Dear Editor,

We appreciate the prompt reviews and would like to thank the reviewer for insightful comments and suggestions on our manuscript entitled "Contributions of meteorology and anthropogenic emissions to the trends in winter $PM_{2.5}$ in eastern China 2013–2018" (MS No.: acp-2022-304). We have carefully considered all comments and suggestions. Listed below are our point-by-point responses to all comments and suggestions of this reviewer (Reviewer's points in black, our responses in blue).

Anonymous Referee #3

This work proposes a different method for the MLR analysis of $PM_{2.5}$. Based on the new interpretation and the comparison with previous studies, the MLR results among different studies were found to be more consistent. In addition, the authors also pointed out that the relationship constrained by long-term data is more reliable. Overall, this is an interesting study and it provides some useful information for other researchers when choosing MLR for air quality trend analysis. However, more explanations, especially for the methodology, are still needed.

Response:

We appreciate the insightful comments and suggestions. More explanations have been added in our revised manuscript, especially in the methodology section.

Specific comments

(1) Line 60, the resolution of the PRD emission inventory is three degrees, which is rather coarse.

Response:

The emission inventory of PRD (PRD-EI) is adopted from Huang et al. (2021) and Zhong et al. (2018). Although the resolution of the PRD-EI is coarser than other two inventories, we can only get the emission information of year 2018 and 2019 from PRD-

EI. In addition, in our study, we mainly focus on the long-term trend and interannual variation in the annual total emissions in each region, which is independent of the resolution of emission inventories.

Figure R1 shows the temporal variation of three emission inventories. They have similar variation for the overlapping period.

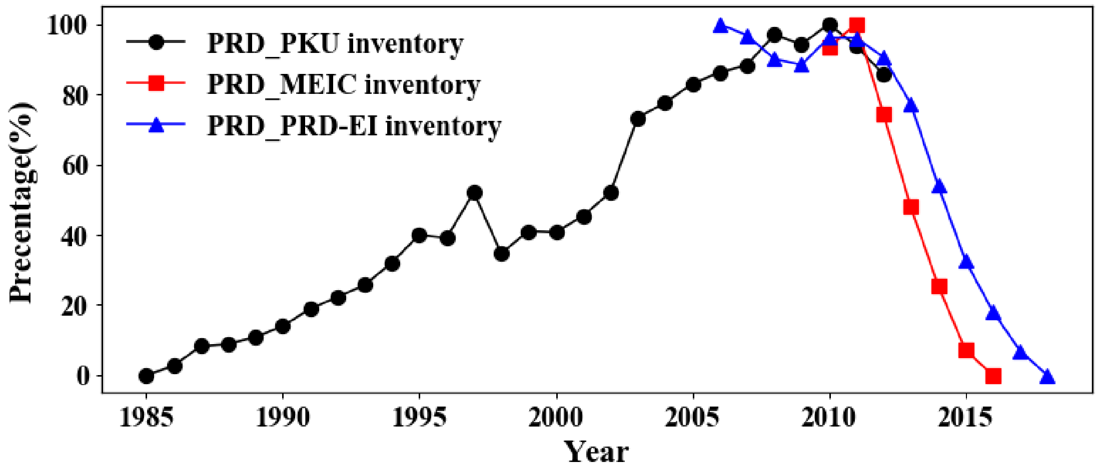

Figure R1. PKU emissions inventory for winter 1985–2012, MEIC emissions inventory for winter 2010–2016 and PRD-EI emissions inventory for winter 2006–2018 for PRD. The raw data is normalized by the difference of the maximum value and minimum value.

(2) This work mainly focuses on the PRD, YRD and Jing-Jin-Ji regions. For YRD and Jing-Jin-Ji regions, the authors combined the MEIC and PKU emission inventories to do the analysis. While for PRD, they combined PKU and PRD-EI to do the scaling. MEIC and PRD-EI are different emission inventories and the methods that used to derive these two emission inventories should be not consistent. Based on the literatures, the MEIC emission inventory should have already covered the PRD region, why not also using the MEIC emission inventory to analyze the PRD region?

Response:

The MEIC inventory does also cover the PRD region, but the time span of MEIC inventory is 2010–2017. The time span of the PRD-EI inventory and PKU inventory is

2006–2019 and 1960–2014, repetitively. Therefore, we combined PRD-EI and PKU inventories in PRD for the winters of 1985–2018.

(3) Please label the scaling factor and Ei equations.

