The authors study the changes of chemical characteristics and health risks for high time-resolved trace elements in PM$_{2.5}$ in a typical industrial city of China and the changes during three years due to clean air actions. They separated the effects due to meteorology from those due to the control measures by the use of a random forest (RF) model. The subject is interesting and the use of 1-hour resolution data is not still widespread. However, the article can be published only after major revisions.

Response: Thank for reviewer's suggestions. These suggestions are beneficial to improve the quality of our manuscript. We have revised the manuscript carefully. Especially, we have reperformed the PMF analysis and recalculated the contribution concentrations and ratios of each source to all of the species. Besides, we also drawn the polar plots to add more proof to reveal the potential sources of trace elements in Tangshan. Furthermore, all of the comments have been responded from points to points. The detailed comments are as follows:

Comments 1: All the work is based on the analysis of the elemental concentration with 1-h resolution. At least in the supplementary material some graphs presenting the trend of the concentration of some elements over the three years should be presented.

Response: Thank for reviewer's suggestions. We have added some figures of time-series result of trace elements during 2017-2020 in SI (Figure S2 and Figure S3).

Figure S2 The time-series variations of Ag, As, Ca, Co, Cr, Cu, Fe, and Ga during 2017-2020.
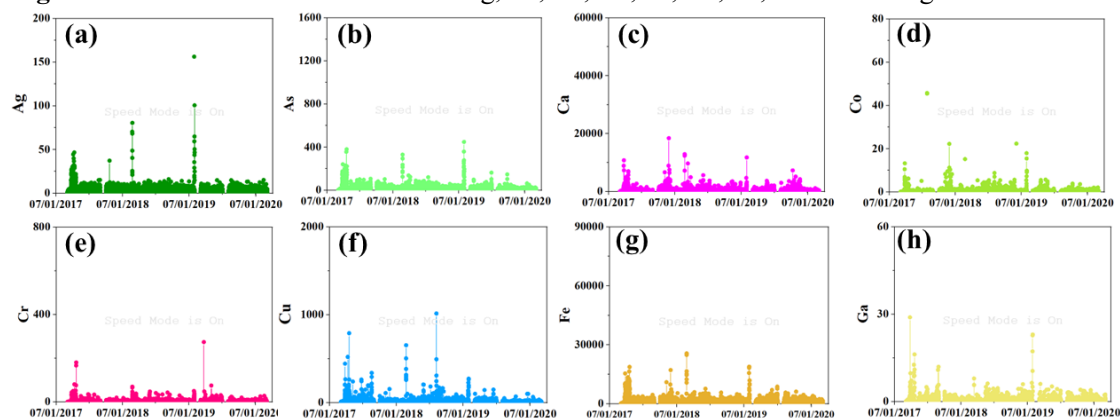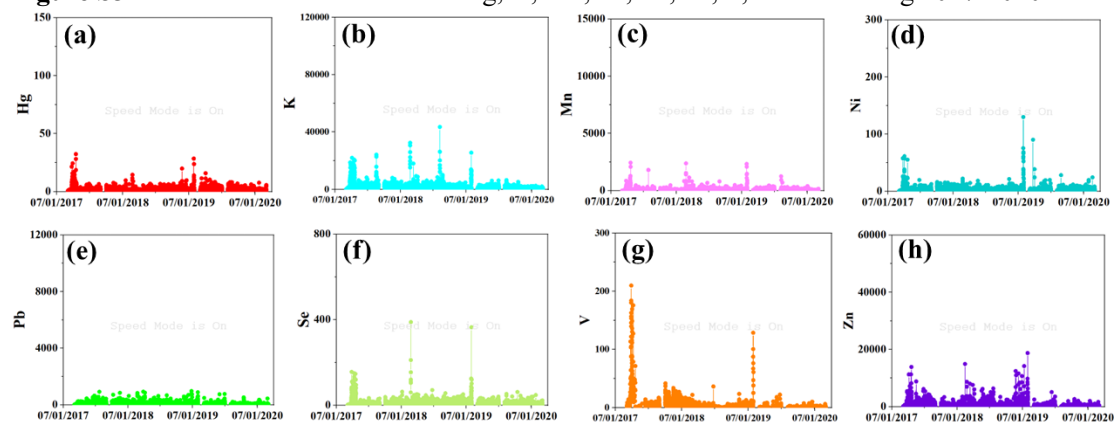

Figure S3 The time-series variations of Hg, K, Mn, Ni, Pb, Se, V, and Zn during 2017-2020.

