Dear Editor,

We thank reviewers for their constructive comments which have greatly improved our manuscript. We have now addressed all comments reviewers raised.

Referee: 1

This study combined the HYSPLIT ensemble with CWT to obtain hourly resolution of pollutant sources and employed machine learning method to quantify the contributions of local emissions and regional transport in Beijing. The article highlights that local emissions were the main cause of pollution events in Beijing from 2013 to 2020 and that the Air Pollution Prevention and Control Action Plan had a more significant effect on reducing emissions through regional transmission. After addressing the following comments, I believe this work has excellent potential for publication.

We are thankful for the valuable comments on our work from the reviewer.

General Comments

1. Line 43: “was gradually increasing”

This is now has been revised.

2. Line 44: “contributors to”

This is now has been revised.

3. Line 122-127: Compared to previous studies that relied solely on CWT analysis with HYSPLIT trajectories to distinguish between local emissions and regional transport, what specific improvements does your study introduce? In other words, after integrating XGBoost models, what are the advantages of your approach in enhancing the analysis? What specific problems or limitations of the previous methods does your study address? These aspects should be clearly articulated to highlight the improvements and contributions of your work.

We thank reviewer to point this out. This is now added in the revision:

Line 125-127: "In this study, we combined CWT analysis with the HYSPLIT trajectory ensemble to obtain hourly-resolution $PM_{2.5}$ source results and used this approach to distinguish between local emissions and regional transport. Solved the problems of traditional CWT methods being unable to obtain hourly time accuracy and models such as NAME consuming a large number of computational resources."

4. Line 125: The sentence "Fig. S2 reveals that ambient pollution events ($PM_{2.5}$>75 μg m^{-3}) in Beijing are primarily influenced by air masses originating from the south and west, particularly under the control of westward air masses." It merely presents the observed phenomenon that ambient pollution events in Beijing are mainly affected by air masses from the south and west, especially under the influence of westward air masses, but fails to provide in-depth analysis or explanation for why the westward air masses have a stronger influence in certain circumstances. It lacks speculation or reference to relevant studies to enhance the understanding of the underlying reasons for this phenomenon.

We thank reviewer to point this out. This is now added in the revision:

Line 229-241: "Numerous studies have indicated that air masses originating from the western region significantly contribute to regional pollution events in Beijing (Streets et al., 2007; Tian et al., 2019; Liu et al., 2020)"

5. Line 142: "which are available every 3 hours."

This is now has been revised.

6. Lines 148 to 160: There is an issue with the formula used to calculate the potential source region airflow trajectory weight concentration using CWT. C_{ij} represents the average weight concentration of the ij-th grid, and W_{ij} is the weight coefficient of grid (i,j) used to reduce uncertainty. Therefore, there is no need to multiply by W_{ij} when calculating C_{ij}; multiplying by W_{ij} is for calculating WCWT.

We thank reviewer to point this out. This is now added in the revision:

Line 156-171: "In the CWT analysis method, each grid point is assigned a weight, and the contribution of each grid point to the pollutant concentration at the target site is calculated using the air mass residence time and pollutant concentration (Hopke et al., 1993; Polissar et al., 1999; Xu and Akhtar, 2010) (equation 1). The grid point resolution was set to 0.25°×0.25° for this study. In equations 1, C_{ij} is the average weighted concentration at grid point (i, j), l is the trajectory index, M represents the total number of trajectories, C_l is the $PM_{2.5}$ concentration corresponding to the target site, and τ_{ijl} is the residence time of trajectory l passing through the grid point. In calculation, the number of trajectories falling on each grid point is used instead of the residence time.

$$C_{ij} = \frac{\sum_{l=1}^{M} C_l \times \tau_{ijl}}{\sum_{l=1}^{M} \tau_{ijl}} \quad (1)$$

To reduce the effect of small values of n_{ij}, the CWT values were multiplied by an arbitrary weight function $W(n_{i,j})$ to better reflect the uncertainty in the values for these grids (equation 2).

$$W(n_{i,j}) = \begin{cases} 1.00, & 3n_{ave} < n_{ij} \\ 0.70, & 1.5n_{ave} < n_{ij} \leq 3n_{ave} \\ 0.4, & n_{ave} < n_{ij} \leq 1.5n_{ave} \\ 0.17, & n_{ij} \leq n_{ave} \end{cases} \quad (2)$$

where n_{ij} represents the number of trajectories that fall within the grid point, and n_{ave} represents the average number of trajectories passing through each grid point."

7. Line 177-180: The authors employed the XGBoost model to predict $PM_{2.5}$ concentrations, using only meteorological, temporal variables and PBLH as input parameters. Considering that the data in this study were obtained from national monitoring stations, which typically provide detailed information on conventional pollutants (e.g., PM_{10}, SO_2, NO_x, O_3, CO), would the exclusion of these pollutant data from the model input impact the model's performance?

This study chose to use only meteorological data to learn $PM_{2.5}$ for two main reasons. Firstly, the learned $PM_{2.5}$ values include both the ambient and the locally emitted $PM_{2.5}$ values. Incorporating ambient PM or AOD values into the machine learning process may impact the local emission results. Secondly, numerous studies have confirmed that PM values can be obtained using meteorological data combined with machine learning method.

