Response to referee #2 (acpd-11-C15296-2012)

J. P. Zhang et al.

qzhang@pku.edu.cn and zhangjp@pku.edu.cn

We thank the reviewer for his/her comments and suggestions for improvement. We have fully considered the comments, and carefully revised the manuscript accordingly. Our point-by-point responses are detailed on the following pages. The reviewer's comments are outlined in italic type and our replies are in regular type.

General Comments:

In this work, the authors investigate how different regional circulation patterns are correlated with different environmental variables over Beijing. The authors used principal component analysis, in connection with reanalysis data of surface pressure variables over East Asia, to look at orthogonal air circulation patterns. They then use this daily average circulation classification to look at how it is correlated with different measurements of some visibility and pollution records.

This is a very interesting and important work, one that has a good potential to help guide further science and policy developments. I would ultimately like to see the results of this effort published. However, given the current issues that I have with many of the details of how these correlations were established, and given that many issues exist with the fundamental statistics and treatment of the modeling environment, at the present time I recommend that the paper be rejected. I would strongly encourage the authors to re-work through the following issues, and if successful, to resubmit their work again.

Reply： We appreciate the reviewer for his valuable comments and efforts for improving our manuscript. The major issues that the reviewer concerned were about the reliability of measurement data (particularly for AOD) and statistics, and the designing of modeling. We have addressed all concerned issues in detail, and made a significant revision of our manuscript. Since there is misunderstanding about the purposes of our study as well as the model simulation and utilization of satellite data, we would like to state them clearly.

This work is to establish long-term daily indices of synoptic-scale circulation types (CTs) for the North China region and their relations to air quality in and around Beijing, with the understanding that synoptic-scale circulation drives the physical processes that control the accumulation, dispersion, and transport of air pollutants. This is the first attempt to study the weather-air quality relationship using an objective procedure of circulation classification in China. This approach and the established correlations could also be useful for developing an operational forecast and warning system for air pollution and for examining the impacts of climate change on air quality. In this study, we focus on the classification of circulation types and their impacts (i.e.

correlations) on air quality.

To deepen the understanding of the CTs associated with air quality in Beijing, FLEXPART-WRF model was used to analyze the characteristics of backward trajectories of air masses (not a specific pollutant) for each CT. For this purpose, the model is competent and can provide valuable information about atmospheric dispersion and transport. Moreover, we used the high resolution data and plume (ensemble) particles to reduce the trajectories error. We do not expect quantitative analysis of the pollutant sources and the details of physical (i.e. transport, coagulation, deposition) and chemical processes (e.g. non-linear transformation). Apparently, the established associations between CTs and air quality are the results of these physical and chemical processes. Due to the limited space of paper, the detailed analyses that may need chemical transport model to resolve can't be included in this study. We would like to present the further analyses in our forthcoming work.

One specific reason for rejection is that many of the correlations are based on inappropriate statistics, and therefore I believe may not be valid. For example, the authors are attempting to correlate the daily average meteorological fields with AOD from AERONET, yet the AERONET daily average data from the stations mentioned is missing from more than half of the days. While it may be possible to look at the correlations on a month by month average, this is not necessarily easy to do, due to gaps in the data. A similar argument is made for comparison with the MODIS data, which has a far lower frequency than that of AERONET over this environment.

Reply: In this study, three different datasets of atmospheric extinction (i.e. visibility, AERONET AOD and MODIS AOD) were used to investigate the relationships between synoptic circulation and optical air quality. While all three of them have their own unique strengths and weaknesses, they complement each other to provide more holistic information about atmospheric extinction. In-situ visibility shows the characteristics of horizontal optical extinction, while the AERONET AOD and MODIS AOD are capable of demonstrating the vertically integrated optical extinction. AERONET AOD is available only during cloudless daytime, and is comparable with respect to different circulation types. At least, the AOD data was useful to assess optical air quality and the aerosol direct radiative forcing associated with CTs during clear-sky conditions. Interpolation was done to get sufficient AERONET AOD data samples for classification analysis. And the distribution of valid measurement days were generally similar to the frequency of CTs duing 2000-2009. Moreover, we have extended the ending time of AERONET AOD data to December 2009 in the revised manuscript.

In spite of the low temporal resolution, the MODIS data can provide valuable information about regional distribution of aerosol optical thickness. For this purpose, we do not expect the day-to-day variations of AOD, but just focus on the averaged state of regional distribution of AOD under a specific CT. On average, about 150 days data are used to calculate the mean of AOD around Beijing. We think these averaged MODIS AODs associated with circulation patterns are justified and meaningful. Similar averaged method, for example, was utilized to study the day-of-week variation of air pollution (e.g. Xia et al., 2008). Further, we recalculate the associations using long-term (2000-2009) MODIS 10-km AOD data instead (Figure N8), which provide more than doubled measurements for type analysis. Consequently, we believe that the established correlations between CTs and regional distribution of AOD are justified and reliable.

In summary, we think that the correlations of MODIS and AERONET AODs with circulation patterns are justified and meaningful. We also have addressed that the correaltions between CTs and AOD were established during the cloud-free conditions in the text. The methods for data quality control and statistics has been addressed more clearly.

A second specific reason for rejection is that nowhere in this effort are considerations for important effects other than transport and meteorology considered. There seems to be no consideration for the state of the boundary conditions, nor does there seem to be consideration of the highly non-linear chemistry and physics that the aerosols and reactive gasses undergo. While some of these are based at least in part upon the temperature, relative humidity, and rates of mixing, others need to be resolved at the grid scale. There have been many efforts to look at this, both in a Lagrangian as well as Eulerian framework, and this must be taken into account, if we are to try to accomplish any type of correlation between the impacts of meteorology and the observed results.