Comments 2: Lines 58-59. More references should be added; many articles dealing with elemental concentrations also in Europe or USA can be found in the literature

Response: I agree with reviewer's suggestions. (Line 58-59) Many references about element concentrations in Europe and USA were also added.

Comments 3: L 77: Dall'Osto et al. (2013) is a good example of application of PIXE, but it was not the first one (see. e.g. P.Prati et al, Source apportionment near a steel plant in Genoa (Italy) by continuous aerosol sampling and PIXE analysis, Atmospheric Environment, 2000, 34(19), pp. 3149–3157 or A. D'Alessandro et al., Hourly elemental composition and sources identification of fine and coarse PM_{10} particulate matter in four Italian towns, Journal of Aerosol Science, 2003, 34(2), pp. 243–259)

Response: I agree with reviewer's suggestions. (Line 77-80) We have added these references in the revised version.

Comment 4: L 173: please use the extended name before introducing the acronym (RMSE, MAE)

Response: Thank for reviewer's suggestions. (Line 183) RMSE and MAE represent Root-Mean-Square Error and Mean Absolute Error, respectively.

Comment 5: L 203: The authors should add in the supplementary material the percentage of the elements reconstructed by PMF; it is necessary to assess the quality of the PMF analysis.

Response: Thank for reviewer's suggestions. We have added the percentage of the elements (factor profile (% of species)) reconstructed by PMF model (Figure S9). At first, more than 26,000 samples were trained to determine the optimal six factors with the lowest values of Q (robust) and Q (true). Besides, we also performed the DISP, BS, and BS-DISP test to assess the quality of PMF analysis (Table S4-6). The result suggested that the largest decrease of Q values in four scenarios were less than 0.1% based on DISP test. In addition, the matching rates for most of the factors reached 100% except the fourth factor based on BS test, indicating the model was robust. At last, the BS-DISP test also suggested the solution has been verified to be useful because the observed drop in the Q value was less than 0.5%. Base on all of these tests, we demonstrated that PMF analysis in our study was robust and reliable.

Comment 6: L 228: How was the total mass measured? The value 5.7% is the average of the hourly or daily values or something else?

Response: Thank for reviewer's suggestions. (Line 238) The total mass of $PM_{2.5}$ was determined by an oscillating balance analyzer (TH-2000Z, China). 5.7% is the average of the hourly values.

Comment 7: L 332: "Such a small change can produce the observed effects?

Response: Thank for reviewer's suggestions. The annual mean air temperature in Tangshan displayed a minor increase during 2017-2020, and the small change cannot affect the variations of trace elements. Then, we found the air temperature in different seasons experienced large changes during 2017-2020. (Line 347-349) During the spring and summer in 2017-2020, the average air temperature decreased from 8.9 and 27 to 7.2 and 26°C, respectively. The decreased air temperature led to the higher water content in the soil and a lower tendency of dust suspension, and might decrease the concentrations of Ca, Fe, and Cu (Manju et al., 2018; Yang et al., 2017; Lyu et al., 2016).

Comment 8: L 338-340: What about WD? It should be even more important than WS

Response: I agree with reviewer's suggestions. Indeed, WD might be more important than WS. During 2017-2020, the dominant wind direction did not experience significant change. However,

the mean wind speed from the southeastern direction displayed slight increase from 1.34 to 1.50 m/s. It was well known that V and Ni were usually emitted from heavy oil combustion of ocean-going ship engines (Zhao et al., 2013). Many coastal ports and ferrous metal smelting industries were located on the southeastern direction of the sampling site, and the enhanced WS might promote the dilution and dispersion of trace elements. Some previous studies have demonstrated similar results. For instance, increased winter monsoon during the negative phase of arctic oscillation promoted the advection and diffusion of air pollutants in North China (Cai et al., 2017 *NCC*). As shown in Figure S8, both of V and Ni concentrations showed gradual decreases along Northwest-Southeast transect.

