



# Global sensitivity analysis of the Indian monsoon during the Pleistocene

P. A. Araya-Melo, M. Crucifix, and N. Bounceur

Université catholique de Louvain, Earth and Life Institute, Georges Lemaître Centre for Earth and Climate Research, 1348 Louvain-la-Neuve, Belgium

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

Received: 4 March 2014 – Revised: 10 November 2014 – Accepted: 26 November 2014 – Published: 12 January 2015

**Abstract.** The sensitivity of the Indian monsoon to the full spectrum of climatic conditions experienced during the Pleistocene is estimated using the climate model HadCM3. The methodology follows a global sensitivity analysis based on the emulator approach of Oakley and O’Hagan (2004) implemented following a three-step strategy: (1) development of an experiment plan, designed to efficiently sample a five-dimensional input space spanning Pleistocene astronomical configurations (three parameters), CO<sub>2</sub> concentration and a Northern Hemisphere glaciation index; (2) development, calibration and validation of an emulator of HadCM3 in order to estimate the response of the Indian monsoon over the full input space spanned by the experiment design; and (3) estimation and interpreting of sensitivity diagnostics, including sensitivity measures, in order to synthesise the relative importance of input factors on monsoon dynamics, estimate the phase of the monsoon intensity response with respect to that of insolation, and detect potential non-linear phenomena.

By focusing on surface temperature, precipitation, mixed-layer depth and sea-surface temperature over the monsoon region during the summer season (June–July–August–September), we show that precession controls the response of four variables: continental temperature in phase with June to July insolation, high glaciation favouring a late-phase response, sea-surface temperature in phase with May insolation, continental precipitation in phase with July insolation, and mixed-layer depth in antiphase with the latter. CO<sub>2</sub> variations control temperature variance with an amplitude similar to that of precession. The effect of glaciation is dominated by the albedo forcing, and its effect on precipitation competes with that of precession. Obliquity is a secondary effect, negligible on most variables except sea-surface temperature.

It is also shown that orography forcing reduces the glacial cooling, and even has a positive effect on precipitation.

As regards the general methodology, it is shown that the emulator provides a powerful approach, not only to express model sensitivity but also to estimate internal variability and detect anomalous simulations.

## 1 Introduction

Since the pioneering studies of Kutzbach and Street-Perrott (1985), modelling efforts with general circulation models have routinely been used to understand, quantify and identify the causes of past changes in monsoon dynamics.

One general approach to this end has been to perform snapshot experiments for specific time slices in the past. The general circulation model is run with a particular set of initial conditions for a perpetual year for a long computational time until equilibrium is reached. The epoch used for defining the astronomical forcing and boundary conditions is one for which specific efforts are being undertaken to collect observations. This is the general spirit of projects such as COHMAP (Anderson et al., 1988) and PMIP (Braconnot et al., 2007). Specifically, the COHMAP project focused on a series of time slices spaced every 3000 years throughout the deglaciation (Kutzbach and Guetter, 1986; Anderson et al., 1988), while PMIP historically focused on the mid-Holocene and the Last Glacial Maximum, though on this basis an increasing number of periods are being considered, including the Eemian (Braconnot et al., 2008) and the last interglacials (Yin and Berger, 2012).

Based on these experiments, it is now well understood that glacial boundary conditions, typical, for example, of the Last

Glacial Maximum, induce a weakening of moisture transport over the Indian subcontinent and a reduction of precipitation in East Asia (see also Kutzbach and Guetter, 1986; Felzer et al., 1998; Yanase and Abe-Ouchi, 2007; Braconnot et al., 2007). On the other hand, an increase in northern summer insolation compared to a reference state strengthens monsoon dynamics, in agreement with general considerations on the dynamics of heat transport and on the location of the Intertropical Convergence Zone. These effects may be combined. For example, Masson et al. (2000) showed the possibility of intense Indian monsoon under glacial conditions, more specifically stage 6.5, when the astronomical configuration is favourable.

These past climate simulations are often complemented with additional sensitivity experiments. One classical experimental setup consists in considering two end-member states, often the pre-industrial and one well-defined past period, and intermediate configurations for which one or several forcing components are “activated” while the others are left as the pre-industrial configuration (e.g. Felzer et al., 1998; Masson et al., 2000; Yin et al., 2009). Such sensitivity studies will be referred here as “local” approaches, in the sense that only a small set of forcing conditions are explicitly considered out of the space of possible forcings.

Palaeoclimate modelers are also concerned with the phase relationship between forcing and climate. In particular, climatic precession may be seen as a quasi-periodic rotation of the point of smallest Earth–Sun distance (it will be referred to here as the perigee because we work in geocentric coordinates) and the vernal equinox. By considering specific periods in the past for GCM experiments one can only already develop a partial understanding of the phase relationships. Specifically, Braconnot and Marti (2003) showed that an “early-phase” configuration (perigee reached in April) produces a stronger monsoon, which occurs earlier in the year than a “late-phase” configuration (perigee reached in September). Alternatively, Kutzbach et al. (2008) (see also Chen et al., 2011) proposed the use of long transient simulations to study the evolutionary response to orbital forcing of global summer monsoon over the past 280 000 years. They showed that north tropical sea-surface temperature leads June insolation by about 40°. This particular work did not consider CO<sub>2</sub> and ice boundary condition effects. At the time of writing, such experimental setups can only be afforded with fairly low-resolution models (these authors used FOAM) with an acceleration technique: one model year actually represents 100 years of simulation time.

Here, we will experiment with an alternative approach that will enable us to simultaneously document the *sensitivity* of a general circulation model (HadCM3); the independent and combined effects of different forcing components on monsoon dynamics, namely astronomical forcing, CO<sub>2</sub> and ice boundary conditions; and, finally, estimate the phase relationship between monsoon response and insolation forcing.

The starting point of this approach consists in performing an ensemble of snapshot simulations. The ensemble is designed such that experiments span the space of possible forcing configurations that the Earth encountered during the late Pleistocene (ca. the last 800 000 years). For this reason the approach will be qualified as “global”; more specifically, this is a *global sensitivity analysis* because we do not explicitly consider a reference state. Thus, a statistical model is used to estimate the state of the system at any input point within the space spanned by the experiment ensemble. To this end, we consider a statistical model that is commonly referred to as an “emulator” in the statistical literature (O’Hagan, 2006). In particular, the term emulator refers to the following properties (O’Hagan, 2006; Petropoulos et al., 2009):

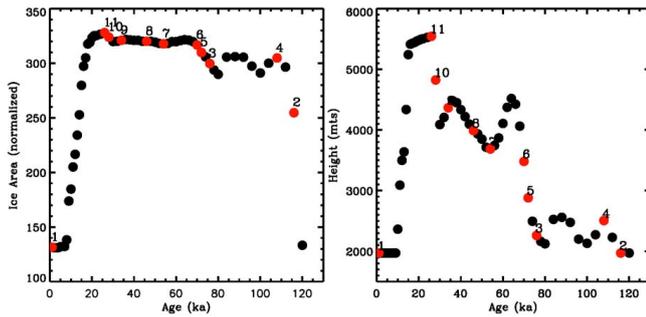
- it is derived from a small number of model runs filling the entire multidimensional input space;
- once the emulator is built, it is not necessary to perform any additional runs with the model.

The emulator is then used to generate visual diagnostics and numerical indices summarising the sensitivity of the model to the different elements of the forcing.

This technique of emulation is beginning to be commonly used to estimate uncertainties on climate model outputs, given probability distributions on uncertain quantities such as model parameters (Lee et al., 2011) or elements of the forcing (Carslaw et al., 2013). Such approaches may also integrate information from observations following a Bayesian formalism in order to construct posterior distributions of model parameters and update current knowledge on predictive quantities such as climate sensitivity (Holden et al., 2010; Schmittner et al., 2011). The inference model may in particular include a statistical quantity called model discrepancy, used to express the distance between the model and the real world (Sexton et al., 2012).

Compared to this series of works the present objective is a bit different. As stated, we are interested in input quantities which we know varied in the past, though we will assume that they varied sufficiently slowly to justify a hypothesis of quasi-stationarity of the ocean–atmosphere system with the forcing. Our purpose is to estimate the contribution of input factors to the temporal climate variance that can be observed in palaeoclimate records. To this end we refer to the statistical theory of global sensitivity analysis with emulation formalised by Oakley and O’Hagan (2002) based on general principles of global sensitivity analysis (Homma and Saltelli, 1996) and experiment design (Sacks et al., 1989), but adapted to our particular objective.

The paper is structured as follows. Section 2 provides a description of the emulator and the simulations used. The section is admittedly technical and contains material that has been published before in the statistical literature. However, following the practice of recent articles of climate literature (e.g. Lee et al., 2013), we choose to walk the reader through



**Figure 1.** Left panel: ice area, in normalized units, and maximum height (in meters) in the region 45–75° N and 240–275° W (Laurentide Ice Sheet), as a function of time in the boundary conditions used in the Singarayer and Valdes (2010) experiment. Right panel: Height (in meters) of the ice sheets. This shows that, although the volume of the ice masses is quite different, their area is not. Red circles indicate the boundary conditions used for this specific study.

the details of emulation design (see also video in the supplementary material). This also gives us the opportunity to document in detail technical statistical modelling choices. The hasty reader may, however, jump to the Sect. 3, where the results of applying the emulator on the Indian monsoon region are discussed. We focus, on the one hand, on the performance of the emulator as such and, on the other hand, on the climatic lessons emerging from this experiment. In particular, the specific influence of ice sheet topographic forcing is quantified. Conclusions follow in Sect. 4.

## 2 Methodology

### 2.1 Experiment design

The first task is to define the space of input configurations to be explored with an ensemble of experiments. We consider five input factors: the three elements of astronomical forcing (eccentricity  $e$ , longitude of perigee  $\varpi$ , where  $\varpi = 0$  when perigee is in March, and obliquity  $\varepsilon$ ), the concentration in carbon dioxide ( $\text{CO}_2$ ), and a variable called the ice or glaciation level, which combines ice and orography forcings associated with the presence of continental ice in the Northern Hemisphere.

