Earth Syst. Dynam. Discuss., 5, C627–C633, 2014
www.earth-syst-dynam-discuss.net/5/C627/2014/

Earth System Dynamics Discussions

Interactive comment on "Global sensitivity analysis of the climate–vegetation system to astronomical forcing: an emulator-based approach" *by* N. Bounceur et al.

N. Bounceur et al.

michel.crucifix@uclouvain.be

Received and published: 12 December 2014

First of all, we would like to thank the reviewers for constructive, detailed, well-informed and at places thought-provoking reviews, of which the manuscript will largely benefit. As customary, a detailed itemized response to all comments is provided in the response to the reviewers. In order to avoid unneeded redundancies we provide here only a common, synthetic response that addresses what we believe to be the most important questions.

The introduction, and the purpose of the article: We agree with the reviewers that there is scope for clarifying our objectives in the introduction. We cannot of course

claim to have developed a *new* kind of emulator, but we believed reasonable to use an experiment design with LOVECLIM, an Earth Model of Intermediate Complexity, as a test case to study the potential applications of emulation in a palaeoclimate context. In addition, we are considering an input space that cannot be considered as "uncertain" but explored predictably by Nature. To some extent, we are advocating a change in current approaches of palaeoclimate modelling: besides previous attempts at "reproducing a rapid climate change at the right time" (e.g., Claussen et al. 1999), we believe that state-of-the-art methodologies coming, say, the "MUCM' paradigm (modelling uncertainties in computer experimetns) may best be used to scan the whole input space in order to detect where and when rapid climate change may occur in response to smooth astronomical forcing changes. In particular, to our knowledge, sensitivity analyses to astronomical forcing published so far rely on a "local" approach (comparing, e.g., two different epochs), and the objective of our introduction was precisely to stress this point. The interest of this exploratory work is perhaps best shown by the reviewers' discussion: cases similar to that of the "experiment 20" (which behaved unexpectedly) or the debate between PCA-emulation vs independent emulation are likely to occur again in the future. These methodological objectives will be better stressed in the revised version. In particular, the introduction was thoroughly revised. Following comments of the reviewer #1, we have also reduced, where appropriate, the repetition of previously material, summarised the more technical questions and moved some of the more technical developments (in particular, the experiment design) to an appendix. We have also toned down the claims regarding the advances in statistical methodology, yet retaining the emphasis on what they could bring to the specific context of palaeoclimate analysis. We also concede that important references were lacking. In this context it is probably best to emphasis the global sensitivity analysis process (Lee et al., 2011-2013) , PCA analysis (Holden et al., 2010, GRL), calibrating an Earth-Model of Intermediate Complexity (Edwards et al. 2011) or inferences about climate sensitivity (Rougier et al., 2009). We note also that Schmittner et al. 2011 have compared different solutions for emulation in a palaeoclimate context.

The notation: We have reduced the number of dummy variables and, where possible, the number of equations and used notations consistently with previous literature on the subject, where possible.

PC vs independent emulators: The PC and the independent emulator approach are two different ways of building multivariate emulators, but there are many other possibilities (coregionalised emulators for example). Certainly the PCA approach is unlikely to be going too far wrong if we are emulating enough of the components, so that the sum of the eigenvalues for these components is close to the total sum. It is likely to produce maps which are smoother and capture spatial structure better than the independent emulator approach (as PCA is designed to exactly capture this spatial variation), as the independent emulator approach will produce produce different sensitivity indices for each grid box, with no requirement for smooth variation in the these indices. But we don't see one method as inherently better or worse than the other, but for any given situation one approach could outperform the other. For example, the idea of using same length-scale for every emulator is questionable. We provided a sensitivity analysis comparing same lenth-scale with per-PC optimum lenth-scales for the PC emulator, but it would be impossible to do detailed checks of 2000 independent emulators, so this assumption is more worrying for independent emulators than for the PC emulator. Furthermore the method of using PC emulators has been around for a while (at least since Higdon et al 2008. This said, we understand the point of the reviewer that we might be among the first ones to use the PCA for the specific purpose of global sensitivity analysis. We implemented the independent emulators approach and found that the results are visually almost identical: in fact it requires a carefull plot inspection to actually realise they are not the same. The diagnositics that we referred to as 'linear-vs-non-linear' are more fragile to the type of emulator used.

We take account of this information in the revised manuscript and briefly discuss this point in new version of the manuscript. Finally, the revised version makes now explicit use of the covariance matrices provided by the PC-emulator.

Estimating "simulator variance" vs "emulator variance" Reviewer #1 made a good point here : we cannot distinguish "simulator" from "emulator" variance. All the discussion now uses the previously noted S_p (now V_p in the new notation). This said, we still use the average emulator covariance averaged over the whole input space, may still be used as an *informal* measure giving the order of magnitude associated with using an emulator rather than the simulator itself.