Reply： We should note that the Lagrangian dispersion model (LPDM; FLEXPART) model also calculated the atmospheric boundary layer height using WRF output data (Fast and Easter, 2006) along with the trajectories calculation. The turbulence, convective processes and atmospheric stability (Stohl et al., 2005) were also included in the modeling. Besides the 3D wind field, some basic meteorological quantities (i.e. temperature, specific humidity, surface pressure and cloud cover) were also used in the simulations (Stohl et al., 2005). FLEXPART-WRF model was used to analyze the characteristics of backward trajectories of air masses (not for a specific pollutant) for each CT. For this purpose, the model is competent and capable of providing valuable information about atmospheric dispersion and transport. Moreover, the high resolution data and ensemble particles were employed in the simulation to reduce the trajectories error and for improving our understanding of atmospheric transport. Various transport patterns could well be described by these footprints analyses, especially for the recirculation and convergence patterns (shown in Fig. S6 in the supplement), which were not resolved via (low-resolution meteorological data or) simple back trajectory calculations. The innovation here is that our study uses the LPDM output which (1) is more quantitative than just trajectory positions, (2) is calculated from high-resolution data, (3) utilizes a new method of releasing particles (in a box with 4 × 4 km^2 horizontal extent and 0–50 m vertical height; considering horizontal and vertical representativeness), and (4) also includes effects of turbulence and convection, which normal trajectories ignore.

We agree with the reviewer that the concentration levels of the observed air pollutants are controlled by the physical (transport, coagulation, wet/dry deposition) and chemical processes (non-linear transformation). The established associations between CTs and air quality are the results of these physical and chemical processes. In other words, those processes have been considered in meteorological perspective. We agree chemical model simulations are useful for improving our understanding of atmospheric transport and chemistry. Each method has its own strengths and weaknesses. It would be dangerous to eliminate the diversity of science by claiming that a given problem should be solved only with one particular approach. It should allow exploring alternatives that will eventually lead to better ways of solving the problem. The chemical model may be useful but it is also not perfect. The LPDM could well describe the dispersion and transport (that is the main purpose of our study), but could not include the chemical reactions. The chemical simulation could not give detailed transport patterns via backward plume trajectory calculations. Since the chemical modeling is not the theme of this study, the LPDM is more appropriate. We would like to present the further analyses about chemical reactions for the CT-air quality relations in our forthcoming work.

Specific Comments

It seems that reanalysis data (at 1ox1o) has been used for downscaling with WRF to higher resolution. I am wondering why it is that the model has not been run using a more typical approach whereby the regional model is run continuously over the entire period from 2000-2009, with updates at the boundaries and some type of nudging applied. I have not previously seen the model run with continuously updating initial conditions, just for 1.5 days, and then reset each day. Perhaps there is precedent for this, but if it is the case, I would like to see this documented. I would be interested to know, in this case, if there are gaps or other differences in the fields observed from day to day?

Reply： Because the skill of limited-area models (LAM; regional models) decreases very rapidly with time, the re-initialization by subdividing a long-term continuous integration into short ones is successful in mitigating the problem of systematic error growth in long integrations (Lo et al., 2008). This sequence of short runs with many re-initializations has been shown to outperform long-term continuous simulations with only one initialization (Pan et al., 1999; ŽAgar et al., 2006; Lo et al., 2008), and is becoming increasingly accepted and adopted in recent years (ŽAgar et al., 2006; Jiménez et al., 2010; Jiménez et al.). Recent papers (e.g. (Bei et al.) also documented the same method of frequently updating initial condition (just 1 day interval) for long-term WRF simulation and plume trajectories calculations.

The nudging scheme (or Newtonian relaxation) is easy to perform and widely used in long-term continuous simulations. However, the nudging approach may have some significant negative side-effects for this study. For the purpose of dynamical downscaling, we hypothesize that such utilization in the LAM can generate realistic regional structures that can't be resolved by the coarse-resolution forcing data. Beijing and its vicinity do have complex topography and land uses, and accordingly complicated air circulations. When the nudging is performed, the simulated state (realistic mesoscale structure) is forced to keep close to the driving state at coarse-resolution NCEP data (unrealistic state). A big difference in solution between our LAM and the global model tends to result in the systematic error. The nudging-performed time points can also potentially cause unrealistic simulated structure, thermo-dynamical imbalance and gaps originated from different data.

In light of the above-mentioned issues, we used the re-initialization approach. The mesoscale meteorological model has high skill in 36 hour forecasting. And the spin-up time is the first 12 hours. Each method has its own strengths and weaknesses. We choose the laborious re-initialization method (1 day interval) to improve the accuracy of the meteorological fields and to limit the gaps. The gaps in the time series of the simulated meteorological parameters (wind direction and speed, temperature, relative humidity and column integrated water vapor) are generally not significant (Figs. S2–4 in the supplement) and acceptable. In summary, this re-initialization approach is justified and appropriate for providing accurate meteorology for trajectory calculations.

The authors state that PM10 and BC are measured from the time period from 2005-2009. Yet their correlations are used against meteorological data from the time period from 2000-2009. Are the results the same when the same time periods are considered? Similarly, with respect to O3, SO2, NO2, and CO measurements, which seem to have been from gathered only from 08/2006-10/2008.

Reply: The correlations between local meteorology and CT are derived based on ten year data set, and have generally have passed the statistical significance test. Moreover, the table S5 in supplementary materials shows the relations of meteorological parameters with circulation types during 2005-2009. Their correlations are very similar to those during 2000-2009 although minor difference remains. Therefore, the correlation is stable. The gaseous pollutants data were only available during the CARE-Beijing campaign (i.e. 08/2006-10/2008). To get more general idea, we think the concluded stable correlations between CT and meteorological parameters during 2000-2009 could be used for interpretation of the variations of gaseous and particulate pollutants in different subset period.

The 500nm AOD data from the AERONET sites Beijing and XH have many individual days missing from their data, when looked at over the period from 03/2001-10/2008. How are these missing data points considered? It seems that the coverage during certain months of this spanning data set are quite small, and in some cases do not exist. Does this allow for a statistically representative correlation with the daily circulation type to be accomplished?

