Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-122-RC1, 2022

Comment on gmd-2022-122

Anonymous Referee #1

Referee comment on "Representing chemical history in ozone time-series predictions – a model experiment study building on the MLAir (v1.5) deep learning framework" by Felix Kleinert et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-122-RC1, 2022

The paper is well written. The authors applied an advanced deep learning architecture and conducted a detailed and interesting sensitivity study on the predictability of ozone using the U-net+LSTM. Analyses on the forecast lead time, the added value of upstream history, and the concept drift could be useful for future studies. The inclusion of reference models enhanced the validity of the results. I recommend the publication of the paper with minor revisions.

General comments:

- At a resolution of ~12 km, could you comment on the potential impact of sampling errors, when your model is applied to real observations in the future? How many sites do you have in each grid box?
- Could you discuss the ozone chemical regime over Germany and the impact of VOC on ozone predictability?

Specific points:

L161: Inconsistent font of “DataHandler”

L278-280: Is it possible that the training/test performances are different because your NN model is slightly overfitted towards regions with more pseudo-stations?

Figure 9: Could you comment on the degraded performance over the coastal regions? Do these coastal sites have something to do with the results of skill scores in Figure 10? For example, the north wind driven by the sea breeze circulation brings less useful information.

Figure 11-14: How would the learned non-linearities compare with the WRF-Chem CTM? Could you comment on that?

L330: I think your NN3s model slightly outperforms OLS at most of the pseudo-sites?

L340: I see measurement uncertainties are not utilized for now. Can you comment on potential utilization of measurement uncertainties in the future (e.g., via Bayesian NN)?