Reply to Referee # 2

Long-term in-situ observations of black carbon aerosols are crucial for studying their environmental and climatic effects. However, in real-world observational studies, there are several inevitable technical challenges, such as data gaps. This paper proposes a machine learning method that elegantly addresses this issue. The method is applied to reconstruct time-series data of elemental carbon (EC) aerosols from four cities in eastern China. The results are also validated by comparing them with other datasets. Furthermore, the paper introduces a novel method for assessing the driving factors of long-term trends in elemental carbon, as well as evaluating the uncertainty associated with this approach. I believe both methods hold significant value for the field of atmospheric monitoring. Overall, the paper is well designed and written. However, I have the following points that the authors should address:

Reply: We would like to express our gratitude to the reviewer for the positive and constructive comments. We have carefully revised the manuscript in accordance with the suggestions provided. Please find our responses below:

The authors introduce MERRA-2 black carbon column concentration data as one of the predictor variables. They also compare MERRA-2 near-surface black carbon concentrations and find that the MERRA-2 data tends to overestimate the site's elemental carbon data. I suggest that the authors conduct a sensitivity test by training the machine learning model without using MERRA-2 black carbon column concentration as a predictor variable and compare the results with the current ones.

Reply: Thank you for your suggestion. We have conducted a sensitivity test by training the machine learning model without including MERRA-2 black carbon column concentration (BCC) as a predictor variable. We compared the results with those of the current model and present the findings in Figure S4. The revised text reads: "*We further evaluated the importance of the MERRA-2 black carbon concentration (BCC) as a predictor by testing the model's performance both with and without this variable (see Figure S4). Inclusion of MERRA-2 BCC significantly improved the model's performance across all evaluation metrics, confirming it as a key contributor to model accuracy.*" More detailed modifications can be found on lines 206 - 209 in the revised manuscript.

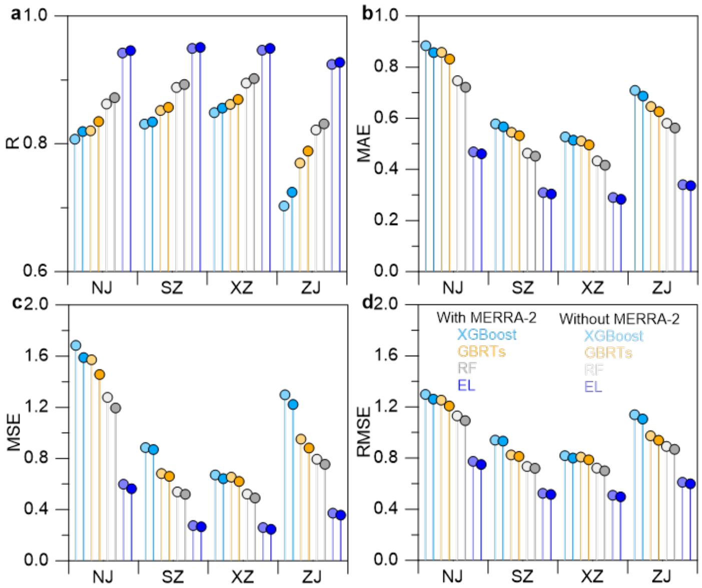

Figure S4. Comparison of the model performance parameters for four cities, training the machine learning model with or without MERRA-2 BCC as a predictor variable (The darker color represents with MERRA-2, while the lighter color represents without MERRA-2.).

The trend changes in EC aerosols are influenced by both meteorological conditions and emissions. In eastern China, the sources of black carbon generally include vehicle emissions and industrial coal combustion. While the paper quantifies the overall anthropogenic emission trend drivers, there is relatively little information on specific emission sectors, which may be a limitation of the method employed. The paper analyzes the daily variation of EC over the years and suggests that the reduction of motor vehicle emissions may be a major factor driving the decline in EC levels. I suggest that the authors could try to extend this analysis by investigating the trend changes of EC during vehicle emission rush hours or by quantifying the driving factors for these peak periods. This could provide a more detailed understanding of the trend changes.

Reply: We appreciate the suggestion to further explore the trend changes in EC during vehicle emission rush hours. We have extended the analysis to include this investigation and quantified the driving factors for these peak periods, which has provided a more detailed understanding of the trend changes. The results are shown in Figures S14 and S15.

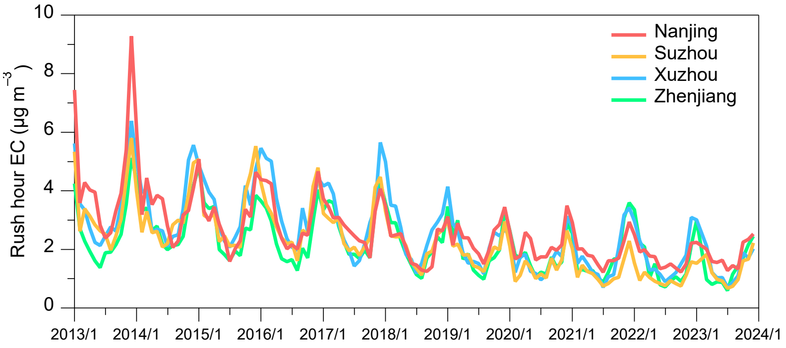

Figure S14. Trends in monthly average EC concentrations during morning rush hours (07:00–09:00) from 2013 to 2023 across the four cities: Nanjing, Suzhou, Xuzhou, and Zhenjiang.

