Response to Referee #1

We thank the reviewer very much for their constructive comments. We respond to each comment individually below, followed by changes to the manuscript.

Reviewer Comment: My first comment is that the present manuscript is missing some description and explanation of the top-down NOx emissions. First, measurements from three satellite instruments are used to estimate the long-term NOx emissions. Are there any instrumental differences among the three satellite products? If so, how do you reconcile them?

Author Response: The referee makes a good point bringing up the instrumental differences between satellite products (a similar concern was raised by another reviewer). The most important differences are the overpass times and the horizontal footprint of individual observations. First, the GOME, SCIAMACHY, and GOME-2 overpass times are roughly 10:30 a.m., 10:00 a.m., and 9:30 a.m. respectively. Second, their horizontal footprints are roughly 320 km x 40 km, 60 km x 30 km, and 80 km x 40 km respectively. We reconciled these three records by consulting published literature and by closely examining the overlapping time periods ourselves, and concluded that a consistent time series is achieved without requiring additional corrections. The reason for this is largely because the daily satellite observations were all gridded to a regular coarse grid of 2° x 2.5° latitude by longitude. Using a comparison with long-term ground-based MAX-DOAS observations, Irie et al. (2012) demonstrated that there is no inherent biases in either SCIAMACHY or GOME-2 that would preclude their combination into a single record. The work of van der A (2008) and Konovalov et al. (2010) show that a self-consistent record can be achieved by downgrading the spatial footprint of the higher resolving instruments (e.g. through smoothing or convolution) to that of the lowest resolving instrument. This is what we have achieved by gridding all the observations to 2° x 2.5°. The combination of these observations is also aided by the fact that the retrieval algorithm for obtaining tropospheric NO_2 column density from all three instruments is the same (http://www.temis.nl/airpollution/no2.html).

We further examined our approach by inspecting the timeseries from individual 2° x 2.5° pixels over selected populated regions. These are shown in Figure D1.

Given the evidence from this figure, and consensus in the literature, we concluded that the instrumental differences between instruments are inconsequential to our analysis.

In response to the referee's comment, we have modified our manuscript to include the additional citations and to elaborate on our reasoning for combining the satellite instrument records despite their instrumental differences:

"We calculate top-down surface NOx emissions from 1996 to 2014 using observations from GOME (1995-2003), SCIAMACHY (2002-2011) and GOME-2 (2007-). The similar overpass time of these three instruments (from about 9:30 a.m. to 10:30 a.m. local time) facilitates their combination to provide consistent long-term coverage (van der A et al., 2008; Konovalov et al. 2010; Geddes et al., 2016; Hilboll et al., 2013). We achieve consistency across all three instruments despite their varying pixel sizes (320 km x 40 km, 60 km x 30 km, and 80 km x 40 km for GOME, SCIAMACHY, and GOME-2 respectively) by gridding the daily observations from each to a regular coarse grid of 2° x 2.5° latitude by longitude."