Response:

Thanks, we have labeled the scaling factor and E_i equations in our revised manuscript.

(4) Line 71, please use data or reference to support this assumption.

Response:

Thanks, we have added Figure S5 in the revised Supplementary Material and revised the Line 71 statement to make it clearer as "Since the ratios of annual emission inventory in PRD to those of YRD and BTH are not expected to change significantly in one or two years (Figure S5)"

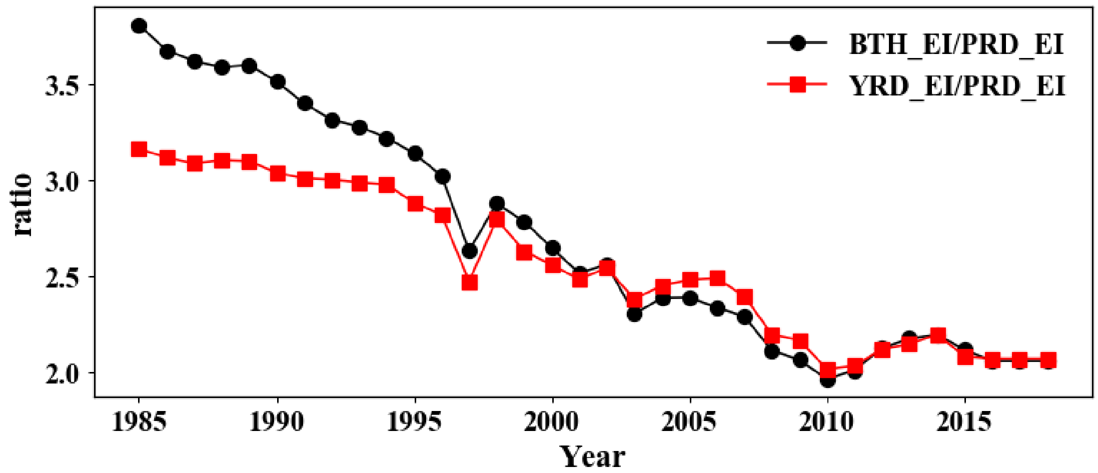

Figure S5. Time series of emission inventory (EI) ratios in the winter of 1985–2018 for the BTH/PRD and YRD/PRD, respectively.

(5) The PRD scaling factor was calculated by the emission sum from 2006 to 2013, while the scaling factors for the other two regions were calculated by the emission sum from 2010 to 2013. Please explain why using different emission sum to derive the scaling factors.

Response:

We calculate the scaling factor based on the overlapping periods of two inventories. As stated in our response to your Point #2, the time spans of these three inventories are different, so we derived scaling factors for PRD from 2006 to 2013, while for BTH and YRD from 2010 to 2013.

(6) The authors applied the nonlinear exponential fitting to retrieve the long-term $PM_{2.5}$ concentration before 2013, because China began to release the air quality observation data since 2013 and it is unlikely to acquire long-term observation data in this nation before 2013. However, based on the figures in the supplemental material, some of the fittings are not acceptable for further analysis, such as BTH-RH (40, 60) and YRD-RH (90, 100). The authors need to analyze and discuss whether such errors can influence their conclusion.

Response:

We believe that our $PM_{2.5}$ retrievals from nonlinear exponential fitting are acceptable for further analysis and do not influence our conclusion for the following reasons: (1) Yes, the correlation coefficient (R) of 0.56 for BTH-RH (40, 60) is a little low, but the fittings for other intervals with more samples are very good, so the overall fitting for BTH remains good. The sample size of YRD-RH (90, 100) is the smallest among all intervals, so its relatively small R (0.36) has negligible effect on the overall fitting for YRD. (2) We compared the retrieved $PM_{2.5}$ concentrations with the observed $PM_{2.5}$ concentration in BTH and YRD (Figure R2), and found that R is more than 0.87, and normalized mean bias (NMB) are 5.9% and 4.6% in BTH and YRD, respectively. These values of R and NMB suggest that the exponential fitting model is capable of reproducing the observed $PM_{2.5}$ concentrations. (3) As you suggested in comment point #7, we use the data of 2015–2019 for the fitting and the 2014 data for the verification, the R (NMB) between the fitted $PM_{2.5}$ concentrations and observed $PM_{2.5}$ concentrations is 0.77 (14.8%), 0.84 (5.5%), and 0.93 (5.1%) in BTH, YRD and PRD, respectively, suggesting that the exponential fitting models are robust. (4) R between the long-term retrieved $PM_{2.5}$ concentrations and observed visibility reached -0.9 in both BTH and YRD, reconfirming that the performance of our exponential fitting model is satisfactory (Figures S4a-b).

