Dear editor,

Here we submitted our updated manuscript for consideration to be published on **Atmospheric Chemistry and Physics**

The further information about our manuscript is as follows:

Topic: Substantial changes of gaseous pollutants and chemical compositions in fine particles in North China Plain during COVID-19 lockdown period: anthropogenic vs meteorological influences

Type of Manuscript: article

Authors: Rui Li[a], Yilong Zhao[a], Hongbo Fu[a, b *], Jianmin Chen[a, b *]

Corresponding author:

Hongbo Fu; Address: Department of Environmental Science and Engineering, Fudan University, Shanghai 200433, China; Tel.: (+86)21-5566-5189; Fax: (+86)21-6564-2080; Email: fuhb@fudan.edu.cn

Jianmin Chen; jmchen@fudan.edu.cn

 # SC1: I found the approach interesting and would like to see this article published. It shows a method to quantify the influence of meteorological parameters on pollutant time series and thus helps to better quantify the impact of COVID-19 lockdown measures on air pollution. There are, however, a few points that require better explanation or change before publication:

Response: Thank for reviewer's suggestions. We have significantly revised the manuscript based on reviewer's suggestions.

Comment 1: A decrease of more than 100 % of a pollutant does not make sense. You cannot remove more than all the pollutant. The authors should reconsider which reference value they use to calculate the percent reduction. This relates to the abstract and section 3.2 (lines 193 ff), and probably also to the figures (see remark 3).

Response: I agree with reviewer's suggestions. The decrease ratio should be estimated based on quotient of (Pre-COVID – Post-COVID) and Pre-COVID rather than Post-COVID (Pre-COVID and Post-COVID represent the pollutant concentrations before COVID-19 lockdown and after COVID-19 lockdown, respectively). We have corrected the errors throughout the manuscript.

Comment 2: The random forest approach (line 118) should be referenced with a citation to the literature. Why do you use this approach instead of other statistical methods?

Response: Thank for reviewer's suggestions. A reference has been cited in the line 122. We selected the random forest (RF) model to distinguish the emission impact from meteorological effect because the model is generally robust for the modelling compared with some traditional statistical models (e.g., GAM, LME). Indeed, some novel models such as XGBoost and LightGBM often showed the better performance compared with RF for the big-data samples. However, the number of the training samples (2184 samples) in our study is not very large, and the RF model was more suitable in our study.

Comment 3: The observed or calculated decreases of pollutants in Fig. 2, 3 and 4 are not sufficiently explained. It is unclear what the trendlines (arrows) mean. The percent values do not explain the slopes of the curves, and it is unclear what the arrows should show. There is definitively more explanation needed to understand which concentrations are compared and used to calculate the reductions, either in the main text, in the figure captions, or in the supplement.

Response: Thank for reviewer's suggestions. We have rewritten the captions of Fig. 2-4 in the revised version. The black solid line and dotted line represent the decrease ratio of observed concentration and simulated concentration from Pre-COVID to Post-COVID, respectively. The white background denotes the changes of gaseous pollutants before COVID-19, while the faint yellow one represents the chemical components after COVID-19 outbreak. Week 1-3 was regarded as Pre-COVID, and Week 4-13 was treated as Post-COVID. The decrease ratio was estimated based on quotient of (Pre-COVID - Post-COVID) and Pre-COVID.

 # Reviewer 1

This paper analyzed changes of gaseous air pollutants as well as chemical compositions of $PM_{2.5}$ based on the observational data in Tangshan in the North China Plain. A random forest model was applied to investigate the contributions of meteorology and anthropogenic emissions to the changes of air quality during COVID-19 period. PMF was further applied to determine changes of source contributions before and during the lockdown. The topic is interesting, however, there are several major concerns to be addressed before it can be considered for publication:

Comment 1: The most important data used in this study is the observation data based on a supersite in Tangshan whereas the title is "North China Plain". NCP covers large area with different topographical and meteorological conditions. Presenting the observation data at a single site is far enough to draw conclusion for a large and heterogeneous region. So the first limitation of this study whether this site is representative of the whole NCP region.