Reply: We extended the period of AERONET AOD data from 03/2001 to 12/2009. The 500 nm AOD was also obtained by interpolating from the 440 nm and 675 nm AODs. Therefore, there were more sufficient samples for classification analysis. But, there were still some missing data of AOD from the two AERONET sites. Time series of valid daily data in each month during 2001-2009 are shown in figure S5 of the supplement. For PKU site, only two periods seriously missed many measurements: 1) from June 2001 to April 2002; and 2) several months around 2008. We also recalculate the correlation using the data from May 2002 to December 2009, and we get very similar result. For XH site, the missing data were more serious. But, after August 2004, the AOD are generally intensively measured. Based on the same period data (August 2004 – June 2009), the relationships between CTs and AOD at Beijing and XH sites were still stable. Therefore, we think the missing data did not have significant impact on the resulted correlations.

What is the top of the domain used in the WRF downscaling at the 36km, 12km, and 4km grids? Is it sufficiently high to capture the effects of the transport of Dust from the Northwest, which are known to make an important contribution to the aerosol loading over Beijing.

Reply: The top height of WRF domains is 50 hPa (about 20 km altitude). This height is a typical option for mesoscale models. It is sufficient to capture the transport of pollutants in troposphere and lower stratosphere.

Some of the comparisons between the WRF fields do not seem to compare well against some of the meteorological variables from the airport. Does this disagreement potentially translate into the EOF decomposition for the circulation types?

Reply: We added further evaluation of the simulated meteorological variables with observations from surface and upper air (Figs. S2-4 in supplementary materials). Meanwhile, the statistical error analyses were performed. The results suggest generally good agreement between the simulated and observed meteorology although some difference remains. We think the disagreement potentially originated from the error of model, initial and boundary conditions, and observation measurements, and the different scales for the comparisons. It dose not relate to the EOF decomposition.

There are three important issues built into the use of FLEXPART that make it problematic in this case, and extremely important to address if it is used in this context. (1) Given that the lifetime of aerosols and some of the gas-phase tracers looked at here are a few days to a week, why has FLEXPART only been run for 2 days? (2) Since some of the pieces of data being compared against are column values, such as AOD, should not FLEXPART be run backwards throughout the atmosphere, or at least up to the middle troposphere. Given that dust is an important source of both PM10 and AOD over Beijing, and given that most is imported by long-range transport, it would seem important to consider this contribution. (3) FLEXPART is a fair model for following relatively unreactive species, such as air molecules and CO, however, when being used for highly reactive species, diffusion to and from surrounding parcels, as well as removal from or addition to the parcel due to wet uptake, evaporation, condensation, coagulation, etc. need to be accounted for. How is that done in this case, and if it is not done, how much error would it add?

Reply: 1) Since the reasons for choosing 48-h integration of trajectories have been addressed in our previous point-by-point response, here we just briefly restate them. The purpose of our paper is to investigate the regional transport pathways of air masse (not for a specific species) under certain circulation types. The 48-hour length of the backward trajectories was chosen as a trade-off in order to sample adequately the history of the air masses over the region of interest, while limiting the trajectories error (Stohl, 1998; Kahl and Samson 1986). When the characteristics of regional transport pathways are concerned, 48-h of integration is a typical choice for backward trajectories calculation (de Foy et al., 2011; de Foy et al., 2009; Guo et al., 2009; Robinson et al., 2011; Subramanian et al., 2010; Wang et al., 2010). For the reasons mentioned above, we think the 48-h retroplume trajectories could generally characterize the regional transport and dispersion, and could provide useful information about the potential emission sources.

2) Two-year air craft measurements showed that the surface level aerosol concentration was a key factor which greatly alters the value of AOD (about ±80% compared to the average value) over Beijing and its surroundings (Liu et al., 2009).

Therefore, the back trajectories released at the bottom 50m could usually explain the most change of AOD. Among the presenting literatures, there are usually two methods to vertically release the particles for the Lagrangian dispersion calculation. One is releasing particle at a fixed height (i.e. 100m (Zhu et al., 2010)); the other is tracing the plume particles at a ranged height (i.e. 0-50m level (de Foy et al., 2007; de Foy et al., 2009; Aiken et al., 2010)). The latter method is apparently more justified, and has been widely adopted in recent years. The reviewer may indicate that releasing particles should span the whole troposphere. While we appreciate the suggestion, the new method is difficult to carry out and may raise some issues. For example, the underestimation of the local sources caused by overly tracing the upper-level air masses at the site. The contributions of high and low level origins to the AOD are different. Many studies suggest that the non-Beijing pollutants are mainly transported from the south and east-south places in low level. Usually, the low level pollutants contribute mostly the integrated AOD, and the air mass origin near surface generally explain the large part of AOD variations. We would carefully discuss the impact of low-level air mass origin on column AOD value.

Dust storm is one of the major non-Beijing natural sources only in spring, and their occurrence frequency decreased significantly in recent years (Zhu et al., 2008). The frequency of dust storms that significantly influence the surface PM10 concentration in Beijing is very low after 2001. About only 3% of very bad visibility (<2km; 7.2% occurrence of total hours) in Beijing during 1999-2007 is associated with dust, respectively (Zhang et al., 2010). As a result, the impact of dust storm on the averaged AOD is limited.

3) In this study, we analyzed the backward transport paths of air masses. As above-mentioned, the diffusion could be well described in the modeling. Even for highly reactive species, the transport pattern also could give valuable information about the difference of emission sources among the CTs because the characteristics of the distribution of anthropogenic sources. The south regions are always characterized with high emissions and high concentration of pollutants, and therefore northerly and southerly pathways have distinct non-Beijing pollutant sources. The removal and transformation processes are associated with atmospheric turbulence, meteorological conditions (e.g. RH), and photochemical activities (hence the solar radiation), all of these are linked to synoptic circulation. The effects of evaporation, condensation, coagulation and wet uptake are difficult to accurately calculate even with a meteorology-chemistry coupled model. All the error caused by not considering these processes are actually include in the deviation of the mean concentration of air pollutants under different types of synoptic circulation.