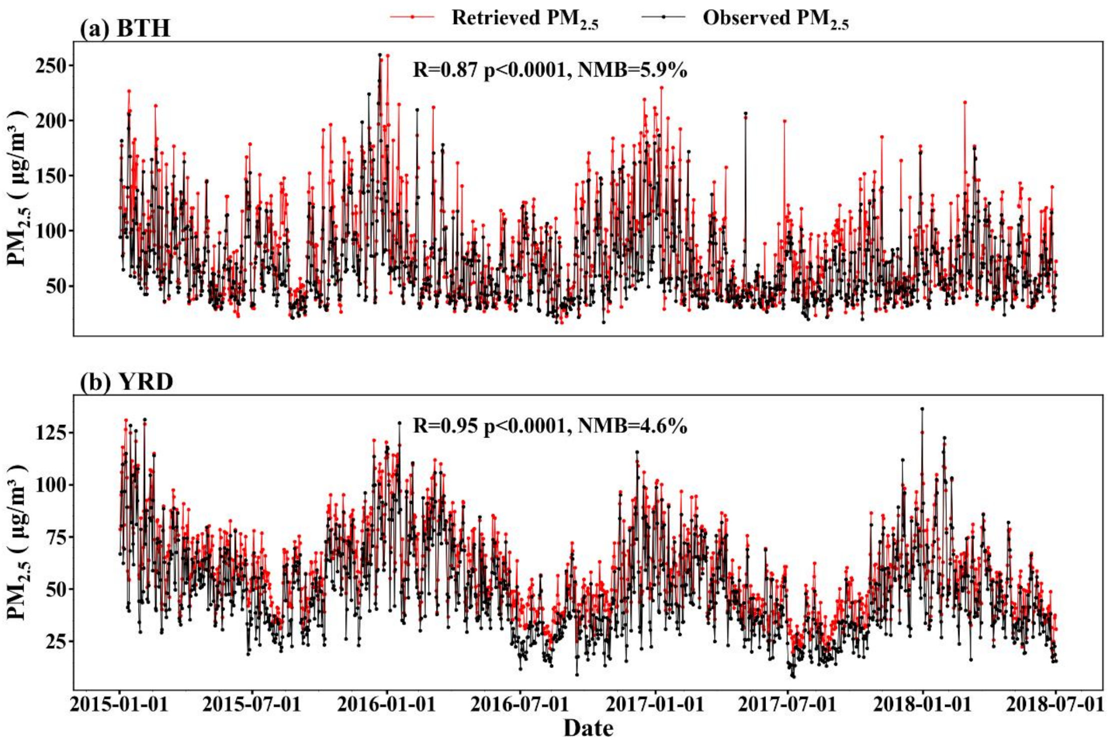

Figure R2. Temporal variation of retrieved $PM_{2.5}$ and observed $PM_{2.5}$ from 2015 to 2018.

(7) For the $PM_{2.5}$ concentration retrieval, I suggest the authors use the data of 2014-2018 for the fitting and the 2013 data for the verification, this can help to verify whether the methods implemented by the authors are reliable or not.

Response:

Since we did not have daily $PM_{2.5}$ concentration data for the three regions in 2013, we used observed daily visibility and $PM_{2.5}$ concentration from 2015 to 2019 to establish the exponential fitting model, as shown in Figures S1–S3. Furthermore, we use 2014 daily fitted $PM_{2.5}$ and observed $PM_{2.5}$ for verification (Figure R3). The R (NMB) between the fitted $PM_{2.5}$ concentration and observed $PM_{2.5}$ concentration is 0.77

(14.8%), 0.84 (5.5%), and 0.93 (5.1%) in BTH, YRD and PRD, respectively, verifying that the exponential fitting models are reliable.

