We thank both reviewers for their careful evaluation of the presented work. Most of their comments and remarks have been processed in the manuscript, which we believe has gained in clarity and scientific soundness. Below is a point-by-point reply (in **black font**) to this round's reviewer's comments (in **blue font**).

Response to Reviewer #1

The paper analyzes not just China but also other countries. So the introduction should reflect this point.

We have adapted the first paragraph of the introduction.

The use of GOME-2 data instead of OMI may be a concern, since OMI provides more data coverage and the results have been validated by Boersma et al. (2011).

Both OMI and GOME-2 tropospheric NO2 retrievals have been compared successfully against MAX-DOAS observations in China and Japan by Irie et al. [2012]. From the slopes of the regression lines for strict coincidence criteria, they estimate biases in OMI and GOME-2 data to be −10%, and +1%, respectively, compared to the MAX-DOAS data. We added this validation reference to the text.

The better spatial resolution of OMI would be a reason to opt for emission estimates based on this instrument. However, the "row-anomaly" caused OMI gradually to loose to 50% of its useable tropospheric NO2 retrievals at the end of the 2007-2011 period. Because most pixels around nadir were lost, the average spatial resolution dropped considerably. GOME-2 is not affected by loss of observations and resolution, and therefore produces a more consistent time series during 2007-2011.

Another reason to use GOME-2 data is that its earlier overpass at 9:30 makes the inversion less sensitive to correct modeling of non-linear NOx photochemistry in the afternoon. This is an important source of uncertainty in the CHIMERE model [Huijnen et al., 2010].

Please give some more analyses of the GOME-2 retrieval, CHIMERE simulation (e.g., model convection, PBL mixing, emission setups), and the mapping between them. Uncertainties in satellite data should be discussed more explicitly. While some of the info may be described elsewhere, an analysis/discussion here will help readers understand the significance and uncertainty of emission results here.

We inserted more information on the GOME-2 retrieval errors, the model set-up, and the mapping between GOME-2 and CHIMERE in Section 2.

It is not clear to me how can a 0.25x0.25 degree map be made appropriately when the footprint of GOME-2 is at least 40x80 km. In addition, due to the footprint of GOME-2 and the limited amount of valid satellite pixels for each month, emission results for some small provinces like Beijing are likely affected by errors in attributing satellite NO2 to individual provinces for pixels around the provincial borders. A brief discussion will be helpful.

As the reviewer correctly remarks, emissions close to the provincial borders can be affected by the spatial detection resolution of the method and be partly attributed to neighbouring provinces. We added a brief discussion to the text. An in-depth analysis is not straightforward (and has not been performed yet), but the observational footprint size is certainly an important factor. The GOME-2 pixel has an area of 3200 km^2, which is smaller than the smallest provinces Shanghai (6430 km^2), Tianjin (11305 km^2), Beijing (16800 km^2), and Hainan (34000 km^2). Because the algorithm takes transport from the emissions source into account, there is no clear relation between the resolution of the satellite measurements and the best meaningful resolution of the emission estimations. The lifetime of NOx, the divergence of the transporting winds, and the sampling density (in space and time) also influence the amount of information which can be inferred for the location and the strength of an emission source.

The use of daily data may have its cons, as daily NO2 data are noisier than monthly data. Please discuss.

Daily data is indeed noisier, but its impact of the observational error on the emission estimates is controlled by the Kalman filter. Noise in the emission estimates is further reduced by taking monthly and regional averages afterwards, as stated in the text.

Working with monthly observational data (as opposed to daily data) would mean averaging the observations for a certain grid. In the averaging process a lot of information is lost on the non-local relations between emission and concentration.

Sect. 3: It is better to make a table to present previous top-down emission estimates.

We inserted a new table (Table 1) with an overview of these data.

The likely time lag in the derived emissions (as shown in Sect. 4.2 and Figs. 4-5) is a concern and potentially points to errors in the underlying assumptions of DESCO. This issue needs to be addressed.

[Added to Section 2:] The observed time lag is inherent to the use of the Kalman filter. At each analysis stage, the new emission estimation is a combination of the true emission and the previous emission estimation. The strength of the coupling to the true emissions depends on the the Kalman gain (determined by the balance of the errors in the observation and the model) and the sensitivity (the relation between the emission and the observable concentration). The weak coupling we have in our inversion setting means that the emission estimations have a certain response time to new emission levels. This results in a spin-up time for the whole system to stabilize from an initial NOx emission inventory, which is in our setting estimated to be at most 3 months.

For periodic emission signals, such as seasonal variability, the response time will show up as a time lag of about 1 month.

[Added to Section 4.2:] For a certain coupling strength, the time lag depends on the sampling frequency of the signal. The Kalman filter acts as a low-pass filter: a higher frequency lowers the amplitude of the response. Emissions peaks will therefore be slightly underestimated, but yearly averaged emissions and the observed linear trend are not affected.

In Sect. 4.2, lightning emissions are comparable to soil emissions (Lin, 2012) and should be analyzed. Lightning emissions also peak in summer (with more convection and precipitation). In addition, with so many assumptions, it will be helpful to discuss uncertainties in the estimated contributions of anthropogenic/natural sources.

Soil and lightning NOx are now taken together as biogenic emissions in the text.

P17524,L5: Should be Lin et al. (2012). In addition, Lin et al. (2012) show significant sensitivity of model NO2 to many other parameters (in addition to clouds and HO2 uptake).

We corrected the reference to Lin et al. (2012). To reflect better the results of this paper, the text has been adapted to “Lin et al. [2012a] show that modelled NO_2 columns are sensitive to several parameters, especially to the correct cloud optical depth, the uptake rate of HO_2 on aerosols, and the rate constant for the reaction between OH and NO_2.”

P17524,L13: please explain why an earlier overpass time can reduce the effect. Also, emission trends may be affected by model biases since the biases are not exactly constant in time.

Huijnen et al. (2010) showed that, compared with OMI observations around 13:30 in Europe, CHIMERE underestimates NO2 columns more in summertime than in wintertime. This bias is partly attributed to errors in the NOx photochemistry. At GOME-2 overpass (9:30) the photochemistry is less active and will contribute less to the error budget.

Seasonal variation in model bias can indeed affect the derived linear trend. We dropped our remark “Note that relative emission trends are unlikely to be affected by biases.”

P17526,L20: A figure is better to present such results.

This paragraph has now been removed; at second glance its content is not interesting enough for the emission comparison analysis.

P17529,L17: what might be the causes of such time lag?

The origin of time lag is now discussed in Section 4.2 (see above for the inserted fragment).

P17530,L14: ‘remarkably constant’ is an overstatement since the standard deviation of the ratio is about 30% of the mean value.

True. Changed to ‘fairly constant’.

P17533,L2: the fractional values may not be 'lower limit values', although the absolute concentration may be.

We found that 22% of the NO2 over North Korea has been emitted locally in the last 24h. 7%, 21%, 5% of the NO2 over this country originates from emissions in South Korea, China, and sea in the last 24h. 45% of the NO2 has been emitted longer than 24h before detection time. The origin of these emissions is not considered explicitly within the DECSO framework, and is therefore unknown. It might originate from North Korea, South Korea, China, sea, or other areas, which makes the fractions 22%, 7%, 21%, 5% lower limit values.