Response: Thank for reviewer's suggestions. Indeed, NCP is large study region, which covers many regions different topographical and meteorological conditions. In general, only a site in Tangshan cannot reflect the overall status in NCP. However, COVID-19 is a unique period and the variations of pollutant concentrations during COVID-19 in Tangshan might reflect the variations of pollutant concentrations in the whole region across NCP. The detailed reasons are as follows: At first, Tangshan is typical industrial cities in China and many energy-intensive industries including cement, steel and iron industries focused on NCP, and thus Tangshan is a representative site across NCP. Moreover, after the COVID-19 lockdown, the energy production by power plants was reduced by one third and the oil refineries and steel industries reached the lowest activity level of the past 5 years in East China (Chang et al., 2020 GRL), suggesting that the industrial activity played a major role on the variation of air pollutants during COVID-19 period in East China. Therefore, we believed that Tangshan is a good site to reveal the effect of COVID-19 lockdown on the temporal variations of air pollutants in NCP.

Comment 2: Nine trace elements observed by Xact 625 are used in this study; however, Xact 625 can observe more than 20 species. Why did the authors only present 9 trace elements? What about the others?

Response: Thank for reviewer's suggestions. Indeed, Xact 625 could observe more than 20 species. In the original model, we also analyzed the effects of meteorological factors on 20 species. However, only Hg, Pb, K, Ca, Cr, Cu, Fe, Ni, and Zn showed the better performance in the RF model. For other species, the worse performance hindered the analysis in our study because analysis based on the worse modelling performance was not reliable. In the future work, we hope to incorporate more variables such as hourly trace metal emission into the model, which might enhance the modelling performance.

Comment 3: The authors used RF to quantify the influences from meteorology and emissions on air quality. Why do the authors use RF instead of other machine learning techniques? This should be clarified. In addition, how can the model results be evaluated, or in other words, how do the authors demonstrate the robustness of their model results?

Response: Thank for reviewer's suggestions. We selected the random forest (RF) model to distinguish the emission impact from meteorological effect because the model is generally robust for the modelling compared with some traditional statistical models (e.g., GAM, LME). Indeed, some novel models such as XGBoost and LightGBM often showed the better performance compared with RF for the big-data samples. However, the number of the training samples (2184 samples) in our study is not very large, and the RF model was more suitable in our study. The R^2 value, RMSE, and MAE were regarded as the major indicator to evaluate the modelling performance of the RF model. In general, for nearly all of the machine-learning models, the 10-fold cross-validation R^2 value should be larger than 0.50 at least because too low R^2 value cannot assure the reliability of the predicted value. Therefore, in our study, we only selected the species with the 10-fold cross-validation R^2 value larger than 0.50 to analyze the impact of emission and meteorology on the pollutant concentrations.

Comment 4: When performing source apportionment with PMF, the major input data does not even include major chemical components like OC/EC. How will this influence the results?

Response: I agree with reviewer's suggestions. Both of OC and EC were incorporated into the PMF model and recalculated the source apportionment of these species. The result suggested that OC and EC were closely linked with K concentration, indicating the contribution of biomass burning. The integration of OC and EC did not significantly affect the final result.

Specific comments:

Comment 5: Abstract should be rewritten; the major findings/answers to the question raised in the title should be present in the abstract. In addition, it is misleading to present results like Cr (-201%) and Fe (-154%). The decrease of more than 100 % of a pollutant is misleading.

Response: I agree with reviewer's suggestions. The decrease ratio should be estimated based on quotient of (Pre-COVID – Post-COVID) and Pre-COVID rather than Post-COVID (Pre-COVID and Post-COVID represent the pollutant concentrations before COVID-19 lockdown and after COVID-19 lockdown, respectively). We have corrected the errors throughout the manuscript.

Comment 6: The NH_3 measurement is made by GAC-IC, while Hg is observed with Xact 625. Quality assurance and quality control procedures and results should be well documented and provided at least as supporting information.

Response: Thank for reviewer's suggestions. The quality assurance and control procedures were added in the text S1 in supporting information.

Comment 7: In "Section 2.2 Deweathered model development", the authors mainly described random forest; however, the deweathered technique with RF is not described. This should be clearly described in detail here. How are the meteorological conditions isolated by RF? How can the results be evaluated?