It is worth noting that many studies also use meteorological data combined with PM or AOD values to learn and obtain actual atmospheric PM results. For example, Xiao et al. used AOD combined with meteorological data to learn $PM_{2.5}$, achieving an r^2 result around 0.8 (Xiao et al., 2021). Similarly, Xu et al. used model-provided $PM_{2.5}$ combined with meteorological data to learn ambient $PM_{2.5}$ values, obtaining an r^2 result around 0.91 (Xu et al., 2023). However, despite the addition of AOD and PM parameters, there is still a significant difference in the r^2 values obtained from these studies, suggesting that sufficient training data is another important factor affecting the learning results.

In this study, the ambient and local $PM_{2.5}$ emissions obtained from meteorological data were compared with actual observations, yielding r^2 values of 0.74 and 0.78, respectively. These learning results are considered acceptable for the purposes of this study.

8. Line 185-195: I'm very interested in how the authors used the XGBoost model to separate local $PM_{2.5}$ from ambient $PM_{2.5}$, as this could be incredibly valuable for work in this field. However, the explanation in this section lacks sufficient detail on how this was achieved. I

believe other readers might have similar questions. It would be both helpful and necessary if the authors could provide more detailed and clear explanations to make the paper easier to understand and more applicable.

We thank reviewer to point this out. This is now added in the revision:

Line 128-133: "By training the XGBoost model with $PM_{2.5}$ dominated by local emissions, which are separately distinguished by CWT, and generalizing the findings to all study periods, the concentration of locally emitted $PM_{2.5}$ (local) can be obtained. Similarly, ambient observed $PM_{2.5}$ (ambient) can be determined by training the XGBoost model with ambient $PM_{2.5}$ data. The contribution of regional transport to $PM_{2.5}$ in Beijing can be quantified by comparing the ambient and local $PM_{2.5}$ concentrations."

9. Line 303: "from the south region"

This is now has been revised.

10. Why does the manuscript divide the year into four seasons (spring, summer, autumn, and winter) instead of four quarters? The commonly understood seasons have time differences, and given the long-time span of this study, this could introduce some error. The study needs to clearly define how spring, summer, autumn, and winter are defined each year.

We thank reviewer to point this out. This is now added in the revision:

Line 143-145: "In this study, a year was divided into four quarters: Spring (March, April, and May), Summer (June, July, and August), Autumn (September, October, and November), and Winter (December, January, and February)."

11. Based on the CWT combined with the HYSPLIT ensemble, the authors distinguished between local emissions and regional transport. However, in the subsequent machine learning process, the authors used XGBoost to derive locally emitted $PM_{2.5}$ and then derived the regionally transported $PM_{2.5}$. Why choose this approach instead of learning regional transmission to calculate local emissions? Please explain the reasoning.

Local emission sources in Beijing have more stable pollution components compared to regional transmission. Thus, the results obtained from learning local emission sources are believed to be more consistent with actual observed values compared to regional emissions, which are influenced by various sources. Therefore, in this study, regional transport contributions are determined by subtracting local emissions from the ambient concentrations, rather than learning regional transport and calculating local emission values.

References

Hopke, P. K., Gao, N., and Cheng, M.-D.: Combining chemical and meteorological data to infer source areas of airborne pollutants, Chemometrics and Intelligent Laboratory Systems, 19, 187-199, 1993.

Liu, D., Hu, K., Zhao, D., Ding, S., Wu, Y., Zhou, C., Yu, C., Tian, P., Liu, Q., and Bi, K.: Efficient vertical transport of black carbon in the planetary boundary layer, Geophysical Research Letters, 47, e2020GL088858, 2020.

Polissar, A., Hopke, P., Paatero, P., Kaufmann, Y., Hall, D., Bodhaine, B., Dutton, E., and Harris, J.: The aerosol at Barrow, Alaska: long-term trends and source locations, Atmospheric Environment, 33, 2441-2458, 1999.

Streets, D. G., Fu, J. S., Jang, C. J., Hao, J., He, K., Tang, X., Zhang, Y., Wang, Z., Li, Z., and Zhang, Q.: Air quality during the 2008 Beijing Olympic games, Atmospheric environment, 41, 480-492, 2007.

Tian, P., Liu, D., Huang, M., Liu, Q., Zhao, D., Ran, L., Deng, Z., Wu, Y., Fu, S., and Bi, K.: The evolution of an aerosol event observed from aircraft in Beijing: An insight into regional pollution transport, Atmospheric Environment, 206, 11-20, 2019.

Xiao, Q., Zheng, Y., Geng, G., Chen, C., Huang, X., Che, H., Zhang, X., He, K., and Zhang, Q.: Separating emission and meteorological contributions to long-term PM_{2.5} trends over eastern China during 2000–2018, Atmospheric Chemistry and Physics, 21, 9475-9496, 10.5194/acp-21-9475-2021, 2021.

Xu, R., Ye, T., Yue, X., Yang, Z., Yu, W., Zhang, Y., Bell, M. L., Morawska, L., Yu, P., and Zhang, Y.: Global population exposure to landscape fire air pollution from 2000 to 2019, Nature, 621, 521-529, 2023.

Xu, X. and Akhtar, U.: Identification of potential regional sources of atmospheric total gaseous mercury in Windsor, Ontario, Canada using hybrid receptor modeling, Atmospheric Chemistry and Physics, 10, 7073-7083, 2010.