EGUsphere, author comment AC1
https://doi.org/10.5194/egusphere-2022-578-AC1, 2022

Reply on CC1

Bo Li et al.

Author comment on "A deep learning approach to increase the value of satellite data for PM2.5 monitoring in China" by Bo Li et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-578-AC1, 2022

Thank you very much for your attention.

1·

Data filtering meets the following conditions:
1.Null value
2.Constant value for more than 6 hours
3.The series were transformed into z scores. The points in the transformed time series meeting one of the three conditions were then rejected and removed from the original time series (1, 2):
(1). having an absolute z score larger than 4 (|zt| > 4),
(2). the increment from the previous value being larger than 9 (zt − zt − 1 > 9),
(3). the ratio of the value to its centred moving average of order 3 (MA3) being larger than 2 (zt/MA3(zt) > 2)
Reference:
M. A. Barrero, J. A. Orza, M. Cabello, L. Canton, Categorisation of air quality monitoring stations by evaluation of PM(10) variability. Sci Total Environ 524-525, 225-236 (2015).

First, we perform a broad filtering scheme, and we only remove null data and constant value for more than 6 hours.
Later, for comparison, we implemented a stricter screening scheme, and the data that met one of the third conditions were eliminated (note that the stricter data screening scheme was only used in the comparison test here, and the rest were all relaxed data screening schemes). The proportion of abnormal data was about 15%. The results are shown in Table S10.S11.

2·

learning rate, Optimization function, batch size all affect the training efficiency of the model. The learning rate has a significant impact on the training efficiency. The larger the batch size, the more memory it occupies (increase time), and the faster the training speed (decrease time). Considering the equipment, we select batch size as 4.

A larger learning rate has a faster training efficiency at the beginning, but with the optimization of the model, it is more and more difficult to train. At this time, it is necessary to reduce the learning rate. The setting of the learning rate is as in equation

(2).

3·

Different optimizers have little effect on the final result, and mainly affect the training efficiency. We find that SGD has faster training efficiency.