Response: Thank for reviewer's suggestions. In our study, the original dataset was randomly classified into a training dataset (90% of input dataset) for developing the RF model and the remained 10% was treated as the test dataset. Based on this method, we can obtained the predicted concentration of each species. After the building of RF model, all of the meteorological data were normalized to remove their effects and then we incorporated the new training dataset into the same RF model to predict the concentration of each species, which meant the concentration without the impact of meteorology. At last, the differences of original pollutant concentrations and deweathered pollutant concentrations were regarded as the concentrations contributed by meteorology.

Many statistical indicators including the R^2 value, RMSE, and MAE were regarded as the major criteria to evaluate the modelling performance of the RF model. In general, for nearly all of the machine-learning models, the 10-fold cross-validation R^2 value should be larger than 0.50 at least because too low R^2 value cannot assure the reliability of the predicted value. Therefore, in our study, we only selected the species with the 10-fold cross-validation R^2 value larger than 0.50 to analyze the impact of emission and meteorology on the pollutant concentrations.

Comment 8: Sulfate and NO_x concentration decreased substantially after lockdown, while nitrate and ammonium increased. The authors explained that this might be due to the adverse meteorological conditions. This explanation is very weak. NO_2 has been reduced by 62.8% while nitrate increased by 2.17%; is this attributable to the adverse meteorology? How did meteorological parameters change before and after the lockdown? Are there any changes in chemical reactions that are responsible?

Response: Thank for reviewer's suggestions. The reasons have been carefully discussed in the section 3.2 and 3.3 from the perspective of primary emission and meteorology, respectively. The relatively stable observed concentrations of NO_3^- and NH_4^+ after COVID-19 lockdown was attributable to that the unfavorable meteorological conditions counteracted the contribution of lockdown measures. Based on the RF model, the deweathered NO_3^- and NH_4^+ concentrations suffered from -27% and -13% decreases after COVID-19 lockdown, respectively. Unfortunately, the high RH enhanced gas- to aqueous-phase dissolution of NH_3 and HNO_3, which was supported by the variable importance of RH. Furthermore, the increased O_3 (160% increase) could promote the secondary aerosol formation and partially offset the decreased $PM_{2.5}$ compositions triggered by the primary emission reduction.

Comment 9: Some elements like Pb, Ca, Cr, Cu, Fe and Zn has been reduced by 7.44~91.5% while Hg, K, Ni increased by 20%, 0.08% and 1.17%, respectively. What's the reason? The authors also mentioned that the slight increase of K might be linked with the unfavorable meteorological conditions. This explanation is still very weak and not convincible. The authors should give detailed data analysis regarding the changes of meteorological parameters to support their explanation.

Response: Thank for reviewer's suggestions. In the first section of results and discussion, we only showed the result of observed concentrations and did not give deep discussion about the reasons for the temporal variations of pollutant concentrations. In the section 3.2 and section 3.3, we analyzed the variations of pollutant concentrations before and after COVID-19 lockdown from the perspective of primary emission and meteorology, respectively. Some metals including Zn, Pb, Cr, and Fe concentrations suffered from dramatic decreases after COVID-19 lockdown. It was well known that Zn, Cr, and Fe originated mainly from metallurgical industry (Sun et al., 2018; Zhu et al., 2018), while Pb might be derived from coal-fired power plants (Cui et al., 2019; Meng et al., 2020). During the COVID-19 outbreak, most of the industries have been shut down and energy production by coal-fired power plants was reduced by one third (Chang et al., 2020). Based on the adjustment factor estimated by Doumbia et al. (2020), the contributions of industrial activity and power sector have decreased by 40% after COVID-19 outbreak, which was close to the decrease ratios of Zn, Cr, Fe, and Pb concentrations. It should be noted that the deweathered Ca concentration also decreased by more than 50%. It was well documented that the Ca was often associated with the dust resuspension (Chang et al., 2018). In fact, the Ca was known as one of the most abundant elements in the upper continental crust, which likely originated from the traffic-related fugitive dust (Chang et al., 2018; Shen et al., 2016). More than 70% reduction of vehicle transportation and domestic flights facilitated the rapid decrease of Ca concentration (Chang et al., 2020).

