

General Comments: First of all, the space was not inserted appropriately in many parts, so it is hard to read and follow. Such a crude revision with low presentation quality should not be sent to reviewers. I have reviewed again this manuscript and found some improvements on the manuscript; however, many replies have not been found in the revised manuscript and/or replied well to my concerns. I feel that the presentation quality is still low as to be published from the high quality journal of Atmospheric Chemistry and Physics. With regret, I have judged to reject this manuscript again.

Response: We are sorry for the space missing problem. We have a double check for our submitted version, but not find this problem. We therefore re-install our software and avoid the problem in the new submission. In the new version, we further improve the quality of the manuscript according to two reviewers' comments. We are confident that it is ready for publishing in a high quality journal.

The authors have made fundamental revisions in this revision, and I have partly understood and agreed the responses. Although this manuscript was improved compared to the first stage, I feel that the current presentation quality and scientific promotion is remained not so high. I have finally judged that this manuscript will not be accepted.

Q1. I have partly understood my concerning issue regarding the definition of dust event. Further concerning issue is the sampling duration of continuous dust event. Even the dust event continued multiple days, how should we consider the representativeness of the sampling? For instance, sample 20080528 and 20080529 (please note that the sampling time of 20080529 will have typo) had approximately one day interval. Was there large temporal variation of PM₁₀ concentration during continuous dust days? If there was large change on PM₁₀ concentration, why the authors collected on the listed time? The authors should state the reason, and should present the representativeness of 4 hrs sampling. In the revised manuscript, it will be kind for readers to explicitly state that '<http://www-cfors.nies.go.jp/~cfors/>' is for forecast model over Asia, and '<http://www.qepb.gov.cn/m2/>' is for observed concentration at Qingdao.

Response: Due to no dust events lasting over 12 hrs (Lee et al., 2015; Su et al., 2017; Zhang et al., 2007), we collected one dust sample with a 4-hr duration in a day. The sampling for dust particles started only when the measured PM₁₀ mass concentration in Qingdao (<http://www.qepb.gov.cn/m2/>) and the forecasted dust mass over Asia (<http://www-cfors.nies.go.jp/~cfors/>) had greatly increased. The samples with ID of 20080528 and 20080529 were subject to two different dust events occurring in two days instead of continuous samples for one dust event. On March 20-21, 2010, two dust events subsequently swept Qingdao. The 4 hr dust samples with IDs of 20100320 and 20100321 may not capture the entirety of the two events. However, the on-line data can allow adequate separation of the two dust event samples.

The same was true for the dust samples with IDs of 20110501, 20110502. The link illustration for these two links has been also added properly in the new version.

From this response, I can partly understand the reason for the representativeness of 4 hrs sampling. However, it seems very hard to follow the revised manuscript on L90-93. What is the evidence of “adequate separation”? The explicit reason is ambiguous here. Are these based on the three sampling category defined in this study (Section 3.2)? If so, it is not appropriate to mention here, because this is the section for methodology.

I have further questioned on the separation of the three categories used in this study really divide the dust samples. This is because three categories were based on the concentration of inorganic nitrogen, which easily decomposed into gas-phase depending on the atmospheric circumstances.

Q2. I have partly agreed, but I have further question on the application of a 3-D air quality model. First, what is the merit of the application of 3-D air quality model? In the revised manuscript, only the spatial distributions of PM10 were shown (Fig. 2 from CFORS model and Figs. S1-S3 with CFORS and WRF-CMA. Can such application reinforce the authors’ discussion points? The behavior of IN were discussed in this manuscript, so what is the purpose to show PM10? The authors stated that ‘The spatial distribution of PM10 concentrations for each dust event was consistent with the model results of dust by the Chemical Weather Forecast System (CFORS) by Uno et al. (2003)’ (L199-201). If the consistency between other models is important, why the author calculated on your own model? I cannot follow this reason from the revised manuscript.

Response: Thank you for the suggestion. We have deleted the results by CFORS. The CMAQ model (v5.0.2) was applied to simulate the concentration of PM10, NO_x, NH₃, NO₃⁻ and NH₄⁺ over the East Asia area for aerosol samples on dust and comparison days. We have revised the discussion on model results. Distribution of PM10 was used to characterize the dust events. Spatial distributions of PM10 during each dust events were consistent with the records in the “Sand-dust Weather Almanac” (CMA, 2009; 2010; 2012; 2013). The model results indicated that CMAQ results reasonably reproduce the mass concentrations of NO₃⁻ (Fig. S6). Simulated NH₄⁺ concentrations in dust samples were severely under-predicted with NMB values at -71%. For reference samples, simulated NH₄⁺ concentrations sometimes can well reproduce the observational values, but sometimes totally off. The external mixing mechanism proposed in this study is urgently needed to be included in the model for accurately predicting the concentrations during dust events.