Why are back trajectories limited to the 0-50m levels? Is there some reason to believe that the local boundary layer is not better mixed than that, with respect to a daily average measurement value (the same scale that is being correlated against)? Is there an issue with local sources being overly sampled at the site that requires only the very bottom of the atmosphere to be considered as being the sample range at the site?

Reply： The purpose is to backward track the air masses around the surface measurements. Therefore, the particles were released from a fixed box with a height of 50 m, which is assured in mixing layer. This method of releasing particles has been widely used and adopted (e.g. de Foy et al., 2007; de Foy et al., 2009; de Foy et al., 2011; (Aiken et al., 2010). This method is believed to have advantages over the way of releasing the particles at one or a few fixed heights since it is more reasonable and well representative of the air masses that mostly influence the measurement site. Moreover, the included turbulence and convection processes in the modeling would improve the simulation of vertical mixing.

The reviewer may indicate that releasing particles should span the whole boundary layer. While we appreciate the suggestion, the new method is difficult to carry out and may raise some issues. The impacts of high and low level origins to the concentration of surface measurement are different although they are both in boundary layer. For the challenges of accurately predicting mixing height, the bias of calculating the mixing height may result in erroneously tracking the particles in free atmosphere. Further studies may be needed to compare the strengths and weakness of the methods for particle releasing, and to design an advantageous procedure.

As far as the surface measurements are concerned, we think our method of releasing particles is justified and would not result to overestimation of local sources. If the motion of released particles can be accurately simulated, especially in PBL, and the sources can be precisely calculated. Accordingly, the error of evaluation of local sources could originated from the calculation bias of turbulence.

Atmospheric visibility is a combination of aerosol concentration, aerosol size and chemical composition, relative humidity, temperature, and other factors. Since the comparisons are being made with strictly the meteorological variables, how relevant are each of these terms? Which meteorological term is more important or the dominating contribution for each correlation?

Reply: The emissions for urban Beijing are relatively stable from a long-term view, so the aerosol size distribution and chemical composition have more slight effect than aerosol concentration on visibility. Relative humidity affects visibility through controlling the hygroscopic growth of aerosol particles and fog droplets. Wind speed affects visibility through transport and dispersion, directly changing the concentration of aerosols. The most two important meteorological variables that significantly influence the visibility are directly associated with meteorological circulation types. The dependence of the visibility in Beijing (Beijing airport) on RH and wind has been documented (e.g. Zhang et al, 2010). Their result shows a clear relation of visibility degradation to southerly wind. The high concentration of air pollutants and fast hygroscopic growth of aerosol due to high relative humidity under CT 5, 8 and 9 could significantly degrade the visibility.

Higher PBL will lead to a change in the surface concentrations, if the boundary layer is well mixed. However, if this is the case, then as mentioned above, the FLEXPART runs should be done at least throughout the entire boundary layer for concentration measurements. Further, the effect of the boundary layer height on column values, such as AOD, is not so straightforward. Your comments on AOD do not match with all of the observations. It is my understanding that at least some of the highest AOD events in Beijing have been when wind blows large amounts of desert dust, which is most similar to CT1. Also, very stagnant or recirculating conditions have led to very high AOD levels. However, since this is an effect of long-range transport in the first example, and longer-lifetime aerosol being recycled in the second example, how can these be incorporated into your modeling simulation?

Reply： We agree that PBL height is crucial to the changes in the surface concentration of air pollutants. Among the presenting literatures, there are usually two methods to vertically release the particles for the Lagrangian dispersion calculation. One is releasing particle at a fixed height (e.g. 100m); another is tracing the plume particles at a ranged height (e.g. 0-50m level). The reviewer may indicate that releasing particles should span the whole boundary layer. While we appreciate the suggestion, the new method may raise some issues. For example, underestimate of the local sources caused by overly tracing the upper-level air masses that may be irrelavent to measurement at the surface site. The impact of high- and low- level origins to the measurement are different although they are both in boundary layer; The bias of calculated the mixing height may result to erroneously tracking the particles in free atmosphere. Further studies may be needed to compare the strengths and weakness of the methods for particle releasing, and to design an advantageous procedure.

The turbulence and convective processes and atmospheric stability and PBL height (Stohl et al., 2005) were also included in the modeling along with the trajectories calculations. Further, our work is to investigate the plume transport patterns that are influenced by CTs, which is characterized by sea level pressure patterns over the limited area (Fig. 1). If the data domain driving the dispersion model is large than the domain characterizing the CTs, it will lose the physical basis. Due to the limitited model domain that did not include main source regions of dust, the mode could not well describe the whole transport paths of dust. But, as the reviewer suggests, the dust storm is potentially associated with CTs with frequent northerly transport pathways (such as CT 1). The dust may contribute some to the AOD of CT 1, but the contribution is limited due to its low frequency occurrence. As a result, CT 1 have relative low AOD on average.

As above-mentioned, the footprint analyses of 48h plume back trajectories could well describe stagnant transport conditions, such as recirculation and low-level wind convergence phenomena, in which the pollutants have very short transport distance and often be recycled (shown in supplement). More longer history information about air masses origin may could be found via investigating the time series of footprint maps. For example, the consistently stagant transport condition occurred during pollution episode 1 (24-28 July), suggesting more aged and recycled aerosol during the later of the episode (i.e. 27-28 July). In general, the 48 h trajectory analysis could well discriminate the transport pattern. We also found those stagnant phenomena are strongly depend on specific CT (e.g. CTs 5 and 8), but it was not included in this paper due to limited space. We are to discuss it in forthcoming work.

AOD is a column value, and is strongly influenced by aerosols throughout the entire column, not just near the surface.