Although the observed K and Ni concentrations did not show marked decrease after the COVID-19 lockdown, the deweathered K (-22%) and Ni (-27%) levels suffered from rapid decreases ($p < 0.05$). It was widely acknowledged that K was considered to be a key fingerprint of biomass burning (Zheng et al., 2020), and thus the result suggested that the open biomass burning was also restricted during the period. Ni was mainly sourced from the traffic-related road dust, which was also significantly affected by the traffic restriction. Unlike the deweathered concentration, the observed K and Ni concentrations remained very stable after COVID-19 lockdown. It was assumed that the unfavorable meteorological condition offset the decreased primary emission. As shown in Figure 9, P was the most important meteorological factor for the K and Ni variations. P is a comprehensive factor and could affect the atmospheric circulation and advection (convection) of air pollutants. Furthermore, the wind speed has decreased from 1.53 to 0.87 m/s during the study period. Overall, the stagnant meteorological condition after COVID-19 lockdown caused the stable observed K and

Ni concentrations after COVID-19, which was different from other trace elements.

Comment 10: Line 194-196, the deweathered Hg concentration still kept stable increase by 18%, which is opposite to other trace elements. What's the reason?

Response: Thank for reviewer's suggestions. Due to our fault, the concentration variation ratio of the species were wrong in the original version. Thus, we have recalculated all of these increase (decrease) ratio in the revised version. The observed Hg concentration increased by 8%, and the deweathered Hg concentration increased by 6%. Indeed, the temporal variation of Hg level was in contrast to most trace elements (e.g., Pb, Zn). Based on the independent sample t-test, the Hg concentration difference before COVID-19 and after COVID-19 was not significant ($p > 0.05$). The minor increase of deweathered Hg level was attributable to that the coal combustion for domestic heating was not restricted during the COVID-19 lockdown period. Based on the updated global anthropogenic emission adjustment factor during COVID-19, the contribution of residential sector to air pollutants did not decrease after COVID-19 lockdown (Doumbia et al., 2020).

Comment 11: Line 207, the deweathered Ca concentration decreased by more than 100%, it is hard to believe. Again, in Line 346, the Pb(-147%), Zn (-219%). This kind of description should be well clarified.

Response: Thank for reviewer's suggestions. These errors have been corrected in the revised version. The deweathered Ca, Pb, and Zn concentrations decreased by 54.2%, 59.4%, and 68.7%, respectively.

Comment 12: Fig5 (C) some bars are not well shown, like Cr, Fe, Zn.

Response: Thank for reviewer's suggestions. The figures have been redrawn in the revised version.

Comment 13: Fig6-8 It does not make any sense to indicate DOY/Year is important or not in the prediction of gaseous pollutants using RF.

Response: Thank for reviewer's suggestion. In the RF models for all the species, year showed the lowest variable importance. However, DOY showed the higher variable importance for most of the species because DOY was closely linked with the implementation of lockdown measures.

Comment 14: There are many English grammar errors. The language should be polished thoroughly. For example, Line 130-131, "…were input into the model" there is grammar error in this sentence. Line 252, "both of " should be "both".

Response: Thank for reviewer's suggestions. The language throughout the manuscript should be significantly revised.

 # Reviewer 2

The authors mainly dealt with the effects of the lockdown measures due to COVID-19 pandemic on gaseous pollutants and fine aerosol particles in North China Plain. The topic is timely and of interest for the research community, and fits into the scope of the journal. The authors also investigated the changes in major inorganic chemical composition and some metal constituents of particles which has been rarely done so far. The MS indicates valuable results and conclusions, which seem worth publishing on one hand. On the other hand, it is difficult to assess their real value since some important information, firm explanations and background discussions are largely missing. They should definitely be complemented. There are also several smaller discrepancies or other issues listed below which are to be improved or corrected for. The present reviewer can arrive at the final suggestion only after all these are clarified and added in a careful and substantial manner.

Response: Thank for reviewer's suggestions. We have significantly revised the manuscript based on reviewer's suggestions.

Major comments

Comment 1: One of the key methods applied to deconvolute the effects of meteorology on the atmospheric concentrations is the random forest approach. The method and its conditions of validity are not described and virtually not discussed (Sect. 2.2). Several important questions can be formulated in the reader with regard to this. For instance, is the investigated time interval sufficient for training the learning method and for its testing as well, in particular when the data sets were shared in a respective ratio of 90%/10% between these phases (lines 86 and 122). Are there any constrains of the method as far as the number of available data and retained variables are concerned? What are the uncertainties or limitations of the modelled results? Are there possibilities to verify or validate the modelled outcomes and were they performed?