First of all, I feel that it is not appropriate to use supporting materials frequently in the main text. The main and supporting texts should be understood with itself as a standalone. Without any speculation for CMAQ modeling results, to mention NMB only seems to be not good.

I am not sure the exact reason can be really attributed the external mixing mechanism in the model.

How about the evaluation for the precursors of NO_x and NH₃? The discussions on L331-L340 were only depends on the model results. Moreover, how can the authors' consider the model assumption of anthropogenic emission status on 2008? As the author's showed in Fig. S3, NO_x and NH₃ emissions in China will largely change from 2008 status to 2011. Again, if the authors' conclude that the external mixing, I have still questioned on the importance of intensity of dust event. In my opinion, Ca concentration will highly depends on the dust intensity, and decide the external status.

Q3. The following specific points also should be revised to clarify the model application.

L189: Centered point is needed because we cannot follow the modeling domain at the current description.

L193: On the INTEX-B emission inventory (Zhang et al., 2009), I suppose that NH₃ emissions have not been provided. If so, this description should be changed.

L195, and Figures 6 and 8: So, all calculations were based on the emission level on 2008? Because the temporal resolution of INTEX-B emission inventory is month, I feel that there are no need to display all emissions on all dust samples. These emissions level should be differed only on month. Therefore, I suppose that the averaged (spring time) emissions of NO_x and NH₃ on each one figure is enough.

Response: We have supplemented the centered point (110°E, 34°N) in the new version. According to the publications of INTEX-B and TRACE-P Asia emission inventories (Zhang Q et al., 2009; Streets et al., 2003.), INTEX-B inventory was developed based on TRACE-P inventory with NH₃ emission considered (the annual emission amount of NH₃ in China was 13.6 Tg). However, due to the low priority and low variability of NH₃ emission during 2000-2006, NH₃ emission was not updated in INTEX-B inventory, and the NH₃ emission in INTEX-B inventory was consistent with TRACE-P.

Agree and revised.

Although I have confirmed this revision, please see Q2 for questions on this assumption.

Q4. Figure S1: What is the purpose to show the difference between (b) and (c)? In this caption, what is 'WRF-CMA'?

Response: We have indicated that one exterior dust sample was collected on 1 March when no dust was recorded in Qingdao by MICAPS. However, the MICPAS information over the whole country indeed showed the dust events in China on 1 March. And the modeled spatial distribution of PM₁₀ and TSP mass concentration for this dust event on 1 March implied that the sample should be classified into dust sample. Therefore we listed all the supporting figures in Fig. S1. Fig. S1 (b) was the weather information from the MICAPS at 8:00 on Mar.2, 2008 and (c) was hourly PM₁₀

concentration modeled by the WRF-CMAQ model at 15:00 on Mar.1, 2008. We guessed that the reviewer maybe refer to the difference between (c) and (d), therefore we deleted (d).

We have revised the caption.

I have confirmed this revision for Fig. S1.

Q5. Figure S3: In the main manuscript, it was stated that 'each dust sampling day are shown' in Fig. S3 (L218-219, L895). However, only the hourly concentration of PM10 concentration at 14:00 on 19 Mar 2011 were shown. Please confirm this supplemental figure.

Response: We really modeled the PM10 concentration on each dust sampling day, but only showed the PM10 concentration at the middle time of the sampling in Fig.S3 (Now Fig.S5 in the new version) due to too many figures. We have revised this section and the sentence has been revised into "The concentrations of PM10 and its major components NO₃⁻ and NH₄⁺ over East Asia on dust days and comparison days were modeled using the WRF-CMAQ model (Fig. S5-6)" in L341-342 in the new version.

I have confirmed this revision for new Fig. S5.

Specific comments:

Q6. L35-36: This conclusion does not match to the manuscript contents. The authors stated that input of nitrogen to the ocean depends on the dust events.

Re-comment: cannot find this revision.

Response: The revision in the last round was prepared after the quick response. After a careful consideration, we agreed with the comment and delete the part in the last revised version.

I have confirmed this point. However, when I have read the abstract at the current version, the readers cannot distinguish the category 2 and 3, but only the category 2 was mentioned on last sentence. The abstract should be re-organized.

