General comments: This manuscript quantified the relative contribution of meteorological conditions and emission control to the decrease of PM2.5 concentrations in Beijing using models. The results suggested that emission control was crucial for the air quality in Beijing, with a contribution of 80% to the decrease in PM2.5 concentrations using KZ filtering and WRF-CMAQ model. The topic is very interesting because the relative importance of influencing factors on air pollutants has been still unclarified. The method and result is helpful to understand the main influencing factors of air pollution and develop effective measures for pollution control and prevention in cities. I would thus recommend this manuscript to be published in ACP after improvement.

R: Thanks so much for your encouragement and useful suggestions. According to your comments, we have fully revised this manuscript. Please feel free to contact us if further revisions are required.

Specific comments:

Q1. L89-92. It is difficult to follow this sentence. Since DingLing Station and the MiYun Station as two background stations, what's purpose of choosing the Qianmen station and the Yufa Station? And the Yufa station can't be found in Fig.1.

R: Thanks so much for pointing this out. This is a good question. Beijing is a mega city with very large area. As a result, $PM_{2.5}$ concentrations and emission factors vary across Beijing, and this is the reason why we selected these urban stations, as well as two background stations (Dingling and Miyun Station), Qianmen (Transport station with intense emissions), and the rural stations with distinct $PM_{2.5}$ concentrations and emission factors. In this case, we can investigate that whether different emission-scenarios influence the relative contribution of meteorological conditions to $PM_{2.5}$ variations. And we are sorry that we did not make this figure clear to demonstrate the difference between different types of stations. The Figure 1 has been reproduced according to your comment in the revised manuscript and the Yufa station can be clearly identified in the updated Figure 1. Thanks so much for this.

Q2. L108. As far as I know, MECI emission inventory is only for 2012, 2014 and 2016, However in this paper, emission inventory in 2013 and 2017 were used. Please clarify more clearly.

R: This is a very good question. We are sorry that we did not make this clear. The Multi-resolution Emission Inventory for China, MEIC 0.5 °×0.5 ° emission inventory (http://www.meicmodel.org/) were updated annually. Therefore, the existing MEIC emission inventory are available from 2013 to 2016. Since 2017 MEIC is not available yet, we updated the 2016 MEIC emission inventory by considering the 2017 emission-reduction scenarios (e.g. the target of coal combustion reduction) required by the local government, a strategy that has been employed by previous

studies (Chen et al., 2019; etc.). We are sorry that we did not make the explanation of emission inventory clear in the previous manuscript. And in the revised manuscript, we explained that the 2017 MEICInventory was updated from 2016 MEIC.

Chen, Z. et al. (2019) Evaluating the "2+26" Regional Strategy for Air Quality Improvement During Two Air Pollution Alerts in Beijing: variations of $PM_{2.5}$ concentrations, source apportionment, and the relative contribution of local emission and regional transport. Atmospheric Chemistry and Physics 19(10):6879-6891.

Q3. L116. Which years' local environmental statistical data and reported emission data were used? From 2013 to 2017? Please clarify it. Did you compare the emission data of this Beijing local-emission inventory with others'? How about the difference? Since the emission is a basic for your research.

R: Thanks so much for this. The local environmental statistical data used for this research were from 2013 to 2017. For your reference, we compared our statistical data with Annual report from National Environmental Statistics Bulletin (http://www.mee.gov.cn/gzfw_13107/hjtj/qghjtjgb/) and the report from "2+26" center for air pollution prevention and control as follows. As you can see, the statistical data used for this research is highly consistent with other official data. The VOC value is very difficult to estimate and our data is very close to the data reported by the "2+26" center for air pollution prevention and control. Through this comparison, we believe the statistical data we collected for this research is valid for the following simulation analysis. Thanks again for pointing this out.