Comment 9: Paragraph 3.4 The impact of clean air policy on source apportionment of trace elements: Do the authors believe that the source profiles change over the years? From the data (fig. S4-S6) it does not seem so. The absolute source contribution both to the elements or to the total mass change because the emissions are reduced. I don't understand why the authors don't perform PMF analysis putting all the years together. This would immediately allow to see how the contribution of the different sources is reduced over the years even without having the total $PM_{2.5}$ mass and how the absolute contribution of the different sources to the elements vary. Instead, it is difficult to assess the influence of clean air policies only by looking at fig.7, where the average contributions of the six sources to the total mass concentrations of metals in $PM_{2.5}$ is reported, which means a percentage contribution. The reduction of the percentage contribution of one source can produce an apparent increase in the percentage contribution of another source, even if the absolute contribution of that source is decreasing! Therefore, in my opinion, it would be more interesting to look at the source trend along the 3 years. The deweathering procedure can be performed also to these data.

Response: I agree with reviewer's suggestions. We have put all of the observations into the single list together to perform PMF analysis. Then, the contribution concentrations and ratios of six sources to deweathered trace elements were calculated. In order to compare the inter-annual trends of major sources directly, the variations of deweathered element concentrations in different years were considered in the model. Based on the result, the deweathered concentrations contributed by multiple sources in different years could be compared directly. Biomass burning and traffic-related dust accounted for the higher ratios because K, Fe, and Ca showed the higher concentrations among all of these elements. Based on our analysis, the contribution concentrations of biomass burning, non-ferrous metal smelting, coal combustion, and ferrous metal smelting to trace elements decreased from 1458, 479, 636, and 569 ng/m^3 to 901, 225, 230, and 345 ng/m^3, respectively. However, the contribution concentrations of heavy oil combustion and traffic-related dust displayed slight increase during 2017-2019, while they decreased rapidly after 2019. The contribution concentrations for nearly all of the sources to trace elements suffered from decreases during 2017-2020 because the total deweathered levels of trace elements experienced decreases in the past three years. However, the contribution ratios of these sources to trace elements did not show similar characteristics. For instance, the contribution ratio of traffic-related dust increased from 25% to 33%. However, the contributions of non-ferrous metal smelting and coal combustion decreased from 11% to 8% and 15% to 8%, respectively. The contributions of ferrous metal smelting, heavy oil combustion, and biomass burning remained relatively stable during this period. The detailed reasons for the variation trends have been introduced in the revised version (Line 432-463).

Comment 10: L 345: The use of the proper experimental error is mandatory in PMF analysis. How

were the experimental uncertainties on deweathered concentrations calculated?

Response: Thank for reviewer's suggestions. The experimental uncertainties were also determined by the equation in the guide of PMF analysis.

$$Unc = \frac{5}{6} \times MDL \ , \ \text{x} \leq MDL$$

$$Unc = \sqrt{(ErrorFraction \times Concentration)^2 + (0.5 \times MDL)^2} \ , x > MDL$$

where MDL denotes the detection limit; Unc represents the uncertainty; ErrorFraction is the relative standard deviation.

We also used MDL during the experiment for the Unc estimates. ErrorFraction was estimated based on the deweathered concentrations.

Comment 11: L 368-371(but the same is true for other sources). The identification of the origin of the sources based only on back-trajectories is too qualitative. Since the authors have wind data, they should produce element/source wind polar plots like the ones e.g. reported in fig 3 c of Y. Chen et al.: Simultaneous measurements of urban and rural particles in Beijing, Atmos. Chem. Phys., 20, 9231–9247, 2020

Response: Thank for reviewer's suggestions. (Figure S6-S8) We have added the element/source wind polar plots to indicate the potential sources of some elements. For instance, both of V and Ni displayed the higher concentrations in the southern wind direction (Figure S8), indicating the major sources of V and Ni focused on the South part of Tangshan. Meanwhile, it was well known that V and Ni were linked with the heavy oil combustion derived from shipping. Therefore, we inferred that both of V and Ni originated from coastal shipping emission.