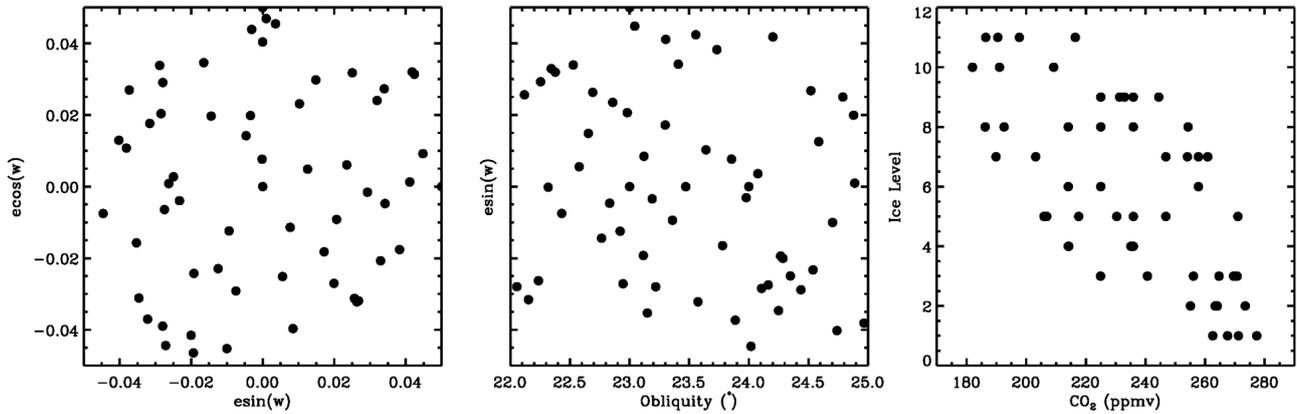
The three elements of astronomical forcing are combined under the form of  $e \sin \varpi$ ,  $e \cos \varpi$  and obliquity  $\varepsilon$ . This choice is justified by the fact that these combinations produce orthogonal patterns in the season–latitude space, and generally insolation at any point and time in year is well approximated as a linear combination of those terms (Loutre, 1993). The factors  $e \sin \varpi$  and  $e \cos \varpi$  are sampled in the range  $[-0.05, 0.05]$ , while  $\varepsilon$  is varied in the range 22–25°. Atmospheric  $\text{CO}_2$  concentration is sampled in the range 180–280 ppm.

The glaciation level is determined as follows. Our purpose is to select 11 realistic boundary conditions representative of glacial–interglacial dynamics. Pragmatically, we sampled these boundary conditions among the series prepared by Singarayer and Valdes (2010), and kindly supplied to us by Prof. Paul Valdes, University of Bristol. Level 1 corresponds to present-day conditions, and levels 2 to 11 are chosen as such to represent approximately 10 equally spaced top altitudes of the North American Ice Sheet, within the glaciation phase. One limitation of this design for the present purpose is that levels 3 to 11 effectively represent similar ice sheet areas – thus similar albedo forcing – even though they sample very different ice sheet volume (see Fig. 1).

The next step is to define an ensemble of experiments to run with the climate model in order to efficiently span the input space. The choice of the number of experiments and, for each experiment, the choice of input parameters is called the design. A *design point* refers in this context to a specific experiment. The construction of the design should conform to rules of good practice explained, for example, in Santner et al. (2003). In particular, we want the design to be *space filling*, and theoretical considerations and experience point to the Latin hypercube design (McKay et al., 1979; Morris and Mitchell, 1995; Sacks et al., 1989; Urban and Fricker, 2010) as a good starting point. The principle for a Latin hypercube design of  $n$  elements is to divide the ranges covered by each input factor into  $n$  distinct categories, each experiment sampling one of the  $n$  categories without replacement. However, many Latin hypercubes could be constructed in this way, and the design most appropriate for emulation should satisfy additional constraints. Following Santner et al. (2003, p. 167) and Joseph and Hung (2008) we combine two criteria. First, we select, among the possible Latin hypercube designs, those maximising the minimum Euclidean distance found between any two design points. This is called the maxi–min criteria. Among those designs, we chose those maximising the determinant of  $X'X$ , so that the resulting design is also near-orthogonal.

For this application, two additional constraints need to be accounted for in order to avoid sampling unrealistic inputs that would be uninformative for the sensitivity analysis of climate over the Pleistocene: exclude forcings with  $e > 0.05$  and exclude combinations of high  $\text{CO}_2$  and high glaciation levels (and conversely), delineated by an ellipse with large and small axes as shown in Fig. 2. To satisfy these constraints, the design points generated by the Latin hypercube sampling procedure lying in the exclusion zone are geometrically projected on the allowed region. This procedure may break some of the original properties of the design (maxi–min and orthogonality), but it offers the practical advantage of enhancing the coverage of the input space near its boundary.

Note that this design is in principle suitable for continuous factor ranges only. The glaciation level used for experiments is an integer obtained by rounding the value obtained by this



**Figure 2.** Experiment plan design, optimised to maximise the minimum distance between points and to achieve orthogonality (maximise the determinant of the covariance of input factors). Right:  $e \cos w - e \sin w$  space distribution; middle:  $e \sin w - \text{obliquity}$  space distribution; right: glacial level– $\text{CO}_2$  space distribution.

**Table 1.** Experiment setup: simulation name and number, astronomical parameters (eccentricity, longitude of the perigee and obliquity),  $\text{CO}_2$  concentration and glacial level.

No.	Name	$e$ –	$\varpi$ (°)	$\varepsilon$ (°)	$\text{CO}_2$ (ppm)	Ice level –	No.	Name	$e$ –	$\varpi$ (°)	$\varepsilon$ (°)	$\text{CO}_2$ (ppm)	Ice level –
1	xadba	0.0527	53.52	23.6	277.3	1	32	xadfa	0.0383	334.53	23.8	257.8	6
2	xadbb	0.0520	211.44	22.9	267.5	1	33	xadfb	0.0417	139.99	24.5	214.1	6
3	xadbc	0.0309	218.44	23.1	262.6	1	34	xadfc	0.0480	215.67	23.2	225.0	6
4	xadbd	0.0201	350.24	23.2	271.2	1	35	xadfd	0.0404	140.60	22.1	225.0	6
5	xadka	0.0282	256.84	24.2	264.1	2	36	xadga	0.0301	194.43	22.4	254.1	7
6	xadkb	0.0466	228.06	24.2	263.4	2	37	xadgb	0.0261	208.55	22.9	189.8	7
7	xadkc	0.0411	88.21	23.3	273.5	2	38	xadgc	0.0503	202.65	24.3	260.8	7
8	xadkd	0.0077	358.66	22.3	255.1	2	39	xadgd	0.0389	122.16	22.3	257.8	7
9	xadaa	0.0403	316.14	22.1	270.6	3	40	xadge	0.0345	97.90	23.4	246.8	7
10	xadab	0.0263	271.85	22.2	270.7	3	41	xadgf	0.0362	299.18	22.2	246.8	7
11	xadac	0.0416	140.71	22.7	269.6	3	42	xadgg	0.0440	355.96	24.0	260.9	7
12	xadad	0.0257	167.54	22.6	256.1	3	43	xadgh	0.0422	287.83	24.7	203.2	7
13	xadae	0.0406	167.95	23.1	240.7	3	44	xadha	0.0436	51.20	22.5	192.6	8
14	xadaf	0.0460	305.89	23.9	224.9	3	45	xadhb	0.0333	26.49	22.7	254.3	8
15	xadag	0.0293	93.07	22.3	264.7	3	46	xadhc	0.0461	205.77	24.3	186.2	8
16	xadda	0.0244	323.78	22.8	214.1	4	47	xadhd	0.0386	246.02	23.1	214.1	8
17	xaddb	0.0421	114.71	23.7	214.2	4	48	xadhe	0.0405	38.22	24.8	225.0	8
18	xaddc	0.0253	23.96	23.6	235.9	4	49	xadhf	0.0491	221.00	23.6	235.9	8
19	xaddd	0.0469	1.20	24.9	235.1	4	50	xadia	0.0150	341.91	22.8	244.4	9
20	xadei	0.0000	0.00	23.0	230.4	5	51	xadib	0.0457	78.40	23.0	235.9	9
21	xadej	0.0500	90.00	23.0	230.4	5	52	xadic	0.0226	113.92	23.0	225.0	9
22	xadek	0.0500	0.00	23.0	230.4	5	53	xadid	0.0400	53.05	22.4	232.9	9
23	xadel	0.0000	0.00	24.0	230.4	5	54	xadie	0.0336	143.57	24.9	231.3	9
24	xadea	0.0155	217.23	23.4	205.9	5	55	xadja	0.0452	260.43	24.0	182.0	10
25	xadeb	0.0527	52.54	24.2	235.9	5	56	xadjb	0.0444	319.59	24.4	209.2	10
26	xadec	0.0456	4.52	24.1	206.6	5	57	xadjc	0.0463	192.48	24.7	191.0	10
27	xaded	0.0135	68.81	24.6	246.8	5	58	xadca	0.0350	305.63	24.1	190.5	11
28	xadee	0.0236	260.39	24.5	217.6	5	59	xadcb	0.0137	145.99	23.9	216.4	11
29	xadef	0.0396	285.78	25.0	246.8	5	60	xadcc	0.0250	136.64	23.3	186.4	11
30	xadeg	0.0251	276.28	24.3	271.0	5	61	xadcd	0.0243	75.55	22.9	197.7	11
31	xadeh	0.0404	359.97	23.5	206.9	5							

process to the closest integer. Designs specifically adapted for input spaces mixing categorical and continuous variables could best be implemented in the future (see, for example, MacCalman, 2013, for an up-to-date review).

Table 1 lists the simulations with their input parameters. The choice of 61 members is a conservative implementation of the recommendation of 10 experiments per input factors (Loeppky et al., 2009). In fact, a first 57-member design was produced using the method above, to which 4 members were added (experiments 20–23). These experiments are idealised orbital changes that were performed during the first phase of this project in order to locally explore the model sensitivity to astronomical forcing.

## 2.2 Climate simulator

The climate model – referred to in this context as the simulator – is the general circulation model HadCM3 (Gordon et al., 2000), using the MOSES2 dynamic land surface scheme (Essery et al., 2003). The atmospheric component dynamics and physics are resolved on a  $3.75^\circ \times 2.5^\circ$  longitude–latitude grid. The oceanic component has a horizontal resolution of  $1.25^\circ \times 1.25^\circ$ .

Initial conditions are the final state of the PMIP2 0K experiment featured in Braconnot et al. (2007). Each simulation is run for 400 years, except for the *xadk#* set. Accidentally, the first 200 years did not account for ice sheet topography. This was corrected for the following 200 years. In the case of the *xadk#* simulations, they were run for 300 years, accounting for ice sheet topography from the beginning. Typical residual deep-ocean temperature trends are of the order of  $10^{-4} \text{ }^\circ\text{C year}^{-1}$ .

The last 100 years of all simulations with orographic forcing were retained for analysis. Over this interval, the top-of-the-atmosphere imbalance ranges between  $-0.2$  and  $-0.1 \text{ W m}^{-2}$ . The last 100 years of the experiment section without orographic forcing are also used for an investigation of the specific effect of the orographic forcing (cf. Sect. 3.6).

