during the mid-Pleistocene epoch, Nature, 418, 65–68, doi:10.1038/nature00886, 2002. 927

Rasmussen, C. and Williams, C.: Gaussian Processes for Machine Learning, Adaptive Computation And Machine Learning, MIT Press, Cambridge MA, 2005. 905

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. 925

Rougier, J.: Efficient emulators for multivariate deterministic functions, J. Comput. Graph. Stat., 17, 827–843, doi:10.1198/106186008X384032, 2008. 908

Ruddiman, W. F.: The early anthropogenic hypothesis a year later: challenges and responses, Rev. Geophys., 45, RG4001, doi:10.1029/2006RG000207, 2007. 927

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P.: Design and analysis of computer experiments, Stat. Sci., 4, 409–423, doi:10.1214/ss/1177012413, 1989. 904, 905, 911

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M.: Sensitivity Analysis in Practite, W. Johx, Sussex, England, 2004. 903, 916

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S.: Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Comput. Phys. Commun., 181, 259–270, doi:10.1016/j.cpc.2009.09.018, 2010. 920

Santner, T., Williams, B., and Notz, W.: The Design and Analysis of Computer Experiments, Springer, New York, 2003. 911

Siegenthaler, U., Stocker, T. F., Lüthi, D., Schwander, J., Stauffer, B., Raynaud, D., Barnola, J.-M., Fisher, H., Masson-Delmotte, V., and Jouzel, J.: Stable carbon cycle-climate relationship during the Late Pleistocene, Science, 310, 1313–1317, 2005. 927

Silverman, B. W.: Smoothed functional principal components analysis by choice of norm, Ann. Stat., 24, 1–24, doi:10.1214/aos/1033066196, 1996. 909

Sobol', I. M.: Uniformly distributed sequences with an addition uniform property, USSR Comp. Math. Math.+, 16, 236–242, 1976. 904

Stein, U. and Alpert, P.: Factor separation in numerical simulations, J. Atmos. Sci., 50, 2107–2115, 1993. 903

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. 911

Climate–vegetation response to astronomical forcing

N. Bounceur et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀ | ▶

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

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. Sc. Lett., 203, 829–843, doi:10.1016/S0012-821X(02)00950-0, 2002. 927

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., John Wiley & Sons, Ltd, Chichester, UK, doi:10.1002/9780470685853.ch10, 195–215, 2010. 908, 909