Comment 12: 400.402: ref. Moreno et al, says something more complex. At the end they say "Thus there is considerable overlap between V/Ni values in natural mineral dusts and those in emissions from the combustion of refinery-produced materials, and this hinders use of this ratio in pollution source identification." Therefore the authors can keep the sentence, but they should be more cautious. Furthermore, is 1.2 the V/Ni ratio within the identified source (which can be easily obtained by the PMF analysis) or the ratio between the average concentrations during the sampling period? It is the former one which must be presented

Response: Thank for reviewer's suggestions. 1.2 is the ratio between the average concentrations during the sampling period. Besides, we also calculated the V/Ni ratio within the identified source using PMF model, and the V/Ni ratio in heavy oil combustion identified by PMF reached 1.7, which was also in the range of 1 and 3.

Comment 13: L411-412: The authors must show in the supplementary material the average day (average concentration for each hour of the day like the ones e.g. reported in fig 3 b of Y. Chen et al.: Simultaneous measurements of urban and rural particles in Beijing, Atmos. Chem. Phys., 20, 9231–9247, 2020) at least for Ca, Fe, Zn, to see the rush hour peaks.

Response: Thank for reviewer's suggestions. We have added the figures in the revised version (Figure S10) to reveal the diurnal variations of Ca, Fe, Cu, and Zn in Tangshan. These elements displayed the higher values during the rush hours.

Comment 14: L 422-423: I would not define the decreases as dramatic. The experimental uncertainties in source contribution should be also taken into account

Response: Thank for reviewer's suggestions. The "dramatic" has been deleted in the revised version. Indeed, the source apportionment showed some uncertainties and we cannot draw arbitrary

conclusions.

Comment 15: L 426: A change of 1% is within the experimental uncertainties due to the fitting procedure, I would not use the word "increase"

Response: I agree with reviewer's suggestions. (Line 452) The "increase" has been replaced by "remained relatively stable".

Comment 16: L 437-443: I agree with the authors that hourly data can give more valuable information regarding the health risk. However, I do not understand how they calculate the risk from the hourly data without averaging their data at least on a daily scale. EF represents the annual exposure frequency (d y−1), ADD is the daily intake (mg/kg/day) of trace metals, all these are quantities averaged at least on a daily base.

Response: Thank for reviewer's suggestions. We calculated the risk from the hourly data with averaging their data at a daily scale. In general, the daily value derived from the average of hourly data might be more reliable to assess the health risk compared with the daily trace element measurement.

#Reviewer 2

This paper reports occurrence levels of elements in $PM_{2.5}$ in Tangshan China, analyzes the temporal variations and evolution of $PM_{2.5}$-associated elements, assesses the contributions of emission and meteorology to these species, apportions sources of elements during the whole period and evaluated influence of pollution control measures on the changes of carcinogenic and non-carcinogenic risks. This manuscript tries to represent the observation results and address relevant scientific questions. The scientific methods and assumptions are almost valid and outlined so that conclusions are reached. The description of experiments and calculations are shown. Observed phenomena as presented in the text and SIs have been described in detail. However, there are some hypotheses in the manuscript lacking of crucial evidential data to support. The link between the observed data and suggested implications is not strong. I do not recommend the publication of this article in its current form except some points.