## 2.3 Emulator

At this stage we suppose that the simulator HadCM3 has been run for all design points. We now show that it is possible to *estimate*, with quantified uncertainty, the output that one would have obtained by running HadCM3 at any input lying within the parameter space spanned by the design.

To this end, we need to develop a statistical model that can interpolate the outputs obtained with the simulator at the design points. The procedure is akin to geospatial interpolation, except that the input field is here five-dimensional, instead of two- or three-dimensional as in most geospatial applications (cf. video in the supplementary material).

In particular, we follow Oakley and O’Hagan (2002) and use a Gaussian process model, with a Bayesian formalism.

Although there is no strict practice, the term *emulator* is often reserved to such Bayesian meta-models.

The calibration of the emulator is mathematically described as follows. Let  $\mathbf{x}_j$  be the set of input values of the  $j$ th member of the design (here: a vector of which the components are the astronomical forcing, ice level and CO2). The output of the climate model is modelled as a stochastic process combining a global response function (the regressors) with a local component. It is fully specified by the mean  $\tilde{m}$  and a covariance  $\tilde{V}$  function, which have the following priors:

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

$$\tilde{V}(\mathbf{x}, \mathbf{x}^*) = \sigma^2 c(\mathbf{x}, \mathbf{x}^*) \quad (2)$$

where  $c(\mathbf{x}, \mathbf{x}^*)$  is the Gaussian process correlation function, and  $\sigma^2$  its variance;  $\mathbf{h}(\mathbf{x})$  is a  $(q \times 1)$  vector of a priori known regression functions; and  $\boldsymbol{\beta}$  is the vector of corresponding regression coefficients. Note that the  $()'$  is used to denote a horizontal vector. The definition of the correlation function is given below.

Let  $f(\mathbf{x})$  denote the climate model output when run at input vector  $\mathbf{x}$ . In Bayesian language, we say that the fact of actually running the model at the design  $n$  points allows us to *update* our knowledge of  $f(\mathbf{x})$  at any input point.

We also need to make a choice regarding the values of  $\boldsymbol{\beta}$  and  $\sigma^2$ . Given that we do not know their true value, we proceed, in the Bayesian way, by defining prior probabilities for these quantities. We would like not to introduce specific information on  $\boldsymbol{\beta}$  and  $\sigma^2$ . Given that  $\sigma^2$  is a scale factor, theoretical considerations show that the prior  $(\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-2}$  is appropriate as a vague prior, i.e., all values of  $\boldsymbol{\beta}$  are a priori equally plausible and the probability density of  $\sigma^2$  decays in a way that preserves independence on unit choices (Berger et al., 2001).

In these conditions, the *posterior* estimate of  $f(\mathbf{x})$  is a Student  $t$  distribution with  $n - q$  degrees of freedom, with the following mean and variance (Oakley and O’Hagan, 2002; Bastos and O’Hagan, 2009):

$$m(\mathbf{x}) = \mathbf{h}(\mathbf{x})' \hat{\boldsymbol{\beta}} + \mathbf{T}(\mathbf{x}^*)' \mathbf{A}^{-1} (\mathbf{y} - \mathbf{H} \hat{\boldsymbol{\beta}}), \quad (3)$$

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

respectively, with

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

$$\hat{\boldsymbol{\beta}} = (\mathbf{H}' \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}' \mathbf{A}^{-1} \mathbf{y},$$

where  $\mathbf{y}$  is a matrix of  $n$  lines, of which each line gathers the input of the respective experiments;  $\mathbf{T}(\mathbf{x})_j = c(\mathbf{x}, \mathbf{x}_j)$ ; and  $\mathbf{P}(\mathbf{x}) = \mathbf{h}(\mathbf{x})' - \mathbf{T}(\mathbf{x}) \mathbf{A}^{-1} \mathbf{H}$ . In the following, we conveniently approximate the Student  $t$  distribution by a normal distribution. Although in principle is true only as  $n \rightarrow \infty$ , is accurate enough in practice for values of  $n - q$  larger than 20.

Remember that  $\mathbf{x}_j$  are the input parameters (astronomical configuration, etc.) of experiment  $j$  of the design. Hence, for example,  $T(\mathbf{x})_j$  is a scalar, obtained by applying the so-called correlation function defined below between the input vector  $\mathbf{x}$  – at which one wants to predict the simulator output – and the input  $\mathbf{x}_j$  of the design. Consequently, the quantity  $T(\mathbf{x})$  is treated as an  $n$ -component vector, of which the respective components are associated with the different elements of the design. With this framework, the choices of the regression functions  $\mathbf{h}(\mathbf{x})$  and the Gaussian process correlation function  $c(\mathbf{x}, \mathbf{x}^*)$  are application-dependent. This is where the user has the opportunity to inject knowledge on the expected response of the simulator.

For this application, linear regression is an adequate choice because the seasonal and annual forcings are almost linear with the input factors, except possibly for glaciation level. Hence,  $\mathbf{h}(\mathbf{x}') = (1, \mathbf{x}')$ .

The correlation function  $c(\mathbf{x}, \mathbf{x}^*)$  is a linear measure of how informative the simulator output at  $\mathbf{x}$  is about the simulator output at  $\mathbf{x}^*$ . It is thus a key component of the emulator. We use here the classical exponential decay (Oakley and O'Hagan, 2002):

$$c(\mathbf{x}, \mathbf{x}^*) = \exp[-(\mathbf{x}'\Lambda^{-2}\mathbf{x}^*)]. \quad (6)$$

The scaling matrix  $\Lambda$  is diagonal, with components  $\lambda_i$  called the length scales. The interpretation is thus that the correlation between the outputs of two experiments decreases exponentially as the normalised distance between two input factors decreases. The normalisation factors are the length scales. Intuitively, the length scale may thus be interpreted as a measure of the roughness of the surface response: the larger the length scale, the smoother the response surface (see video animation in the Supplement).

There is a further correction to be accounted for before using this function. The quantity we are interested in emulating is the hypothetic mean of an infinitely long experiment that has perfectly reached the stationary state. In practice, we have to be content with the mean of a finite-length experiment, obtained for a specific set of initial conditions and which may not have perfectly reached the stationary state. The difference between the output of an experiment and the ideal experiment average is expected to be small yet impossible to predict exactly because it may chaotically depend on initial conditions. It may effectively be accounted for in the emulator as follows. Observe that the function  $c(\mathbf{x}, \mathbf{x}^*)$  always appears as filling the elements of a matrix (Eqs. 2 and 4). This matrix is further modified by adding a small element along the diagonal called the nugget  $\nu$ , which will absorb the effects mentioned about the experiment sample being only an estimate of the stationary state. The error tolerance will be of the order of  $\hat{\sigma}^2\nu$ .

The nugget has another benefit: it regularises the problem for large length scales, and it may in particular be shown that posterior means converge to the solution of a linear re-

gression problem for  $\lambda_i \rightarrow \infty$  (Andrianakis and Challenor, 2012).

The remaining problem is to estimate the hyperparameters  $\lambda_i$  and  $\nu$  completely. Following Kennedy and O'Hagan (2000), we maximise the emulator likelihood (the expression used here is from Andrianakis and Challenor, 2012):

$$\log L(\nu, \Lambda) = -\frac{1}{2} \left( \log \left( |A| |H^T A^{-1} H| \right) + (n - q) \log(\hat{\sigma}^2) \right).$$

In order to guarantee that the emulator is at least no less informative than would be linear regression, Andrianakis and Challenor (2012) recommend the use of a *penalised likelihood* as follows:

$$\log L^P(\nu, \Lambda) = \log L(\nu, \Lambda) - 2 \frac{\overline{M}(\nu, \Lambda)}{\epsilon \overline{M}(\infty)}, \quad (7)$$

where  $\overline{M}(\nu, \Lambda)$  is the mean squared error between the training points and the emulator's posterior mean at the design points, and  $\overline{M}(\infty)$  is its asymptotic value at  $\lambda_i \rightarrow \infty$ . We use  $\epsilon = 1$ .

It is worth noting that, in our case, using the normal likelihood or the penalised one has practically no effect on the results.

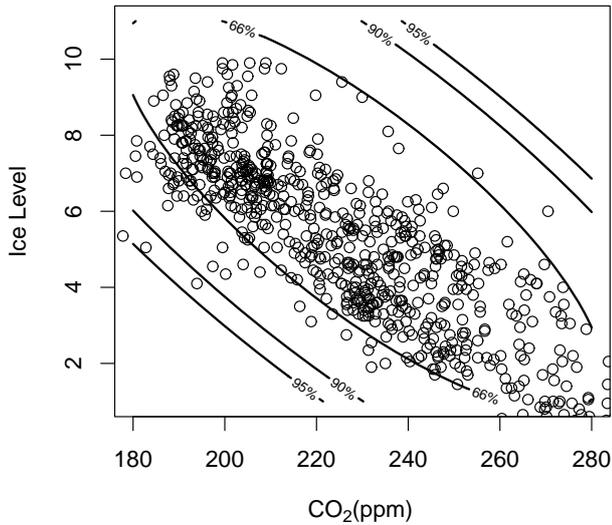
## 2.4 Sensitivity measures

We are now in a position to estimate the simulator output at potentially any input point spanned by the design. It is now possible to develop indices, of which the purpose is to summarise the sensitivity of the simulator to individual or combined factor throughout the whole input space. This is the general idea of global sensitivity analysis.

In particular, one of the early applications of Bayesian emulators (as we use here) was to estimate sensitivity measures to quantify the uncertainty on a simulator output arising from the fact that the inputs are themselves uncertain (Oakley and O'Hagan, 2004). In this context, the uncertain inputs may be quantified by means of a multivariate probability density function  $\rho(\mathbf{x})$ . The problem of interest here is slightly different because we know how the inputs varied in the past. The theory of global sensitivity analysis may, however, be recycled by giving  $\rho(\mathbf{x})$  a frequentist interpretation. In other words, we use  $\rho(\mathbf{x})$  to describe the time-wise occupation density of the input space estimated by considering the history of the late Pleistocene.