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. 903

Yin, Q.: Insolation-induced mid-Brunhes transition in Southern Ocean ventilation and deep-ocean temperature, Nature, 494, 222–225, doi:10.1038/nature11790, 2013. 927
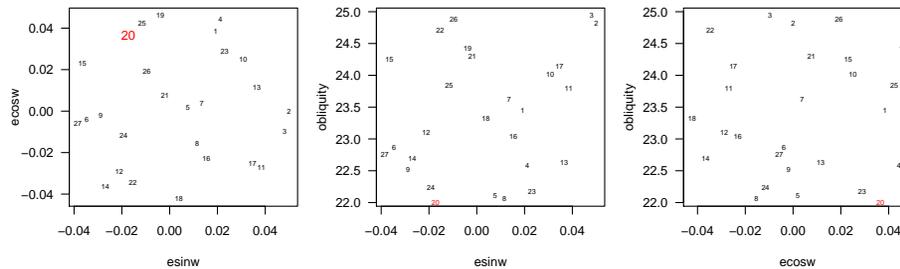
Figure 1. Experiment design. The experiment marked in red was discarded from the analysis (cf. the Supplement).
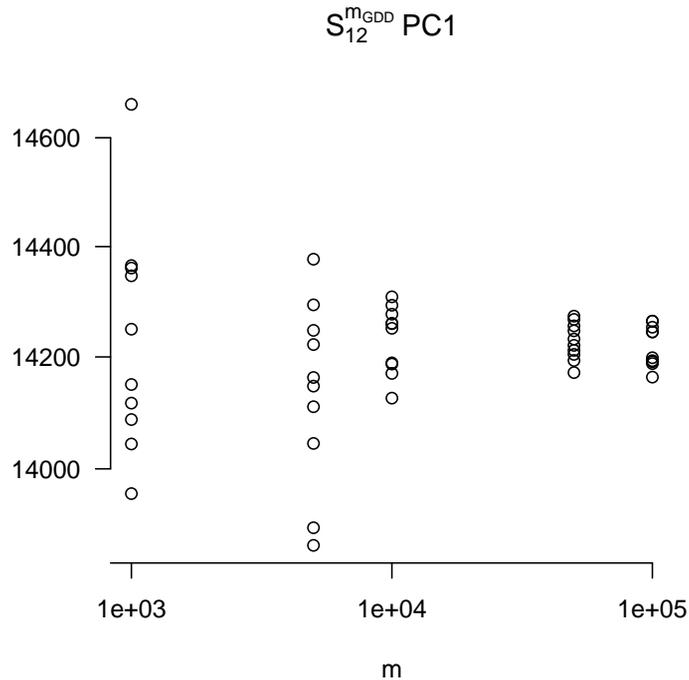
Figure 2. Monte Carlo estimates of $\overline{S}_{1,2}^{m}$ associated with GDD, averaged over all grid cells, as a function of the sample size of the Monte Carlo estimator. We hereafter use 10 000 members.
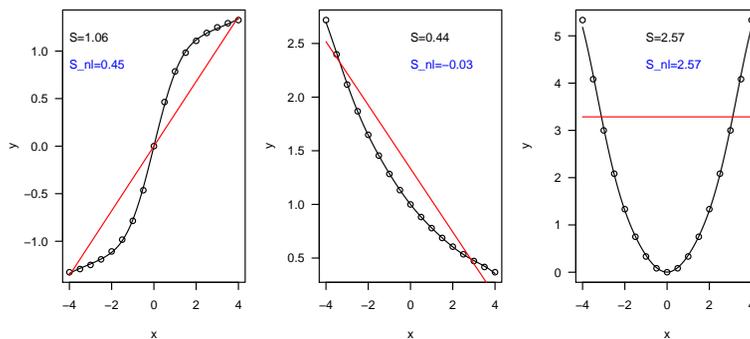
Figure 3. Estimate of linear and non-linear variance indices (S^m and $S^{m,nl}$) in three toy examples, where an emulator is calibrated on three distinct one-dimensional datasets (open circles). Full line is the emulator prediction and red lines its linear component. Note that the sigmoid response produces a positive non-linearity index, and the exponential, a negative one.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
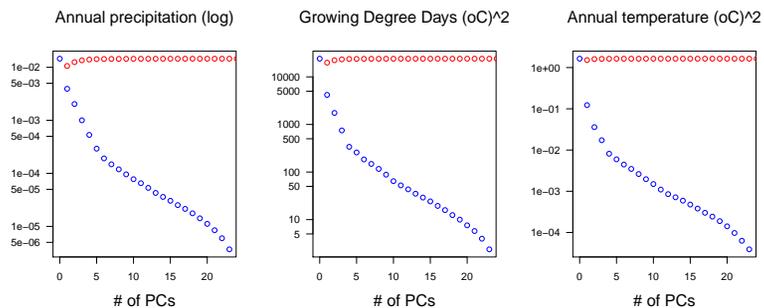
Figure 4. PCA decomposition of the logarithm of annual precipitation, GDD and annual temperature fields based on the experiment design. Shown are the cumulated variance resloved by the PC (red) and the left variance (blue), which is modelled as white noise in Eq. (10), as a function of n'. Quantities are grid-cell averages.