Reply: We agree. But, the near surface pollutants usually can mostly explain most variation of AOD, especially in megacities with frequent air pollution. Two-year air craft study of vertical aerosol and AOD suggested that the surface level aerosol concentration was a main factor which greatly alters the value of AOD (about ±80% compared to the average value) over Beijing and its surroundings (Liu et al., 2009).

MODIS AOD at 1km is a model product. Since there are not too many cloud-free or uncontaminated clouds, it has a very low reliability in terms of AOD. Furthermore, even where it exists, it tends to have quite poor statistics even on a monthly average value. It would not be considered trustworthy on a day-to-day type of basis at all.

Reply: Due to the limitation of spatial resolution, in-situ measurements can not represent regional distribution of air pollution. In spite of the low temporal resolution, MODIS AOD data can provide valuable information about regional distribution of aerosol optical thickness during clear-sky conditions. For this purpose, we do not discuss the day-to-day variations of AOD, but focus on the averaged state of regional distribution of AOD under a specific CT. On average, about 150 days data are used to calculate the mean of AOD around Beijing. We think these averaged MODIS AODs associated with circulation patterns are justified and meaningful. Similar averaged method, for example, was utilized to study (day-of-week variation of air pollution) weekend effect (Xia et al., 2008).

The MODIS 1-km product has been validated (Li et al., 2005a; Li et al., 2005b) and widely used. Since the reviewer indicates that the cloud-free detection for the high-resolution retrieval has low reliability, we recalculate it using long-term (2000-2009) MODIS 10-km AOD data instead. The data are the standard Level-2 aerosol products released by NASA, and have been validated widely over the global in many literatures. This standard satellite product of AOD has longer time duration, and provides more sufficient measurement samples in a specific CT. On average, about 390 days data were used to calculate the mean of AOD in a specific CT, which is similar to average one year mean AOD. The result (Figure N8) suggests generally good agreement between the two data sets although some difference remains. Consequently, we believe that the correlations between CTs and regional distribution of AOD are justified and reliable.

Although your mean PM10 values look different, their error bounds seem wide. Are the differences statistically significant?

Reply: The significant test of the difference was shown in Table S2 of the supplement. For example, the PM_{10} loadings in CT1 are significantly different from those of any other CTs, while the PM_{10} concentration of CT 8 and CT 9 does not significantly differ.

BC is more closely correlated with dispersion, but under heavily polluted urban regions, the chemistry is still quite important. This is all the more important since chemical and physical ageing of BC alter its lifetime and atmospheric properties, and cannot be determined from local conditions alone (eg: Kim et al 2008; Cohen et al. 2011). It seems quite interesting that CT1 has a relatively high amount of SO2, although overall it tends to have lower amounts of other species. Is there a difference in the correlation found between the high average SO2 days and BC in CT1, as compared to the low SO2 days and BC in CT1?

Reply: We agree. The recommended papers apply a new way for our future research. But, we think the most important reason is that CT 1 has low BC loading because of the good dispersion conditions, higher PBL height and clean air masses. SO_2 was mainly from coal burning and could be transported via long distance to Beijing, while BC could come from coal, traffic, and biomass burning, the later two have very different source regions and temporal variation than coal. As a result, the concentration of SO_2 has larger difference between heating and non-heating period than that of BC loadings. Therefore, the frequent very high concentration of SO_2 in winter could significantly elevate the averaged value in CT 1. As Figs.9 and 10, CT 1 has more low BC days than low SO_2 days.

Why is the SO2 so much higher, as compared with most of the other species, in the clean case CT1, unless either the emissions of SO2 are higher, or unless the chemical destruction of SO2 is lower? If the emissions are assumed to be constant, then the chemistry must be important. This is an important point that needs to be addressed.

Reply: Because coal burning is the main source of SO_2, the concentration of SO_2 in Beijing has large difference between heating and non-heating period. CT 1 occurred most frequently in winter when fossil fuel combustion for heating is intensive in North China. Therefore, the high averaged and baseline concentration of SO_2 in winter could significantly elevate the mean value in CT 1. The large seasonal variation of emissions could diminish the correlations between SO_2 and CTs. Moreover, the long transport could be a major source of SO_2, and SO_2 could have longer lifetime in CT 1 because of the less removal processes and lower chemical transformation rate due to the cold and dry airmass and less precipitation.

High levels of sulfate indicate that the air has had a somewhat long residence time in the atmosphere, since it has to have time to chemically form. High levels of SO2 indicate a short residence time in the atmosphere, or special conditions under which the SO2 has not been able to oxidize to form sulfate. How do you resolve the fact that some of the CTs seem to have considerably different ratios of sulfate to SO2, unless either or both of long-range transport and chemistry are considered?

Reply: Agree. We have noticed that some CTs have considerably different ratios of SO_2 to sulfate, due to the chemical conversion of SO_2 to sulfate by both gas and liquid phase processing. This clearly demonstrated the idea that the associations of CTs with concentrations of air pollutants are the combining results of chemical and physical processes. To identify the importance of individual processes, one need the help of other methods, such as by adding more chemical measurements, or using chemical transport model.

How much of this correlation between the southerly airflow into Beijing is related to a net mass flux of pollutants into Beijing, as compared with the fact that the warmer air will cause secondary aerosol chemistry to go faster?

Reply: With our methods, it is difficult to quantitatively estimate the contributions of individual physical and chemical processes to the observed concentrations of air pollutants. One possibility is to combine our CT analysis results with regional quality model, such as CMAQ or WRF-CHEM, to do the analysis. However, this model analysis should cover a large time span same as our CTs analysis. This is out of the scope of this paper but would be interested to try in the forthcoming research.

The Beijing site AERONET data is not sufficient, on a daily resolution, to make such a broad statement. The total number of individual daily measurements made at this site, from 2002 through 2010, during August are not sufficient to cover even a reasonable fraction of the daily average AOD values. How such a strong conclusion can be drawn seems unlikely. The conclusions of the daily variability cannot be supported using this set of data for this case.