Q7. I have confirmed and understood the meaning. However, is this revision corresponded to L226-230? If so, this revised sentence seems to contain many errors (NOT Table S2 but Table S1?). For example, we can find 410 $\mu\text{g}/\text{m}^3$ on dust day sample on 20080315. What is the value of 80-1303%? These increased value were not corresponded to 'Ratio of DD to CS' shown in Table S1.

Response: Yes, this revision corresponded to L226-230 in last version. And the times of dust to non-dust day samples were replaced by the ratio according to the former suggestion. To avoid the confusion, we have revised Table S1 to give the increased ratio.

I have confirmed this revision for Table S1.

Q8.L171: Again, I cannot follow the calculation of “1.7-21.9 times (mean: 6.9)”.

L175: I cannot follow “10.3 times” for Fe. It can be calculated as 7.90 from the values in Table 2.

L176: “3.6-fold” will not be followed from Fig. 2. It should be listed in Table 2.

Re-comment: So, in this revised manuscript, these statements of the increment ratio on dust-day compared to non-dust day have not been explicitly appeared. In L243, the authors stated ‘Table S1’, but Table S1 contained not only the information of inorganic nitrogen but also TSP, Al, Fe, and nss-Ca. So, it is appropriate to mention on Table S1 in Section 3.1.

Response: Agree and revised.

In the revised Table S1, the averaged value was not shown; hence it is partly hard to follow L202 and L204. In the last part of Table S1, it is explained as “Mean ratio of all samples on dust days”, however, in L205, it is stated that “a median value of 403%”. In the analysis, mean and median will cause important differences. This should be clearly used.

Q9.L175: So, please state explicitly regarding this point to the readers. In the current form, nss-Ca was suddenly shown in Fig. 3 without any introduction.

Response: Agree and revised.

I have confirmed this point in new Fig. 2.

Q10. First of all, I cannot find the revision of ‘The concentrations of ammonium were increased by 20’ anywhere. Is this corresponded to Table S1? I suppose that the authors discussed regarding this point in L240-L244. Although ratio was shown in Table S1. percentages are discussed here. So it is hard to follow the manuscript. Why the discussion point have not been arranged on the uniformed unit?

Response: Agree and revised.

I have confirmed in Table S1.

Q11.L194-L195: In this sentence, the authors stated “the effect of dust on inorganic nitrogen differed during different types of dust events”. Why the authors suddenly focused on inorganic nitrogen here? In L192-193, it was mentioned “inorganic ion SO₄²⁻ exhibited concentration variations that were similar to those of nitrate”.

L197: The figures for inorganic nitrate will be helpful information here, if the authors focused on inorganic nitrogen.

Re-comments: I cannot find this revision.

Response: The revision in the last round was prepared after the quick response. After a careful

consideration, we completely rewrote the part to avoid confusion in the last revised version.

I still cannot find the explicit reason to focus on the inorganic nitrogen in the revised manuscript.

Why the inorganic nitrogen was focused in this study? This will be an important statement.

Q12. L207: (respectively less than 50 ug/g and 6 ug/g) will be the correct expression for ammonium.

Re-comments: I cannot find this revision.

Response: The revision in the last round was prepared after the quick response. After a careful consideration, we completely rewrote the part into "The ratios of mass concentrations of nitrate and ammonium to the total mass of sand particles were very low, i.e., less than 81µg/g, which are approximately three orders of magnitude less than the corresponding values in our dust samples." at L230-232.

I have confirmed.

Q13. I am wondering that the differences of IN concentration between Duolun and Zhurihe. Both are Hunshandake Desert, however, as is shown in Table 5, IN concentration was much higher in Duolun. Are there some emission source?

Response: Sand samples were collected at a remote site in Zhurihe desert. Little anthropogenic influence is expected. Atmospheric aerosol samples were collected at an urban site in Duolun on dust days for comparison. It is not surprised for a strong signal for anthropogenic sources. This has been clarified in the new version.

I have confirmed in Table 5.

Q14. Again, only from the dust spatial distribution, it is hard to state the dust intensity.

Response: We had made a substantial revision on the part in the last round revision and didn't consider dust intensity as an important factor for our unique results.

Q15. L214-L216: Without more information of the intensity of dust, the discussion on 'dilution effect' seems to be lacked in scientific understanding. This part should be fully revised based on not only dilution effect but also dust intensity.

Re-comments: Again, only from the dust spatial distribution, it is hard to state the dust intensity.