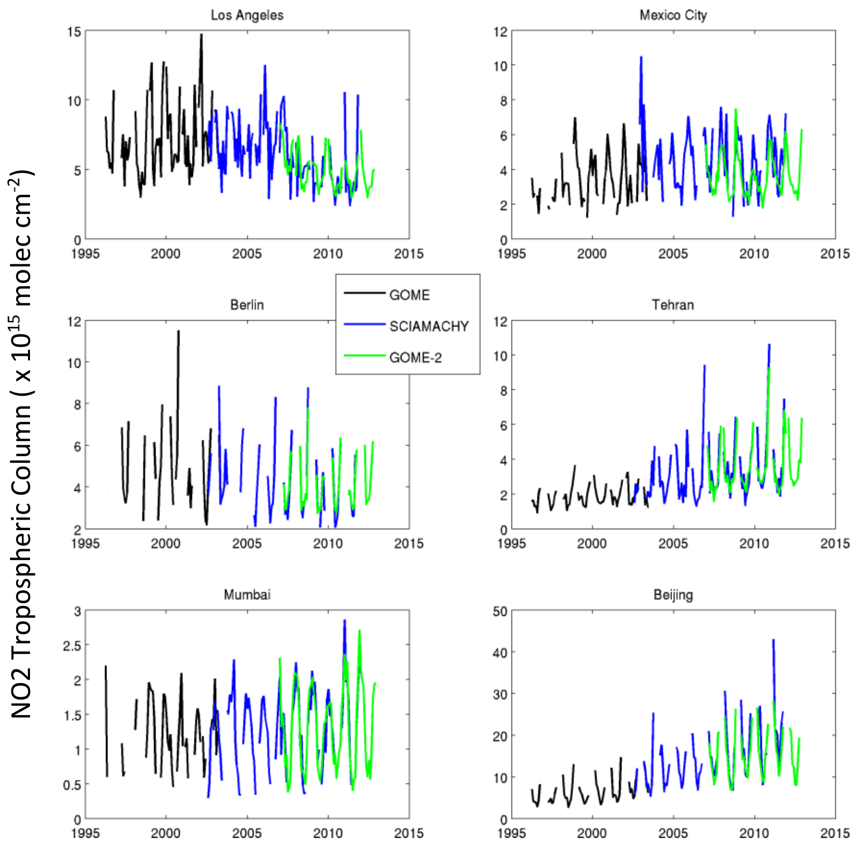

Figure D1: Monthly mean tropospheric NO2 from GOME, SCIAMACHY, and GOME-2 showing consistent agreement during overlap between instruments.

RC: Second, the three satellite instruments have some overlapping time periods (e.g. 2007-2011 for SCIAMACY and GOME-2). How do you use them for emission estimates during these periods? While these may have been presented in another study, a brief description here will help readers to better understand the method.

AR: In the case of GOME and SCIAMACHY, their period of overlap is small (June 2002 – April 2003). We use GOME observations alone for 2002, and SCIAMACHY observations alone for 2003. During the overlap between SCIAMACHY and GOME-2 (2007-2011), we use the GOME-2 observations given its more frequent global coverage (roughly every day) compared to SCIAMACHY (roughly every 6 days). We therefore use SCIAMACHY from years 2003-2006, and GOME-2 from years 2007-2014.

In response to the referee's comment, we have added the following to our manuscript:

"In our study, we use GOME observations for years 1996 to 2002, SCIAMACHY observations for years 2003 to 2006, and GOME-2 observations for years 2007 to 2014".

RC: And third, I strongly suggest present more information and analyses on the global top-down NOx emissions and their trends during the focused period by adding more text and a figure. This is not clear at present, and will be really helpful to understand the trends in NOy deposition as presented in Figure 6 and 7.

AR: We thank the referee for their suggestion, and agree that including details on the top-down emissions would be useful. We believe a table would be the optimal way to share this information. We have also included anthropogenic HNO_3 fluxes (from shipping emissions) to our calculation of total emissions to demonstrate full mass balance with NO_y deposition.

In response to this comment, we have added the following material to our manuscript:

"Table 1 shows the annual global top-down NOx emissions from our calculations. We derive global mean satellite-constrained NOx emissions from 1996-2014 of 55.6 $\pm$ 3.4 Tg N yr^{-1}. Our top-down global NOx emissions for 2001 of 52.3 Tg N are consistent with the mean from over 20 models used in the Coordinated Model Studies Activities of the Task Force on Hemispheric Transport of Air Pollution (HTAP) for the same year of 46.6 $\pm$ 7.8 Tg N (Vet et al. 2014)."

Table 1: Global top-down NOx emissions calculated using the finite mass balance inversion approach in combination with observations from GOME, SCIAMACHY, and GOME-2.

