Nonlin. Processes Geophys. Discuss.,
https://doi.org/10.5194/npg-2019-49-RC1, 2019
© Author(s) 2019. This work is distributed under
the Creative Commons Attribution 4.0 License.

Nonlinear Processes in Geophysics
Discussions

Interactive comment on "Remember the past: A comparison of time-adaptive training schemes for non-homogeneous regression" *by* Moritz N. Lang et al.

Anonymous Referee #1

Received and published: 8 October 2019

This manuscript compares the effect of different schemes to compose training data for statistical post-processing methods (here: non-homogeneous regression) on the performance of the resulting forecasts. It is well written and highly relevant to operational forecasting where availability of reforecast data may be limited and the consequences of changes in the NWP model on forecast calibration must be understood in order to decide whether forecasts from an older NWP model version can be used to fit the parameters defining the post-processing model. This last point is the only one where I feel the manuscript could benefit from a more detailed discussion. Specifically: The CRPS skill scores in Fig. 5 h) suggest that the regularization scheme struggles with the adjustment to the NWP model upgrade and to the annual cycle, but also the SW

plus and the smooth model have an overall neutral effect on skill even though these schemes increase the training sample size significantly. It would be interesting to better understand the causes of this result. Figure 3 gives some good idea about the problems with the regularization scheme (parameters adjust very slowly to changes) but it is not ideal to illustrate problems with 'SW plus' and the 'Smooth model' since no NWP model upgrade happens in data set A. Wouldn't it be better to use data set B for this figure, where we can expect some adjustment during the first days/weeks of the validation period? Also, is Innsbruck the best location to illustrate the effect of a NWP model upgrade? As 'best' alpine location in this context I would consider the one that is most strongly impacted by the horizontal resolution change (this could be studied by considering changes in biases in the raw ensemble forecasts) in the ECMWF model and therefore presents a worst case scenario in terms of adjustment to a NWP model upgrade. I would encourage the authors to provide some more discussion along these lines, since NWP model upgrades have been the main argument to justify the need for reforecasts, and I am not aware of any previous study that looks at the effect of NWP model upgrades on the performance of post-processed ensemble forecasts in a quantitative way.

Minor comments:

244-245: While it's possible (even likely) that a larger slope coefficient is due to higher skill of the EPS temperature forecasts, one cannot be sure if at least to some extent the larger slope coefficient is due to an amplitude bias of the raw ensemble forecasts, i.e. the ensemble underpredicts high temperatures and overpredicts low temperatures, and increasing the slope coefficient compensates for that.

Interactive comment on Nonlin. Processes Geophys. Discuss., https://doi.org/10.5194/npg-2019-49, 2019.