Response: Thank for reviewer's suggestions. I have significantly revised the manuscript based on reviewer's suggestions. We have added more proof to identify the potential sources of some trace elements and analyzed the reasons for the temporal variations of trace elements carefully. Many methods including PMF analysis, machine-learning algorithms, and polar plots were utilized simultaneously to support our conclusion. For instance, we identified the sources of V and Ni. At first, we found V and Ni showed the higher loadings in the factor, and the factor did not show good correlation with other trace elements. Meanwhile, it was well known that V and Ni were linked with the heavy oil combustion or shipping emission. As shown in Figure S8, both of V and Ni showed the higher concentrations in the southeastern part of Tangshan and the concentrations displayed gradual decreases along the Southeast-Northwest transect, which further demonstrated that both of V and Ni in the sampling site could be derived from coastal shipping emission in Southeast direction. In addition, we also used multiple techniques to identify the trace element sources or demonstrate some inferences. The detailed comments are as follows:

Comment 1: The authors should go beyond reporting the measurements, present informative interpretations of the data and provide worldwide implication rather than local attention.

Response: Thank for reviewer's suggestions. Although we submitted the manuscript to Measurement Report in ACP, it did not mean our study only focused on the data report. In our study, we selected Tangshan as a study region to assess the impact of clean air policy on trace elements in

PM$_{2.5}$ in China. In fact, the study region was well chosen because Tangshan showed good spatial representativeness for the assessment of clean air actions in NCP even in China. First of all, Tangshan possesses many energy-intensive industries including coal-fired power plants, non-ferrous smelting industries, textiles, building materials, chemical engineering, and papermaking industries (Ren et al., 2011). Second, coal to gas project has started to be implemented in North China (including Tangshan) and the energy structure suffered from significant change (Wang et al., 2020b) since 2017. Therefore, our study might reflect the impact of clean air actions on trace elements in PM$_{2.5}$ at the national scale rather than local attention.

Moreover, many conclusions in our study also have worldwide implications. For instance, we found that the large-scale control measures including the strict implementation of "coal to gas" strategies, optimization of industrial structure and layout, and improvement of oil quality significantly decreased the concentrations of Pb, Zn, and As. However, these measures cannot largely decrease the concentrations of some dust-related elements such as Ca and Fe. Other measures such as afforestation and watering for dust suppressing should be enacted. Besides, the trace metal measurement/control standard of vehicle exhaust should be implemented as soon as possible. Our study could provide many useful implications for policy makers around the world.

Comment 2: Please introduce clean air actions and pollution control measures in detail. Which control measures are used to improve air quality?

Response: Thank for reviewer's suggestions. I have added detailed clean air actions and pollution control measures in supporting information (SI). "*Since 2013, Chinese government imposed many emission controls to alleviate air pollution across China. First of all, the government department enhanced comprehensive air pollution control on industrial enterprises, non-point source control, and vehicle pollution control. Second, many industrial enterprises were required to upgrade industrial structure, to control new capacity with high energy consumption and to accelerate the elimination of backward production capacity. Third, the energy structure adjustment and utilization of clean energy were also accelerated. The fossil fuel consumption also has been strictly controlled and the energy efficiency were tried to improve. In addition, Chinese governments tried to optimize industrial layout. Chinese government also established the market mechanisms to improve pricing and tax policy. At last, many monitoring sites were constructed across China to investigate the near-time air pollution and the joint prevention and control measures were also enhanced*" has been added in SI to introduce the detailed air pollution control measures enacted in China since 2013.

Comment 3: Atmospheric trace metals include nutrient elements and hazard elements. The nutrient elements are beneficial for ecosystem safety.

Response: Thank for reviewer's suggestions. Indeed, a trace of nutrient elements such as Cu, Fe, and Zn are beneficial for ecosystem safety because most these elements were necessary for some biological organisms. However, the excessive accumulation of some biological essential elements such as copper (Cu), iron (Fe), and zinc (Zn) could initiate activation of inflammatory cascades in tissues and the induction of biochemical synthesis pathways by catalyzing the generation of reactive oxygen species (ROS) (Alies et al., 2013; Lopez-Cruz et al., 2016; Saffari et al., 2014). Hormesis effect is also suitable to these elements.

Comment 4: A comparative analysis was performed on the concentration of the trace metals in PM$_{2.5}$ observed in different cities. A scientific summary is necessary. Moreover, considering that the analytical methods and sampling duration were probably different for the available data set from the measurement and the literature review and uncertainty existed, the difference in the

concentration of trace elements should be obtained based on statistical analysis.