Response: Thank for reviewer's suggestions. We selected the random forest (RF) model to distinguish the emission impact from meteorological effect because the model is generally robust for the modelling. In our study, many statistical indicators including the R^2 value, RMSE, and MAE were regarded as the major criteria to evaluate the modelling performance of the RF model. In general, for nearly all of the machine-learning models, the 10-fold cross-validation R^2 value should be larger than 0.50 at least because too low R^2 value cannot assure the reliability of the predicted value. Therefore, in our study, we only selected the species with the 10-fold cross-validation R^2 value larger than 0.50 to analyze the impact of emission and meteorology on the pollutant concentrations. The detailed evaluation method has been added in the revised version.

The time interval of training data should be daily at least for the distinguishing of emission and meteorology because the meteorological conditions varied greatly with the time. To date, most of the studies concerned the issue used the hourly data to train the model. Thus, we also used the hourly data to training the model. In most of the current machine-learning models, 10-fold cross-validation was a mainstream method to evaluate the accuracy of the developed model though some studies also used 7-fold or 8-fold cross-validation. In our study, we still used 10-fold cross-validation to assess the robustness of this model because the 10-fold cross-validation was more serious than 7-fold or 8-fold cross-validation, and thus predicted result based on 10-fold cross-validation might be more convincible. Although no strict regulation about the number of available data was required for RF modelling, the less training dataset led to the low accuracy of RF model because RF model was more suitable to the big-sample training. Based on our experience, the number of training dataset should be larger than 1000, or the modelling performance might be not satisfied. The training variables should include all of the meteorological data including wind speed (WS), wind direction (WD), air temperature (T), relative humidity (RH), precipitation (Prec), and air pressure (P) observed by the supersite system. Besides, some indicators reflecting the time variable should be also incorporated into the model, which might be associated with the lockdown intensity. Actually, the daily emission data might be more suitable to train the model compared with time variable. Unfortunately, the hourly emission data was not available in Tangshan.

The limitations of our study focused on the variable selection. Although the time variables were closely linked with the lockdown intensity, they were not less suitable to train the model compared with hourly emission data. Unfortunately, the hourly emission data was not available in Tangshan. To date, we only used the statistical indicators to confirm the reliability of the RF model. Some previous studies also used other statistical models and chemical transport models (CTMs) to remove the impact of meteorology. Overall, the results of statistical models and CTMs were broadly similar, and thus we believed that the result based on machine-learning approach was robust. For the study concerned about the COVID-19 lockdown, I think the machine-learning model might be more accurate than CTMs because the uncertainty of emission inventory was very large during COVID-19 lockdown period.

Comment 2: If I understand it correctly, the deweathered results can still contain the effects of changes from other natural or anthropogenic sources of the investigated chemical species or some substantial dependency of the variables on larger (e.g. seasonal or annual) time scales. Cf. also line

116. How was the possible interannual variability considered? The authors are requested to explained and discussed these items. The related interpretations of the atmospheric concentration changes and their quantifications are not convincing.

Response: Thank for reviewer's suggestions. In our study, the year was incorporated into the training model as a variable. However, the year showed the lowest variable importance for nearly all of the species as shown in Figure 6, Figure 7, and Figure 8. The result suggested that the impact of inter-annual variation of meteorology on pollutant concentrations was not significant. It was assumed that all of these training data were determined in 2020, and thus we did not need to remove the inter-annual variability in our study.

In our study, the original dataset was randomly classified into a training dataset (90% of input dataset) for developing the RF model and the remained one was treated as the test dataset. Based on this method, we can obtained the predicted concentration of each species. After the building of RF model, all of the meteorological data were normalized to remove their effects and then we incorporated the new training dataset into the same RF model to predict the concentration of each species, which meant the concentration without the impact of meteorology. At last, the differences of original pollutant concentrations and deweathered pollutant concentrations were regarded as the concentrations contributed by meteorology.

Comment 3: A large part of the results concerns secondary aerosol components. Their formation can depend sensitively on the solar radiation, which is missing from the set of meteorological variables. The uncertainties related to this missing factor should be discussed separately. Similarly, O_3 as a main representative of secondary pollutants is missing from the list, although it would be very exciting to see its modelled changes. At many urban or polluted locations, O_3 concentrations increased or stayed constant during the restriction intervals. The authors may want to explain why O_3 is not among the investigated gases, and to discuss the enhancement from the aspects of their results. (It is noted that the MS could gain from citing some earlier references in the field and published in this journal.)