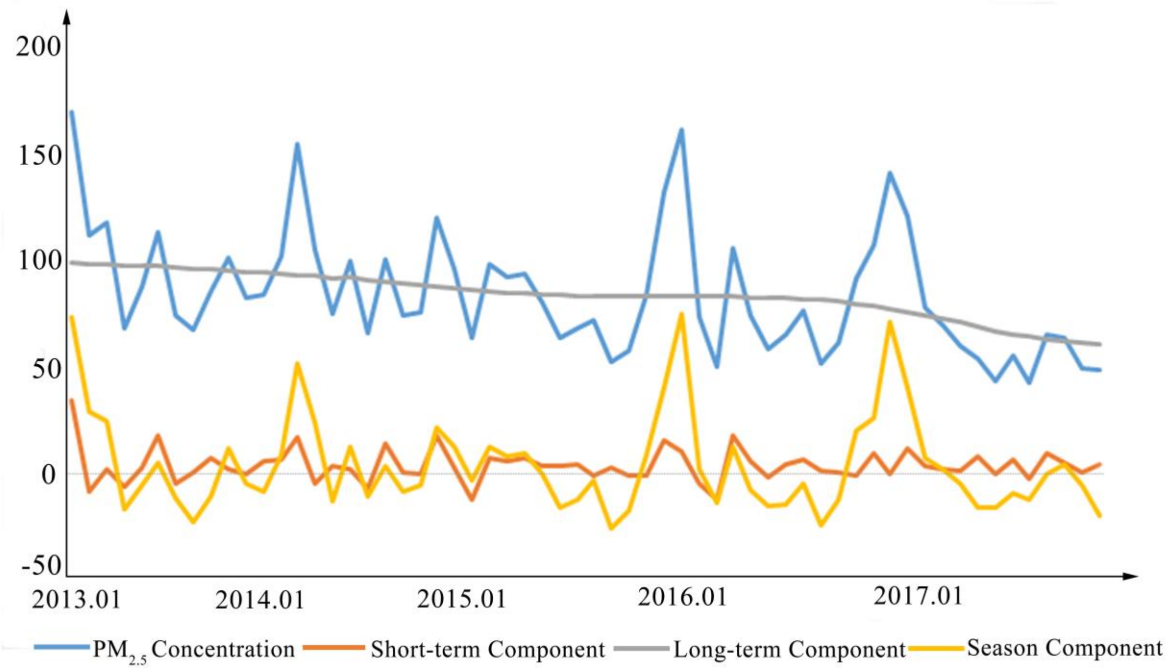
The comparison of The local environmental statistical data used for this research and other official statistical data in 2017 (unit: 10k tons)

	SO_2	NO_x	CO	VOC	NH_3	PM_{10}	$PM_{2.5}$	BC	OC
Statistical data for this research	1.38	10.15	49.54	13.47	3.20	14.74	3.92	0.17	0.44
National Environmental Statistics Bulletin	1.38	12.16	52.03	24.24	3.26	14.68	3.91	0.22	0.41
"2+26" center for air pollution prevention and control	0.89	9.24	48.98	13.93	3.16	13.82	3.72	0.19	0.46

Q4. Section 3.1. Filtering is a key research method for this study, which decomposes the original signal into trend signal and seasonal signal and the disturbance. Although the contribution rate in Table 3 partially reflects the composition of the decomposition, a time series diagram is still necessary to show that the components are correct after filtering.

R: Thanks so much for this comment. Yes, the accuracy of KZ filtering is closely related to the reliability of our research and how to judge whether the result of KZ filtering is satisfactory. According to your comments, we again reviewed the general principle of KZ filtering and other papers that employed KZ filtering, and found that no other criteria is supplied to verify KZ filtering. The commonly used criterion is that the closer the sum of three components to 1, the better filtering accuracy is. This is because a large value of the sum of the three components indicates that a majority of meteorological influences to $PM_{2.5}$ variations have been considered in

the KZ filtering. And for our research, the sum of three components for all stations are very close to 1, indicating a majority of meteorological influences has been filtered through KZ filtering. According to your comments, here we presented a KZ filtering curve as follows. Although we cannot easily judge the quality of KZ filtering according to the decomposed KZ components, we can see that the long-term component demonstrates a smooth curve whilst the trend of season component and short-term component is highly consistent with that of the original $PM_{2.5}$ time series, especially for some simultaneous peaks. This means the extracted seasonal component and short-term component made a significant contribution to seasonal and short-term variations of original $PM_{2.5}$ concentrations in Beijing from 2013 to 2017, indicating a satisfactory KZ filtering result. Thanks again for pointing this out and we have added relevant explanation to the revised manuscript.



Long-term, seasonal and short-term component extracted from original $PM_{2.5}$ time series in Beijing from 2013 to 2017 using KZ filtering.

Q5. Section 3.2.2. Model evaluation is the key point in this paper. If the model data is not consistent with observation, contribution of emission control is out of the question. It seems that lots of data are far from the observation especially during the heavy air pollution days. So it is better to convert Fig 2 to time series plots, which can tell us more detailed information about the model evaluation.