In particular, the occupation density along the components of the astronomical forcing can be estimated with histograms of long time series generated with known astronomical solutions, such as those presented by Berger (1978). We then consider the following empirical distribution to broadly capture the observed covariance between CO<sub>2</sub> and glaciation level (see Fig. 3):

$$\rho(c^*, i^*) \propto \begin{cases} \mathcal{N}\left(0.5, \frac{3}{80} \begin{pmatrix} 1 & \frac{1}{3} \\ -1 & \frac{1}{3} \end{pmatrix}^2\right) & \text{where } 0 < c^* < 1, \quad 0 < i^* < 1 \\ 0 & \text{elsewhere,} \end{cases} \quad (8)$$



**Figure 3.** Lines: 66, 90 and 95 % percentiles of the empirical distribution used to describe the probability distribution in the CO<sub>2</sub>–ice space (Eq. 8). Dots: observations of CO<sub>2</sub> (Luethi et al., 2008; Siegenthaler et al., 2005; Petit et al., 1999) and estimates of ice level assuming a linear relationship with the LR04 stack of benthic foraminifera δ<sup>18</sup>O (Lisiecki and Raymo, 2005) over the last 800 000 years. Based on these observations, the empirical distribution appears to be slightly biased towards high ice level at low CO<sub>2</sub>.

where  $c^*$  and  $i^*$  are inputs standardised as follows:

$$c^* = (\text{CO}_2 - 180 \text{ ppm}) / (100 \text{ ppm}), \quad (9)$$

$$i^* = (\text{glaciation level} - 1) / 10. \quad (10)$$

In order to relate output variances with input variances, we first define what is known in the global sensitivity literature as the *main effect* associated with an input  $p$  (e.g. Saltelli et al., 2004, Chapter 1):

$$\eta(\mathbf{x}_p) = \int_{\mathcal{X}_{\bar{p}}} f(\mathbf{x}) \rho(\mathbf{x}_{\bar{p}} | \mathbf{x}_p) d\mathbf{x}_{\bar{p}}, \quad (11)$$

where we have denoted  $\mathcal{X}_{\bar{p}}$  as the space spanned by all the components of  $\mathbf{x}$  but  $p$ , and  $\rho(\mathbf{x}_{\bar{p}} | \mathbf{x}_p)$  is the density of occupation of the space  $\mathcal{X}_{\bar{p}}$  given the vector  $\mathbf{p}$ . The main effect is thus the expected mean of the simulator output, given a known value of  $\mathbf{x}_p$  but no more information than the prior on the other components of  $\mathbf{x}$ .

Given that we cannot run the model at every point of the space  $\mathcal{X}_p$ , this quantity is uncertain, but its mean and variance may be estimated with the emulator:

$$m_p(\mathbf{x}_p) = \mathbb{E}_f(\eta(\mathbf{x}_p)) = \int_{\mathcal{X}_{\bar{p}}} \rho(\mathbf{x}_{\bar{p}} | \mathbf{x}_p) d\mathbf{x}_{\bar{p}}, \quad (12)$$

$$V_{pp}(\mathbf{x}_p, \mathbf{x}_p^*) = \text{Var}_f(\eta(\mathbf{x}_p)) = \iint_{\mathcal{X}_{\bar{p}} \times \mathcal{X}_{\bar{p}}} V(\mathbf{x}, \mathbf{x}^*) \rho(\mathbf{x}_{\bar{p}} | \mathbf{x}_p) \rho(\mathbf{x}_{\bar{p}}^* | \mathbf{x}_p) d\mathbf{x}_{\bar{p}} d\mathbf{x}_{\bar{p}}^*, \quad (13)$$

where  $\mathbb{E}_f$  and  $\text{Var}_f$  denote mean and variance due to using the emulator instead of actually running the simulator at all points. On this basis, it is possible to define two measures of sensitivity of the outputs to input  $\mathbf{x}_p$ :

$$S_p = \mathbb{E}_f \text{Var}(\eta(\mathbf{x}_p)) \quad \text{and} \quad (14)$$

$$\bar{S}_p = \mathbb{E}_f [\text{Var}(\eta(\mathbf{x})) - \text{Var}(\eta(\mathbf{x}_{\bar{p}}))]. \quad (15)$$

The quantity  $S_p$ , called the main effect index<sup>1</sup> is the loss in output variance that would occur assuming that  $\mathbf{x}_p$  is known and constant, compared to a situation where all factors vary. More precisely, this is the *expected* loss, averaged over all possible values of  $\mathbf{x}_p$  (e.g. Saltelli et al., 2004, Chapter 1). On the other hand,  $\bar{S}_p$  is the output variance that occurs when factor  $p$  is variable; all other factors assumed to be known and constant. This is the total effect index<sup>2</sup>. The distinction between main and total effect is particularly important when there is a covariance between input factors. This is the case here: CO<sub>2</sub> and ice volume co-vary. More precisely, the main effect index associated with, for example, ice volume, includes an implicit contribution associated with the fact that CO<sub>2</sub> co-varies with ice level. The total effect index does not include this contribution. Therefore, we use the total effect index.

In order to compute  $S_p$  and  $\bar{S}_p$ , we define the auxiliary quantities:

$$\Sigma_p = \int_{\mathcal{X}_p} [m_p(\mathbf{x}_p)^2 + V_{pp}(\mathbf{x}_p, \mathbf{x}_p)] d\rho(\mathbf{x}_p), \quad (16)$$

$$\Sigma_0 = [m_0(\mathbf{x})^2 + V_{00}(\mathbf{x}, \mathbf{x})], \quad (17)$$

$$\Sigma = \int_{\mathcal{X}} [m(\mathbf{x})^2 + V(\mathbf{x}, \mathbf{x})] d\rho(\mathbf{x}), \quad (18)$$

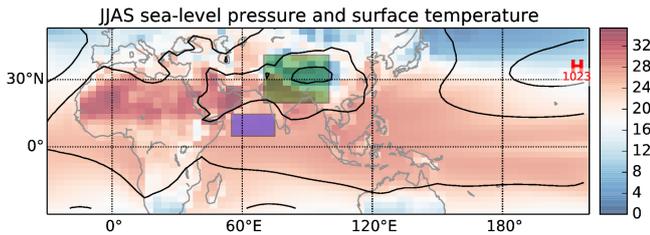
where the subscripts 0 and 00 imply that the space  $\mathcal{X}_{\bar{p}}$  referred to in the integrals (12) and (13) is the full input space. It may then be shown that (Oakley and O’Hagan, 2004)

$$S_p = \Sigma_p - \Sigma_0, \quad (19)$$

$$\bar{S}_p = \Sigma - \Sigma_{\bar{p}}. \quad (20)$$

<sup>1</sup>Strictly speaking, the word *index* applies when this quantity is divided by the total output variance.

<sup>2</sup>As above.



**Figure 4.** JJAS sea-level pressure and surface temperature of the two regions depicted: NI and IO. Units are in °C.

### 3 Results

In order to study the Indian monsoon, we define two regions: northern India (NI), with coordinates 70–100° E, 20–40° N, and the northwestern Indian Ocean (IO), with coordinates 55–75° E, 5–15° N (see Zhao et al., 2005). The chosen regions are depicted in Fig. 4, in which the sea-level pressure and surface temperature of one of the simulations are shown. The NI region covers the Indian subcontinent and part of the Tibetan Plateau (which is dry today), while IO covers the northwestern part of the Indian Ocean. In the supplementary material we explore another continental region which does not include the Tibetan Plateau (Chen et al., 2011).

We focus specifically on four physical variables representative of the summer Indian monsoon process: June–July–August–September (JJAS) temperature and precipitation on the continental box, and JJAS sea-surface temperature (SST) and mixed-layer depth on the Indian Ocean box. Over the experiment design, continental temperature varies between 15 and 21 °C. Precipitation varies between 72 and 230 mm month<sup>-1</sup>, SST between 25 and 31 °C, and mixed-layer depth between 29 and 59 m. For emulation, the logarithms of precipitation and mixed-layer depth are used, because these distributions are more Gaussian than those of the absolute values.

#### 3.1 Emulation validation

An emulator using all 61 experiments is calibrated using the procedure given in Sect. 2.3, with scales  $\lambda_i$  (with  $i = 1, \dots, 5$ ) and nugget determined by maximisation of the penalised likelihood. The performance of the emulator is then assessed following a leave-one-out cross-validation approach, that is, we construct 60 emulators to predict the experiment being left out. Figure 5 shows the result of this leave-one-out cross-validation procedure for SST and mixed-layer depth only, the other variables being discussed later.

This leads us to the following observations:

1. For  $e \sin \varpi$ ,  $e \cos \varpi$  and ice volume, the length scales  $\lambda$  are of the same order of magnitude as the range covered by the input factors. This is the ideal scenario: the space between two experiments is consistent with the decorrelation length of the simulator.

2. There are some instances where length scales are much greater than the scale of the variables: this is observed on all output variables for the response to CO<sub>2</sub> and, to a lesser extent, for obliquity. A large covariance scale implies that response is linear with respect to the factor, which is indeed a realistic outcome for CO<sub>2</sub>, in the range considered. This is not a problem on its own. It simply informs the user that a sparser sampling of this factor would have worked as well.

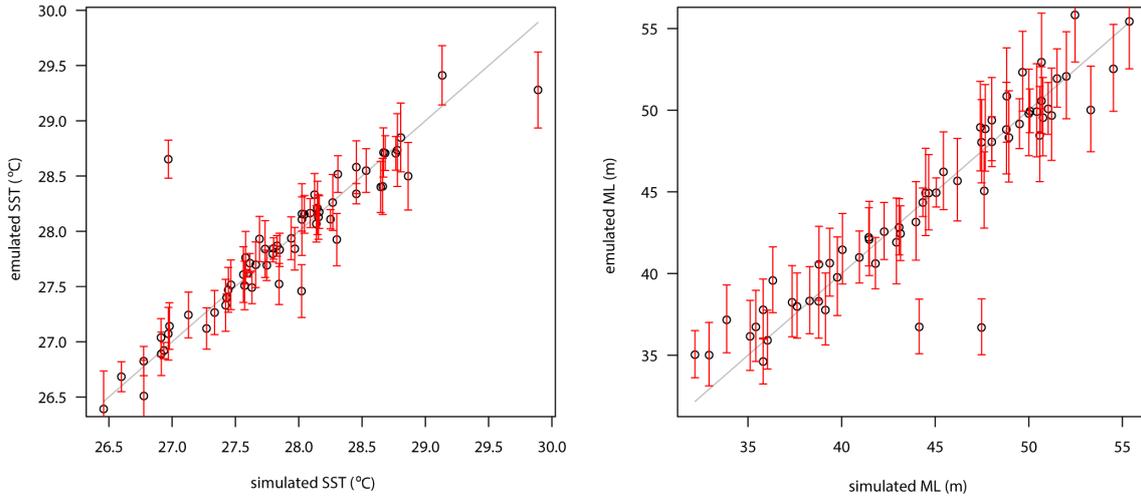
3. The leave-one-out cross-validation plot shows that two experiments are not well captured by the Gaussian process model for SST (experiments 11 and 40), and one for mixed-layer depth (experiment 40). The emulator fails to predict the outputs within an error of less than 3 standard deviations when they are left out of the calibration procedure. The effects of these experiments on the emulator output are well visible in Fig. 6 (top panels). These plots, which will be commented on in more detail in Sect. 3.5, represent the mean model response (Eq. 11) as a function of glaciation level and  $e \sin \varpi$ , and assuming CO<sub>2</sub> fixed. The figure reveals departure from smooth gradients contours, most notably the 26.25 and 26.5 °C isotherms on the SST plot and the 38.5 m iso-depth that conflict with our expectation of a smooth response structure.