Distinguishing "linear" vs "non-linear" effects: One fair point brought by the reviewer #1 is that there was confusion between the idea of 'slow vs fast' changing (in response to slow astronomical changes) vs other effects that may be associated with non-linear behaviours such as, e.g., non-monotonous response. Let us think again about the objectives of the paper: demonstrate the potential of an emulator-based approach of sensitivity analysis to astronomical forcing — possibly to be used later with much more computationally demanding models — to address typical problems of palaeoclimatology. We adopt here a hypothesis of time-scale separation, i.e., assume that the ocean-atmosphere system is quasi-stationary with respect to the astronomical forcing and thus, as correctly identified by the reviewer, our interest eventually is in detecting regions of the input space that correspond to steep gradients in the climate response. Following these considerations and this reviewer's comment we chose to focus on simulating the "maximum rate of change" of climate variables, estimating by using as input realistic successions of astronomical forcing that can be computed from the available astronomical solutions.

The case of "experiment 20", not properly simulated and discarded from the calibration set: Experiment 20 is not correctly predicted by the emulator. In a word, it behaves against our *expectations*, in the sense that Gaussian process emulator using linear relationships to the forcing plus a smooth stochastic component reflects our prior beliefs about the simulator response. We agree with the reviewer that this situation calls for a careful inspection of experiment 20, and that we were falling short of this discussion in the original version. This is now better done. It is shown that experiment 20 oscillates at low frequency, at first sight similar to the dynamics of Dansgaard-Oeschger events. The oscillation period varies as the deep ocean temperature adjusts to the forcing and at the end of the experiment it is roughly 800 years. Consequently, a 500-year average (as used here) may randomly produce different fields. The emulator, founded on the hypothesis that the data supplied for calibration represent a stationary climate, has no chance to capture this. In fact, the climate system oscillates in this experiment 20 between two meta-stable states, which may be termed as "warm North Atlantic" and a "cold North Atlantic", and the emulator predicts the warm equilibrium. After the original submission of the manuscript we tried a couple of more experiments with nearby input configurations, and realised that the oscillator behaviour occurs for these parameter choices as well. The response to be given this state of affairs is a matter of judgement. We could, for example, start an iterative experiment plan procedure, the objective of which would be to delineate the 3-D contours of the experiment design within which the oscillatory behaviour occurs, in order to build an emulator specific to that zone. This is a paper on its own, and the added value of this work for palaeoclimate science is questionable, for two reasons. The first one is that the region of the parameter space where this occurs is in fact almost never reached by Nature. In particular, the obliquity of 22° used for experiment 20 is in the lower 0.013th lower percentile (i.e., values lower that 22 ° occur less than 0.013 % of the time) according to the Laskar et al. 2004 solution over the last 10 million years. So it has in fact no weight on the global sensitivity measures. The second reason is that the oscillator behaviour has been documented before in Goosse and Renssen 2002, and Loutre et al. 2014

and generally judged to be 'non-robust' : i.e., the parameter region where this oscillation could occur is sensitive to physical parameter changes well within uncertainties. Therefore, we prefer to acknowledge the behaviour, consider that it *could be of significance* (i.e. we are definitely *not* ignoring the experiment or pretending that it never occurs) but use discard it from the emulator design procedure to the benefit of the performance of the emulator on the rest of the experiment domain, which covers quasi 100 % of the actual astronomical forcing region. Then, the potential implications of this choice are discussed in the last part of the article.

Using Gaussian or Student-t statistics: We reproduced barplots using the Student distribution (will be shown in the revised version). Visually the results are hardly distinguishable from the original version. In fact, due to the nature of the PC emulator, the variance at any grid point will be the *sum* of several Student-t distributed quantities (the distribution of which has to be estimated by Monte-Carlo), plus a variance associated with residual PCs that is Gaussian distributed. The result is really very close from being Gaussian, and it might well be that the original barplots were more exact, since they did not require Monte-Carlo simulations.

The revised version of the article includes

- Comparision of Monte-Carlo integrals with analytical integrals,
- Discussion of the precipitation response phase of African precipitation to precession,
- Fingerprints, that is, eigenvectors of output covariance matrices,
- A comparision between grid-point emulation and PC emulation, in the form of a brief discussion, and figures available in supplementary material,

- A discussion of the oscillatory behaviour of experiment 20, and its consequences for paleoclimatological interpretation,
- More detailed mathematics about the calculation of variance indices with the PC emulator,
- Thoroughly revised notation.

Once again, we are grateful to the reviewers and the editors for careful consideration of this article.

Interactive comment on Earth Syst. Dynam. Discuss., 5, 901, 2014.