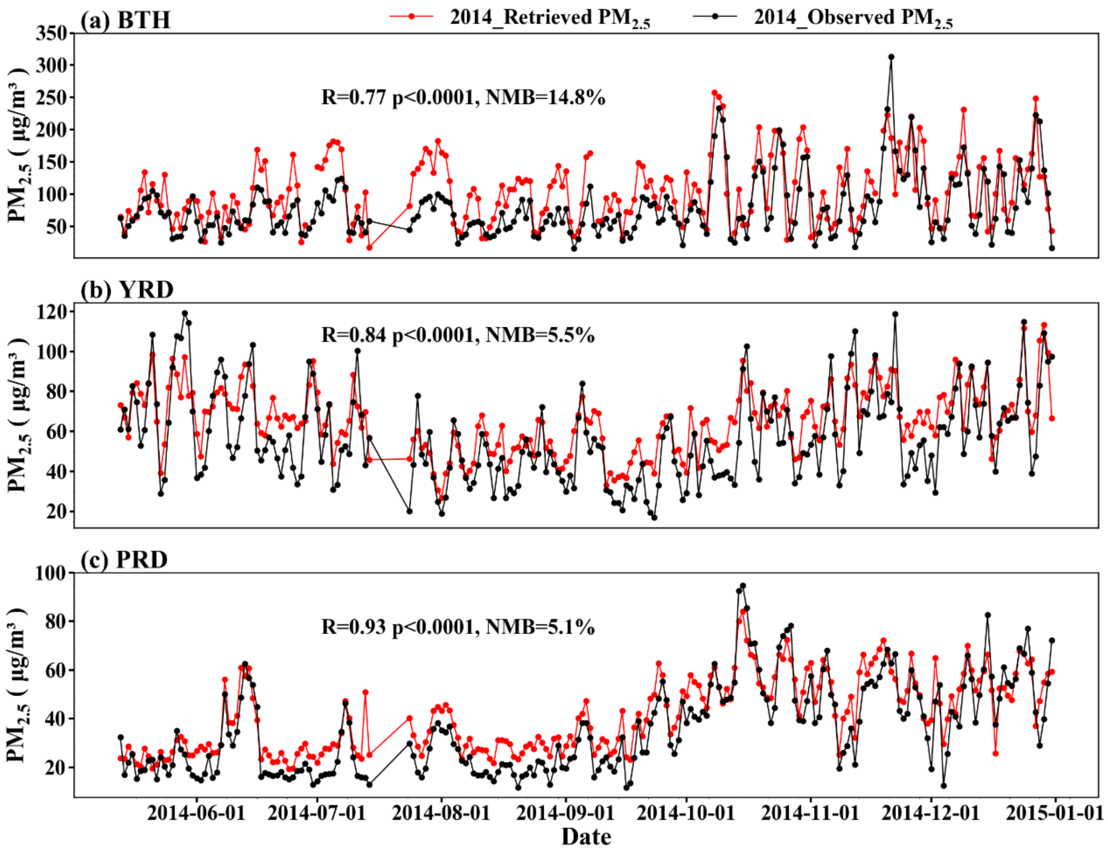

Figure R3. Temporal variation of retrieved $PM_{2.5}$ and observed $PM_{2.5}$ in 2014.

(8) Please introduce about the data source of RH and visibility in section 2.2. Generally the locations of the meteorological stations are not the same with those of the air quality stations. Did the authors use the nearest matching to pair the data? If so, what is the mean distance between the meteorological station and air quality station?

Response:

The visibility data we used is obtained from Global Summary of the Day (GSOD) database from National Environmental Satellite, Data, and Information Service (NESDIS) of the US Department of Commerce. Besides visibility, GSOD also provides daily average temperature and dew point, sea level pressure, wind speed and other meteorological elements and records of weather phenomena such as fog, rain and snow (http://www.ncdc.noaa.gov/cgi-bin/res40.pl). The GSOD data undergo extensive automated quality control by the Air Weather Service, and over 400 algorithms are applied automatically to correctly 'decode' the synoptic data, and to eliminate many of the random and systematic errors found in the original data. Data are generally available from 1929 to the present.

The RH used in this study is derived from dew point temperature and local air temperature following the approach proposed by Lawrence (2005). Therefore, RH and visibility data come from the same location.

We have added more information about the data source of visibility and RH in our revised manuscript as follows:

"Winter visibility data in 1973–2019 are obtained from Global Summary of Day (GSOD) provided by the National Climatic 50 Data Center (NCDC) (https://gis.ncdc.noaa.gov/maps/ncei/cdo/daily, last access: 10 March 2022). The relative humidity (RH) is derived from dew point temperature and air temperature of GSOD following the approach proposed by Lawrence (2005)."

(9) Line 120, combined with other studies and this work, we understand that the emission is the major factor that influences the $PM_{2.5}$ trend when compared to the meteorological variables. However, in Chen et al. (2019), the meteorological factors can still account for 21% of the contribution, which is much larger than the values reported by the authors in Line 117. I do not think this is an 'agreement'.

