We thank the referee #3 for giving valuable comments. We respond to each specific comment below. The comments and questions from referee #3 are in italic font.

One concern is the 1-month lag between derived emission reduction (in September) and the actual emission regulation (in August and prior). The lag has been attributed to the issue of Kalmar-based inversion and the lack of OMI data. While this may be the case, a more detailed analysis is needed. This is because the lifetime of NO_x is short (several hours in summer), and thus there should not be any obvious lag (i.e., more than 1 day) between regulation and reduction. How many days are missing in August compared to other months (e.g., February, when there is no lag)? In addition, all NOx-relevant emission regulations have been implemented by Aug 15, therefore it is not clear why emission reduction is not shown in August and previous months. In particular, Fig. 10 shows that June-August 2014 have higher emissions than June- August 2013, in contrast to the emission regulation starting from May 2014.

We understand the concern of the referee and have added more explanations to the text.

Reductions in emissions at the end of August or the following months can appear with a time lag in the Kalman filter results (see e.g Brunner et al., 2012). This time lag is not fixed but depends on the amount, interval, accuracy and distance of the observations and it is therefore difficult to quantify. Although the strong point of a Kalman Filter is its detailed error analysis, this time lag is not incorporated in it is formalism. In future research we intend to reduce this time lag by using a Smoothing Kalman Filter technique.

The question how many days are missing cannot be answered straightforwardly, because the DECSO algorithm is using all measurements in the neighborhood that has been transported to or from the Nanjing region. This most important feature of DECSO has been emphasized more in the text. In February no time lag appears because the Spring Festival is usually in the beginning of February, thus a time lag of several days will have no effect on the results.

The difference between 2013 and 2014 are all within the error bars of the emission estimates, so it is difficult to draw any conclusions from that. Due to the fact that we use monthly mean values and the Olympic Games took place at the end of the monthly period, its effect will be less obvious in August because of the first half of the month having normal NO_x emission levels. Since many regulations for NO_x had a more permanent character the emission reduction is better visible in September.

At the end of page 3 we added:

"In addition, several technical improvements have been implemented to reduce pollution from heavy industry and power plants."

We changed the text at page 17, line 8-10 into:

"This reduction is <u>probably</u> caused by <u>the more permanent</u> air quality regulations taken by…"

And line13:

"This is partly a consequence of the use of monthly means, while the regulations became active at the end of August. It is also a consequence of the lack of……."

The authors make efforts to filter out aerosol-affected data. The criterion is based on NO2 comparison (OmF) rather than on the aerosol amount directly. There is a concern with this practical choice – whether a scene is filtered out or included depends on how "bad" the comparison is rather than depending on the underlined physical reasons (i.e., aerosols). As lots of "outliers" are filtered out, the criteria certainly affect the subsequent emission inversion. Is it possible to compare the chosen filter to an alternative filter where days with aerosols higher than some threshold are excluded? Are there high-aerosol days in the chosen days?

As shown by Lin et al. (2014) the relation between aerosols and cloud retrievals are complex and non-linear. It is therefore not straightforward to filter deviating NO2 retrievals based on aerosol information. It depends on the type of aerosols and the concentration as shown by both Lin et al. (2014) and Leitao et al. (2010). The study of the effect of aerosols on NO2 retrievals and how to filter or to improve these retrievals is topic of future research. On the other hand, filtering of outliers in data assimilation by using an OmF filter criterion is not uncommon. The reason why we have not applied such a filter in previous versions of DECSO is already explained in the text. However, in this case we have applied a very cautious version of the OmF filter, which avoids building a complicated filter based on aerosol information of type, concentration and its interaction with clouds.

In the text on page 6349 we have added: "The effect of high aerosol concentrations on the NO_2 retrieval is non-linear and depends strongly on both the type of aerosol and its concentration (Lin et al., 2014, Leitao et al., 2010). It is therefore difficult to filter out outliers in the observed NO_2 based on aerosol data."

How many data are available for each month?

There are about 1500 OMI observations used in DECSO each day for the whole domain of East Asia.

The DECSO algorithm is using all measurements in the neighborhood that have been transported to or from the Nanjing region. This most important feature of DECSO has now been emphasized more in the text.

The OmF filter is based on absolute value rather than percentage value. Considering the seasonality in NO_2, applying the filter means stronger filtering in summer and weaker filtering in winter. This will affect the derived emission seasonality. Please discuss.

We considered to use a percentage value during the research. However, it didn't work well when we used the percentage of the forecast or observation value. For example, if the forecast value is very high, the OmF absolute value could also be very high but the percentage can be small. It still causes the problem that the derived NO_x emissions drop to zero in one day and then slowly increase again to the previous emission levels in the following days. On the other hand, if the forecast or observation value is very low, the OmF absolute value are low enough and doesn't cause any problem in DECSO but the percentage can be higher than the criterion. In general, when we used a percentage value as a criterion, many data causing a wrong emission update were still used, while many data with good quality were filtered out.

We add "Not losing sensitivity to new emission sources is also the reason we do not choose a relative filter criterion" in the paper.

It looks like stronger filtering in summer and weaker filtering in winter. However, the amount of observations are much higher in summer than in winter. After using the same filter for summer and winter, there are still more observations used in summer than in winter.

The choice of asymmetric filtering could be better discussed. Line 24-27 of P6349 is not clear.