Year	Global NOx Emissions (Tg N yr^{-1})[a]
1996	60.1
1997	58.4
1998	59.2
1999	59.6
2000	53.4
2001	52.3
2002	55.1
2003	50.1
2004	51.5
2005	51.2
2006	50.0
2007	54.7
2008	56.1
2009	55.9
2010	57.5
2011	58.9
2012	59.3
2013	58.5
2014	54.0
Mean	55.6 $\pm$ 3.4

[a] Includes anthropogenic HNO_3 flux of 2.3 ± 0.1 Tg N yr^{-1}.

RC: Monthly mean simulated NO2 columns are calculated using days with coincident satellite observations". How do you select coincident days when using monthly-mean gridded satellite NO2 observations (Line 11)? And how do you sample the model simulation? Please clarify.

AR: First, we grid the daily NO_2 tropospheric column retrievals to a regular 2° x 2.5° grid, then calculate monthly means ourselves. Therefore, we are able to keep track of which 2° x 2.5° grid boxes have satellite observations that pass quality control on each individual day. From our simulations we output late morning mean vertical NO_2 column profiles every day. Using this information, we calculate monthly mean tropospheric NO_2 columns from the model simulation only on days that are coincidently sampled by successful satellite observations.

In response to the referee's comment, we have clarified this in our manuscript in the following places:

"We achieve consistency across all three instruments despite their varying pixel sizes (320 km x 40 km, 60 km x 30 km, and 80 km x 40 km for GOME, SCIAMACHY, and GOME-2 respectively) by gridding the daily observations from each to a regular coarse grid of 2° x 2.5° latitude by longitude."

"In all cases, monthly mean simulated NO_2 columns are calculated using days with coincident satellite observations. The simulated NO_2 vertical column is output daily for late morning."

RC: Do you mean you do not change the seasonality of NO_x emissions in the model? Please clarify. And what is the NO_x emission seasonality in the model? This is not described in the Appendix

AR: The referee is correct. We have not changed the seasonality of NO_x emissions in the model. We assume that the scaling factor determined from coincidently sampled model and satellite NO_2 tropospheric columns applies uniformly to the model prior emissions all year long regardless of whether successful satellite observations are available.

In response to the referee's comment, we have modified our manuscript to clarify our approach:

"In all cases, monthly mean simulated NO_2 columns are calculated using days with coincident satellite observations. The simulated NO_2 vertical column is output daily for late morning. We calculate scaling factors for every month with available satellite observations, then calculate an annual mean scaling factor that is used to infer annual mean from the mean monthly top-down emissions. Our top-down emissions retain the same seasonality as the prior emissions to mitigate concerns about seasonally missing data (such as from snow or monsoonal clouds)."

We have also added the following details regarding the emission seasonality to the Appendix:

"Monthly scaling of NO_x emissions are included in North America (based on the VISTAS inventory), Europe (based on the EMEP inventory), and Asia (based on the Zhang et al. (2009) inventory). Monthly scaling of EDGAR emissions is based on the seasonality from the Global Emission Inventory Activity (Benkovitz et al 1996)."

RC: Is there any trends in the export efficiency or changes in the export fraction during the period 1996-2014 over the US and Asia? From Line 11 below, it appears that the export fractions over Europe have a decreasing trend.

AR: We thank the reviewer for their comment, and have evaluated the statistical significance of trends in export fraction over each region in the same manner as our evaluation of long term trends in deposition.

In response to the referee's comment, we have added the following details to our manuscript:

"We estimate a similar fraction of NO_x export from the continental US using our observationally-constrained simulation (34% $\pm$ 2% from 1996-2014), with a small declining trend from a maximum of 38% in 1999 to a minimum of 31% in 2013.

"We calculate mean export of NO_x emissions from western European countries to be 45% $\pm$ 4%, with a notable decreasing trend from a maximum of 50% in 1997 to a minimum of 39% in 2014."