At this stage one could consider an alternative emulator, calibrated on a 59-member experiment design in which the two problematic simulations are omitted.

This new emulator with new scales  $\lambda_i$  and nugget (see Table 2) presents a much more satisfactory performance (Fig. 8):

1. All ancillary emulators constructed for the leave-one-out diagnostic capture between 38 (mixed-layer depth) and 43 (continental temperature) of the leave-one-out experiments within 1 standard deviation, and between 56 and 58 within 2 standard deviations, which roughly correspond to the 66 and 95 % ratios expected for a normal distribution.
2. The normalised errors are compatible with a normal distribution based on the Shapiro–Wilk normality test, except for continental temperature (normality rejected with 97 % confidence).
3. There is no error exceeding 3 standard deviations.
4. Finally, the suspicious anomalies generated on the glaciation/precession plots are cleared (Fig. 6, bottom panels).

Based on our experience with HadCM3 we are inclined to give more credit to this new emulator as a predictor of HadCM3 outputs, rather than the one obtained with simulations 11 and 40. Of course, this choice leaves us with the task



**Figure 5.** Diagnostic of emulator performance considering experiments 11 and 40. Shown are the mean and standard deviations of sea-surface temperature (left panel) and mixed-layer depth (right panel). Clearly seen are the two bad predictions, especially in the case of sea-surface temperature.

**Table 2.** Emulator scales for the different fields under study. In general, scales are commensurate with the range covered by the input factors. However, for CO<sub>2</sub> and sometimes obliquity, the scales are much larger than the fields’ scale. This simply indicates that the response is linear with respect to the factor.

	Length scales					Nugget
	$\lambda_{e \cos \varpi}$ –	$\lambda_{e \sin \varpi}$ –	$\lambda_{\varepsilon}$ (°)	$\lambda_{\text{CO}_2}$ (ppm)	$\lambda_{\text{ice}}$ –	
Land temperature	0.0704	0.0914	3.191	940	3.348	0.0047
Land precipitation	0.1153	0.3037	20.221	12 588	2.2807	0.0188
Sea surface temperature	0.1118	0.1142	600.	9786	7.307	0.0035
Mixed-layer depth	0.0767	0.0308	3.7724	411	10.6960	0.0439

of explaining what went wrong with these two simulations. It seems that we have to leave it as an open case. Further inspection of these particular experiments reveals a clear warm–cold–warm pattern in the North Atlantic, and cooling over the rest of the ocean, exemplified here by comparing experiments 11 and 15 (Fig. 7). This pattern has been seen before in HadCM3, most notably in early experiments of the Last Glacial Maximum (Hewitt, 2003). It was associated with an enhancement of the North Atlantic Overturning Circulation cell, and can be annealed by addition of freshwater in the North Atlantic (Hewitt et al., 2006). Experiments 11 and 40 have, however, low to moderate glaciation levels, and reasons why their behaviour should differ from the other experiments are far from clear. Based on further inspection of time series as well as that of longer experiments, we are left with the speculation that the particular 100 years used to construct climatic averages correspond to some meta-stable state of the ocean circulation, possibly excited by the spin-up procedure.

Although we appreciate the difficulty, from a statistical inference prospective, of rejecting problematic experiments for the calibration of the emulator, we find it in fact positive that

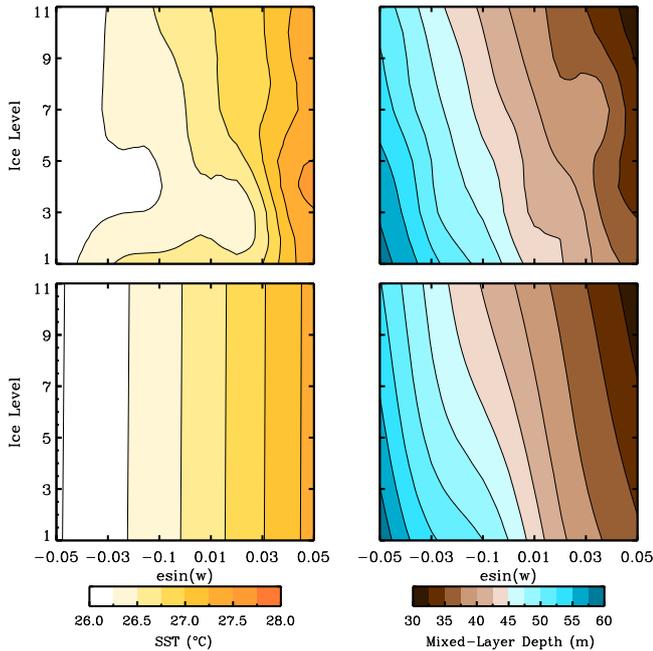
the emulator is effective in identifying experiments that behave unexpectedly compared to the bulk of the design.

Let us now consider the nugget.

As explained, this quantity quantifies the uncertainty of the simulation, i.e. how representative of the mean model state are the 100-year simulations.

The residual error in the emulator is of the order of  $\hat{\sigma}^2 v$ , but it can be estimated precisely by looking at the posterior variance at design points. Here, the obtained nuggets induce residual errors with standard deviations of 0.04 °C on continental temperature, 2.3 % on precipitation, 0.05 °C on SST, and 0.7 % on mixed-layer depth. All these values are consistent with the 100-year variances of the corresponding quantities in HadCM3.

Thus, remarkably, the emulator calibration has successfully estimated model internal variability using only 100-year means, which we take as one more argument to use the recalibrated emulator.



**Figure 6.** Sensitivity to glaciation level and  $e \sin \varpi$  for sea-surface temperature and mixed-layer depth. Top panels: the contour plots include the experiments 11 and 40. The effect of these experiments are clearly visible in both cases, ice level 3 in the case of sea-surface temperature and glaciation level 7 for mixed-layer depth. Bottom panels: the removal of these experiments a smooth response of the emulator, as clearly seen in the contour plots.

### 3.2 Sensitivity measures

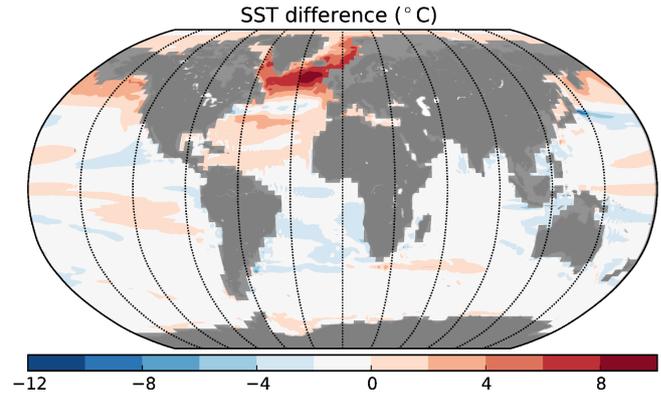
Figure 9 summarises the sensitivities of the four different variables to the external factors.  $e \cos \varpi$  and  $e \sin \varpi$  are grouped together under the term “precess”, for climatic precession.

The figure shows that continental summer temperature is primarily determined by precession,  $\text{CO}_2$  and, to a lesser extent, ice volume. It shows no significant sensitivity to obliquity. Continental precipitation is also mainly driven by precession and less to ice volume. In contrast to temperature, it exhibits no sensitivity to  $\text{CO}_2$ .

Similar to continental temperature, SST is primarily driven by precession and  $\text{CO}_2$  and, to a lesser extent, ice volume. It also shows a larger response to obliquity. Finally, mixed-layer depth shows a pattern similar to precipitation, except that the response to obliquity is not significant compared to the sources of uncertainty induced by the emulation and sampling variance.

### 3.3 Sensitivity to precession

Figure 10 displays the effects of precession on the four variables retained for analysis. The choice here is to show the effects by fixing ice and  $\text{CO}_2$  concentration at three distinct levels representative of the course of glaciation (from top to



**Figure 7.** Sea surface temperature difference between simulations 11 and 15 (see Table 1). There is a clear warming pattern in the North Atlantic, which affects the mean sea-surface temperature.

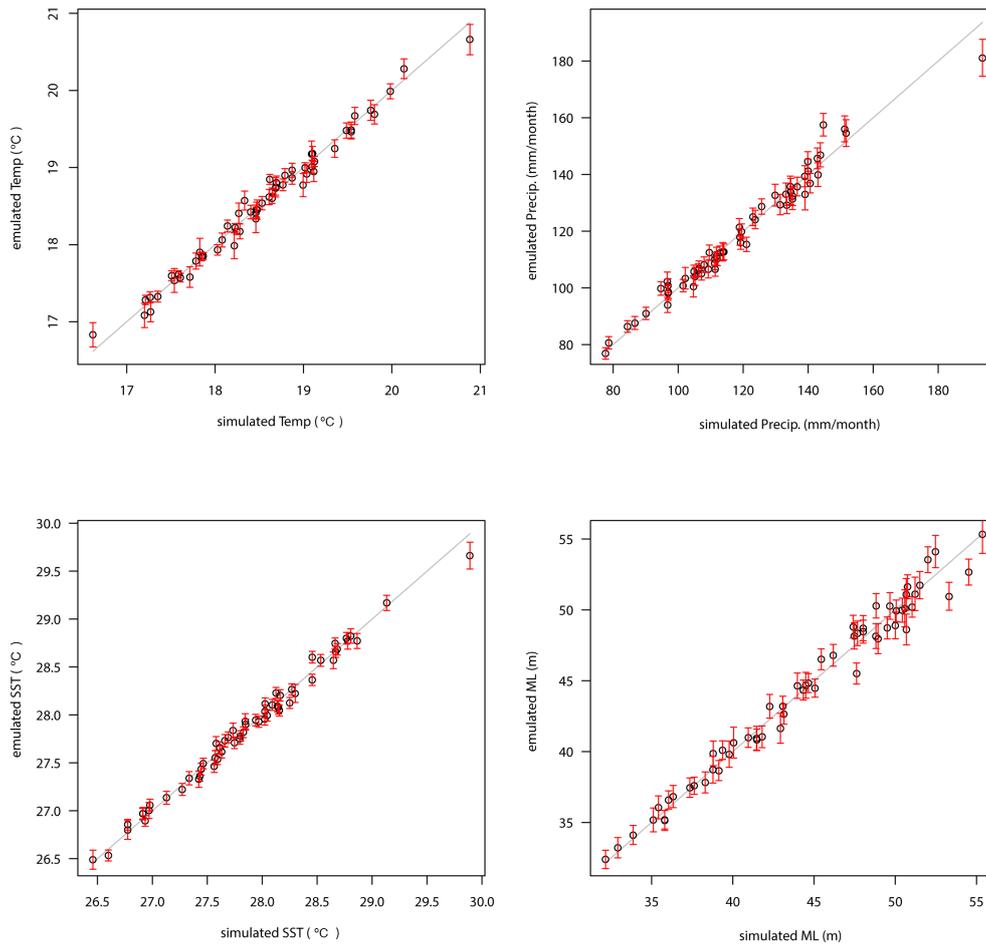
bottom): glaciation level 1/ $\text{CO}_2 = 280$  ppm, glaciation level 5/ $\text{CO}_2 = 230$  ppm and glaciation level 11/ $\text{CO}_2 = 180$  ppm. Quantities are further averaged over obliquity. In order to ease the interpretation, the months representing the time at which perigee is reached are written on the plots: June for  $\varpi = 90^\circ$ , September for  $\varpi = 180^\circ$ , etc. That is, neglecting slow transient effects that could be associated with the deep ocean response, this graphical representation provides an indication of the phase lag between the climate response and the precession forcing of insolation.