Assuming the relative error of observation will be more or less equal. Then during the data assimilation process in DECSO, the error of high observations will be relatively high, thus the weight of this high observation is low. But the low NO_2 observation value with low error has more weight. This will favor the low observations

We add the following sentences at the end of *Line 24-27 of P6349:*

"The observation with low error have more weight in the data assimilation process."

A looser cloud screening is used to include more OMI pixels. CRF > 50% means less than half of TOA radiance comes from 'clear-sky' portion of the pixel. The looser criterion may increase the data noise, which is especially relevant to the daily-based emission inversion. Some relevant figures and quantitative analysis would be welcome.

We have checked the distribution of the selected satellite data, which remained unchanged. Indeed, the error on the monthly mean data changes. Therefore, we have replaced line 9-11 on page 6345 with:

"From our analysis of the satellite data we conclude that as a result of this new limit on the cloud fraction the error on the measurements increases with less than 20% and without introducing biases. Yet this effect is compensated by the advantage that more data becomes available. The number of observations increases with about 37 % over the whole domain"

On the same issue – the 'back-of-the-envelop' calculation should be cautiously interpreted, as the lifetime of NO_x could vary dramatically from one day to another, leading to large changes in NO_2 VCD even with the same emissions.

The purpose of back of the envelop calculation is to give readers an impression on the effect of the large OmF on emission estimates. We use seasonal average to reduce the effect of horizontal transportation. The lifetime of NO_x could vary from one day to another. We give the difference of the NO_2 concentration caused by different emissions on an average value. We add this to the paper.

The effect of Chinese Spring Festival has been studied by previous research on satellite NO_2, so as the effects of lifetime versus emission on the seasonality of NO_2. Please give more discussions on or comparisons with these relevant studies.

Wang et al. (2011) showed the monthly NO_2 ground observations of the three cities Beijing, Shanghai and Guangzhou in 2005. There is also a small decrease of NO_2 concentrations in February. But they didn't discuss this small decrease.

Most recent studies we find about the Chinese Spring Festival and NO_x emissions only focus on the emissions from fireworks (usually aerosols) during the Spring Festival and from traffic during the travelling before and after the Spring Festival.

In line 17 page 6351, we add:

"The monthly averages of NO_2 in-situ observations shown by Wang et al. (2011) for Beijing, Shanghai and Guangzhou in 2005 were also reduced by around 10% in February."

At the end of line 7 page 6353, we add:

"Ran et al. (2009) explained high NO_x concentrations in are caused by slower chemical processes and shallow boundary layers contributing to accumulation of NO_x. The table of Wang et al. (2012) annual and summer NO_x emissions from coal-fired power plants in 2005-2007 for different provinces in China showed that the NO_x emissions in Jiangsu Province in summer is higher than mean seasonal emissions."

The model seems to exclude lightning emissions. Lightning emissions are very important in summer, especially considering the much increased sensitivity of OMI to aloft NO2. In addition, lightning emissions vary significantly from one year to another, and a climatological adjustment does not affect the interannual variability. Please discuss the implications of the lightning treatment on the emission inversion here.

Opposite to regions more to the South in China the vertical column of NO_2 from lightning (max. $0.5*10^{15}$ molecules/cm^2) in the region of Nanjing is much less than the NO_2 from anthropogenic sources ($10\text{-}30*10^{15}$ molecules/cm^2) as can be seen in studies of e.g. Schumann and Huntrieser (2007), van der A et al. (2008), and Belmonte Rivas, et al., (2015). Thus 2% of the total column NO_2 can be attributed to lightning and variations in lightning activity will be much lower than

2%. Therefore, we have ignored the lightning contribution in this study, although in other regions this can have important consequences.

Is 'biogenic' emission of NO the same as soil NO_x? If not, does the model consider soil emissions that also peak in summer?

Yes. Biogenic emission of NO is soil emission. We clarified this in the revision.

How is the model error treated?

This is described in the paper of Mijling and van der A (2012) and we have added an additional reference to the paper.

The choice of Eobs in Eq. 1 needs more justification. Also, what is the unit of Csat, Eobs and Esat?

The C_{sat} is a normalized value from 10^{15}. We add the explanation in this part (Page 6343 line 18).

"[…] the satellite observations. The unit in this formula is 10^{15} molecules cm^{-2}. The modified […]".

This aim of this formula is to keep the original observation error for low NO_2 observation values and half the original error for high NO_2 observation values. So the high observation values will have more weight during the data assimilation process.

We add the reason after the description of the effects by tuning the satellite error on Page 6343 (line 21). "[…] absolute error for low values (typically around 0.5 10^{15} molecules cm^{-2}). Thus, DECSO can better capture new emission points or high emission episodes."

The paper cites H. Zhang et al. (2009) for explanation of the large emission seasonality. What is the quantitative result from Zhang et al. and is their result comparable to here?

They showed the trends of electricity consumption in Nanjing from 2000 to 2006. It kept increasing and has strong seasonality during these 7 years. The value of electricity consumption in summer is at least two times higher than in winter every year as shown in their figure. They studied the relation of temperature and electricity consumption and concluded there is a significant correlation between them. We can't quantify the effect on the amount of NO_x emissions from the value of electricity consumptions, since there are also other sources involved. We add "The value of electricity consumption in summer is at least two times higher than in winter every year and keeps increasing during those 7 years." in the text.