"We estimate that an average of 24% $\pm$ 4% emissions from China are exported, varying over time from as little as 15% of emissions in 1998 to a maximum of 31% of emissions in 2011 (an overall increasing trend)."

RC: A recent study on atmospheric nitrogen deposition over China reported a NOy export fraction of 36% (Zhao et al., 2017), not that different from the values for Europe and the US, compared with 24% in this study. Can you explain why? different NOx emissions, inclusion of adjacent oceans, or model horizontal resolution?

AR: We thank the reviewer for this important reference, which we have added to our manuscript. In our opinion, the most obvious explanation for the discrepancy is in model resolution or perhaps the rapidly changing satellite-constrained emissions over this period of time, which peak in China in 2011.

In response to the referee's comment, we have made the following changes to our manuscript:

"We estimate that an average of 24% $\pm$ 4% emissions from China are exported, varying over time from as little as of 15% of emissions in 1998 to a maximum of 31% of emissions in 2011 (an overall increasing trend). Zhao et al. (2017) used a higher resolution (0.5° x 0.667°) GEOS-Chem simulation and estimated that 36% of China's NOx emissions over 2008-2012 are exported. We calculate an export fraction of around 27% for the same time period. The discrepancy between the two estimates may be attributed to the coarser horizontal resolution of our simulation (2° x 2.5°), pointing to important resolution-dependent effects in global simulations of deposition. Other factors may include the use of different NOx emissions (our satellite-constrained emissions indicate rapid change over this period of time), and the treatment of adjacent oceans."

RC: Please explain "perturbing NH3 emissions everywhere". Increase or decrease? Do you change all anthropogenic and natural NH3 emissions, including the oceanic NH3 emissions?

AR: In response to the referee's comments, we have clarified the approach in our manuscript:

"Contemporary emissions of NH3 are highly uncertain (Reis et al., 2009), so we perform a sensitivity experiment by perturbing (increasing) all anthropogenic and natural NH3 emissions in the model by 25% for the year 2012."

RC: The unit "kg N ha^{-1} yr^{-2}" here might be confusing. Suggest add here "at a rate of…" or use annual deposition changes during the period.

AR: We thank the referee for their suggestion. In response, we have modified our manuscript to use the following wording:

"NOy deposition declined most steeply throughout the northeastern United States at a rate of up to -0.6 kg N ha^{-1} yr^{-2}"

"In Europe, statistically significant declines at a rate of up to -0.1 kg N ha^{-1} yr^{-2} are seen over some western countries."

"On the other hand, NO_y deposition has increased substantially throughout East Asia, exceeding a rate of +0.6 kg N ha^{-1} yr^{-2} in some parts."

RC: Energy statistics are used to scale emissions between 1996 and 2010. How about emissions after 2010? Please clarify.

AR: We clarify this in our manuscript:

"For other species and for emissions beyond 2010, the closest available year is used."

RC: Please state in the figure caption that the sensitivity test is for the year 2012.

AR: In response to the reviewer's comment, we have added this to the figure caption.

REFERENCES:

Benkovits et al. (1996), Global gridded inventories of anthropogenic emissions of sulfur and nitrogen, *Journal of Geophysical Research-Atmospheres,* 101, 29239–29253

Irie et al. (2012), Quantitative bias estimates for tropospheric NO2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia, *Atmospheric Measurement Techniques*, 5, 2403-2411

Konovalov et al. (2010) Multi-annual changes of NOx emissions in megacity regions: nonlinear trend analysis of satellite measurement based estimates, Atmos. Chem. Phys., 10, 8481–8498

van der A et al. (2008), Trends, seasonal variability, and dominant NOx source derived from a ten year record of NO2 measured from space, *Journal of Geophysical Research-Atmospheres, 113, doi:10.1029/2007JD009021*

Zhao et al. (2017), Atmospheric nitrogen deposition to China: A model analysis on nitrogen budget and critical load exceedance, *Atmospheric Environment,* 153, 32-40