

Geosci. Model Dev. Discuss., referee comment RC2 https://doi.org/10.5194/gmd-2021-402-RC2, 2022 © Author(s) 2022. This work is distributed under the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-402

Anonymous Referee #2

Referee comment on "Conservation laws in a neural network architecture: enforcing the atom balance of a Julia-based photochemical model (v0.2.0)" by Patrick Obin Sturm and Anthony S. Wexler, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-402-RC2, 2022

The authors consider neural network surrogate models with physics constraints to enforce conservation laws, they demonstrate the effectiveness of their method using a photochemistry model. Experiment results show perfect atom conservation. The presentation is clear. It is an important research topic to enforce conservation laws (e.g., mass conservation) in neural network surrogates.

Some comments on the paper:

The authors should refer to and discuss the paper: Beucler et al. 2019 "Achieving Conservation of Energy in Neural Network Emulators for Climate Modeling" (arxiv.org/abs/1906.06622). In that paper, two options to implement physics constraints were discussed: 1. Constraining the loss function; 2. Constraining the architecture. The approach of this paper belongs to option 2 (figure 2, Beucler et al. 2019). It presents a detailed derivation and implementation of the constrained layer with fixed weights for a photochemical model.

Line 105, it is stated that a key difference with the work of Beucler et al. 2021 is "1. Our entire output is calculated under the constraints, rather than only a portion of the output." In my opinion, this statement is not correct. In the general architecture (Beucler et al.), parts of the output can be constrained and some parts of the output don't need to be constrained. It's a choice, to constrain the entire output is just a particular case.

The authors compared their constrained NN with a simple one-layer NN (naïve NN). The constrained NN has two layers (though the 2nd layer has fixed weights). A two-layers NN could (possibly) improve the performance and it is also more comparable to their constrained NN, it would be fair to compare the constrained NN with a 2-layers naïve NN.

Some minor textual comments:

- (1), use a different symbol for the function C (C already refers to concentrations);
- (2), use a different symbol for function S (the 2nd S)
 Line 221, 1.2 million samples should be: 0.12 million samples
- Line 232, "126 days with continuous observations", data in each day are generated/initialized separately (see Line 220), so they are not continuous?
- Line 240, Svector à S vector