R: Thanks so much for this comment. Generally, model-simulation cannot perfectly fit the actual curve of $PM_{2.5}$ concentrations due to the deficiency of emission-inventories, the incomplete descriptions of reactions mechanisms for secondary formation of $PM_{2.5}$ and other uncertainties. For long-term simulation based on unified parameter setting, the model simulation outputs can demonstrate notable difference with the observed

$PM_{2.5}$ concentrations, especially during heavy pollution episodes (Li et al., 2011). This is because commonly employed CTMs do not fully consider heterogeneous/aqueous reactions, which significantly deteriorate $PM_{2.5}$ pollution (Chen et al., 2016). So yes, you are right. It is a common challenge that long-term CTM simulation may significantly underestimate $PM_{2.5}$ concentrations during heavy pollution episodes. And this is also the situation of our simulation. However, despite the relative large bias during the heavy pollution episodes, the general simulation accuracy: the correlation coefficient R, normalized mean bias (NMB), normalized mean error (NME), mean fractional bias (MFB) and mean fractional error (MFE) between observed and simulated data was 0.69~0.74, 11%~17%, 20%~27%, -21%~-17%, and 15%~27% respectively is satisfactory (EPA, 2005; Boylan et al., 2006). Following your constructions, we have convert the Fig 2 to time series plots. In addition, we acknowledged that the model simulation produced some large bias during heavy pollution episodes, caused by the common limitations of CTMs.

Li, G., Zavala, M., Lei, W., Tsimpidi, A.P., Karydis, V.A., Pandis, S.N., Canagaratna, M.R., Molina, L.T., 2011. Simulations of organic aerosol concentrations in Mexico City using the WRF-CHEM model during the MCMA-2006/MILAGRO campaign. Atmos. Chem. Phys., 11: 3789-3809.

Chen, D., Liu, Z., Fast, J. , Ban, J. 2016. Simulations of sulfate–nitrate–ammonium (sna) aerosols during the extreme haze events over northern china in october 2014. Atmospheric Chemistry and Physics, 16(16), 10707-10724.

Q6. Section 3.2.2. You verified the accuracy of the WRF-CMAQ model using the data of three stations. How about other urban stations? This does not mean that the figures of all stations should be supplemented, but it does require that extrapolation to difference between observed and WRF-CMAQ simulated PM2.5 concentrations.

R: Thanks so much for this comment. In the previous manuscript, we verified the data of three stations and following your comments, we further added the verification of another three urban stations to the revised manuscript to prove the reliability of the model simulation. Thanks again for this valuable comment.

Q7. L339. How did you get the conclusion "KZ filtering provides a more reliable method"? Just because the KZ filtering was station-based and WRF-CMAQ model was a grid-based? The averaged relative contribution of meteorological variations to PM2.5 reduction using the WRF-CMAQ model was very similar to that using KZ filtering. Verification is very important for the model results. So what's the criteria for judging reliability of your model?

R: This is a very good question. The advantage of KZ filter is that this statistical method is based on the observed meteorological data and $PM_{2.5}$ concentrations and predicts the variations of airborne pollutants on the hypothesis of unchanged meteorological conditions. In this case, by comparing the original and filtered time series of airborne pollutants, the relative contribution of meteorological conditions to long-term variations of airborne pollutants. Since KZ filtering is based on observed data, and simply consider the influence of time-series meteorology data on $PM_{2.5}$ time series, less uncertainty is involved in this analysis, KZ is influenced mainly by the variations of meteorology-$PM_{2.5}$ interactions in different areas and seasons. On the other hand, CTMs, e.g. WRF-CMAQ or WRF-CAMx considers both meteorological conditions (which is large-scale meteorological data, not as accurate as local observed meteorological data) and anthropogenic emissions for estimating $PM_{2.5}$ concentrations under different emission scenarios. Therefore, the accuracy of these models are not only decided by proper understanding of meteorological data, but also the reliability of emission inventories and proper descriptions of reaction mechanisms for $PM_{2.5}$ production, especially during heavy pollution episodes, which is a major challenge for current model simulation. For instance, without consideration of heterogeneous/aqueous reactions between sulfate, nitrate, and ammonium (denoted as SNA) in high-humidity environment, WRF-CAMx failed to simulate maximum $PM_{2.5}$ concentrations during extreme haze episodes (Chen, D. et al., 2016). And the emission inventories, no matter how fine they produced, are quite different from actual emission situations. Therefore, model simulated $PM_{2.5}$ concentrations, especially the relative contribution of anthropogenic emissions to $PM_{2.5}$ concentrations, are influenced by much more factors than the KZ filters. In this case, KZ filtering is most suitable for quantifying the relative contribution of meteorological conditions to long-term variations of airborne pollutants and recently been increasingly employed for this type of research.

On the other hand, since the emission inventory includes different emission sources, CTMs, e.g. WRF-CMAQ or WRF-CAMx, are suitable for quantifying the relative contribution of different sources to $PM_{2.5}$ variations, though large variations remained.

Q8. L398-399. Supplement the correlation coefficient between wind speed and PM2.5. And how about the influence of the other meteorological parameters (such as T, RH, wind direction) on PM2.5?