Response: We had made a substantial revision on the part in the last round revision and didn't consider dust intensity as an important factor for our unique results.

I have checked that the discussion on dust intensity was fully removed. Please see Q2.

Q16. I have confirmed the revision, but if the authors discussed on average (L300-303), the averaged values were also needed.

Response: We really had given the average of TSP in form of $\text{average} \pm \text{standard deviations}$ at L300-303 in last revision. Now we had made a substantial revision on this part and didn't discuss TSP average concentration.

I have checked that the statement on TSP with Table 4 was fully removed.

Q17. L227-L228: The favorable condition to form ammonium cannot be discussed without the information of NH₃. In addition, Table 3 indicated the aerosol samples in the coastal region of the Yellow Sea. How about the status over air mass path? Is it sufficient to conclude only from the downwind information to the formation of inorganic nitrogen?

Re-comments: Again, I cannot understand the model application results.

Response: We modeled the emission and concentration of NO_x and NH₃ over East Asia on the dust and comparison days. The model results showed that the calculated trajectories of the entire dust air mass passed over those highly polluted regions with strong emissions of NO_x and NH₃ shown in Fig 6 and experienced different residence times therein. The average concentration of NO_x and NH₃ during transport were calculated and discussed according to Categories 1 and 2. The air masses in Category 1 took over 11-39 hrs to cross over the highly polluted area with appreciable concentrations of NO_x (5.7 ± 1.4 ppb) and NH₃ (7.6 ± 3.3 ppb). Except for the exterior samples, air masses in Category 2 took less than 10 hrs to cross over the polluted areas with lower concentrations of NO_x (3.6 ± 3.4 ppb) and NH₃ (4.7 ± 4.7 ppb) and the mixing layer height along the route was 916-1194 m (on average) for each dust event. This further led to the external mixing of anthropogenic particulate matters and dust.

Although I can partly agree these kinds of analysis, please see Q2.

Q18. L230: "strong dust storm" cannot be discussed without any information on dust intensity here.

Re-comments: Again, from the additional information of CFORS, the spatial distribution pattern was found; however, how can we estimate the intensity?

Response: We had made a substantial revision on the part in the last round revision and didn't consider dust intensity as an important factor for our unique results.

Why the dust intensity is not an important factor? What results indicates this conclusion? Please see Q2.

Q19. L233-L234: But NO_x concentration was high in Case 3. I cannot follow why the authors concluded on Case 3.

Re-comments: So where did the authors discussed the NO_x concentration in the manuscript?

Response: We had made a substantial revision on the part in the last round revision after the quick response. The NO_x concentration was discussed in Section 4.3 in the new revision.

I have confirmed this revision in Section 4.3.

Q20. L254-L255: The authors simply mentioned “local emissions” here. Because the samples were collected on downwind regions in the coastal region of the Yellow Sea, I guess that the discussion on emission characteristics of each (or, at least, some categorized) air mass should be discussed in detail. The inorganic nitrogen concentrations are highly related to the local conditions both on emissions strength and meteorological parameters, so the discussion only on air mass speed and air mass path over ocean are insufficient.

Re-comments: Again, model is used only for spatial distribution and not inform the chemical production process.

Response: We had made a substantial revision according to the suggestion. The chemical production process was discussed in Section 4.1 “Theoretical analysis of the three categories”. In Category 1, ammonium salt aerosols may externally exist with dust aerosols in these dust day samples and NO₃⁻ and SO₄²⁻ were almost completely associated with NH₄⁺ in these dust day samples; whereas a larger fraction of NO₃⁻ +SO₄²⁻ may exist as metal salts due to reactions of their precursors with dust aerosols in Category 2. The simulated NO₃⁻ and NH₄⁺ concentrations was compared with the observation in Qingdao, and the results indicated that the external mixing mechanism proposed in this study is urgently needed to be included in the model for accurately predicting the concentrations during dust events.

I have confirmed this revision in Section 4.1.

Q21. I have rechecked the discussion of backward trajectories discussed on Section 3.4. There are many points should be clarified.

Figure 5: Please add the explanation of the trajectory of 20110415 was excluded based on the discussion on Fig. 2. Why the authors displayed “non-dust samples”? What were the differences between non-dust and dust samples trajectories? I feel that these were similar.

Response: Agree and revised.

I have suggested that the reason or short notice is needed in Fig. 5.

Q22. L314: What is the ‘remaining one’? Please specify the trajectory data. In my opinion, two trajectories of 20110418 and 20110501 originated from northeast China.