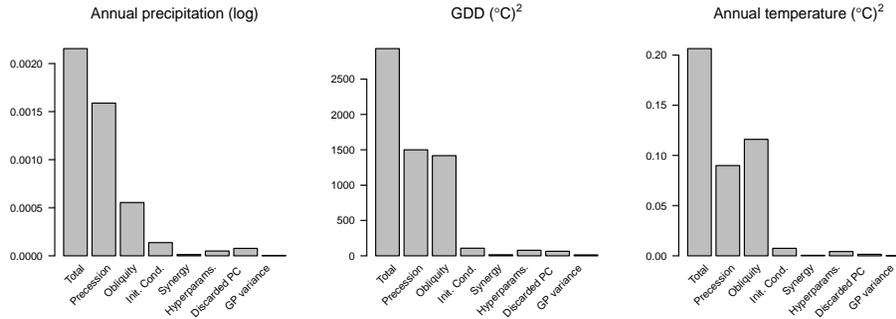
Figure 5. Total variances associated to inputs ($\overline{S}^m_{1,2,3}$), averaged over all grid points, and its contributions associated with precession ($\overline{S}_{1,2}$), obliquity ($\overline{S}_{1,2}$), and synergy ($\overline{S}^m_{1,2,3} - \overline{S}^m_{1,2} - \overline{S}^m_{1,2}$). These variances are then compared to different elements of variances associated with the PC emulator, namely: output dependency to initial conditions (exp. 20 excluded), the sensitivity of $\overline{S}^m_{1,2,3}$ to the choice of hyperparameters (constant for all PC, or PC-dependent), and the field variance associated to the emulator $\overline{S}^v_{1,2,3}$, split into the contribution assiated with the discarded PC ($V^{(pc)}$ in Eq. 10) and to the co-variance of the Gaussian process ($V^{(gp)}$) of the PCs explicitly resolved.
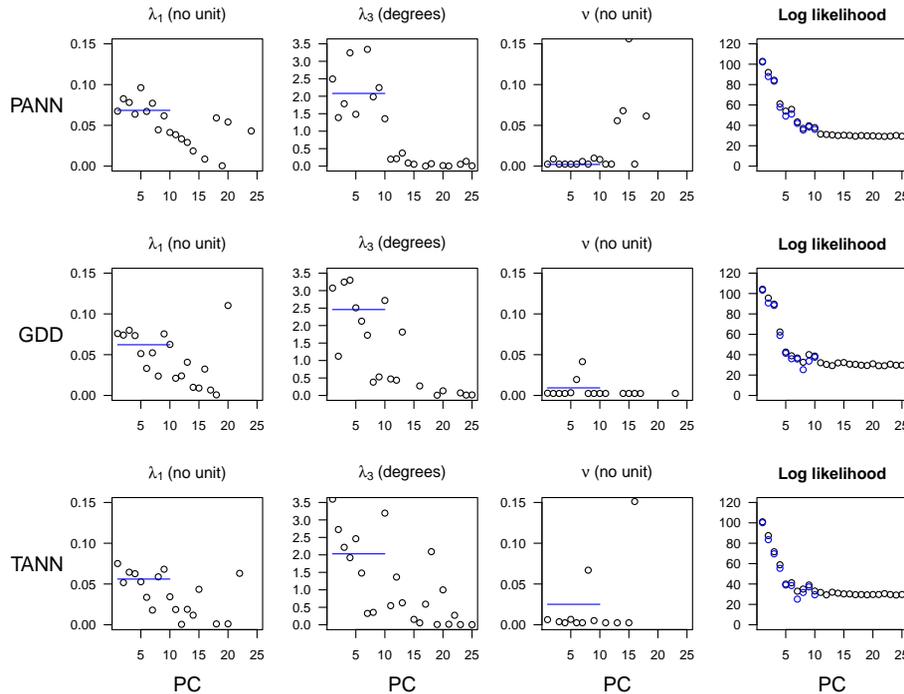
Figure 6. Hyperparameters maximising the penalised log-likelihood (Eq. 8) for three variables: GDD, log(Annual Precipitation) and Annual temperature, either *(black)* optimised for each PC independently or *(blue)* optimised based on the product of the likelihoods of the first 10 PC, assuming that the same hyperparameters are used on all PCs. The log-likelihood associated with each PC is given for reference.