Reply: We should note that interpolation is performed to get sufficient data for classification analysis. AERONET AOD at the wavelength of 500 nm is interpolated on a log–log plot assuming linearity between 440 nm and 675 nm (Vucetic et al., 2008) if the measurement at 500 nm is missing. Consequently, there are 6 days during the Olympics for Beijing site. The quality-assured daily data set of level-2.0 also indicate the same valid data numbers in August 2008. Considering the representativeness during the Olympics, the AOD data have been excluded from the quantitative evaluations.

Since policies enacted lead to a different condition in 2008, looking at the average of non-2008 data and then making a comparison with the 2008 data could be justified. However, in terms of the meteorology, this is not true, unless people had an impact on the large-scale circulation. Therefore, it is not statistically correct to look at the anomaly of the meteorology excluding 2008 as compared with 2008, and instead the value in 2008 should be compared with the baseline value over the entire time span being looked at.

Reply: We agree. We did compare the frequencies of circulation types in 2008 with the baseline (averaged) value over the entire time (2000-2009). We should address this part more clearly in the revised paper.

Reductions or increases in NO2 are much more difficult to understand as compared with NOx or NOy. Emissions reductions could lead to an increase or a decrease on NO2, based on the various different changes in other gas-phase chemicals, and there- fore, are not very useful when making such comparisons.

Reply: Agree. We have added the analysis of NO_x. It is better to use NO_y, but in Beijing NO_y is only measured with research-based campaigns. NO_x contains both NO and NO_2. NO_2 is also an important species of gaseous pollutant, and is reserved as a comparison.

You mention that your result is for visibility is 50%. Yet you mention that Q. H. Zhang et al. 2010 find that RH contributed 24%. However, visibility is a combination of both changes in aerosols as well as changes in water uptake. If they found, as you state, that there was a lower amount of hydroscopicity, then this would imply a higher ratio of BC and other hydrophobic aerosols to hydrophilic aerosols. This is a very different conclusion: it means that the aerosol concentrations were not kept in the same proportion. I know that you have mentioned sulfate in the paper, but I do not see a careful exam-

ination of how it has correlated with the different CTs. I think that this is an important point that must be addressed, if the conclusion is to be valid whether it is circulation, aerosol physics, or hydroscopic effects. Just stating that it is all due to meteorology is an oversimplification.

Reply: Agree. As the reviewer indicated, the two evaluations by different methods are not comparable. We removed this paragraph of comparison.

Figure 7: How many different days are being compiled? What is the fraction of the total days?

Reply: There are 1563 and 1007 valid daily data at AERONET Beijing and XiangHe sites, respectively. We extended the data duration. The averaged, minimum and maximum valid days at Beijing site in the CTs are 205, 105 and 402 days, respectively, during 2001-2009. For Xianghe site, the numbers are 136, 63 and 316 days, respectively. We have added this in the text.

Figure 9 shows that the concentrations of PM10 and BC are generally lowest in CT1 and CT6, as claimed in the paper, although with a large amount of variability. Yet, in Figure 10, it shows that for CO, the best of the chemicals present in this report to use as a "pure tracer" (due to its relatively longer chemical lifetime) that the lowest concentrations are in CT6 whereas CT1 seems quit similar to many of the other types. SO2 seems to be relatively higher in CT1 than the other types. These conclusions lead me to believe that chemistry and aerosol physics are important processes occurring here. On the other hand, it could also be an artifact of how gaps in the data were dealt with. Are these values presented, as mentioned here, daily averages, and if so, how were they computed? Are there specific times of day or days of the year that were not systematically measured less frequently?

Reply: Because coal burning is the main source of SO_2, the concentration of SO_2 in Beijing has large difference between heating and non-heating period. CT 1 occurred most frequently in winter when fossil fuel combustion for heating is intensive in North China. Therefore, the high averaged and baseline concentration of SO_2 in winter could significantly elevate the averaged value in CT 1. The large seasonal variation of emissions could diminish the correlations between SO_2 and CTs. Moreover, the long transport could be a major source of SO_2, and SO_2 could have longer lifetime in CT 1 because of the less removal processes and lower chemical transformation rate due to the cold and dry airmass and less precipitation. CO was also significantly influenced by the long range transport, and have relatively high concentration in winter.

The daily averaged data for the particulate and gasous pollutants were used to study the relationship between air quality and CTs. Hourly data from LT 00:00 to 23:00 were averaged as the daily mean if there are more than 16 valid hourly measurements in the day. The instruments at PKU site were well maintained and continuously (systematically) measured for the data periods (e.g. Chou et al., 2011).

Measurements significantly influenced by the precipitation were excluded. The missing data is very limited and would not significantly influence the result.

Technical Corrections:

33779(6-7): Mixing of air does not reduce the amount of aerosols, what it does is increases aerosols in some places while reducing them in others, smoothing out gradients in concentration. An increase in wind speed also does not reduce the total amount of aerosols, it merely causes them to be transported elsewhere.

Reply: Thanks. The phrase "to reduce the amount of aerosols" has been replaced with "to reduce the concentrations of aerosols". However, under the CT 6, the higher average wind speed benefited the dispersion of local urban emissions and more frequent precipitation rate (Table 1) helped to reduce the concentrations of aerosols through wet scavenging processes in the atmosphere.

33480(3): AOD values of 0.43 are not that low. These are already somewhat polluted condition, where the chemistry and physics of the aerosols present cannot be neglected.

Reply: Agree. The AOD value of 0.43 is over the region of Beijing megacity and its vicinity, which have relatively higher baseline concentration of air pollutants than that of the rural. The dust storms that associated with CT 1 may also contribute to the loadings of the total column AOD.

33483(3): Why during the Olympic period are only 24 hour back-trajectories used, as compared to the 48 hour back-trajectories used during the other times?