Response: Yes, trajectory 20110501 was really from northeast China, however it then passed over the Inner Mongolia, and arrived at Qingdao from north, just like 20110502. Therefore, we grouped the trajectory into the air mass originated from Inner Mongolia, China. However, it was really easy to mislead the readers. Therefore, we accepted the suggestion, and revised the sentences into “The calculated air mass trajectories for 13 out of 14 samples showed that the air mass originated from North and Inner Mongolia, China (Fig. 5), generally consistent with the results of Zhang and Gao (2007). The remaining one, with ID of 20110418 originated from Northeast China.”.

Q23. L317: What is the ‘one exterior sample’? Please specify.

Response: Agree and revised.

I have confirmed this revision.

Q24. L319: I cannot see ‘the air masses crossed over the sea for 94-255km’ from Figure 5, because Fig. 5 showed the whole view of trajectories across China. More detailed figure or explanation will be required.

Response: The distance over sea of the air mass for each sample was measured from the trajectory using TrajStat software (Wang et al., 2009). We have added the explanation in Section 2.3.

Wang, Y. Q., Zhang, X. Y., and Draxler, R. R.: TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data, *Environ. Modell. Softw.*, 24, 938-939, 2009

I have understood this methodology.

Q25. L328-329: What is the definition of the ‘average mixing layer’. I suppose that the altitude of backward trajectories were so high because most of trajectories were originated outside China on 72 hrs. So, where is the averaged region to calculate ‘900m’ in this sentence?

Response: The average mixing layer was calculated as an average of all points on the air mass back trajectory of each sample. This has been clarified in section 2.4 in the new version.

I have understood this methodology.

Q26. I cannot still understand the authors’ conclusion here. As was discussed on L355-361, Fig. 8, and Table 7, although the coal combustion have increased on dust days, the contributions of local anthropogenic sources (especially secondary aerosols) have decreased on dust days. According to the discussion on Section 3.3, the authors concluded that ammonium salts were externally co-exist with dust aerosols in Category 1. So, why the contribution of secondary aerosols were decreased from PMF analysis. I feel that these results have contradicted. More careful discussion is required

for this conclusion.

Response: The source of coal combustion on dust days became complicated. "mixture of coal combustion and other pollutants" means these compounds present contemporaneously, because that PMF model can't show the mixing or existing state. We have revised the sentence into "The source profile for coal combustion in dust day samples showed a high percentage of K⁺, Cl⁻, Ca, Mg, Co, Ni, As, Al and Fe, indicating coal combustion presenting contemporaneously with other pollutants emitted along the transport path on dust days.". Ammonium salts were externally co-exist with dust aerosols in Category 1, but showed lower concentrations in Category 2 likely due to unfavorable conditions for forming ammonium salts. Here the conclusion was a result of source appointment for all dust samples including Category 1 to 3. And we have revised the sentence "In these dust samples, including Categories 1-3, oil combustion, industry, soil dust, secondary aerosols, and coal combustion/other sources were identified as five major sources (Table 6).

I have confirmed this revision.

Q27.L306: I cannot follow the calculation of "a factor of 1.1-5.8" and "a factor of 1.8-6.3".

Re-Comments: I cannot find this revision.

Response: We had made a revision on these sentences in the last round revision after the quick response. According to the former suggestion, we changed the factor to ratios. And this sentence was revised to "In Category 1, the dry deposition fluxes of $\text{NH}_4^++\text{NO}_3^-$ increased by 9-75% with increased TSP flux by 86-252% (Table S3)" at L371-372 in the new revision. And we also revised Table S3 to give increased proportion and the calculation method.

I have confirmed this revision in Table S3.

Q28. L309: What is the calculation method of "63%" and "46%"?

L310: What is the calculation method of "14%" ?

Re-Comments: I cannot find this revision.

Response: We had made a revision on these sentences in the last round revision after the quick response. We have revised Table S3 to give increased proportion and the calculation method.

I have confirmed this revision in Table S3.

Q29. L317: For Fe, it seems that the increased ratio were 2.81-11.08 from Table S3.

Response: This sentence has been revised to "However, the dry atmospheric deposition fluxes of Fe increased by a factor of 124-2370% in dust day samples." at L383-384 in new revision.

I have confirmed this revision in Table S3.

Specific comments:

Q30. Table 6: Missing the note of a and b.

Response: Done.

I have confirmed this revision in new Table 7.

Q31. Table S2: Please align the right-column, it is hard to follow. What is the meaning of *?

Response: Agree and revised.

I have confirmed this revision in Table S3.