Anonymous Referee #3

This manuscript describes results of a regional air quality model evaluation over China using the new ground-based NO_2 network that was recently installed. The paper is of scientific importance because this is the first significant publication of data from this network. The model was run with both bottom-up and top-down emission inventories.

Response: We thank Referee #3 for the encouraging comments. All comments and suggestions have been considered carefully and well addressed below.

I have two major comments concerning the paper:

1) The paper states that the NO_2 monitoring stations are located at least 50 m from stationary sources and at least 10 to 100 m from roadways. I don't think these restrictions are stringent enough to use stations this close to sources to evaluate a model at quarter-degree horizontal resolution (or even a few kilometer resolution!). If information is available concerning the distances of each station from these types of sources, the stations should be screened to use only those a more substantial distance away from the sources. Model vs. observation result may change significantly as a result of such screening.

Response: We agree that stations located far away from emission sources are more suitable for model validations. However, it is pity that the information on distances of each station from different types of sources is not available to us. The general description about distance in the manuscript is derived from the placement criteria of urban assessing stations laid down in the legislation (MEP, 2013).

The ground-based air quality monitoring network of MEP in China is a valuable database to validate model simulations, however it lacks provision of detailed information about the stations, e.g., the information on distances mentioned by the review. This motivates us to explore a reasonable scheme to classify MEP in-situ stations over China, aiming to select measurements that better represent the mean of a grid box of the simulations. The category of "main sample" selected by the scheme is expected to better represent the grid-mean.

2) The paper concludes that the satellite-based DECSO emissions are too low, and the MIX emissions are too high. But, applying the MIX total emission with the DECSO spatial

distribution yielded better results. The conclusions of the paper fail to address some "bigger picture" questions. What do these results imply about emissions derived from satellite data? Are the magnitudes to be trusted? What about air quality trends from satellite data? Are they meaningful? The conclusions need to be augmented to include more general implications of the results of this study.

Response: We thank you for the suggestions. We have included the implications of the results of this study in the conclusion section of the revised manuscript, as follows:

"Nevertheless, the good performance of the satellite-derived emission inventory, in particular the spatial distribution of emissions, has been confirmed by the widespread in situ measurements over China for the first time in this study. The magnitude of satellite-derived emissions show slightly negative bias by taking the negative representativity offset of in-situ measurements into account, which is attributed to biases in the OMI tropospheric NO_2 column densities, or representation errors introduced by the projection of the CTM onto the measured NO_2 satellite footprint. In addition, satellite-derived NO_x emissions succeed to detect the emission trend for the period of 2010–2015, which is consistent with that in bottom-up emissions (Liu et al., 2016a; van der A et al., 2017)."

Minor comments:

p. 3, lines 22- 23: Annual mean simulated surface NO_2 is compared....

Response: Thanks. We have corrected it in the revised manuscript.

p. 5, line 20: what concentration is the grade II air quality standard for NO_2?

Response: The annual mean concentration is 40 μg/m^3 in the grade II air quality standard for NO_2. We have added it in in the revised manuscript.

p. 6, lines 29-30: the correction factor is largest (ideal value of 1.0) over polluted regions, where NO_2 is a larger fraction....

Response: Thanks. We have corrected it in the revised manuscript.

p. 7 line 1: Hourly correction factors derived from CHIMERE.... Are the hourly factors for each hour of each day during the year, or are they means for each hour of the day computed over a season?

Response: They are the hourly factors for each hour of each day during the year. We have clarified this in the revised manuscript.

p. 8, lines 32 - 33: "negative representivity" may be able to reduce this by being more restrictive on the stations used with regard to proximity of sources.

Response: We agree that the negative bias may be reduced by using stations located more far away from emission sources. However, as mentioned previously, the information on distances of each station from different types of sources is not available to us. We have tried our best to select the measurements with better representativity of the grid-mean in this study.

p. 9, lines 12 -13: can you give any reasons why this is the case?

Response: The MIX inventory downscales local emissions from regional totals and distribute them to grid cells using spatial proxies (e.g., population density and GDP). However, the spatial proxies may not match the locations of the individual emitting sources, especially for industrial plants located far away from urban centres that tend to have a larger population density and GDP (Liu et al., 2017). Such a decoupling will result in an overestimation of emissions over urban areas, which has been proven by the comparison of proxy-based regional inventory with high-resolution urban inventories developed from the extensive use of information of individual emitting sources (Zheng et al., 2017). We have clarified this in the revised manuscript.

p. 11, line 27: “But the modelled NO_2 is generally low compared to ground measurements.” This is true for DECSO, but not for MIX.

