Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2016-308-AC1, 2017

Geoscientific Model Development Discussions

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

John Backman et al.

john.backman@fmi.fi

Received and published: 30 August 2017

The authors would like to thank the reviewer for the constructive comments on how to improve the manuscript.

General comments

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

Reply: This is a good point and should definitely be included in the revised manuscript. The comparison was added to the revised manuscript as: "The source transformed

computer program was thus used to construct full Jacobian matrices and took just 4.5 times longer to run than the original program."

Comment: 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?

Reply: It would indeed be feasible to hand code the derivative information into the original code and compare with the AD code. Although feasible, hand coding is, however, in the authors' opinion quite a tedious task for a code of this length. Since the present study is not focused on AD development or verification, hand-coding the derivative information was not pursued.

Comment: 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.

Reply: Again, a very good point. In the revised manuscript, these points are discussed as follows:
"The evaluation of finite differences is further complicated if input variables differs by orders of magnitude. By choosing the AD method, the tedious and imprecise evaluation can be avoided."
What is not visible from the figures, but discussed in writing, is that the stability parameter L^{-1} can be very close to zero when the wind speed is high; hence, good accuracy is needed in those cases.

Comment: 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 ?

Reply: This is a valid point. Since the code is not computationally that expensive the "vector" tangent mode was not initially used. In the revised manuscript, the differentiated code (and the wrapper) is done by exploiting the "vector" mode as suggested. The vector tangent mode is 2.4 times faster than the non-vector code in this case.

Comment: 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.

Reply: The Jacobian is not sparse which is why the full Jacobian was constructed for each data point. This is now explicitly stated in the revised manuscript. Furthermore, it was not worth exploiting the sparsity that existed since the code is so quick to run anyway.

Other punctual remarks

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

Reply: It was not left out because of technical complications. The radiosonde data was left out because it does not affect the calculations of friction velocity nor the Obukhov length. The code that deals with radiosonde data is essentially a lookup procedure to find the temperature-inversion height from temperature and relative humidity data and is not interesting from a sensitivity point of view.

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

Reply: OpenAD is now also mentioned as an alternative.

Comment: 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.

Reply: The authors had missed the possibly misleading sentence which was also picked up by the other referee. To avoid confusion a more comprehensive explanation is now given which reads:
"A source transformation tool approaches the differentiation by analysing the source code of a given computer program and generating an augmented source code which contains, in addition to the original operations, instructions that carry out their chain rule differentiated versions. As these differentiated statements accompany each relevant mathematical operation in the source code, they propagate the derivative information across the entire program, providing exact sensitivity information (to machine precision) on how the inputs of the program influence its results."

Comment: 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"

Reply: This is of course true and was changed accordingly.

Comment: 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.

Reply: Yes, the optimism needs to be downplayed. This relates to the earlier comment with the need to exploit sparse matrices to speed things up. The sentence was change

to
“The sensitivities could be analysed for a wide range of input conditions both accurately and effectively.”

Typos:

Comment: Line 92: the the **Reply:** Corrected.

Comment: Line 105: covarince **Reply:** Corrected.

Comment: Line 149: a sequence of **Reply:** Corrected.

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