

Interactive comment on “Global sensitivity analysis of Indian Monsoon during the Pleistocene” by P. A. Araya-Melo et al.

T.L. Edwards (Referee)

tamsin.edwards@bristol.ac.uk

Received and published: 15 July 2014

This is an excellent study and an important area of research in both method and results. To my knowledge it is the second published study emulating a climate simulator in a palaeoclimate application (after Schmittner et al., 2011, which needs citing somewhere), and the first for a coupled atmosphere-ocean GCM. Most of my comments relate to clarity, language, or minor errors: the manuscript needs careful proofreading and clarification of several points, and some figures need improving.

//// Minor scientific points

1623/25 Appears to contradict earlier statement that parameter dependence is smooth... [It's also not clear to me at this stage why only one outlier appears to distort

C1042

each of SST and mixed layer depth, not both for both.]

1624/10 How about other reasons for oceanic change - e.g. precip?

1624/13-15 This is important but not very well justified or well written.

1624/25 Er. . .because you removed them! Not relevant.

1626/27 As far as I can tell from this figure your signal direction is consistent with Zhao et al. but not magnitude (approx. 3 times larger) - can you comment? (And what is the magnitude of Braconnot and Marti result?)

1627/15 "linear" - How is this quantified - same reasoning as 1628/2? Refer to values in Table 2 for both?

1628/21 "reverse" - do you mean the response is not monotonic? Is this shown somewhere? Is it strange that they *exactly* cancel?

//// Comments and questions for improving clarity

At the moment the Emulator and Sensitivity Measures sections are quite heavy going for someone a little familiar with emulators, and not sufficiently intuitive or descriptive for someone new to them. They should either be moved to an appendix or SI, or expanded with more explanatory text. I've assumed the latter in my comments.

1610/3 Useful to give Pleistocene dates

1616/4 if you use the terms maxi-min and orthogonality here, mention them when describing them on p1615 too

1617/13 "input vector associated with the jth component of the experimental design" -> I think this would be clearer as "input vector (set of parameter values) of the jth ensemble member"

1618 Is it possible to add some more explanations of choices, terms and notation? Some suggestions:

C1043

- better explanation of the meaning of various correlation functions $c(x, x^*)$, $c(x, x_j)$ and $c(x_i, x_j)$, i.e. the intuitive meaning of T and A compared with V
- describing components of L3 in Bayesian and/or data assimilation and/or least squares terms
- perhaps show a 1D or 2D version and/or diagram first?
- might be worth "borrowing" ideas from this page in the MUCM toolkit: <http://mucm.aston.ac.uk/MUCM/MUCMToolkit/index.php?page=MetaFirstExample.html>
- 1619/5 experiment length - better to say something like "degree to which the simulation has reached equilibrium" to make this point?
- 1619/13 Can you explain why you want to (i.e. why A&C recommend to) restrict the nugget amplitude? Again, an intuitive explanation of this section would aid the reader.
- 1620 The empirical distribution part is not clear. Please give more information on how rho was obtained, what the resulting values indicate, and perhaps add some kind of figure.
- 1620/13-15. Every time I try to read this sentence I fail. . .and then you lose me for the rest of the section. I think the concepts here are not difficult, but they are currently a bit impenetrable in the explanations. This section needs to be longer and with shorter sentences...
- 1621/17 This could be better explained: e.g. "These two are equivalent if the covariance between all input factors does not depend on the value of factor p (x_p). This is true for [example factors] . . . However, if the covariance does depend on x_p , as it does for [. . .], then ...".
- 1622/13 IND could stand for India or Indian Ocean. How about using NI and IO, or NIND and INDO? or just call them the land and ocean boxes/regions?
- 1623/3 Can you add a recap of the meaning of the scales (e.g. "length scale of the cor-

C1044

relation between outputs from different simulations, for each of the five input factors") and nugget (e.g. something like "error due to simple modelling of this correlation" or "model structural uncertainty")? And maximisation of the penalized likelihood? (e.g. "i.e. found by the best fit of the emulator to the ensemble outputs"?)

1623/10 First sentence could be clearer. How about "The correlation length scales are generally similar to the input factor ranges [give percentages?]. This .." And replace "may thus be suitably calibrated by the experiment design" with something more specific e.g. "and thus the ensemble size is sufficient and a Gaussian process emulator is appropriate."?

