Geosci. Model Dev. Discuss., 4, C362–C365, 2011
www.geosci-model-dev-discuss.net/4/C362/2011/

Geoscientific
Model Development
Discussions

Interactive comment on "iGen: a program for the automated generation of models and parameterisations" *by* D. F. Tang and S. Dobbie

VTJP Phillips (Referee)

vaughanp@hawaii.edu

Received and published: 14 June 2011

General comments:

This is a seminal paper, establishing a new approach in the problem of parameterizing complex algorithms. It may lead to better statistical representation of sub-gridscale processes that are too small to represent explicitly in large-scale atmospheric models and other models. It may prove useful for financial modeling.

For this reason, I think the authors should invest effort in improving the overall style of the paper. Phrases like "notoriously" or "fear" or "successfully" or "the authors firmly believe" are not really scientific. There needs to be a "Description of iGen" section, with an overview of the general structure of it and perhaps a flow-chart. At present crucial

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

details are in the introduction instead of being in such a description section.

Also, practical details of implementation of the algorithm would be good to include, perhaps in a new section. It would be good to compare the proposed method with existing alternatives for creating cheap algorithms, such as neural networks. Perhaps when there are too many inputs for iGen to cope with, neural networks or some other approach, would be better.

The validation section is fine. It would be good to show how iGen converges as the specified error to be accepted is diminished and the sparse grid resolution is enhanced. It would be useful to see how iGen performs when given a 1D, or 2D lookup table (e.g. for droplet-droplet collision efficiencies). Can it cope with hundreds of if statements in a lookup table ?

A topic to tackle concerns accessibility to the iGen algorithm and reproducibility of the results. Where and when is iGen going to be available to the community, if at all ?

Finally, the limit on the number of input variables for iGen to function properly seems quite problematic. It would be good to comment on how implementation of iGen could circumvent such difficulties, (for example by selective use of iGen for certain parts of a complicated program that are identified as being more tractable with more inputs).

Detailed comments

Page 844, line 9-10: A key issue about the utility of this proposed technique is whether the bounds on the error converge as iGen is altered towarsd a higher resolution. Does the approximate solution tend towards the original exact solution as the specified error is reduced ? It would be good to show this on the validation plots.

A potential problem is the number of variables of the parameterising model and whether the approximate solution has a realistic sensitivity with respect to perturbations of the input variables. Could the approximate solution exhibit unrealistically wide transitions from one state to another in sensitivity tests ?

GMDD

4, C362–C365, 2011

Page 845, Line 6-9: This is an important point and could be developed more in the paper. The ability to simplify equations of a high-resolution model by making approximations is part of the process of understanding a complex system (See Yano's papers that do this for deep convection; see Isaac Held's philosophy in BAMS about the necessity to create analytical simplified equations in order to understand a system). By creating approximate formulae, iGen can advance scientific understanding of the essence of a system.

Page 845, line 11: The description of the overall structure of iGen is not as lucid as it could be. It needs to be stressed that the input to iGen is the source code of the high-resolution model and the output from iGen is more simplified source code, to replace the original source code. A flow-chart or similar diagram would summarize clearly the overall structure of iGen.

The paper needs to be structured better. At this point, the introduction needs to be wrapped up and the description section should start.

I think towards the end there need to be some notes about practical details on implementation of iGen. What programming language must the original source code be in ? Do any intricate parts of the original source code need to be simplified (E.g. by iGen) to prepare it in advance ? What to do if the error bounds are large or if there are too many input variables ? Is there ever advantage to combining use of iGen with neural networks or other techniques for creating simplified algorithms, and how does iGen complement these alternatives ?

Try this:

1 Introduction

2 Description of iGen program for automatic param. 2.1 Overview [with flow chart indicating the very general sequence of steps that iGen takes] 2.2 Symbolic analysis applied by iGen: operations defined 2.3 random numbers 2.4 fixed loops etc

GMDD

4, C362–C365, 2011

3. Validation of iGen

4. Practical Implementation of iGen

5. Discussion

6. Conclusions.

Page 857, line 13: "without fear" is not a scientific phrase in this context; try to use the passive tense of the verb instead of "we", so as to focus on what is being described.

Page 861, line 10: I think the full equation for the scattering cross-section of this droplet size distribution needs to be written.

Interactive comment on Geosci. Model Dev. Discuss., 4, 843, 2011.

GMDD

4, C362–C365, 2011