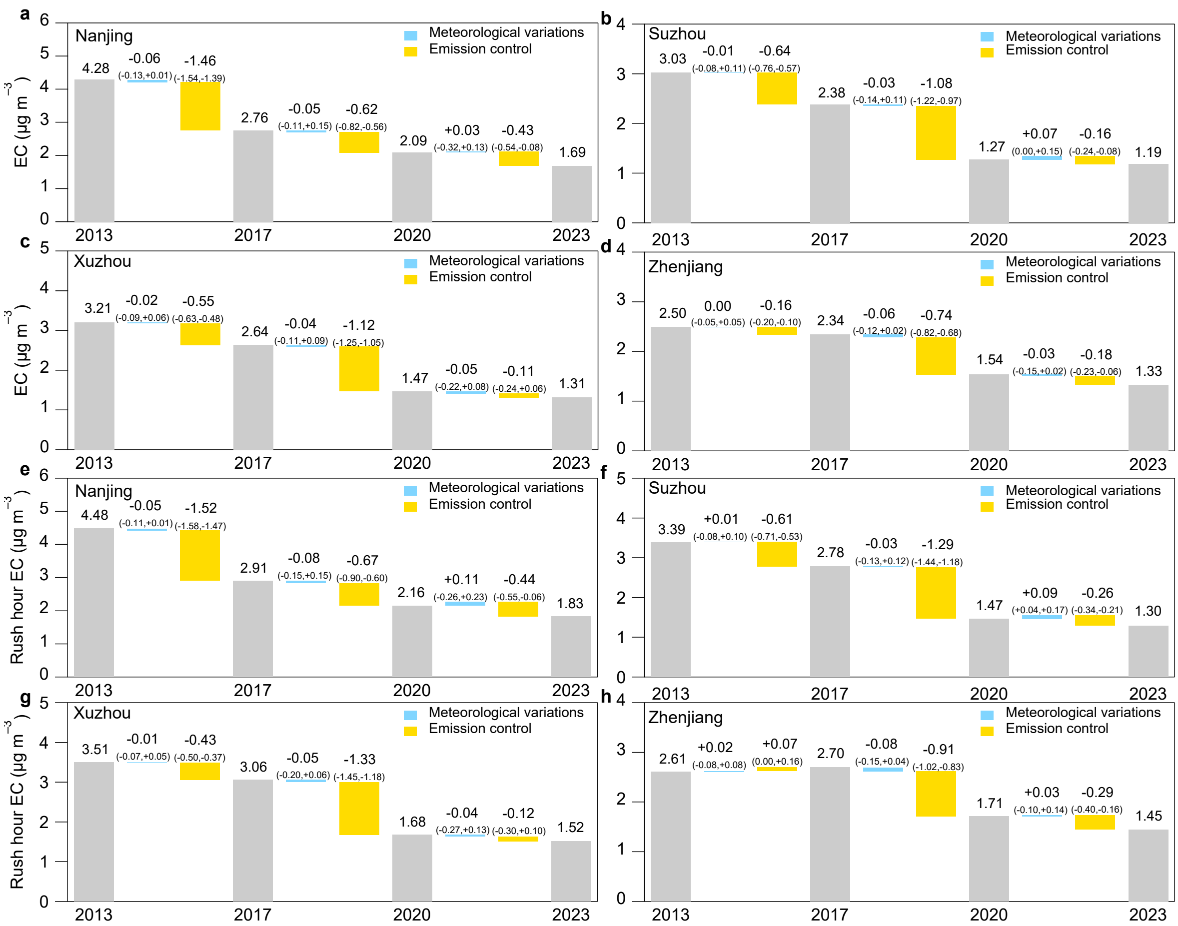

Figure S15. Drivers of EC trends from 2013 to 2023. (**a–d**) Contributions of anthropogenic emission control and meteorological variations to EC concentration trends across four cities. (**e–h**) Same as (**a–d**), but specifically for rush-hour periods (7:00 – 9:00) in the four cities.

The revised text reads: "*Figure S6a–d and Table S7 further demonstrate strong correlations between EC and NO_2 concentrations, suggesting that vehicle emissions are a major contributor to urban BC levels. The observed reductions in NO_2 in all four cities underscore the impact of transportation-related emission control measures on declining EC concentrations. We further examined rush-hour EC trends in Figures S14 and S15 and found more pronounced declines during these periods compared to the interannual average, indicating that targeted control strategies for mobile sources have been particularly effective. This finding aligns with previous work (e.g., Zheng et al., 2018), which reported substantial reductions in BC emissions from the transportation sector in China between 2010 and 2017.*" More detailed modifications can be found on lines 372 - 378 in the revised manuscript.