We see that the temperature response is in phase with June insolation at low glaciation levels, and in phase with July insolation at mid- and high-glaciation stages.

This feature may physically be understood by considering the summer precipitation response. Precipitation enhances latent heat cooling when perigee is around July. This effect gradually weakens as glaciation takes place and the total amount of precipitation declines, hence the drift towards a more linear response. At higher glaciation levels the JJAS temperature response phase also aligns with July insolation.

The maximum precipitation is obtained when perigee is reached in early July. Among the series of experiments shown by Braconnot et al. (2008), it is indeed the 126 000 year BP experiment (i.e. July perigee) experiment that shows the strongest precipitation response over India.

Furthermore, continental precipitation and mixed-layer depth show opposite response phases to precession. This result is consistent with the earlier findings of Zhao et al. (2005), who identified a shoaling of the mixed-layer depth in this region by about 6 m, consistent across different models, in 6000-year experiments (September perigee). Braconnot and Marti (2003) examined also two nearly opposite precession configurations with the IPSL model, corresponding to perigee in April and October, respectively, and they found a shoaling of the mixed-layer depth compared to the present-day (perigee in January) in both cases.



**Figure 8.** Diagnostic of emulator performance. Shown are the mean and standard deviation of the simulated and the emulated data points for the all the simulations with the exception of simulation number 11 and 40. Top left panel: continental temperature; top right panel: continental precipitation; bottom left panel: sea-surface temperature; bottom right panel: mixed-layer depth.

Zhao et al. (2005) attributed the mixed-layer depth shoaling to a stratification effect involving the response of SST. On this point, our analysis reveals that the maximum SST response occurs when perigee is reached in May. This is not so surprising given that the ocean thermal inertia generally imposes a lag of a few months between the forcing and the response. This response, however, induces an asymmetry between perigee in April and perigee in October, the first one only showing anomalously high SSTs. This is consistent with the analysis of seasonal cycle response provided by Braconot and Marti (2003).

### 3.4 Sensitivity to obliquity

The response of obliquity is mostly linear, as we can infer from the high values of the length scales (see Table 2).

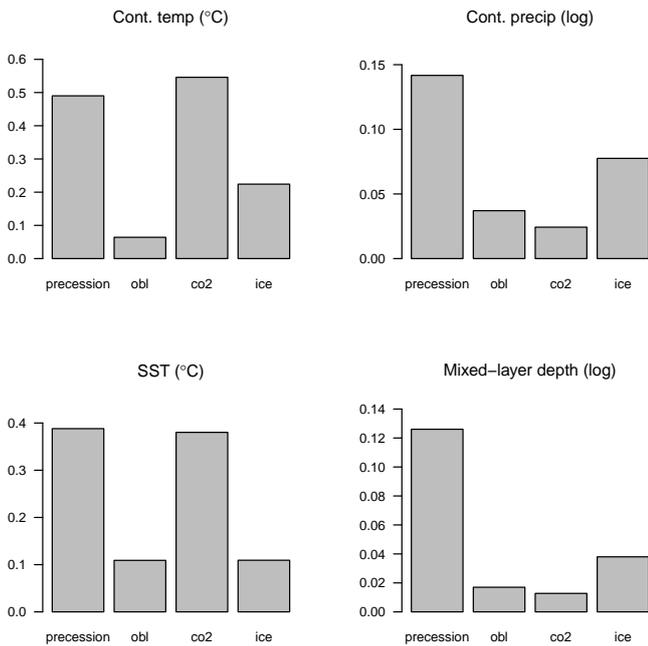
The range of obliquity covered during the Pleistocene induces negligible continental temperature response over the west Indian box. It also induces a slight increase in precipitation. Regarding the Indian Ocean box, there is a somewhat

larger effect on SST compared to continental temperature, but not significant. As for the mixed-layer depth, the response to obliquity is negligible.

In order to better understand the effect of obliquity, we considered the four *idealised* experiments (simulations 20–23; see Table 1). In particular, we discuss here experiments 22 and 23, termed OBL23 and OBL24. They use zero eccentricity, the same CO<sub>2</sub> concentration and glaciation level, and differ by the configuration of obliquity (24 and 23°, respectively). The temperature difference map for JJAS reveals the signature of obliquity-induced insolation changes, with a warming of Northern Hemisphere continents, and slight cooling of significant areas of the tropical oceans (see Fig. 11).

### 3.5 Sensitivity to CO<sub>2</sub> and glaciation level

The response of all variables to CO<sub>2</sub> is best captured by linear processes (optimal  $\lambda_i$  largely exceeds the range covered by the experiment design). Hence, the contribution of CO<sub>2</sub>



**Figure 9.** Sensitivity analysis: shown is the standard deviation of model outputs ( $\sqrt{S}$ ) of each variable, induced by variations in input factors during the Pleistocene. From left to right, top to bottom: continental precipitation, continental temperature, sea-surface temperature and mixed-layer depth.

to the climate response may be estimated straightforwardly from the coefficients  $\hat{\beta}$ , given by Eq. (5). Specifically, the continental temperature and SST responses to the 100 ppm range covered by the experiment design are 2.03 and 1.40 °C, respectively. This corresponds to CO<sub>2</sub> doubling sensitivities of 3.20 and 2.21 °C, in line with the reported HadCM3 sensitivity in CO<sub>2</sub> doubling experiments (see, for example, Fig. 5 of Williams et al., 2001) The responses of precipitation and mixed-layer depth are, again, opposite and very moderate: +6% of precipitation over 100 ppm and –0.5% of mixed-layer depth.

Figure 12 shows the response of continental temperature (left panel), sea-surface temperature (middle panel) and mixed-layer depth (right panel) to the variations of CO<sub>2</sub> concentration and glaciation level. The temperature ranges covered by CO<sub>2</sub> and glaciation levels are of the order of 1 and 2 °C for the continent and ocean surface, respectively. The continental ice effect is mainly present between glaciation levels 1 and 3. With the ice sheet reconstructions used here, the ice area extent which is responsible for the short-wave forcing almost reaches its maximum value at glaciation level 3. Further increasing the glaciation levels affects climate predominantly through the orography forcing (cf. Sect. 3.6).

### 3.6 Orographic effect

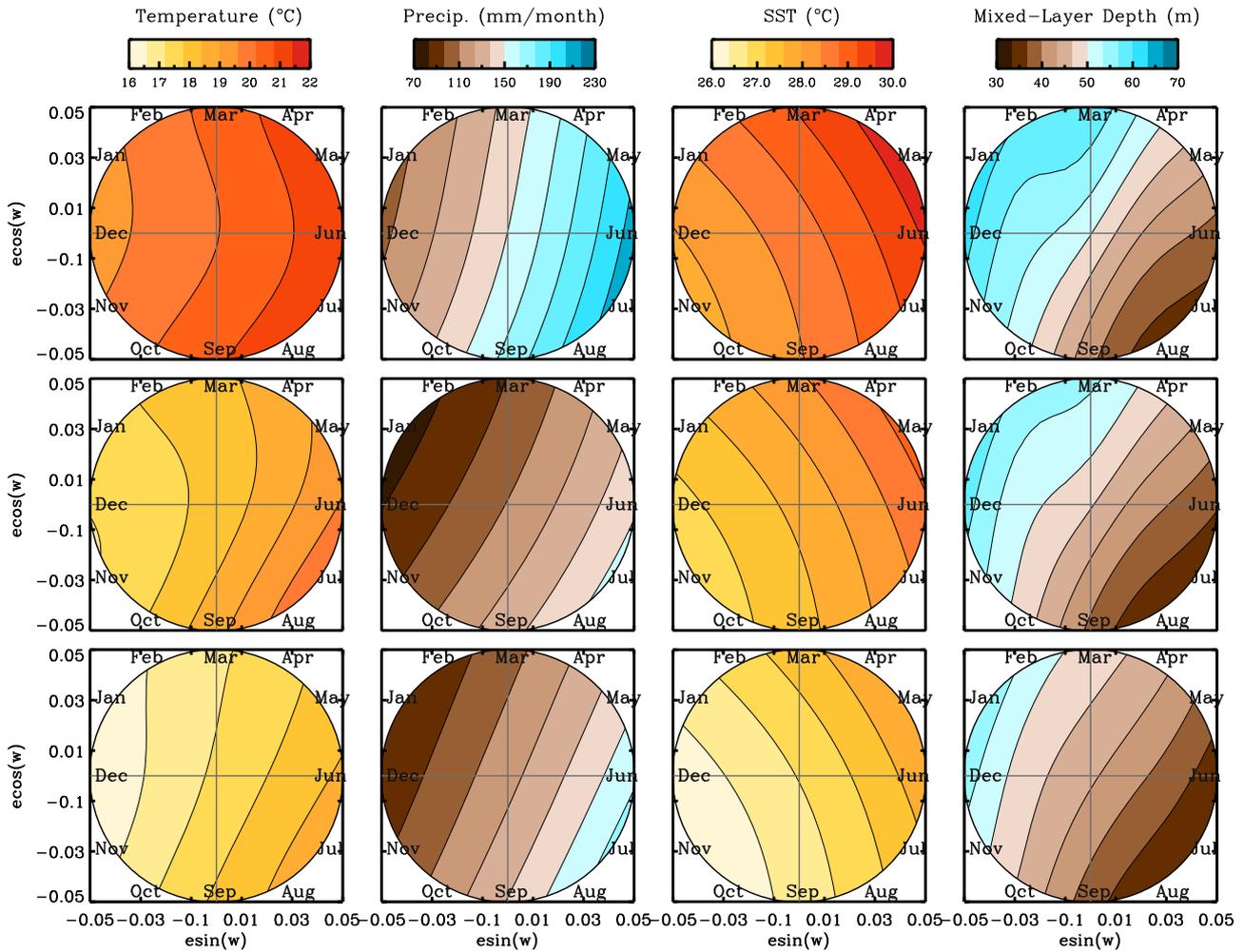
Finally, we consider the differences between the simulations with and without orography forcing of the ice sheets. The latter is potentially important given that mountains and elevated land masses affect the atmospheric circulation and precipitation patterns, and then the whole climate system. To this end, an emulator was calibrated on the available present-day orography experiments.