Reply: A part of 48 hr back-trajectories are out of domain, especially for northerly transport paths. This would lead to some error or missing data if the main positions of the 48 hr backward trajectories were calculated. Therefore, we present the 24 hr back trajectories instead. It also clearly shows the large temporal variation of air mass origins.

References:

Cohen, J. B., R. G. Prinn, and C. Wang, 2011: The impact of detailed urban-scale processing on the composition, distribution, and radiative forcing of anthropogenic aerosols, Geophys. Res. Lett., 38, L20808, doi:10.1029/2011GL047417.

Kim, D., C. Wang, A.M.L. Ekman, M. C. Barth, and P. Rasch, 2008: Distribu-tion and direct radiative forcing of carbonaceous and sulfate aerosols in an interactive size-resolving aerosol-climate model, J. Geophys. Res., 113, D16309, doi:10.1029/2007JD009756.

References:

Aiken, A., de Foy, B., Wiedinmyer, C., DeCarlo, P., Ulbrich, I., Wehrli, M., Szidat, S., Prevot, A., Noda, J., and Wacker, L.: Mexico city aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0)–Part 2: Analysis of the biomass burning contribution and the non-fossil carbon fraction, Atmos. Chem. Phys, 10, 5315-5341, 2010.

Bei, N., Li, G., Zavala, M., Barrera, H., Torres, R., Grutter, M., Gutiérrez, W., García, M., Ruiz-Suarez, L. G., Ortinez, A., Guitierrez, Y., Alvarado, C., Flores, I., and Molina, L. T.: Meteorological overview and plume transport patterns during Cal-Mex 2010, Atmos. Environ., doi:10.1016/j.atmosenv.2012.01.065, in press.

Chou, C. C. K., Tsai, C. Y., Chang, C. C., Lin, P. H., Liu, S. C., and Zhu, T.: Photochemical production of ozone in Beijing during the 2008 Olympic Games, Atmos. Chem. Phys., 11, 9825-9837, doi:10.5194/acp-11-9825-2011, 2011.

de Foy, B., Lei, W., Zavala, M., Volkamer, R., Samuelsson, J., Mellqvist, J., Galle, B., Mart nez, A. P., Grutter, M., and Retama, A.: Modelling constraints on the emission inventory and on vertical dispersion for CO and SO2 in the Mexico City Metropolitan Area 20 using Solar FTIR and zenith sky UV spectroscopy, Atmos. Chem. Phys, 7, 781–801, 2007.

de Foy, B., Zavala, M., Bei, N., and Molina, L.: Evaluation of WRF mesoscale simulations and particle trajectory analysis for the MILAGRO field campaign, Atmos. Chem. Phys, 9, 4419-4438, 2009.

de Foy, B., Burton, S., Ferrare, R., Hostetler, C., Hair, J., Wiedinmyer, C., and Molina, L.: Aerosol plume transport and transformation in high spectral resolution lidar measurements and WRF-Flexpart simulations during the MILAGRO Field Campaign, Atmospheric Chemistry & Physics, 11, 3543-3563, 2011.

Fast, J. D., and Easter, R. C.: A Lagrangian particle dispersion model compatible with WRF, 2006,

Guo, H., Jiang, F., Cheng, H., Simpson, I., Wang, X., Ding, A., Wang, T., Saunders, S., Wang, T., and Lam, S.: Concurrent observations of air pollutants at two sites in the Pearl River Delta and the implication of regional transport, Atmos. Chem. Phys., 9, 7343-7360, 2009.

Jiménez, P., González-Rouco, J., Montávez, J., García-Bustamante, E., Navarro, J., and Dudhia, J.: Analysis of the long-term surface wind variability over complex terrain using a high spatial resolution WRF simulation, Clim. Dyn., 1-14, doi:10.1007/s00382-012-1326-z,

Jiménez, P. A., González-Rouco, J. F., García-Bustamante, E., Navarro, J., Montávez, J. P., de Arellano, J. V.-G., Dudhia, J., and Muñoz-Roldan, A.: Surface Wind Regionalization over Complex Terrain: Evaluation and Analysis of a High-Resolution WRF Simulation, Journal of Applied Meteorology and Climatology, 49, 268-287, doi:10.1175/2009jamc2175.1, 2010.

Li, C., Lau, A. K. H., Mao, J., and Chu, D. A.: Retrieval, validation, and application of the 1-km aerosol optical depth from MODIS measurements over Hong Kong, Geoscience and Remote Sensing, IEEE Transactions on, 43, 2650-2658, 2005a.

Li, C., Mao, J., Lau, A., Yuan, Z., Wang, M., and Liu, X.: Application of MODIS satellite products to the air pollution research in Beijing, Science in China Series D(Earth Sciences), 48, 209-219, 2005b.

Liu, P., Zhao, C., Zhang, Q., Deng, Z., Huang, M., Ma, X., and Tie, X.: Aircraft study of aerosol vertical distributions over Beijing and their optical properties, Tellus B, 61, 756-767, doi:10.1111/j.1600-0889.2009.00440.x, 2009.

Lo, J. C.-F., Yang, Z.-L., and Pielke, R. A., Sr.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, J. Geophys. Res., 113, D09112, doi:10.1029/2007jd009216, 2008.

Pan, Z., Takle, E., Gutowski, W., and Turner, R.: Long Simulation of Regional Climate as a Sequence of Short Segments, Monthly Weather Review, 127, 308-321, doi:10.1175/1520-0493(1999)127<0308:lsorca>2.0.co;2, 1999.

Robinson, N., Newton, H., Allan, J., Irwin, M., Hamilton, J., Flynn, M., Bower, K., Williams, P., Mills, G., and Reeves, C.: Source attribution of Bornean air masses by back trajectory analysis during the OP3 project, Atmospheric Chemistry & Physics Discussions, 11, 15157-15226, 2011.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys, 5, 2461-2474, 2005.

