Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2020-38-AC2, 2020

Nonlinear Processes in Geophysics
Discussions
Open Access

Interactive comment on "Training a convolutional neural network to conserve mass in data assimilation" *by* Yvonne Ruckstuhl et al.

Yvonne Ruckstuhl et al.

yvonne.ruckstuhl@lmu.de

Received and published: 15 December 2020

article xcolor

Review response

December 15, 2020

This is a well-written manuscript with very interesting results. My major comment is that this manuscript is rather short and that it could be extended to give more insightful results.
Thank you for taking the time to thoroughly read our manuscript and for the positive feedback. We have taken your comments into account and have adjusted the manuscript accordingly, see below.

Major comments:

1. I would be in favour to see how the conclusions change depending of the grid size and the ensemble size.
   The relative improvement of the QPEns over EnKF for different DA settings (including ensemble size) has been covered in Ruckstuhl and Janjic (2018). We added a sentence at the end of section 2.2: "We refer to Ruckstuhl and Janjic (2018) for a comparison of the performance of the EnKF and the QPEns as a function of ensemble size for different localisation radii, assimilation windows and observation coverage."

We are confident that as long as the CNN can remove the bias in h, the CNN can match the performance of QPEns for any DA setting. One could then try to compare the differences in training process of the CNN's (does one setting require more data than the other?). However, a clean comparison among the different settings would require rigorous tuning of the architecture, the amount of training data needed, and the training process of the CNN. And this tuning is very tricky because we are interested in the performance of the CNN in DA context, not the value of the loss function. Since we are working in a highly idealised setup, we want to be economical with time spent on fine tuning the CNN's. We therefore feel that we have exploited the modified shallow water model on this specific topic. Any further experiments should be done on more complex models. That being said, we did investigate in addition the trade off between mass and RMSE as you suggested in the next point and performed additional experiments that test the relation of the kernel size of the CNN to localisation.

2. There is a trade-off between mass conservation and low RMSE for u and h. What happens if in the experiments with the additional penalty term for mass conservation instead of a linear activation function for u and h, the "relu" activation function is used for both u and h as well as for r? Is the trade-off smaller then?
The reason we use the relu function for r is that the rain cannot be negative. This does not hold for u and h, so using the relu function for these variables is not an option. We did perform some additional experiments to investigate the trade off between mass conservation and RMSE, which is now summarized in Figure 4.

3. Authors remove the climatological mean from u and h. What happens if the climatological mean is not subtracted? Is the bias too high for the methods to handle?
We want to clarify that we only subtract the mean to make the problem better conditioned for the training process. Since the difference between input and output data for the 3 variables differ in at most 1 order of magnitude, we do not expect huge problems if the mean is not subtracted. However, as far as we know, there

is no disadvantage to normalizing the training data, which is why we have not tried training the CNN on the raw data.

Minor comments:

1. l.8: The last sentence of the abstract is rather vague. Please elaborate.
We have removed this sentence.

2. l.146: Does the loss function JËĘ γ account for the mass twice: in J and in the penalty term?
J is the RMSE averaged over the 3 variables. This means that J accounts for the mass error indirectly (as the RMSE goes to zero, the mass error also goes to zero). The penalty term directly accounts for mass errors by first averaging the h field over the 250 grid points for y^p and y separately, and then squaring the difference.

3. Please change γ to something else, since it is already reserved for the gravity wave speed.
Yes, you are right. We changed it to η.

4. Why is the penalty term chosen in such a way, namely L1 norm and not L2 as in J?
Note that J takes the norm of a vector of size 250, whereas the penalty term takes the norm of a scalar (namely the difference of the spatial mean of h). Therefore the L1 and L2 norm are equivalent for the penalty term.

5. If I look at Fig. 2(a) I see that NN is performing slightly better than QPEns. Is there an explanation for that?
The QPEns is also not perfect, so it is possible that the CNN performs better. It is indeed then interesting to speculate why that is. Most of the last paragraph before the conclusion is dedicated to this (from line 174 to 184 in the new manuscript).

6. l.92: “For the EnKF negative values for rain are set to zero if they occur”. This is the variable r, if I understand correctly. However, if I look at Figure 7, I see negative values of r for EnKF. Could authors please explain?
   The fields are shown before negative values are set to zero. We have clarified this in the caption

7. A table consistent of wall-clock time for different methods would be insightful for the computational cost gain.
   The costs of applying the NN are negligible with respect to the costs of the EnKF, as mentioned in line 50 of the new manuscript. So it is about the difference in computational costs between the EnKF and QPEns. Since we are working with a cheap model, no effort has been made in the implementation of the algorithms to make them computationally efficient. Therefore wall-clock times may be misleading. However we agree this is an important point and we actually have a paper under review that thoroughly discusses the computational costs of the QPEns. We therefore added a sentence in the introduction: "For a detailed discussion on the computational costs of the QPEns we refer to Janjic et al. (under review)".

8. I do not want to be self-promoted but authors could have a look at Dubinkina 2018 and decide if they would like to refer to it in their manuscript. Thanks for mentioning this paper. We added now a reference to this manuscript.