Geosci. Model Dev. Discuss.,
doi:10.5194/gmd-2016-308-RC1, 2017

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Sensitivity analysis of the meteorological pre-processor MPP-FMI 3.0 using algorithmic differentiation" *by* John Backman et al.

L. Hascoet (Referee)

laurent.hascoet@inria.fr

Received and published: 24 February 2017

The submitted article deals with sensitivity analysis of a meteorological model, provided as a Fortran program. The derivatives used for this analysis are obtained by Source-Transformation Algorithmic Differentiation of the source of the program. The article provides an extensive discussion and physical interpretation of the obtained sensitivities. The discussion eventually leads to some recommendations on the measurements that are usually fed to the meteorological model. The article also provides a few suggestions on future extensions of the sensitivity study, for example by including more meteorological modules in the study.

Before giving my opinion, I'd like to point out that I'm not in a position to judge the

Printer-friendly version

Discussion paper

meteorological or physical aspects of the work. On the other hand, I can give my opinion on the computer science aspects and the use of Algorithmic Differentiation.

Overall, the paper is well organized and well written. The long part (section "Results") on the physical interpretation of the sensitivity results is probably obscure for outsiders but certainly meaningful for specialists. Still, I appreciate the visible effort that was devoted to present these results as clearly as possible. To my eyes, the interest of this paper is its illustration of the use of AD to sensitivity analysis in an Earth Sciences application.

From the viewpoint of AD tools, the paper could give more answers to a few questions, such as:

– Technical data: How does the runtime of the (tangent) differentiated code compare with the runtime of the original/primal code?

– The primal code being relatively short, did someone consider hand coding, and in that case how does the automatic AD code compare with hand-coded derivatives ?

– You mention that AD gives you machine accuracy (compared with divided differences), but the later discussion is based on figures 2,3,4 and probably doesn't need this accuracy all that much. Maybe the "accuracy" argument can be made stronger by pointing out that the choice of the "good" epsilon perturbation for divided differences is difficult and costly, especially when the orders of magnitude of the inputs are very different.

– I understand you selected tangent mode rather that reverse/adjoint mode, as you have 11 independents and 10 dependents. Your argument is slightly weakened by the fact that the results section concentrates only on two dependent outputs instead of 10. Nevertheless, your choice is still ok.

– Still, using the tangent mode, you need to run it 11 times at each data point, as you explain on page 7. I see from the provided files that you didn't use the "vector" tangent

mode, that could save 10 out of the 11 redundant executions of the primal instructions. Why is that ?

– Classically, when people want to compute full Jacobians (admittedly yours are a small enough 10*11) they try to exploit known sparsity of the Jacobian to compute it in a compressed way. Why didn't you do that? Maybe your Jacobian is not sparse? Then you might want to state that.

Other punctual remarks:

– Why was the radiosonde code not considered? Did it pose a problem to the AD tool ?

– You might reword slightly line 49: Tapenade is not the "only": OpenAD also pretty much fits.

– Line 51 is slightly misleading: readers might understand that AD produces the set of differentiated equations of the original math equations. We agree that if we consider the computer program as an alternative, roughly equivalent set of equations, then AD can be presented as producing the derivative equations of those alternative equations.

– On line 172: in fact the derivative instructions are always performed *before* the primal. The reason is quite anecdotical: think of the tangent diff of "y = x*y"

– Your statement on line 324 seems slightly optimistic: with or without AD, studying sensitivities at a large number of input data points is proportional to this number of points, and therefore not cheap. Not being a specialist, I suppose there might be ways to make it cheaper (surrogate models?) but they are clearly outside the scope of your study.

Typos:

Line 92: the the

Line 105: covarince

Line 149: a sequence of

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-308, 2017.

Printer-friendly version

Discussion paper