Response:

Sorry for the confusion! You are right that at this stage of the paper (Lines 111–125), "we understand that the emission is the major factor that influences the $PM_{2.5}$ trend when compared to the meteorological variables". Hence any number that shows (emission >> meteorology) is considered an agreement, so is the 21% meteorology because it is much less than the 79% emission. Nevertheless, you are quite right about "the meteorological factors can still account for 21% of the contribution, which is much larger than the values reported by the authors in Line 117". We will clarify this point in the revised manuscript.

(10) Lines 162-164, based on the analysis performed by the authors, if there exists any method that can compensate the shortcomings of the MLR and prognostic model?

Response:

Very important question!

As a start we believe that the alternative interpretation of MLR results proposed in Section 3.3: "The correlation coefficient should be interpreted as the maximum contribution of an independent variable to the dependent variable and the residual should be interpreted as the minimum contribution of all other independent variables" can help compensate some shortcomings of the MLR.

In regard to prognostic models, we are quite optimistic because some innovative studies have already appeared. For instance, Dang and Liao (2019) made a 33-year (1985–2017) model simulation study of severe winter haze days in BTH (purple line in Figure R4). There is an excellent agreement between the purple line and $PM_{2.5}$ concentrations observed by the US Embassy in Beijing (blue line, 2009–2018). The agreement with $PM_{2.5}$ concentrations observed by CNEMC in BTH (red line, 2013–2018) is also very good. For the entire period of 1985–2017, there are moderate mismatches near 1997–2002 and 2010 between the purple line (Dang and Liao, 2019) and green line (Li et al., 2021), but still has an acceptable overall correlation coefficient of 0.4. As cited in lines 201–202 of our paper, Dang and Liao (2019) "found that meteorology contributed significantly more than emissions to the linear trend", which is consistent with the result of our study.

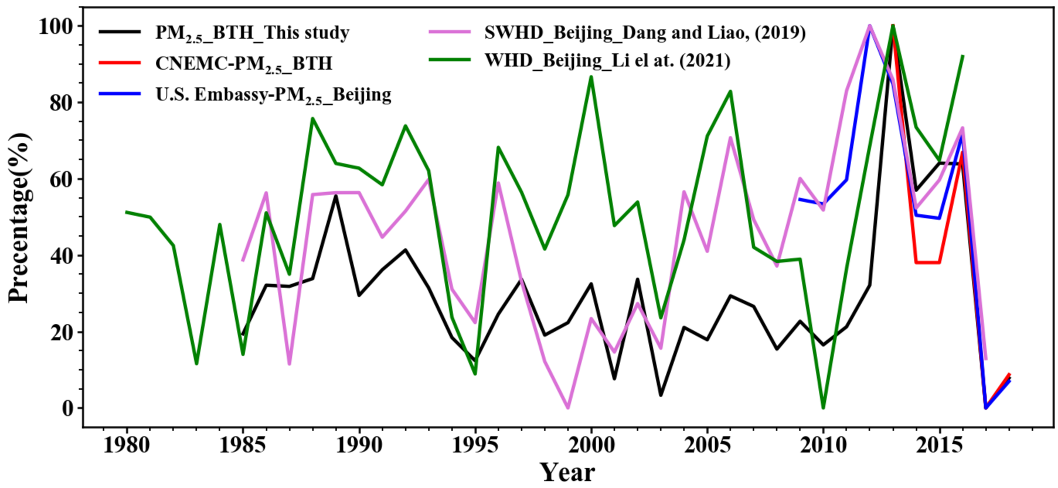

Figure R4. Temporal variations of winter inversed $PM_{2.5}$ concentrations in BTH of this study (black, 1985–2018), simulated $PM_{2.5}$ concentrations in Beijing by Dang and Liao (2019) (purple, 1985–2017), $PM_{2.5}$ concentrations observed by the US Embassy in Beijing (blue, 2009–2018) and those observed by CNEMC in BTH (red, 2013–2018).

(11) Lines 180-185, whether this means that previous studies that focused the ASI harbor relatively large uncertainty?

Response:

This part of the analysis mainly emphasizes that the MLR results are highly sensitive to the length of study time. Any short term MLR study, including those involving ASI, can harbor large uncertainty.

(12) Lines 166-169, I suggest the authors to provide some theoretical foundations to support this interpretation.