Subramanian, R., Kok, G., Baumgardner, D., Clarke, A., Shinozuka, Y., Campos, T., Heizer, C., Stephens, B., De Foy, B., and Voss, P. B.: Black carbon over Mexico: the effect of atmospheric transport on mixing state, mass absorption cross-section, and BC/CO ratios, Atmospheric Chemistry and Physics, 10 (1): 219-237, 10, 2010.

Vucetic, S., Han, B., Mi, W., Li, Z., and Obradovic, Z.: A Data-Mining Approach for the Validation of Aerosol Retrievals, Geoscience and Remote Sensing Letters, IEEE, 5, 113-117, 2008.

Wang, T., Nie, W., Gao, J., Xue, L., Gao, X., Wang, X., Qiu, J., Poon, C., Meinardi, S., and Blake, D.: Air quality during the 2008 Beijing Olympics: secondary pollutants and regional impact, Atmos. Chem. Phys, 10, 7603-7615, 2010.

Xia, X., Eck, T. F., Holben, B. N., Phillippe, G., and Chen, H.: Analysis of the weekly cycle of aerosol optical depth using AERONET and MODIS data, J. Geophys. Res, 113, D14217, 2008.

ŽAgar, N., ŽAgar, M., Cedilnik, J., GregoriČ, G., and Rakovec, J.: Validation of mesoscale low-level winds obtained by dynamical downscaling of ERA40 over complex terrain, Tellus A, 58, 445-455, doi:10.1111/j.1600-0870.2006.00186.x, 2006.

Zhang, Q. H., Zhang, J. P., and Xue, H. W.: The challenge of improving visibility in Beijing, Atmos. Chem. Phys, 10, 7821-7827, 2010.

Zhu, C., Wang, B., and Qian, W.: Why do dust storms decrease in northern China concurrently with the recent global warming, Geophys. Res. Lett, 35, 2008.

Zhu, L., Song, Y., Shi, H., Huang, X., and Cai, X.: Transport pathways and potential sources of PM10 in Beijing, Atmos. Environ., 2010.

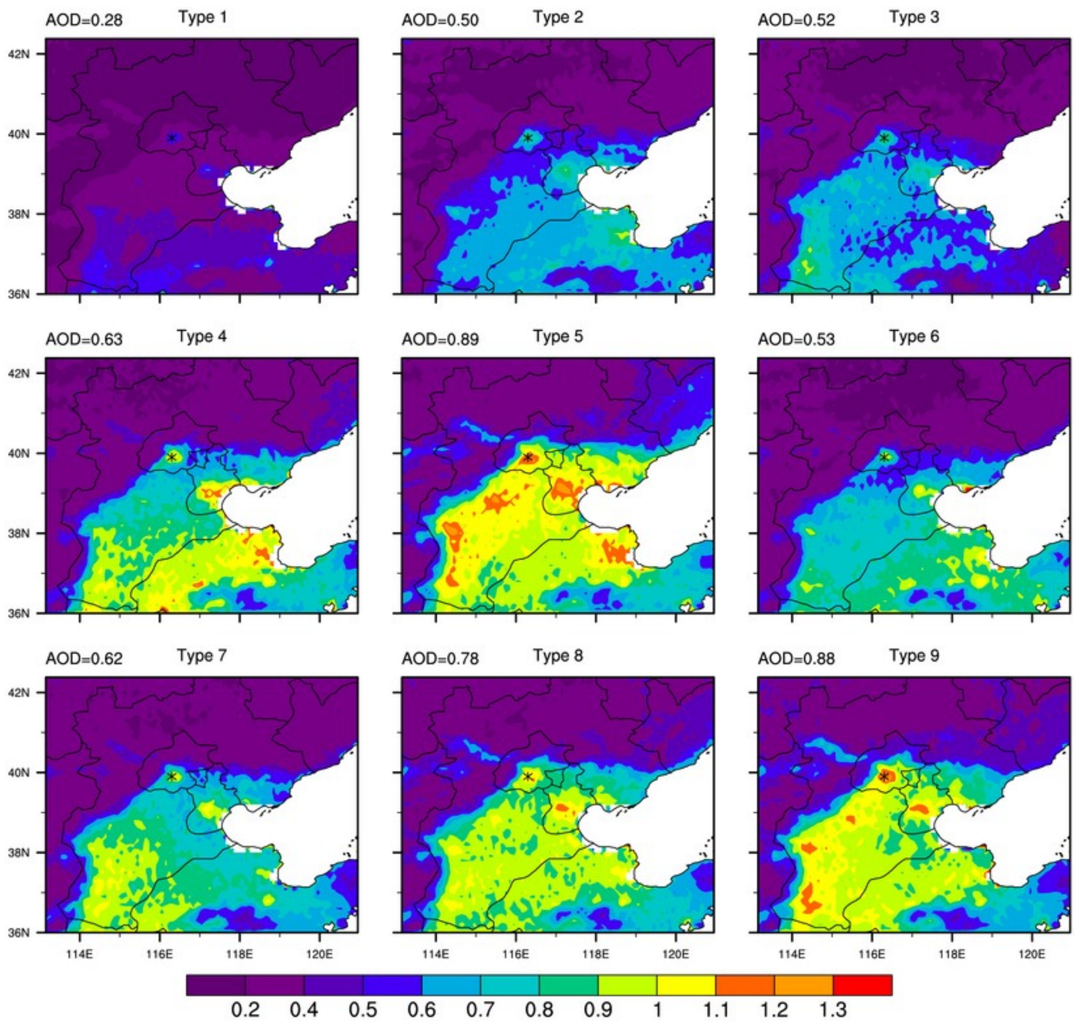

Fig. N8. Averaged 10-km AOD (MODIS Level-2 550 nm product) maps for each circulation type from March 2000 to December 2009 (the upper left number denotes the area mean over Beijing).