Geosci. Model Dev. Discuss., 3, C662–C664, 2011
www.geosci-model-dev-discuss.net/3/C662/2011/

Geoscientific
Model Development
Discussions

Interactive comment on "Automated continuous verification and validation for numerical simulation" *by* P. E. Farrell et al.

Anonymous Referee #2

Received and published: 8 January 2011

The paper provides an introduction to the verification and validation (V&V) of numerical modeling software, and describes in detail the implementation of the automated testing of the Fluidity-ICOM package. My only concern however is that the paper doesn't sufficiently address the particular needs of geoscientific model development. I feel that the paper would be significantly improved if more emphasis were devoted to describing how Fluidity-ICOM is verified in it's role as an ocean general circulation model (GCM). The Stommel test case presented is far from adequate for this purpose.

The comments below reflect my own experience as a developer of an atmospheric GCM. The V&V process that we use is very similar to the one employed at the Met Office Hadley Centre as described in an article by Easterbrook and Johns (2009).

The paper under review begins with the claim that typically the V&V process is seen as

a discrete event performed once after the development is complete. This is certainly not the case for validating GCMs. As pointed out by Easterbrook and Johns, the continuous validation process is really just part of "doing science." But there is some truth to the assertion relative to the code verification part of V&V.

This paper states that the most reliable method of code verification is generally considered to be order of convergence testing, and in instances where no exact solutions exist the only technique suggested for this is the method of manufactured solutions (MMS). It's far from clear how this technique might be applied to a GCM in it's full complexity. A GCM contains dozens of subgridscale physics parameterizations, and they typically involve very complex functional representations. But suppose the forcing provided by the parameterizations could be derived analytically for a desired manufactured solution. There is evidence in a recent paper by Williamson (2008) indicating that there is a lack of convergence with horizontal resolution of the CAM-3 model even when simplified to an "aqua-planet" mode which removes the complications of topography and different surface types from the lower boundary conditions. Williamson then asks the question "Are there reasons to believe simulations if they do not converge?". His answer seems to be that there are scales in the solutions that do converge and can be believed, and others that are responsible for the non-convergence which cannot be believed. The complexity of this situation leads me to wonder whether testing using smooth manufactured solutions would lead to similar conclusions.

The fact that the subgridscale physics parameterizations are such a critical part of GCMs implies that their verification should play a role in the overall V&V process. But this issue is not directly addressed by this paper.

A common practice in code verification for the dynamical core of atmospheric GCMs is to replace the full parameterization suite by an idealized forcing. The analysis focuses on the long term statistical properties of a fully developed general circulation. Convergence properties are evaluated in the context of statistical climate states (see for example Wan et al., 2006). This is a very different approach to studying convergence from the method of manufactured solutions. Is it possible that the method of manufactured solutions might provide a more efficient method of acquiring these results?

In the presentation of the first fluid dynamics test case the statement is made that "MMS provides an easy way of generating analytical solutions against which to verify model code." I think it's an open question whether this statement is true for an important class of geoscientific models, namely GCMs. But I think there is value in exposing the GMD community to these ideas (reviewing this paper was my first exposure to MMS) and so I'd like to recommend that this paper be published, but with modifications to at least acknowledge the uncertainy in their application to GCMs.

References:

Easterbrook, S. M., John, T. C.: Engineering the Software for Understanding Climate Change. Computing in Science & Engineering, 11(6), 64-74, doi:10.1109/MCSE.2009.193, 2009.

Wan, H., Giorgetta, M. A., Bonaventura, L.: Held-Suarez Test with ECHAM5. Reports on Earth System Science, 20, Max Planck Institute for Meteorology in Hamburg, 2006.

Williamson, D. L.: Convergence of aqua-planet simulations with increasing resolution in the Community Atmospheric Model, Version 3. Tellus, 60A, doi:10.1111/j.1600-0870.2008.00339.x, 2008.

Interactive comment on Geosci. Model Dev. Discuss., 3, 1587, 2010.