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.

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

(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.

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

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