R: As detailed explained in our previous studies (Chen, Z et al., 2017; 2018), the causal influence of individual meteorological factors on $PM_{2.5}$ concentrations cannot be precisely quantified using correlation analysis, as the complicated interactions between different meteorological factors. Instead, a robust model CCM, which can remove the influence of other influencing meteorological factors, has been employed in our research to extract the dominant meteorological factors for $PM_{2.5}$ concentrations in Beijing and other mega cities across China. The ρ value, similar to

correlation coefficient, is a quantitative and more reliable indicator of meteorological influence on $PM_{2.5}$ concentrations. Detailed information concerning the influence of many meteorological factors on $PM_{2.5}$ concentrations can be found in our research (Chen, Z. et al., 2017; 2018). Here, we listed part of the table here as below for your reference. Meteorological influences on $PM_{2.5}$ concentrations vary across seasons, and SSD (sunshine duration), wind speed, humidity and temperature are major influencing factors for $PM_{2.5}$ concentrations in Beijing and other mega cities in China, especially the North China plain. Meanwhile, wind direction was not significantly correlated with $PM_{2.5}$ concentration in Beijing (Chen et al., 2017, 2018). The major reason is that the influence of wind direction on $PM_{2.5}$ concentrations is subjected to geographical conditions and not strongly correlated to $PM_{2.5}$ conditions (Chen et al., 2017).

Therefore, for this research, based on our previous studies on $PM_{2.5}$-meteorology interactions in Beijing, the major meteorological factors temperature, humidity, wind speed and solar radiation are used for the KZ filtering.

Chen, Z.Y., Cai, J., Gao, B.B., Xu, B., Dai, S., He, B., Xie, X.M., 2017. Detecting the causality influence of individual meteorological factors on local $PM_{2.5}$ concentrations in the Jing-Jin-Ji region, Scientific Reports, 7:40735.
Chen, Z.Y., Xie, X., Cai, J., Chen, D., Gao, B., He, B., Cheng, N., Xu, B., 2018. Understanding meteorological influences on $PM_{2.5}$ concentrations across China: a temporal and spatial perspective, Atmos. Chem. Phys, 18: 5343-5358

The correlation coefficient and ρ value of different meteorological factors (Temperature, humidity and wind speed) on seasonal $PM_{2.5}$ concentrations in Beijing (Chen et al., 2017)

City	Spring	Summer	Autumn	Winter
Beijing	RHU** (0.532, 0.490)	RHU** (0.648, 0.546) SSD** (−0.447,0.324) TEM** (0.554, 0.455),	RHU** (0.587, 0.555), SSD** (−0.509, 0.410), WIN** (−0.468,0.223),	RHU** (0.738,0.738), SSD** (−0.715,0.577), WIN** (−0.558, 0.531)

****Correlation is significant at the 0.01 level (2 tailed); *Correlation is significant at the 0.05 level (2 tailed).**
The first value in the brackets presents the correlation coefficient between the meteorological factor and PM2.5 concentration.
The second value presents the quantitative influence of individual meteorological factors on local PM2.5 concentration (ρ value), whilst the feedback effects of PM2.5 on these meteorological factors are not listed here.
NA indicates that no significant correlation exists between the meteorological factor and PM2.5 concentration.

Q9. Contribution of local emission-reduction measures was discussed in this paper. Please describe the reduction amount of each pollutant (SO_2, NO_x, PM) from each measure (e.g. coal boilers, residential use, industrial restructuring). It is better to discuss the contribution of reduction of different pollutants, which could reflect the effect of primary emission and secondary formation.

R: Thanks so much for this suggestion. In the past several months, including our research, there are some recent publications to discuss the variations of emission factors from 2013 to 2017 in Beijing and the underlying drivers for this. Specifically, UN published a formal report on air pollution in Beijing in the past two decades and released some official statistical data for the emissions of different pollutants in Beijing from 2013 to 2017. Therefore, in the revised manuscript, we fully reviewed these relevant studies and conducted an in-depth discussion on how emission-reduction measures have changed the relative contribution of different sources to $PM_{2.5}$ concentrations in Beijing. Thanks again for this valuable comment, which improved this manuscript significantly.

Technical corrections:

L40. I suggest that authors change keyword "PM2.5 reduction" to "PM2.5".
R: Corrected.

L144. Supplement the time period for "a historical record".
R: Actually, there are multiple haze episodes in December 2012 and January 2013, and the historical high record was observed during these episodes, no specific time period given by previous studies (e.g. Zhang, R., 2013).

L184. Supplement the link for "the website PM25.in".
R: Corrected.

P183. Check and revise the Formula (9).
R: Corrected.