Response:

In the following we present a simplistic idea about a possible theoretical foundation to support our alternative interpretation of "the maximum possible contribution of the independent variable to the dependent variable". As an example, Figure R5 below depicts an MLR analysis of the contribution of emission to the linear trend of $PM_{2.5}$ in

BTH. It can be seen in Figure R1 that the MLR analysis is, in effect, performing the best-fit between the red line (emission) and the black line (observed $PM_{2.5}$). In other words, the best-fit enables the red line to attain the "maximum possible contribution" to the variability (including the linear trend) of the black line, where the "maximum" is established because all factors-other-than-emission that may contribute to the variability are excluded in the best-fit process. We propose the argument above as a possible theoretical foundation to support our alternative interpretation.

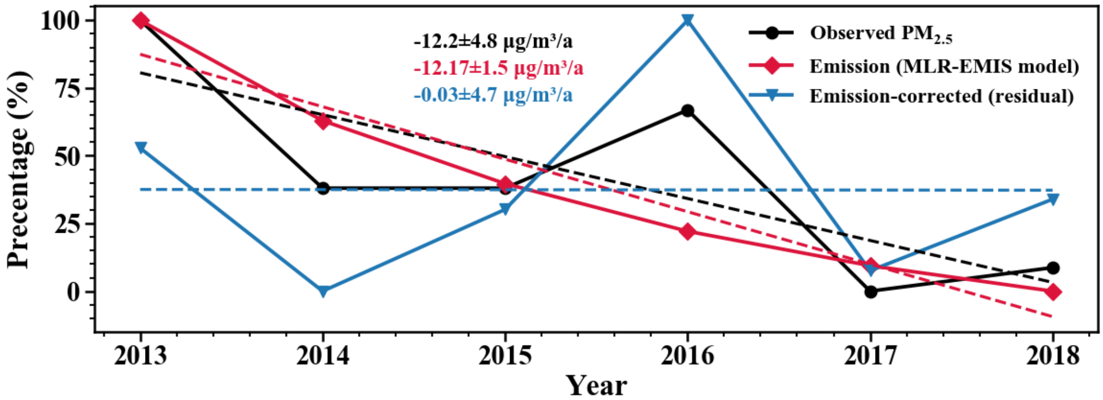

Figure R5. Results of MLR-EMIS analysis for 2013–2018 in BTH. Temporal variations of observed winter $PM_{2.5}$ concentration are shown in black, contributions of anthropogenic emissions to the $PM_{2.5}$ trend are shown in red, and the residual is shown in blue. Values inset in each panel are the ordinary linear regression trends, with 95% confidence intervals obtained by the student's t test.

References

Dang, R. and Liao, H.: Severe winter haze days in the Beijing-Tianjin-Hebei region from 1985 to 2017 and the roles of anthropogenic emissions and meteorology, Atmos. Chem. Phys., 19(16), 10801–10816, https://doi.org/10.5194/acp-19-10801-2019, 2019.

Huang, Z., Zhong, Z., Sha, Q., Xu, Y., Zhang, Z., Wu, L., Wang, Y., Zhang, L., Cui, X., Tang, M. S., Shi, B., Zheng, C., Li, Z., Hu, M., Bi, L., Zheng, J. and Yan, M.: An updated model-ready emission inventory for Guangdong Province by incorporating big data and mapping onto multiple chemical mechanisms, Sci. Total Environ., 769, 144535, https://doi.org/10.1016/j.scitotenv.2020.144535, 2021.

Lawrence, M. G.: The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications, Bull. Am. Meteorol. Soc., 86(2), 225–233, https://doi.org/10.1175/BAMS-86-2-225, 2005.

Li, H., Yang, Y., Wang, H., Li, B., Wang, P., Li, J. and Liao, H.: Constructing a spatiotemporally coherent long-term $PM_{2.5}$ concentration dataset over China during 1980–2019 using a machine learning approach, Sci. Total Environ., 765, 144263, https://doi.org/10.1016/j.scitotenv.2020.144263, 2021.

Zhong, Z., Zheng, J., Zhu, M., Huang, Z., Zhang, Z., Jia, G., Wang, X., Bian, Y., Wang, Y. and Li, N.: Recent developments of anthropogenic air pollutant emission inventories in Guangdong province, China, Sci. Total Environ., 627, 1080–1092, https://doi.org/10.1016/j.scitotenv.2018.01.268, 2018.