Earth Syst. Dynam. Discuss., 1, C189–C192, 2011
www.earth-syst-dynam-discuss.net/1/C189/2011/

Earth System
Dynamics
Discussions

Interactive comment on "Emulating Atlantic overturning strength for low emission scenarios: consequences for sea-level rise along the North American east coast" *by* C. F. Schleussner et al.

J. Rougier (Referee)

j.c.rougier@bristol.ac.uk

Received and published: 7 March 2011

This paper opens promisingly, with a sentence on the importance of estimating uncertainty ranges. I would like to comment on some aspects of this, making it clear that I am a statistician and not a climate scientist, although I do do a lot of work jointly with climate scientists, and I am familiar with the type of experiment being run here.

One point to make on section 2, concerning terminology. The word 'emulator' tends to be reserved (in statistics, and more widely) for stochastic representations of deterministic functions, for example as fitted using Gaussian processes. Simpler deterministic models fitted to more complicated deterministic models tend to be called 'surrogates'.

The box model would be a surrogate for the AOGCM.

Now to section 4. I would like more information about how the box model was initialised for each AOGCM. I would also like a lot more information concerning the statement "the parameter set was optimised to reproduce each AOGCM output". Inspecting Fig 2, I guess that the authors have minimised the sum of squared residuals between the box model output and the AOGCM. Why would this be a good idea? Well, the resulting parameter estimate will be the maximum likelihood estimate if it were the case that the residuals are IID Gaussian with mean zero. With the exception of GFDL_R30 this would seem to be a reasonable assertion about the residuals. I would like to make two further points. First, the marginal variance of the residuals appears to be increasing with AMOC strength. This suggests to me that a better criterion would be to minimise the sum of squared residuals of the logarithms. Second, the authors have missed an opportunity to use the very detailed theory of maximum likelihood estimators to provide an uncertainty estimate for the parameter.

Let me expand on this second point. The standard deviation of the residuals for each AOGCM is itself uncertain; let's call it sigma. Its value is not important for estimating the box model parameters for the AOGCM, but it /is/ important for estimating their uncertainties. If sigma is incorporated into the log-likelihood the resulting penalty function to be minimised (which is the negative of the log-likelihood) is

-log L(theta, sigma) = c + n log sigma + (1 / 2 sigmaˆ2) SSR(theta)

where SSR(theta) is the sum of squared residuals as a function of the box model parameters theta. It is easy to see that sigmahat, the ML estimate of sigma, is the square root of RSS(thetahat) / n, where thetahat is the ML estimate for theta, which minimises the RSS. For small n, caution suggests replacing sigmahat with the square root of RSS(thetahat) / (n - k), where k is the number of parameters.

Standard asymptotic theory can now be used to approximate the standard errors of thetahat and sigmahat, and to give confidence intervals for each parameter. The estimated standard errors are the square root of the diagonal of the inverse of the hessian of the negative log-likelihood. It's a mouthful but found in any statistics textbook. I like Davison, 2003, Statistical Models (ch4). These standard errors are crucial for quantifying uncertainty; both sigmahat and the standard errors should have appeared in Table 2. Without them, the authors have not done justice to their first sentence.

The diagnostic in Fig 3 is reassuring.

Section 5 also raises some interesting questions. It appears as though the five surrogates, standing in for the five AOGCMs, have been averaged together. What is the epistemic principle which justifies this? One can see that the five AOGCMs vary widely (eg Figs 2 and 3), so it follows that some of them match the recent observations better than others. So, on this basis it would have been more prudent to show each AOGCM separately.

I am also concerned the Fig 4a underestimates the uncertainty. If the tuned box model is standing in for the AOGCM, then the error in its approximation must be incorporated into the simulation. But the text does not suggest that this has been done; and, indeed, because it needs estimates of sigma, I doubt that it has. The authors need to add a Gaussian IID error to the box model output with standard deviation sigmahat, in order for their simulation to account for the difference between the tuned box model and the AOGCM. This could add a few Sv to the range of uncertainty at 2100.

Continuing in the same vein, the authors should not use the point estimate thetahat for each run of their box model, but should sample from the uncertainty about thetahat, in order to incorporate parametric uncertainty. I expect this will also add a few Sv to the range at 2100.

Overall, I like this paper because I like the idea of using simple models to understand complex systems, and AOGCMs qualify as complex. But what we have here seems to be a missed opportunity to quantify uncertainty in a statistically defensible manner and, more concerning, a consequence of this is an underestimate of predictive uncertainty.

Interactive comment on Earth Syst. Dynam. Discuss., 1, 357, 2010.