Response: I agree with review's suggestions. Indeed, the solar radiation is an important factor especially for the O_3 formation. Unfortunately, the solar radiation was not determined by the instrument in the supersite of Tangshan. Thus, we have added the discussion about the uncertainty associated with the missing factor.

In the revised version, we have added 8-h O_3 concentrations and distinguished the emission and meteorology contributions to 8-h O_3 concentrations. The result suggested that the observed 8-h O_3 concentrations increased by 160% after COVID-19 lockdown. The increase of observed 8-h O_3 was contributed by emission (80%) and meteorology (80%) equally. More rapid decrease of NO_x than VOCs led to the increase of surface O_3 concentration because urban is generally VOC-limited region, which has been supported by many previous studies (Liu et al., 2020 ACP; Jiang et al., 2020 ACP; Liu et al., 2020 EST). The excessive decrease of $PM_{2.5}$ from primary emission significantly increased the HO_2 radical concentration on the surface of aerosol, thereby promoting the O_3 formation (Shi and Brasseur, 2020). In addition, the meteorological conditions played crucial role on the 8-h O_3 increase. As shown in Figure 7, air temperature (T) was also treated as the most important variable. On the one hand, the higher T generally enhanced biogenic isoprene emissions, which was the most abundant biogenic VOC and showed the highest ozone formation potential (Liu and Wang, 2020). On the other hand, high T often increased chemical reaction rates and accelerated the O_3 formation (Shi et al., 2020). Besides, wind speed (WS) also played an important role on the 8-h O_3 concentration. Shi et al. (2020) have demonstrated that weaker winds often slowed down the advection and convection of NO_x and VOCs, which was beneficial to O_3 formation.

Comment 4: Many parts of the MS are too descriptive/fuzzy or too long (e.g. Sect. 1, 3 pages) and more importantly, real interpretations and discussion are often missing. This should be improved, e.g., metal components and secondary inorganic aerosol constituents are to be discussed from the point of view of their regional sources/source sectors.

Response: I agree with reviewer's suggestions. Metal components and secondary inorganic aerosol constituents have been discussed from the point of view of their regional sources/source sectors. The detailed revision has been shown in the revised version (section 3.2).

Comment 5: Atmospheric concentrations can usually be described by log-normal distribution and, therefore, median descriptive statistics is preferred to mean (which is called average in the text). Furthermore, the authors should rethink their rounding off strategies all over the MS; e.g. the value of 63.5 (line 157) implies a relative uncertainty <1%, which is unusual in analytical chemistry or should be explained separately; or similarly, a value of 93.5% is questionable (line 163). Are the differences between some corresponding data pairs of the results, e.g. 0.33 and 0.28 (line 285) significant?

Response: Thank for reviewer's suggestions. The median was generally suitable to the dataset with large variation coefficient and subjective to log-normal distribution. However, in our study, the hourly concentration for all of the species did not show large variability. Moreover, most of pollutant concentrations in our study suffered from normal distribution rather than log-normal distribution. Therefore, we hope to remain the mean value instead of the median.

I agree with the suggestion about the rounding off strategies proposed by reviewer. Indeed, some redundant digits was meaningless. Thus, all of the percentage were kept in integer (e.g., 94%). For the pollutant concentrations, we kept 2 decimal places in most conditions in the revised version (e.g., 15 or 1.3).

Minor comments are listed as examples

Comment 6: Unusual citing practice in line 33.

Response: Thank for reviewer's suggestions. The error has been corrected.

Comment 7: Which transport sector is specifically meant?

Response: Thank for reviewer's suggestions. The transport sector mainly denotes the railway transport and highway transport.

Comment 8: Abbreviation $PM_{2.5}$ expresses the particles in the specified size fraction, while it is used in the sense of $PM_{2.5}$ mass in many places of the text (e.g. lines 44, 62, 101).

