Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., author comment AC1
https://doi.org/10.5194/gmd-2021-420-AC1, 2022

Reply on RC1

Weichao Han et al.

Author comment on "A comparative analysis for a deep learning model (hyDL-CO v1.0) and Kalman filter to predict CO concentrations in China" by Weichao Han et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-420-AC1, 2022

The paper compared the performances of a deep learning model and a chemical transport model with data assimilation in predicting surface carbon monoxide in China. It is of high scientific significance, as the results have implications for understanding the roles of machine learning and numerical methods in the prediction of air pollution. The manuscript is well-written. I recommend the publication of the paper with minor revision. I have the following comments that need the authors to address.

Answer: Thank you for the comments! The manuscript has been revised based on the comments.

Question: (1) How does the deep learning model, hyDL-CO, developed in this study perform when compared with other deep learning models?

Answer: A new paragraph (at the beginning of Section 2.3) and more discussions were provided in the revision: "Our hyDL-CO model is a modified version of the U-net model used in He et al. (2022), where the model shows a superior capability in predicting surface summertime O3 in North America. The U-net architecture is a variant of autoencoder and was originally proposed for biomedical segmentation applications. In the first U-net paper, Ronneberger et al. (2015) conducted three experiments and showed that the U-net model outperforms other DL models. Since the proposal of U-net, it has become one of the most popular choices in the DL community and is compared with other ML models in many studies. For example, Korznikov et al. (2021) used several ML models for tree recognition using satellite images and the U-net model shows the highest accuracy. Ravuri et al. (2021) used U-net as a baseline model and compared against their Generative Adversarial Network (GAN) in precipitation nowcasting. Andersson et al. (2021) showed that their IceNet, which is an ensemble of similar U-Net networks, has outstanding performance in seasonal forecasts of Arctic Sea ice".

Question: (2) I recommend the authors use more indices to evaluate the performances of the models. The Figures and Tables show that the Kalman Filter performs better than the deep learning model in the test period of 2019-2020. I think this is important that need to be mentioned in the abstract.

Answer: Thank you for this suggestion! Root Mean Square Error (RMSE) was added in Table 2. The following sentence was added in the abstract: "we find a weaker prediction

capability of DL model than KF in the test period".

Question: (3) How do models perform in simulating the spatial variability of surface carbon monoxide?

Answer: As shown in the revised manuscript: "Although we find broadly good agreements in the spatial distributions between predicted CO concentrations by DL and KF and MEE CO observations, there is still a noticeable discrepancy. The DL model suggests the highest CO concentrations in the Shanxi province, by more than 1200 ppb, and background CO concentrations by about 400 ppb over remote areas. By contrast, the CO concentrations in the KF (Fig. 5C-D; Fig. 6C-D) are lower, and the highest CO concentrations are found in NCP rather than Shanxi province". "because most MEE stations are urban sites, the good agreement between DL model and MEE CO observations may not be able to ensure the accuracy of predicted CO concentrations over remote rural areas, as well as the high CO concentrations over mountain areas around urban basins in the Shanxi province".

Question: (4) Probably, in the title, 'A comparative analysis for a deep learning model …' is more appropriate, as only one deep learning model is used in this study.

Answer: The title has been changed.