Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-53-RC1, 2019

Geoscientific Model Development Discussions

Interactive comment on "Weather and climate forecasting with neural networks: using GCMs with different complexity as study-ground" *by* Sebastian Scher and Gabriele Messori

Peter Düben (Referee)

peter.dueben@ecmwf.int

Received and published: 21 March 2019

This is a very interesting paper that provides an honest presentation of new results on the use of neural networks to learn the equations of motion of the atmosphere. The paper is relevant for GMD and should be published. However, a minor revision that is addressing the comments below could improve the paper.

- As we argue in Dueben and Bauer 2018, I am very surprised that you get away with 1-day timesteps. I would guess that a T21 model with a 1-day timestep would be un-stable when using explicit time-stepping schemes (maybe I am wrong?)

  and it is hard to believe that a neural network is learning something like an implicit scheme that would allow for larger timesteps. Can you provide the timestep that is used in Plasim for comparison? If this is much smaller, can you comment why you think that the neural network model may get away with this, in particular towards the pole? (I assume that you are using a regular Gaussian grid where grid-spacing will become smaller towards the pole).

- I am also (positively) surprised that you are having no problems with loss of stability when using neural networks while other papers report problems when using neural networks to represent physical systems. Can you speculate why this is? Maybe due to the convolutional layers? Can you diagnose the change of global energy in climate simulations?

- When using convolutional networks, will the stencil of gridpoints around a gridpoint in the neural network have exactly the same weights for all gridpoints? Or will the weights be changing for each gridpoint? If they are the same, how can you justify that gridpoints towards the poles (that will have a very different resolution for the stencil of surrounding points) use the same weights as for points at the equator?

- The paper would benefit if you would provide more discussion how to improve neural network configurations in future studies at the end of the conclusion (speculations are welcome).

- Figure 4: Why are the PUMA results not symmetric in zonal direction?

Really minor:

Figure 2: The figure should be made bigger to improve visibility.

Page 2: A "local model" as suggested in Dueben and Bauer 2018 would increase the amount of training data significantly since you would train for 40 (or 100) years of data

for N grid points. Maybe worth mentioning? But only if you think this would fit.

Page 2: "...ambitioun has shifted from..." not really. This is rather a different application.

Page 4: "(...see Scher(2018)." add ")"

Page 6: Why zg500 (ta800) and not z500 (t800)?

Page 6: "has higher error at longer lead times" At long lead times and close to zero correlation a bias in the model can cause interesting behaviours in rmse plots. It can also reduce the error.

Page 8: Can you comment how important it was to include the day of year as input?

Figure 5: I do not see the shading in my printout.

Page 11: "to to"

Page 12: Personally, I am very sceptical that the use for different networks for different months would produce reasonable results.

Please feel free to contact me if some of the comments are unclear.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-53, 2019.