The authors use the ridge regression algorithm for the multivariate regression analysis but do not employ the traditional multiple linear regression algorithm. I recommend that the authors clarify this choice. Additionally, regarding Equation 1, the expression may cause confusion because GBRTs, XGBoost, and RF are abbreviations for different machine learning algorithms, yet they are presented as variables in the formula. I suggest the authors optimize the notation for clarity.

Reply: Thank you for the comment. We chose ridge regression for multivariate regression analysis due to its ability to handle multicollinearity, which is often present in complex models. Ridge regression introduces a regularization term that enhances model stability and reduces the risk of overfitting. This approach ensures more reliable estimates when dealing with correlated predictors. The revised text reads: "*For multivariate regression analysis, we chose ridge regression over traditional multiple linear regression to account for multicollinearity among the three model outputs. Ridge regression is particularly effective in handling multicollinearity by introducing a regularization term that improves computational stability and reduces the risk of overfitting (Kidwell and Brown, 1982; Hoerl and Kennard, 1970).*"

We have also revised the notation to clarify that GBRTs, XGBoost, and RF are machine learning algorithms, not variables, to avoid any confusion. The revised equation reads: "

$$f_{EL} = m_1 f_{GBRTs} + m_2 f_{XGBoost} + m_3 f_{RF}, \quad (1)$$"

More detailed modifications can be found on lines 130 – 133 and 135 in the revised manuscript.

Line 148 – 149: Appropriate references should be cited to support the use of these pollutants as tracers for source characterization.

Reply: Thank you for the comment. We have added references to support the use of these pollutants as tracers for source characterization. More detailed modifications can be found on lines 142 - 143 in the revised manuscript.

Line 239: The phrase "Reconstruction of missing data of EC and trend analysis" should be revised to "Reconstruction of missing data of EC and comparison".

Reply: Thanks for the suggestion. Modified.

Line 336 – 337: The discussion on the impact of COVID-19 lockdowns on EC trend changes is well noted as a factual observation. Could the authors further discuss or quantify such impact?

Reply: Thank you for your valuable suggestion. We have further discussed the impact of the COVID-19 lockdowns and quantified their contribution to EC trend changes. Our analysis, as shown in Figure 5, indicates that the COVID-19 lockdown contributed approximately 3%-10% to the total reduction in EC concentrations between 2017 and 2020, taking into account both meteorological variations and changes in anthropogenic activities.

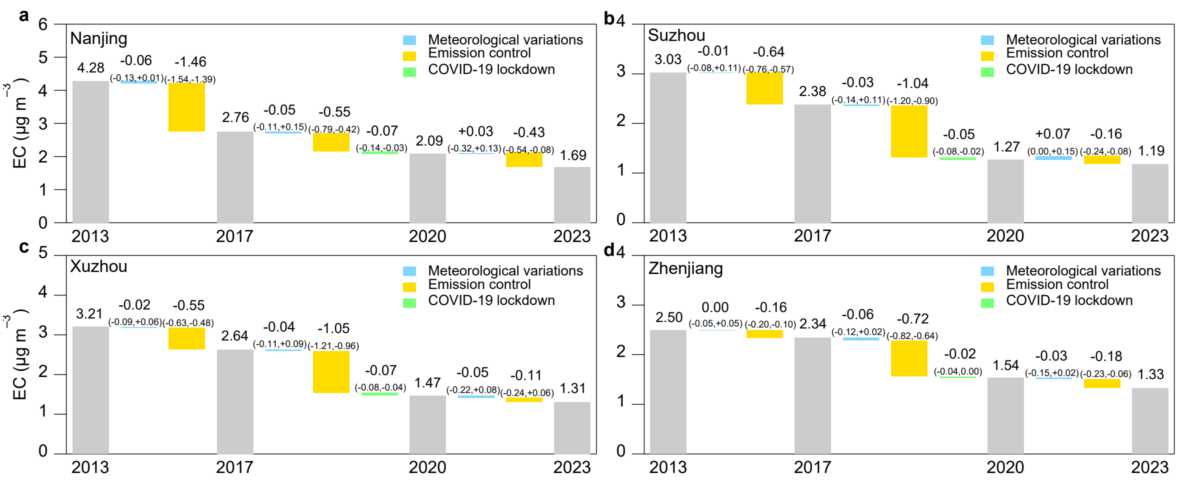

Figure 5. Drivers of the EC trend from 2013 to 2023. a-d the contributions of anthropogenic emission control, meteorological conditions and COVID-19 lockdown on the trends in EC concentration in the four cities.

The revised text now reads: "*As shown in Figure 5, we quantified the impact of the lockdown on EC levels, estimating that the pandemic contributed 3–10% of the total EC reduction from 2017 to 2020—an effect stemming from both emission reductions and changes in human activity. This is consistent with Zheng et al. (2021), who reported large-scale BC emission reductions during the lockdowns of early 2020, especially from industrial, residential, and transportation sources.*" More detailed modifications can be found on lines 181 – 190 and 398 – 402 in the revised manuscript.