Response: Thank for reviewer's suggestions. We have added a table S3 in SI to introduce the detailed sampling duration, sampling site, and analytic method of trace element measurement due to the element concentration might be affected by these reasons. Then, we have rewritten the discussion about the element concentration comparison (Line 236-253). As shown in Table S3, we selected the trace element dataset mostly determined by XRF technique. Some dataset measured by ICP-MS or ICP-OES were not included into the model because the difference of analytic method might affect the data comparison.

Comment 5: Uncertainty on results of the random forest (RF) model (deweather) should be further shown in detail. Detailed input should be introduced into the model in detail.

Response: Thank for reviewer's suggestions. The detailed input has been introduced in details (Table S1). As shown in Table S1, the dependent variables (e.g., trace elements) and independent variables including meteorology and time variables were incorporated into the model. Besides, the unit and time resolution were also added in the table. For each trace element, time variables and meteorological parameters were utilized to predict the element concentration, and then the meteorological parameters were normalized to estimate the deweathered element level. The sum of deweathered element level and the meteorology-related element level equaled to the total element concentration.

(Line 505-513) It should be noted that our work is still subject to some limitations. At first, some elements such as Cr (0.5) and Ga (0.5) showed relatively lower CV R^2 values though they were generally higher than 0.5. These elements might show relatively higher uncertainties during the meteorology-normalization compared with other elements such as Cu (0.85) and K (0.85). Besides, few variables were applied to deweather the element concentrations, which might be responsible for the lower CV R^2 value for some elements. Due to the lack of available hourly emission inventory of each element, we only used the time variable to train the model. This method still suffered from some uncertainties, which should be improved by the establishment of near-time emission database.

Comment 6: Sufficient information was not provided by authors to validate the quality of the data. Being a field/experimental study, I am surprised by the lack of details on the data quality assessment and/or quality control methods, instrument calibration, how uncertainties are estimated. How to obtain meteorological parameters and $PM_{2.5}$?

Response: Thank for reviewer's suggestions. Hourly averaged meteorological parameters including air temperature (T), relative humidity (RH), air pressure (P), wind direction (WD), and wind speed (WS) during the sampling period were measured by a weather station with sonic anemometer (150WX, Airmar, Milford, NH, USA). The hourly mass concentration of $PM_{2.5}$ was determined by the particulate monitor (Thermo, FH62C-14). The routine procedures, including the daily zero or standard calibration, span and range check, station environmental control and staff certification, followed the Technical Guideline of Automatic Stations of Ambient Air Quality in Tangshan based on the national specification HJ/T193-2005, which was revised based on the technical guidance established by the US EPA. Quality Assurance and Quality Control (QA/QC) for the Xact measurements was implemented throughout the sampling period. The internal Pd upscale value was recorded after daily programmed test for the instrument. The limit of detection (LOD) for each element is depicted in Table S2.

Comment 7: No legends in some figures were given. Figure s has too crowded X-axe. Please modify them.

Response: Thank for reviewer's suggestions. We have modified these figures in the revised version. The legends have been added in the figures. Besides, the crowed X-axe in figure 4 has been also improved in the revised version.

Comment 8: Why As had relatively high contribution in factor 1? Why was no vehicular emission identified in the results of source apportionment? As well known, dust (fugitive dust, dust storm, soil dust and road dust) have high or moderate loading of Ca, Cu and Fe.

Response: Thank for reviewer's suggestions. As could be also released from the combustion of biomass fuels (Wiinikka et al. 2013 *Energy Fuels*). Therefore, As displayed moderate loadings in factor 1 (biomass burning). In our study, vehicle emission and dust resuspension were integrated into the same source in factor 6. The factor 6 displayed the higher loadings of Ca, Cu, and Fe, which could be sourced from the soil dust resuspension or brake wear.

