Atmos. Chem. Phys. Discuss., 11, C9579–C9582, 2011
www.atmos-chem-phys-discuss.net/11/C9579/2011/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters" *by* L. A. Lee et al.

L. A. Lee et al.

l.a.collins@leeds.ac.uk

Received and published: 30 September 2011

The authors wish to thank Jonty Rougier for reading our paper and providing some valuable comments. We have replied to the comments below where the comments are made in Italics. Changes to the manuscript are shown in bold.

My only technical comment is on p20443, "Independence of the emulator inputs". At the point where you run the ensemble and build the GP emulator, you do not need any probabilistic concepts for the simulator (I prefer 'simulator' to 'model' or 'computer model') at all – they only enter at the point where you want to do a variance-based sensitivity analysis. So 'independence', which might be taken to be probabilistic independence, has to be carefully specified.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

In Frequentist statistics, where one is not allowed to treat the parameters as uncertain, the common construction here is to require the parameters to be 'variation independent'. This means, effectively, that the setting of any subset of the parameters does not affect the range of the others. Ie the parameter space is a box. My reference to hand is David Cox, 2006, Principles of Statistical Inference (p2). Emulators do not need the inputs to be variation independent, although it definitely helps.

The situation to be avoided is where two or more inputs are not separately identifiable in the simulator output. This has got nothing to do with independence, but reflects the nature of the parameterisation: for example if x and y always enter the simulator in the form (xy) then they are not 'dependent', but clearly there is no need to vary them both. Generally, we will not know the identifiability structure of the full set of inputs, but we will have judgements about interactions which can inform us about inputs which, for reasons of identifiability, we choose not to vary.

The authors are glad this point has been discussed and sorry the point was not clear enough in the original manuscript. The second assumption in Section 2.2.1 (page 20443) has therefore been reworded in order to highlight the meaning of independence in our setting:

"Separately identifiable emulator inputs

The emulator inputs (the model parameters under investigation) should be separately identifiable. The identifiability of the inputs may not be known before the emulator is built but when there is some prior knowledge of an identifiability issue between parameters then only one or some function of them should be varied. Using separately identifiable inputs also keeps the necessary model runs to a minimum. "

At the top of p20449 you might have found it helpful to reference J.C.Rougier and D.M.H. Sexton (2007), Inference in Ensemble Experiments, Philosophical Transactions of the Royal Society, Series A, 365, 2133-2143.

ACPD

11, C9579–C9582, 2011

where we explicitly use an emulator to experiment with different distributions over the parameters.

This is a good reference for the point being made and has been added to the text.

Finally, we do now have the technology to handle multivariate emulation of functional output, as is done in section 4.2.2. Two papers which describe this are J.C. Rougier, S. Guillas, A. Maute, A.D. Richmond (2009), Expert Knowledge and Multivariate Emulation: The Thermosphere-Ionosphere Electrodynamics General Circulation Model (TIE-GCM), Technometrics, 51(4), 414-424. doi:10.1198/TECH.2009.07123

and

J.C. Rougier (2008), Efficient Emulators for Multivariate Deterministic Functions, Journal of Computational and Graphical Statistics, 17(4), 827-843. doi:10.1198/106186008X384032. It will be interesting in due course to see whether the extra structure that can be incorporated into a multivariate emulator has a benefit in terms of reduced prediction uncertainty.

The authors would be keen in the future to explore the use of multivariate emulation. Multivariate emulation may be useful to help reduce prediction uncertainty in different aspects of aerosol modelling. It may be used to build an emulator for multiple outputs to describe specific aspects of global aerosol distribution determined by multiple model outputs, or to describe the vertical or regional distribution of aerosol by allowing structure between model output in neighbouring grid boxes. Our current work has not allowed for structure between model outputs but we have instead built univariate emulators of individual model outputs. Given the nature of the global aerosol model we would be in an ideal position to compare the univariate emulation study we have completed so far with the multivariate emulation suggested and really test the benefit in terms of reduced prediction uncertainty. A paragraph has been added to Section 5:

"A future direction is multivariate emulation which may be useful to help reduce

ACPD

11, C9579–C9582, 2011

prediction uncertainty in different aspects of aerosol modelling. It may be used to build an emulator for multiple outputs to describe specific aspects of global aerosol distribution determined by multiple model outputs, or to describe the vertical or regional distribution of aerosol by allowing structure between model output in neighbouring grid boxes. Our current work has not allowed for structure between model outputs but we have instead built univariate emulators of individual model outputs. Given the nature of the global aerosol model we would be in an ideal position to compare the univariate emulation study we have completed so far with the multivariate emulation to investigate the reduction in prediction uncertainty."

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 20433, 2011.