The net effect orography can then be seen in Fig. 13, where all four variables are plotted as a function of the glaciation level. Black solid lines show the respective variables obtained with the standard experiment design, while red solid lines show the response obtained with the experiment design assuming pre-industrial orography, regardless of the presence of ice sheets. The value plotted here is also implicitly a function of CO<sub>2</sub> concentration, which enters Eq. (11) via the factor  $\rho(\mathbf{x}|\mathbf{x}_{\text{ice}})$ . Dotted lines indicate a 1 $\sigma$  deviation, in both cases, based on Eq. (13), using  $\mathbf{x}_p = \mathbf{x}_p^*$ .

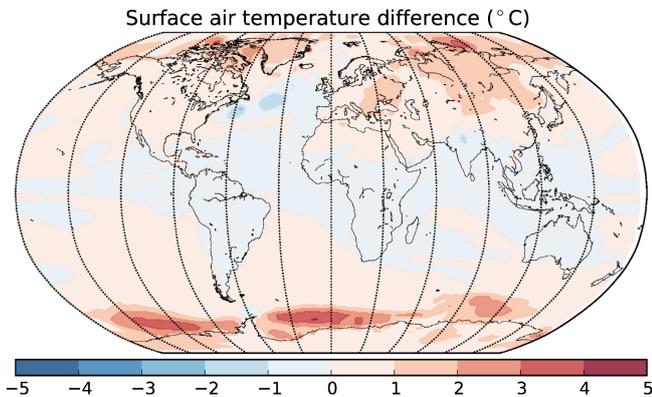
A clear deviation is seen around glaciation level 3. This effect is due to the fact that, as explained in Sect. 2.1, levels 3–11 represent effectively similar ice sheet area, but significantly higher orography (see Fig. 1). Hence, the albedo forcing dominates over the lower range of glaciation levels (1–3), with decreasing temperatures, precipitation and mixed-layer depth shoaling. The orography–no-orography differences appear more markedly above index 3: orography reduces the cooling trends by as much as 1 °C on the continent at glaciation level 11, and even reverses the precipitation trend. As stated in the Introduction, it is known that ice orography forcing may impact monsoon precipitation regimes, but to our knowledge the specific effect of Northern Hemisphere ice sheet orography on the Indian monsoon is yet to be documented. The warming signal caused by orography may be understood by considering the increase in surface potential temperature over elevated regions, similar to what is seen today over the Tibetan Plateau. Because of these high potential temperatures, down-sloping air is effectively warmer than it would be in the absence of orography forcing, and contributes here to increasing the Northern Hemisphere continental surface temperatures. Orographic forcing generally induces atmospheric circulation anomalies and effects on ocean circulation and stratification. For example, Fig. 13 suggests a weak positive effect on mixed-layer depth, quite small compared to the astronomical forcing effects. An in-depth analysis of these effects falls beyond the scope of the present contribution.

## 4 Conclusions

We present a first application of a global sensitivity analysis theory to study the climate response of the Indian monsoon to the climate factors which evolved during the Pleis-



**Figure 10.** Sensitivity to  $e\cos(\varpi)$  and  $e\sin(\varpi)$  for all fields. Each panel, from top to bottom, shows the four fields with a different configuration of glaciation level –  $\text{CO}_2$  concentration. Top panels: glaciation level = 1 and  $\text{CO}_2 = 280$  ppmv. Middle panels: glaciation level = 5 and  $\text{CO}_2 = 230$ . Bottom panels: glaciation level = 11 and  $\text{CO}_2 = 180$ . All fields were integrated over obliquity.

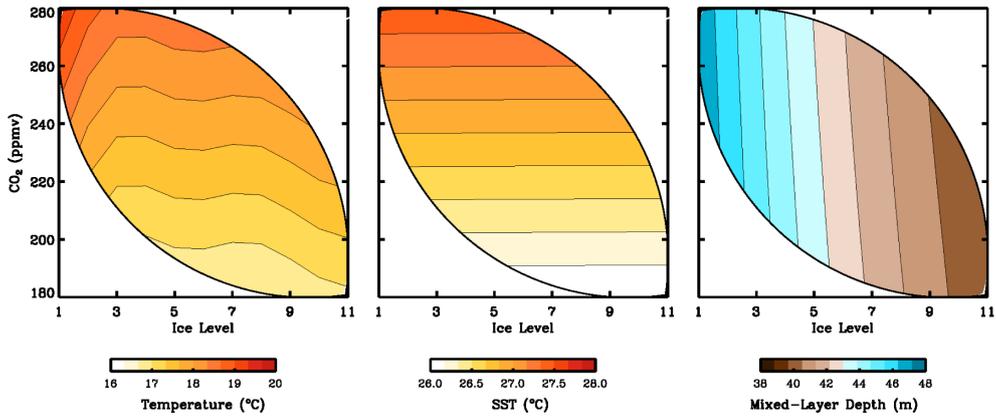


**Figure 11.** Sea surface temperature difference for two *idealised* simulations.  $\text{CO}_2$  concentration, glaciation level and precession remained fixed, the only difference being obliquity (23 and 24°).

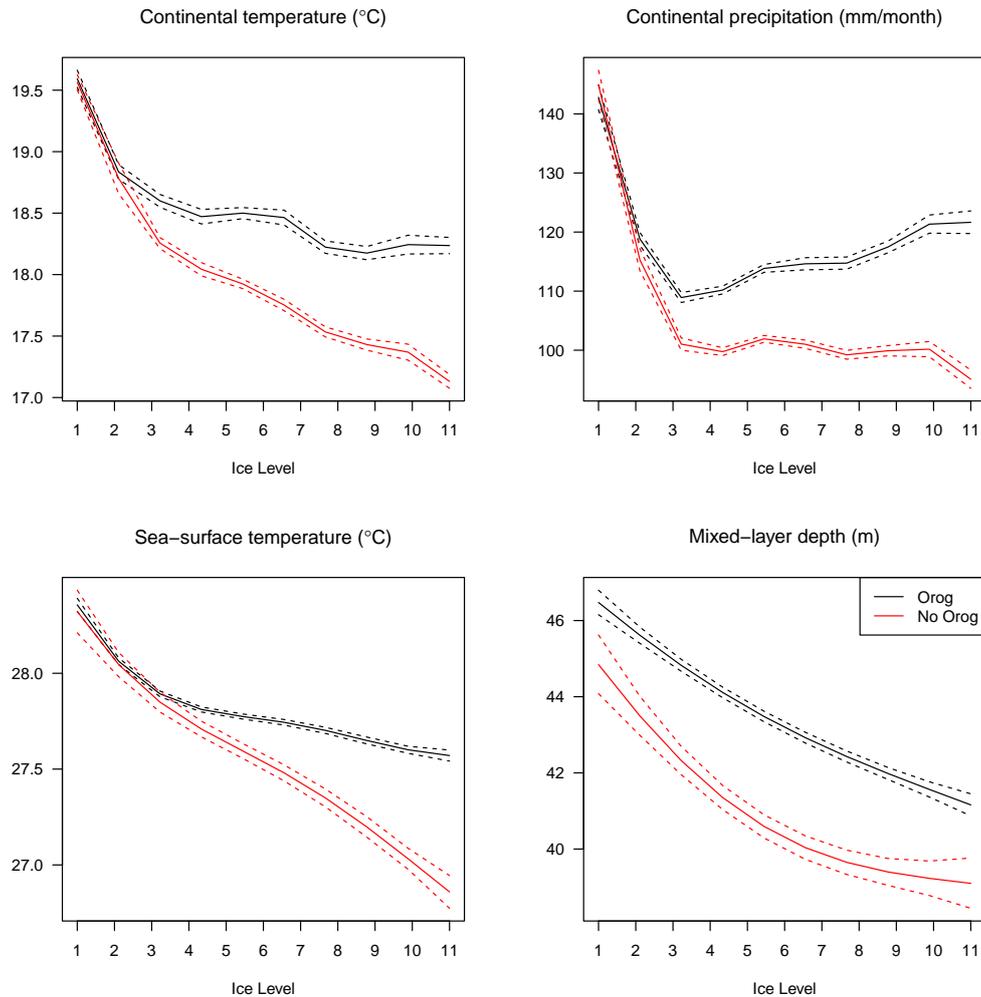
tocene, namely the astronomical forcing ( $e\sin(\varpi)$ ,  $e\cos(\varpi)$ ,  $\varepsilon$ ),  $\text{CO}_2$  concentration and glaciation level.

We focus, in particular, on four variables: continental temperature, continental precipitation, sea-surface temperature and mixed-layer depth. These variables were averaged for the JJAS season over northern India and northwestern Indian Ocean.

Similar to a number of recent studies based on statistical modelling for global sensitivity analysis of computationally expensive simulators, the technical implementation follows a three-step methodology:



**Figure 12.** Sensitivity to CO<sub>2</sub> and glaciation level. From left to right: continental temperature, sea-surface temperature and mixed-layer depth. Fields were integrated over  $e \sin(\varpi)$ ,  $e \cos(\varpi)$  and obliquity.



**Figure 13.** Orography–no-orography difference. From top to bottom, left to right: effect on continental temperature, precipitation, sea-surface temperature, and mixed-layer depth, with orography forcing (black) and without (red). The dotted lines show one standard deviation of the emulator prediction. One may see a departure point from glaciation level 3 in all four fields, as this is the point at which orography forcing becomes the most significant.

- *Designing an experiment plan.* We adopted a Latin hypercube design, optimised following two constraints: maximisation of the minimum distance between two points in the input space – this is called the maxi–min property – and maximisation of the determinant of the matrix of covariance between the input factors – this is a constraint of orthogonality. In addition, the design excludes configurations with excessive eccentricity and unrealistic combinations of CO<sub>2</sub> and glaciation level.
- *Calibration and validation of the emulator.* The validation was performed following a leave-one-out cross-validation approach. Two experiments were excluded of the design as presenting an anomalous North-Atlantic SST patterns. The emulator calibrated on the remaining 59 experiments overall validates the present statistical modelling choices.
- *Quantifying and visualising the individual and combined effects of the different factors on the summer Indian monsoon,* based on sensitivity measures and cross-section plots.

