Review of Manuscript

'When ancient numerical demons meet physics-informed machine learning: adjoint-based gradients for implicit differentiable modeling'

By Y. Song et al.

Dear Editor,

I have reviewed the aforementioned work. My conclusions and comments are as follows:

1. Scope

The article is within the scope of HESS.

2. Summary

In their paper, the authors introduce an adjoint-based method to allow for efficient training of hydrological models via backpropagation, even if they use implicit numerical solvers. At first, the authors give an overview on the current state of hydrological modeling with a focus on the different approaches (conceptual modeling based on physical process understanding, purely data-based such as LSTMs, and hybrid combinations thereof) and on how formulating these models in a differentiable manner is key for their efficient training through backpropagation. They also point out that typical physics-based conceptual hydrological models rely on explicit, non-iterative numerical schemes, which potentially introduces numerical error and hampers efficient model training and unambiguous parameter identification. In this context, the main goals of the paper are therefore to **i)** introduce the adjoint-based method ('discretize-then-optimize'), **ii)** compare and evaluate models with implicit and explicit numerical schemes, and **iii)** to assess whether local and regional parameter distributions differ between the two schemes.

The authors use the well-known CAMELS-US data set to set up and compare various models: An LSTM-only and the SAC-SMA model as benchmarks, and several variants of a hybrid LSTM-HBV model, where the LSTM provides the parameterization for the HBV model. The variants comprise the four possible combinations of two numerical solvers (explicit and implicit), and two model architectures (without/with capillary rise).

With respect to i) and ii), the authors compare the LSTM-only and the LSTM-HBV hybrids with explicit and implicit numerical schemes for a range of performance measures and conclude that the implicit model improves streamflow simulation (lines 412 pp) and that this can be attributed to the numerical errors of the explicit numerical scheme (lines 413-415).

The authors then discuss the effects of the changed HBV model structure (addition of capillary rise) to help the model produce (near-)zero base flow in accordance with observations, and conclude that the added process is helpful in low-flows, but comes with some deterioration for high-flows.

With respect to iii), the authors compare spatial parameter patterns of the HBV hybrids with different numerical schemes, and conclude that the patterns largely agree, indicating the robustness of the involved parameter regionalization scheme.

3. Evaluation

First of all, I acknowledge the work of the authors to further merge data- and physics-based approaches to modeling by showing how models using implicit solvers can be integrated into a typical machine-learning workflow with backpropagation at its core. This is a valuable contribution to the hydrological modeling sciences, but unfortunately the authors do not convincingly prove in their paper the immediate benefit thereof, and they obscure their point by adding aspects to the study that are not related to the main message. I will explain this in the following:

A) The authors are correct that mainstream conceptual hydrological models like HBV have traditionally - and often without much reflection – been used with simple explicit numerical schemes, and a pre-set order of process execution, and that this may cause substantial problems (see Clark and Kavetski, 2010 as cited by the authors), and that implicit schemes can solve these problems. Therefore, in this manuscript, in addition to the description of how to include implicit schemes in ML-workflows, I was expecting a demonstration of how this actually solves a problem. That is, showing that for a particular hydrological modeling task (here: modeling streamflow in daily resolution of the CAMELS-US basins) i) the standard explicit scheme introduces problems and ii) that an implicit scheme solves them. The authors mention this point in the paper (line 324-326), but unfortunately do not address it. For example, one could operate HBV models for some of the CAMELS catchments with various execution orders and extremely fine-grained time-stepping, thus effectively removing the detrimental effect of the explicit scheme, and then compare to a standard time stepping and execution order, and to a model using an implicit scheme. The authors conclude in their study that the (small) model improvements between the HVB-hybrid variants using explicit and implicit schemes are due to problems introduced by the explicit scheme (lines 412-415), but because they do not provide a proof for a cause, the conclusion based on an effect is not convincing.
In this context, it might also be interesting to analyze if decreasing negative effects of explicit schemes by higher time stepping (or other changes to the model computational setup) might be more efficient than shifting to implicit schemes. The authors mention that computational costs for the latter increased by a factor of 5-10 (line 581). Increasing the time stepping of the explicit scheme from daily to 6 hours would only mean a factor of 4, but would already resolve diurnal cycles, which might be relevant additional information for the model.

B) Motivated by problems of the HBV model to simulate (near-)zero base flow during extended dry spells, the authors integrate a detailed study about the effect of adding an additional capillary rise process to the HBV model. This is a valid question and analysis, but it does not at all support the main argument of the paper about how and why implicit schemes can be integrated into modern hybrid modeling workflows. I therefore suggest presenting this analysis in another paper, and removing it from this one.
In this context, it is interesting that the authors provide a range of possible adjustments to the HBV model to help it achieve (near-)zero flow (strategies 1-5 in lines 258-261; and lines 479-482). These adjustments touch very different physical subdomains and processes of the model, and one may wonder about the limitations of a supposed key advantage of physics-based models – realism and interpretability – if it remains mainly up to the user's preference which one is chosen. In particular, I wonder why capillary rise from the lower subsurface should bypass the upper surface and directly connect to the surface soil moisture storage, and why the authors chose it this way.

Based on the above points, my overall recommendation is that the key topic of the paper is worth publication, but also that the required changes will require time. Therefore I recommend rejecting the paper in present form, but strongly encourage a resubmission.

Yours sincerely,

Uwe Ehret