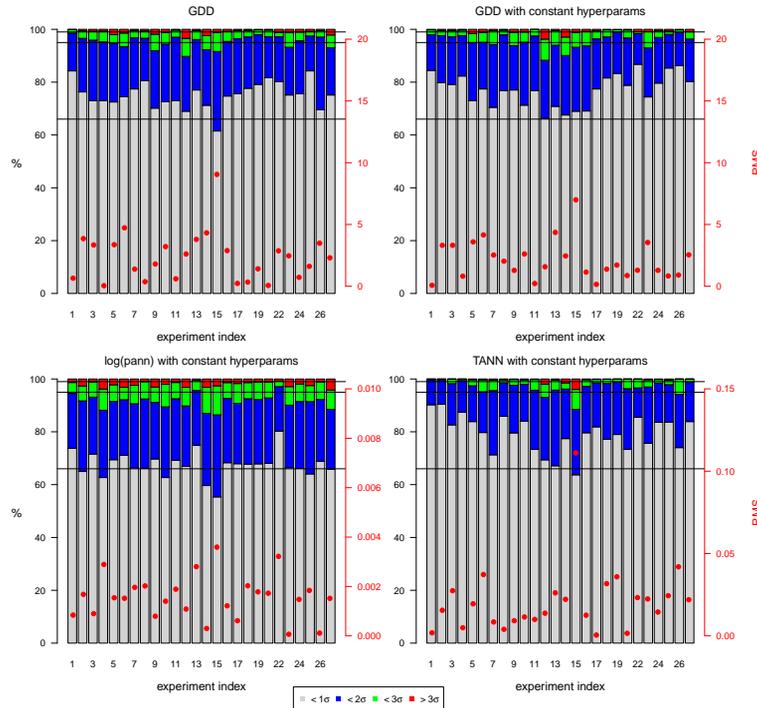
Figure 7. Evaluation of the PC emulators. The bars give the fractions of grid points for which the emulation correctly predicts the value of the experiment left out of the calibration, within 1, 2, 3 standard deviations, or more. The light and dark blue lines correspond to the fractions corresponding to 1, 2, and 3 standard deviations in the case of perfect Gaussian distributions. Dots are 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. Remark that experiment 20 is omitted (see the Supplement.)
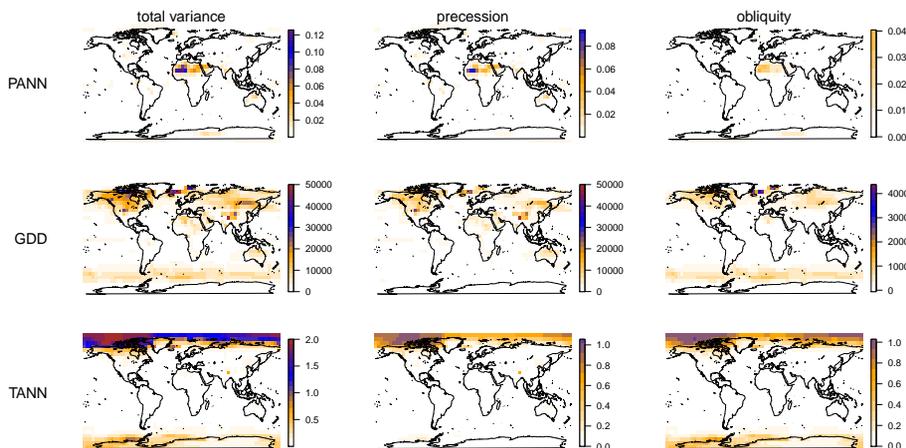
Figure 8. Geographical distribution of total variance $\overline{S}_{1,2,3}^{m}$, split into contributions from precession $\overline{S}_{1,2}^{m}$ and obliquity $\overline{S}_{3}^{m}$. Color schemes are consistent accross each variable and the right-hand side scale gives the range of values covered.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion
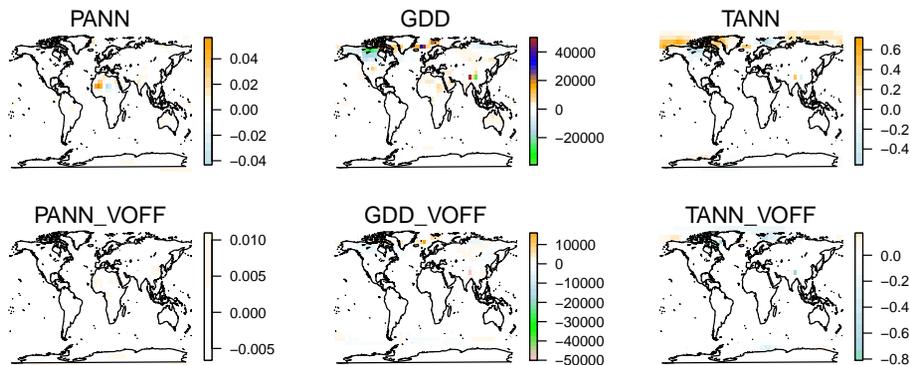
Figure 9. Geographical distribution of total variance associated with non-linear effects $\overline{S}_{1,2,3}^{m,\text{nl}}$, with the current surface scheme using interactive vegetation (the current LOVECLIM standard), and with the original ECBILT surface scheme, with fixed vegetation (deprecated). Color schemes are consistent with Fig. 8.