Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-213-RC1, 2019

Geoscientific Model Development Discussions
Open Access

Interactive comment on "SICOPOLIS-AD v1: an open-source adjoint modeling framework for ice sheet simulation enabled by the algorithmic differentiation tool OpenAD" *by* Liz C. Logan et al.

Laurent Hascoet (Referee)

laurent.hascoet@inria.fr

Received and published: 3 September 2019

General comments:

This article describes a new development in the SICOPOLIS glaciology simulation code, to introduce sensitivity/adjoint/gradient computations. The article describes why adjoint capability is a significant improvement to a simulation code, allowing for sensitivity studies and solution to inverse problems and parameter estimation. The article discusses these new possibilities specifically for glaciology. This new adjoint capability was introduced through the use of an Algorithmic Differentiation tool: OpenAD. The article describes the amount of work that this AD tool required, and the amount of work that it saved. The article also points to a few difficulties where AD tools still require the help of the end-user. Global performance of the adjoint-enabled code is described shortly. The article gives an in-depth discussion and interpretation of the obtained gradients for the glaciology and climate specialist, and points at further exploitation of this adjoint capability as further work. The article also provides some discussion about some observed deviation in the computed gradients.

Please note that I am not able to comment on the glaciology-specific parts of the text, although their general music seems completely reasonable.

The article is well structured and well written. It is easy to read, although some parts are obviously directed at true specialists of glaciology.

Like I write in the specific comments, I am slighty worried by the deviation observed for some gradient values, between adjoint and Divided Differences. I am only partly convinced by the explanation about numerical noise. The text also evokes the cases of non-smoothness of the implemented function. Could this be part of the explanation? I think this part of the discussion might be developed a bit, as some readers may really take it as an argument against AD.

I recommend publication of this article. It describes a solid work on an important code, it is useful as a clear example of what AD adjoints can do, it promotes AD towards the glaciology community, and it seeds for further work.

Specific comments:

P3, L12 : True, the adjoint propagation runs backward in time. More generally it runs backward the original simulation order (which happens here to be forward in time). Maybe it would be useful to stress that?

P7, L1 : Rather than "can be conceived as ...", I would advocate writing that "as soon as a numerical model is implemented as a code, it is in fact translated as ..."

P9, L15 : Are the preprocessor options used (to exclude or include arts) at "compilation time" or at "differentiation time" ? I take it that you mean "differentiation time". Does this imply then that there will be one particular adjoint model of SICOPOLIS for each model configuration. If so, your text presents this as an advantage but you understand some people might consider this as a drawback, not having a unique adjoint SICOPOLIS source at hand. (Here I'm playing the devil's advocate, as I think any source-transformation AD tool will face the same drawback)

P10, L3 : Does this raise the question about why, in the adjoint, some time steps are more stable than others? In other words, why can't one take the same time step sizess than in the forward simulation. That can be an interesting question for a Numerical Analysis specialist (not me...)

P10, L31: I'd replace "sufficently approximates" with "is sufficiently consistent with", because I tend to think that it is divided differences that is an approximation of the other.

P11, Table 1: I would swap rows 7 and 8 for consistency with rows 3 and 4. Deviations on rows 5 to 8 seem surprisingly high. Are they discussed in the text ? I read in P14 L20 that the finite difference chosen is around 5%, which can explain the high deviation. And yes, the explanation in P15 L1 may be right. But does the deviation decrease when the divided difference is smaller e.g. 0.5% instead of 5% ? As it is, a deviation of 57% is still worrying.

P15 L3: The question that comes immediately is how do these times compare with the primal simulation ? It might also be appropriate to describe the checkpointing scheme used in this experiment, on time-stepping: is it multi-level, or binomial, how many checkpoints are used, how many duplicated forward steps. Are these questions left for future work? Oh, I see it is in the appendix, P24 L24. Could you just, in the main text, point out that the appendix mentions that ?

P17 L17: The question of best practices makes me think that you may want to cite the Utke-Hascoet paper on that "Programming language features, usage patterns,..."

(OMS 2016)

Technical corrections:

P2, L25 : "construed"

P4, L13 : "more confident projections" -> "more faithful" ?

P5, L30 : "substantive ... to" -> "substantial ... over" ?

P6, L1: "ever" or "even" ?

P8, L15: why comma after warmer?

P9, L25: "appendix" is repeated

P10, L13: "instantaneously" ? "instantly"

P13, L3: Is it January? Figure 3 caption writes July. Or did I miss something?

P16, L18: "introduces" or "-introduced" ?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-213, 2019.