Response: Thank for reviewer's suggestions. Indeed, $PM_{2.5}$ denotes the size fraction, while the $PM_{2.5}$ concentration or the concentration of $PM_{2.5}$ represents $PM_{2.5}$ mass. In order to avoid the ambiguous expressions, these sentences have been changed into "the concentrations of $PM_{2.5}$" in the revised version.

Comment 9: Check the order of the words in lines 52-53

Response: Thank for reviewer's suggestions. The sentence has been changed into "Huang et al. (2020) employed the chemical transport models (CTMs) to infer that these extraordinary findings might be attributable to enhanced secondary pollution." (Line 52-54)

Comment 10: Unusual expression: neutralized the decreases (line 60).

Response: Thank for reviewer's suggestions. The "neutralized" has been changed into "counteracted" (Line 61)

Comment 11: The distribution of semi-volatile components between gaseous and condensed phases is completely missing from the explanations in line 68.

Response: Thank for reviewer's suggestions. The sentence has been changed into "Zhang et al. (2020a) also revealed that the release of primary pollutants and the generation of reactive semi-volatile products partitioned between gas and aerosol phases were strongly dependent on the temperature and relative humidity (RH)." (Line 69-70)

Comment 12: Change "knowledge was critical… in the future" to "knowledge is critical… in the future" in line 75.

Response: I agree with reviewer' suggestions. "was" has been replaced by "is".

Comment 13: Change "might take more sense" to "more sensible" in line 84.

Response: I agree with reviewer' suggestions. "sense" has been changed into "sensible".

Comment 14: What is meant by: standard sample. Is it a sample collected for off-line analyses, and if yes, what were its collection conditions? What does the expression "agreed well" specifically mean? All in lines 111–114.

Response: Thank for reviewer's suggestions. Trace elements in $PM_{2.5}$ were measured using a Xact multi-metal monitor (Model Xact 625, Cooper Environment Service, USA). Ambient $PM_{2.5}$ was collected using a cyclone inlet at a flow rate of 16.7 L min^{-1} and deposited on a filter tape. The mass concentrations of trace elements were determined using X-ray fluorescence (XRF). In each measurement, the detector energy gain was automatically adapted using pure Pd as an internal standard. The XRF response of the interest element was calibrated using the standard thin-film provided by the manufacturer. The measured mass agreed well with standard mass for each element and the deviation was less than 5%.

Besides, we also performed the off-line sampling of fine particles and then determined the concentrations of trace elements in $PM_{2.5}$ using Inductively Coupled Plasma Mass Spectrometry (ICP-MS) and Inductively Coupled Plasma-Atomic Emission Spectroscopy (ICP-AES). The main objective was to confirm the reliability of the online data. The detailed collection conditions are as follows: The daily $PM_{2.5}$ samples were also collected at the same site using a four-channel aerosol sampler (Tianhong, Wuhan, China) on 47 mm cellulose acetate and glass fiber filters at a flow rate of 16.7 L min^{-1}. A total of 60 filter samples were collected. In the laboratory, the elemental analysis procedures strictly followed the latest national standard method. The concentrations of Hg, Pb, K, Ca, Cr, Cu, Fe, Ni, and Zn determined by Xact showed the better correlation with those measured by ICP-MS (ICP-AES), indicating the higher accuracy of the trace element dataset.

Comment 15: Meteorological parameters were not resolved at their first occurrence (in lines 120-121).

Response: Thank for reviewer's suggestions. The full names of meteorological factors have been added in the revised version (Line 125-127).

Comment 16: Check the order of words in line 191.

Response: Thank for reviewer's suggestion. Throughout the manuscript in the revised version, all of the numerical values were rearranged from low to high or from high to low. All of the chaotic digital expression have been corrected.

Comment 17: Check Figures S1–S3 in lines 185, 188-189, etc. and similar typos.

Response: Thank for reviewer's suggestions. All of these typos have been corrected.

Comment 18: Captions of the figures are not self-explanatory or descriptive enough.

Response: Thank for reviewer's suggestions. The captions of all the figures have been significantly revised and many self-explanatory information were added in the revised version.

Comment 19: The reference list is often deficient in required details.

Response: Thank for reviewer's suggestions. We have revised the reference list based on the reviewer's suggestions and the required details have been added in the revised version.

Comment 20: The language and grammar should be improved.

Response: I agree with reviewer's suggestions. The language and grammar throughout the manuscript have been significantly revised.