Response: Thanks for pointing out this. We planned to state that the mean of a grid box of the simulations is expected to be lower than in-situ measurements, due to a combination of preferential placement of monitors in polluted locations and the limitation of model resolution to resolve large NO_2 gradients over urban areas. However, we agree that the sentence is confusing and a repeat of the following sentence. We deleted it in the revised manuscript.

p. 12, lines 9 - 10: Did you try a simulation with the MIX spatial distribution, but scaled to the DECSO total emission? Given the greater spatial correlation with MIX, this might be worthwhile.

Response: The correlation coefficient is slightly higher in the MIX based simulations as compared to the DECSO simulations. However, it would be rash to conclude that the spatial distribution of MIX emissions is better based on the minor difference in the correlation coefficient (0.85 vs 0.73), ignoring the significant positive bias in the MIX based simulations. On the contrary, the spatial distribution of bottom-up emissions over China has been reported to be highly uncertain and may result in an overestimation of emissions over urban areas (e.g., Zheng et al., 2017; Liu et al., 2017). We have clarified this in the revised manuscript, as follows:

"On the other hand, we also show that the correlation coefficient of the simulated NO_2 concentrations versus the in situ measurements is slightly higher in the MIX based simulations as compared to the DECSO simulations. However, this does not necessarily contradict the findings that the spatial distribution of NO_x emissions is more reasonable in DECSO, considering the difference in correlation coefficient is minor but the bias in the MIX based simulations is significant."

The goal of developing a scaled inventory is to better compare the spatial distributions of the two inventories and identify the sensitivity of simulated NO_2 on spatial distributions of NO_x emissions. Comparing simulations using the inventory with the DECSO spatial distribution, but scaled to the MIX total emission, to those using MIX and DECSO has reached this goal. The comparison confirmed the reported bias in the spatial distribution of bottom-up emissions and suggested a better spatial distribution in DECSO. Meanwhile, DECSO emissions are reported to be underestimated over northern part of China (Ding et al., 2017). Thus we did not consider to construct an inventory based on MIX spatial distribution and DECSO magnitude, as we already know that the total DECSO emissions are underestimated and the spatial distribution of MIX is not very accurate.

Reference

Ding, J., Miyazaki, K., van der A, R. J., Mijling, B., Kurokawa, J. I., Cho, S., Janssens-Maenhout, G., Zhang, Q., Liu, F., and Levelt, P. F.: Intercomparison of NO_x emission inventories over East Asia, Atmos. Chem. Phys., 17, 10125–10141, doi: 10.5194/acp-17-10125-2017, 2017.

Liu, F., Zhang, Q., Ronald, J. v. d. A., Zheng, B., Tong, D., Yan, L., Zheng, Y., and He, K.: Recent reduction in NOx emissions over China: synthesis of satellite observations and emission inventories, Environ. Res. Lett., 11, 114002, 2016.

Liu, F., Beirle, S., Zhang, Q., van der A, R. J., Zheng, B., Tong, D., and He, K.: NOx emission trends over Chinese cities estimated from OMI observations during 2005 to 2015, Atmos. Chem. Phys., 17, 9261–9275, doi: 10.5194/acp-17-9261-2017, 2017.

Ministry of Environmental Protection of the People's Republic of China (MEP): Technical regulation for selection of ambient air quality monitoring stations (on trial), available at http://kjs.mep.gov.cn/hjbhbz/bzwb/dqhjbh/jcgfffbz/201309/t20130925_260810.htm (last access: 1 May 2017), 2013 (in Chinese).

van der A, R. J., Mijling, B., Ding, J., Koukouli, M. E., Liu, F., Li, Q., Mao, H., and Theys, N.: Cleaning up the air: effectiveness of air quality policy for SO_2 and NO_x emissions in China, Atmos. Chem. Phys., 17, 1775–1789, doi: 10.5194/acp-17-1775-2017, 2017.

Zheng, B., Zhang, Q., Tong, D., Chen, C., Hong, C., Li, M., Geng, G., Lei, Y., Huo, H., and He, K.: Resolution dependence of uncertainties in gridded emission inventories: a case study in Hebei, China, Atmos. Chem. Phys., 17, 921–933, doi: 10.5194/acp-17-921-2017, 2017.