1623/14 The "however" is confusing in this context. And "response to" makes more sense to physical scientists than "response in". . .? Why write "This is not a problem on its own"? Better: "This indicates that a sparser sampling of this factor would have been sufficient".

1623/25 "mean effect" is not explained here nor in its first appearance (1621/18). Something like "mean effect, which here shows the response to x and y while keeping z constant" would help. Convoluted is not really the right word. Waves / ripples / distortions in the contours? Departure from smooth linear gradients?

1624/22 Either insert "standard" before "normal" or else remove "normalised by standard deviation". Give context to p value, e.g. "null hypothesis rejected at. . .".

1625/3 "observation" is confusing -> how about "uncertainty of the simulation, i.e. how representative the 100 year simulation is of the mean model state".

1625/7 Better to use residuals errors than residual uncertainties, I would say, because you are talking about the differences between known numbers here.

1625/10-12 Not particularly well-written. How about: "Thus, remarkably, the emulator calibration has successfully estimated model internal variability using only 100 year means."

C1045

1625/16 The final sentence of Fig. 7 caption text is useful and could/should be re-used here. For clarity, reorder figures / text to match each other (precip first or temp first).

1626/14 Low glaciation looks the most symmetric in terms of contours - can you please clarify?

1627/9 I must be missing something but I'm confused - you seem to describe this as if it is a lag between maxima within the same year, but it is a phase shift across many years?

1627/10-13 I don't understand this bit.

1627/16-19 Delete (poorly written repetition of earlier discussion).

1628/8 Seems odd to only cite such an old paper - add something more recent (e.g. emulated..?) too

1629/8-20 I think these explanations need expanding.

Figure 4 caption confusing and has errors: suggest "Diagnostic of emulator performance for sea surface temperature (left panel) and mixed-layer depth (right panel), showing mean and standard deviation of emulator prediction versus actual HadCM3 output.

Figure 6. Suggest rewriting, e.g. replace second sentence onwards with: "The response to the input factors is generally smooth, except in the regions near experiments 11 and 40 (glaciation level = 3 and 7)."

//// Corrections

1610/2 "the" Indian Monsoon

1610/22 control

1611/11 Delete "Namely"

1611/15 changes "in"

C1046

1611/16 affect<-> significantly

1612/7 Consistently (or, better, Consistent)

1612/14 "of" insolation

1612/15 put "6000 years ago" earlier

1612/26 exercises

1612/27 allowed further testing of

1613/5 depend

1613/6 coupled; parameter

1613/12 individual

1613/ delete "influences"

1613/21 GCM's or GCM

1613/23 acts

1614/9 (Gordon et al., 2000)

1614/19 input

1615/4 On the other hand, CO2 is -> Atmospheric CO2 concentration is

1615/14 effectively -> efficiently; factor

1615/25 to be -> is

1616/20 associated with the MOSES2 surface component -> using the MOSES2 dynamic land surface scheme

1616/23 delete "On the other hand"

1617/1 ran -> run for

C1047

1617/13 input -> output?
1619/2 (Gramacy and Lee, 2012)
1619/3 delete "Indeed"
1619/20 of *a* Bayesian approach
1620/2 missing "."; delete "For CO2 and glaciation level"
1623/22 mixed layer *depth* (and various other places elsewhere)
1624/19 left -> leave
1625/11 years -> year
1625/22 Contrarily - Contrary / In contrast
1627/20 ideal cases -> idealised
1628/16 insert "(which is...)" around responsible...forcing
1629/23 temperature -> temperatures. Is down-sloping the right/best word?
1629/28 No need for "though"

Figures are very out of order with respect to their reference in the text, which makes it harder to read.

Units "mts" should be "m".

Fig. 2 Poorly annotated and explained. How does it show both SLP and temp? Is the temp SAT or SST? Missing label on palette. What is IL? Why are the regions coloured green and purple instead of outlined in black - is that the SLP? If so where is the scale? Why does the caption say it only shows SLP and temp for the boxes when in fact it shows it for a much larger region?

Fig. 7 Improve or remove titles.

C1048

Fig. 9 Palette range needs larger / infinite bounds (saturated areas white). And ideal -> idealised

Fig. 12 These plots could be very nice, but need tweaking. Colour scales and layout great in principle, but: palette increments horrible (74, 103, ... etc); contour increments similarly horrible; palette range often too large (dark blue and white never used); contour numbers often half missing.

Interactive comment on Clim. Past Discuss., 10, 1609, 2014.

C1049