Comment 9: The overall carcinogenic properties of mass PM itself should be taken into account. Emissions from gasoline and diesel engines and metal smelting and refining are considered Class 1A carcinogens by IARC; numerous populations studies have demonstrated increased (lung) cancer risk in communities with high urban PM exposures. Thus, the overall cancer risk from mass-based PM is likely to outweigh the risk of a limited set of elements. The authors seem to overinterpret their own risk assessment. If a realistic assessment of the actual health risks was involved, the results would be more meaningful.

Response: Thank for reviewer's suggestions. Indeed, the traditional method used in our study showed some weaknesses though the method has been widely used to assess the health risks of trace elements. Thus, we have reviewed many references to search some new health risk assessment methods and tried to re-estimate the potential health risks of trace elements in our study. Recently, Xie et al. (2019 EP) considered the impact of speciation on the health risk. Besides, Lyu et al. (2017 EP) assessed the health risks of trace elements in different sizes because the fine particles generally possessed the higher health risks compared with the coarse ones. Unfortunately, the high time-resolution trace element dataset in our study cannot consider the impact of speciation on health risk and cannot assess the contribution of particle size to toxicity. Therefore, these new methods are not suitable to our study. I also agree with reviewer's views that the emissions from gasoline and diesel engines and metal smelting and refining might show the higher carcinogens risk. However, we cannot quantify the risk difference of metal smelting and other sources. Moreover, the toxicity weights of different sources were not revealed based on previous studies. Hence, we cannot establish a new and reliable health risk assessment method of trace metals based on limited materials. In the future work, we have to perform some experiments to assess the toxicity of trace elements derived from various sources, and then to determine the risk weight. I think the improvement of health risk assessment method is a new topic to study.

Comment 10: Since RfD and HQ typically have (considerable) uncertainty/safety factors included, a HQ> 1 does not indicate "occurrence of adverse n-c effects" as claimed.

Response: I agree with reviewer's suggestions. (Line 488) Both of RfD and HQ showed the higher uncertainties. HQ > 1 only suggested that the trace metal showed the higher potential health risk rather than the definite health damage. We have corrected the improper expressions in the revised version.

Comment 11: Open biomass burning has been banned in China after the execution of clean air act since 2013, in particular after 2018. How authors judged that biomass burning or open waste incineration in NCP contributed to the concentrations of trace elements in Tangshan?

Response: Thank for reviewer's suggestions. In the revised version, we used the biomass burning instead of open biomass burning. Indeed, Chinese government imposed strict control measures on open biomass burning since 2013. However, the stealing emission still cannot be banned. Ke et al. (2019) (Spatial and temporal distribution of open bio-mass burning in China from 2013 to 2017) revealed the open biomass burning still released a large number of air pollutants since 2013. Indeed, the number of fire points in China showed dramatic decreases after 2018 due to more strict control measures for open biomass burning. Some small-scale biomass burning events cannot be forbidden especially in autumn harvest season and winter. In addition, the biomass burning in our study included open biomass burning, domestic biomass (biofuel) burning, and wildfire rather than open biomass burning alone. Both of domestic biomass burning and wildfire could also release trace elements to the atmosphere and the contributions of these sources might offset the decreases in the contributions derived from open biomass burning. Moreover, some key fingerprints such as K showed the higher loadings in the factor 1. Meanwhile, As and Se displayed the moderate loadings in this factor. It was well known that K was a typical element emitted from biomass burning (Cui et al., 2019 *JGR*). As could be also released from the combustion of biomass fuels (Wiinikka et al. 2013 *Energy Fuels*). Meanwhile, K displayed poor correlation with some crustal elements such as Ca ($p > 0.05$), and thus K in $PM_{2.5}$ of Tangshan might be not derived from dust resuspension. Therefore, we believed that these elements in $PM_{2.5}$ in Tangshan could be derived from biomass burning. The domestic biomass burning could make great contribution to the trace elements in the atmosphere of Tangshan.

Meanwhile, I agree with reviewer's suggestions. We cannot distinguish the accurate sources (e.g., open, domestic biomass burning, or wildfire) of K, As, and Se in our study. Therefore, we have replaced the open biomass burning and waste incineration by biomass burning in the revised version.