This analysis yielded the following conclusions:

- precession controls the response of four variables: continental temperature in phase with June–July insolation; high glaciation favouring a late-phase response; sea-surface temperature in phase with May insolation; and continental precipitation in phase with July insolation, and mixed-layer depth in antiphase with the latter.
- The effect of CO<sub>2</sub> on continental temperature and SST is of similar size to that of precession on summer continental temperature and SST.
- Obliquity is a secondary effect, negligible on most variables except sea-surface temperature.
- The effect of glaciation is dominated by the albedo forcing, and its effect on precipitation competes with that of precession.
- The orographic forcing reduces the glacial cooling induced by the albedo forcing, and even has a positive effect on precipitation.

The present study confirms the high potential of emulation for exploring and understanding the response of climate models. One originality of the present work was to consider, as inputs, several elements of the climate forcing that (have) varied in the past, and the emulator was used as a method to help us quantify the link between forcing variability and climate variability. The methodology may naturally be applied to other regions of focus and other climate models.

**The Supplement related to this article is available online at doi:10.5194/cp-11-45-2015-supplement.**

*Acknowledgements.* Paul Valdes and Gethin Williams (University of Bristol) are thanked for their assistance in setting up HadCM3 on our systems, and for providing us with ice topography boundary conditions. Richard Wilkinson (University of Nottingham) provided important clarifications to statistical notations. Constructive comments on the discussion version of this paper provided by Tamsin Edwards, Guangshan Chen, Yongqiang Yu and Alexander Kislov are warmly acknowledged. This research is a contribution to the ITOP project, ERC-StG grant 239604. M. Crucifix is funded by the Belgian National Fund of Scientific Research (FRS-FNRS) and P. A. Araya-Melo and N. Bounceur are funded through the ERC ITOP project. Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the FRS-FNRS. The data used in this work, together with the emulator code, are available in the supplementary material.

Edited by: P. Braconnot

## References

- Anderson, P. M., Barnosky, C. W., Bartlein, P. J., Behling, P. J., Brubaker, L., Cushing, E. J., Dodson, J., Dworetzky, B., Guetter, P. J., Harrison, S. P., Huntley, B., Kutzbach, J. E., Markgraf, V., Marvel, R., McGlone, M. S., Mix, A., Moar, N. T., Morley, J., Perrott, R. A., Peterson, G. M., Prell, W. L., Prentice, I. C., Ritchie, J. C., Roberts, N., Ruddiman, W. F., Salinger, M. J., Spaulding, W. G., Street-Perrott, F. A., Thompson, R. S., Wang, P. K., Webb III, T., Winkler, M. G., and Wright Jr., H. E.: Climatic changes of the last 18 000 years – observations and model simulations, *Science*, 241, 1043–1052, 1988.
- 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, 2012.
- Bastos, L. S. and O’Hagan, A.: Diagnostics for Gaussian process emulators, *Technometrics*, 51, 425–438, 2009.
- Berger, A. L.: Long-term variations of daily insolation and Quaternary climatic changes, *J. Atmos. Sci.*, 35, 2362–2367, 1978.
- Berger, J. O., De Oliveira, V., and Sansó, B.: Objective Bayesian analysis of spatially correlated data, *J. Am. Stat. Assoc.*, 96, 1361–1374, 2001.
- Braconnot, P. and Marti, O.: Impact of precession on monsoon characteristics from coupled ocean atmosphere experiments: changes in Indian monsoon and Indian ocean climatology, *Mar. Geol.*, 201, 23–34, Asian Monsoons and Global Linkages on Milankovitch and Sub-Milankovitch Time Scales, 2003.
- 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., Laine, 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.
- Braconnot, P., Marzin, C., Grégoire, L., Mosquet, E., and Marti, O.: Monsoon response to changes in Earth’s orbital parameters: com-

- parisons between simulations of the Eemian and of the Holocene, *Clim. Past*, 4, 281–294, doi:10.5194/cp-4-281-2008, 2008.
- Carlsaw, 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, 2013.
- Chen, G.-S., Liu, Z., Clemens, S., Prell, W., and Liu, X.: Modeling the time-dependent response of the Asian summer monsoon to obliquity forcing in a coupled GCM: a PHASEMAP sensitivity experiment, *Clim. Dynam.*, 36, 695–710, 2011.
- Essery, R. L. H., Best, M. J., Betts, R. A., Cox, P. M., and Taylor, C. M.: Explicit representation of subgrid heterogeneity in a GCM land-surface scheme, *J. Hydrometeorol.*, 4, 530–543, 2003.
- Felzer, B., Webb III, T., and Oglesby, R. J.: The impact of ice sheets, CO<sub>2</sub>, and orbital insolation on late quaternary climates: sensitivity experiments with a general circulation model, *Quaternary Sci. Rev.*, 17, 507–534, 1998.
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Clim. Dynam.*, 16, 147–168, 2000.
- Gramacy, R. and Lee, H. H.: Cases for the nugget in modeling computer experiments, *Stat. Comput.*, 22, 713–722, 2012.
- Hewitt, C. D.: The effects of ocean dynamics in a coupled GCM simulation of the Last Glacial Maximum, *Clim. Dynam.*, 20, 203–218, 2003.
- Hewitt, C. D., Broccoli, A. J., Crucifix, M., Gregory, J. M., Mitchell, J. F. B., and Stouffer, R. J.: The effect of a large freshwater perturbation on the Glacial Atlantic Ocean using a coupled general circulation model, *J. Climate*, 19, 4436–4447, 2006.
- Holden, P., Edwards, N., Oliver, K., Lenton, T., and Wilkinson, R.: A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1, *Climate Dynamics*, 35, 785–806, 2010.
- Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, *Reliab. Eng. Syst. Safe.*, 52, 1–17, 1996.
- Joseph, V. R. and Hung, Y.: Orthogonal-maximin latin hypercube designs, *Stat. Sinica*, 18, 171–186, 2008.
- Kennedy, M. C. and O’Hagan, A.: Predicting the output from a complex computer code when fast approximations are available, *Biometrika*, 87, 1–13, 2000.
- Kutzbach, J. E. and Guetter, P. J.: The influence of changing orbital parameters and surface boundary conditions on climate simulations for the past 18 000 years, *J. Atmos. Sci.*, 43, 1726–1759, 1986.
- Kutzbach, J. E. and Ruddiman, W. F.: Model description, external forcing, and surface boundary conditions, in: *Global climates since the last glacial maximum*, 12–23, 1993.
- Kutzbach, J. E. and Street-Perrott, F. A.: Milankovitch forcing of fluctuations in the level of tropical lakes from 18 to 0 kyr BP, *Nature*, 317, 130–134, 1985.
- Kutzbach, J., Liu, X., Liu, Z., and Chen, G.: Simulation of the evolutionary response of global summer monsoons to orbital forcing over the past 280,000 years, *Clim. Dynam.*, 30, 567–579, 2008.
- 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.
- Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic  $\delta^{18}\text{O}$  records, *Paleoceanography*, 20, PA1003, doi:10.1029/2004PA001071, 2005.
- 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 insolation, Ph.D. thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 1993.
- 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, 2008.
- MacCalman, A. D.: Flexible space-filling designs for complex system simulations, Ph.D. thesis, Naval Postgraduate School, Monterey, California, US, available at: <http://hdl.handle.net/10945/34701> (last access: 29 March 2014), 2013.
- Masson, V., Braconnot, P., Jouzel, J., de Noblet, N., Cheddadi, R., and Marchal, O.: Simulation of intense monsoons under glacial conditions, *Geophys. Res. Lett.*, 27, 1747–1750, 2000.
- 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, 1979.
- Morris, M. D. and Mitchell, T. J.: Exploratory designs for computational experiments, *J. Stat. Plan. Infer.*, 43, 381–402, 1995.
- Oakley, J. and O’Hagan, A.: Bayesian inference for the uncertainty distribution of computer model outputs, *Biometrika*, 89, 769–784, 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, 2004.
- O’Hagan, A.: Bayesian analysis of computer code outputs: a tutorial, *Reliab. Eng. Syst. Safe.*, 91, 1290–1300, 2006.
- Pepelyshev, A.: The role of the nugget term in the Gaussian process method, in: *mODA 9 – Advances in Model-Oriented Design and Analysis*, Springer, 149–156, 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, 1999.
- Petropoulos, G., Wooster, M. J., Carlson, T. N., Kennedy, M. C., and Scholze, M.: A global Bayesian sensitivity analysis of the 1d SimSphere soil–vegetation–atmospheric transfer (SVAT) model using Gaussian model emulation, *Ecol. Model.*, 220, 2427–2440, 2009.

- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P.: Design and analysis of computer experiments, *Stat. Sci.*, 4, 409–423, 1989.
- Santner, T., Williams, B., and Notz, W.: *The Design and Analysis of Computer Experiments*, Springer, New York, 2003.
- Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M.: *Sensitivity analysis in practice*, John Wiley and Sons, Ltd, Thichester, W. Sussex, England, 2004.
- 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, 2011.
- Sexton, D. M., Murphy, J. M., Collins, M., and Webb, M. J.: Multivariate probabilistic projections using imperfect climate models part I: outline of methodology, *Clim. Dynam.*, 38, 2513–2542, 2012.
- 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.
- Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr, *Quaternary Sci. Rev.*, 29, 43–55, 2010.
- 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, 2010.
- Williams, K. D., Senior, C. A., and Mitchell, J. F. B.: Transient climate change in the Hadley Centre Models: the role of physical processes, *J. Climate*, 14, 2659–2674, 2001.
- Yanase, W. and Abe-Ouchi, A.: The LGM surface climate and atmospheric circulation over East Asia and the North Pacific in the PMIP2 coupled model simulations, *Clim. Past*, 3, 439–451, doi:10.5194/cp-3-439-2007, 2007.
- Yin, Q. and Berger, A.: Individual contribution of insolation and CO<sub>2</sub> to the interglacial climates of the past 800 000 years, *Clim. Dynam.*, 38, 709–724, 2012.
- Yin, Q. Z., Berger, A., and Crucifix, M.: Individual and combined effects of ice sheets and precession on MIS-13 climate, *Clim. Past*, 5, 229–243, doi:10.5194/cp-5-229-2009, 2009.
- Zhao, Y., Braconnot, P., Marti, O., Harrison, S. P., Hewitt, C. D., Kitoh, A., Liu, Z., Mikolajewicz, U., Otto-Bliesner, B., and Weber, S. L.: A multi-model analysis of the role of the ocean on the African and Indian monsoon during the mid-Holocene, *Clim. Dynam.*, 25, 777–800, 2005.