Author comments in reply to the anonymous referee on "Global lightning NOx production estimated by an assimilation of multiple satellite datasets" by K. Miyazaki et al.

We want to thank the referee for the helpful comments and suggestions. We have revised the manuscript according to the comments, and hope that the revised version of the manuscript is now suitable for publication. Below are the referee comments in italics with our replies in normal font.

Reply to Referee #3

1. *The a posteriori lightning NOx product will reflect corrections to convolved errors in the model representation of both flash activity and NOx yields per flash. The lightning flash rate was not assimilated (satellite coverage is poor; global ground networks have low detection efficiencies). However, the flash rate parameterization was also not adjusted to match the satellite climatology from LIS/OTD, as is done for most global models. This is surprising, because the global lightning flash rate distribution is the best-known aspect of the lightning NOx source. If the authors wish to maintain discussion of the assimilated LNOx emissions in the individual context of the unconstrained flash rate (Section 6.2.1) versus NOx yields per flash (Section 6.2.2) –Âa both of which have very large uncertainties in models – then the flash rate distribution of the model should be shown and quantitatively evaluated against the spatial and seasonal distribution from LIS/OTD. The authors seem to suggest that the lightning flash rate parameterization performs very well when unconstrained, which would be a very surprising result in the context of the literature (e.g., Tost et al., 2007), and therefore should be documented.*

Evaluation results using the LIS/OTD measurements are shown in Tables 1, 7, 8, and in Figs. 2, 14 in the revised manuscript. To discuss the results, the following sentences have been added in Section 3.2.1:

"Table 1 and Figure 2 compare the global flash rate between the LIS/OTD high resolution monthly climatology (HRMC) data (Cecil et al., 2014) and the model parameterisation. Compared with the observations, the global distribution of the total flash rate is generally reproduced by the model."

"Mainly because of the low bias over central Africa, the model underestimates the annual flash rate in the tropics (20S-20N) by about 27 %, leading to about 13 % underestimation in the global total flash rate."

The relevant sentence in Section 7 has been rewritten as follow:

"First, errors in flash rates can explain only a small fraction of the uncertainty in LNOx estimates, as the main observed features of the annual global flash rate are generally reproduced by the model, except for the large low bias over central Africa."

The evaluation results for the flash rate and the NO production efficiency using the LIS/OTD observations are discussed in Section 6.2.1 and Section 6.2.2, respectively, in the revised manuscript. Please also see my reply below.

2. *The technique used here should not be able to distinguish between co-located NOx emission sources in a grid cell (e.g., surface lightning and anthropogenic sources, free tropospheric aircraft and lightning), and assumably depends on the a priori fraction of emissions for source attribution. If this is the case, some discussion should be included as it pertains to the results presented here. E.g., if the Ott et al. (2010) vertical probability distributions for lightning emissions were used instead of Pickering et al. (1998), which had a much smaller fraction emitted in the boundary layer, then the assimilation would attribute more of its surface NOx corrections to anthropogenic sources than lightning, which would influence the total lightning NOx value. Corrections of biases in surface sources in strongly polluted but lightning-prone regions (e.g., Gulf Coast, Congo) may be erroneously ascribed to lightning. Similarly, it is unclear to me how this technique could be used to differentiate between IC and CG flash yields, unless they have very separate spatiotemporal signatures from one another.*

The combined use of the multiple datasets with different vertical sensitivities will provide some information on the vertical LNOx profile (see section 6.2.3) and allows the assimilation to distinguish between the surface NOx emissions and LNOx sources as described in Section 3.2.2. Furthermore, transport from the source region is different for sources at different altitudes. In the boundary layer, the LNOx fraction can in principle be constrained in a meaningful way if it is the dominant source. In case of strong, simultaneous surface sources the emission adjustments will be distributed according to the assumed errors in the surface and lightning sources, and the a priori fraction of the NOx emissions (see Section 6.1.3). As suggested by the referee, use of a different a priori LNOx profile may indeed affect the LNOx source analysis.

The relevant sentence in Section 6.1.3 has been written as:
"Therefore, the estimated LNOx sources could have large uncertainties, especially where the surface emissions are large and variable."

The importance of separately estimating for IC and CG flashes is discussed in Section 6.2.2. The following sentence has been added in Section 3.2.2:
"The data assimilation optimizes the multiplication factors for the total LNOx sources, and does not separately optimise for IG and CG flashes.

The following sentence has been added in Section 6.2.3:

"When the observational constraints are insufficient to adjust the vertical profiles, changes in a priori LNOx source profiles (e.g., from the profiles of Pickering et al. (1998) to those of Ott et al. (2010)) or changes in the vertical structure of the covariance matrix will affect the analysed profiles."

Specific Comments

p29206 l25-27 - Does it not also have the potential to introduce larger errors if uncertainties are large in the additional constraint? e.g., the bias in TES UT ozone as shown in Fig. 8?

Additional error sources can be introduced by simultaneous data assimilation. This point is discussed in Section 6.1.1.

p29206 l29 - I suspect the "while" is erroneous?

Replaced by "whole".

p29207 l13-17 - Equation 1 would be better placed in Section 3.

Moved to Section 3.1.2.

p29207 l23 - remove subjective term "strong," perhaps replace with "useful"

Replaced.

p29208 l23-25 - There appears to be a missing word after "halfway"?

Replaced by "in the middle of the clouds".

p29210 l6-10 - Version and access date should be given for the OMI/MLS product, which has changed over time.

Added.

p29211 l25 – p29212 l4 - What is meant by "based on"?

Replaced by "obtained from".

p29212 l3-4 - The authors should compare the aircraft emissions used here in the context other estimates from the literature (e.g., Wilkerson et al., 2010, http://doi.dx.org/10.5194/acp-10-6391-2010). The interpretation of the assimilated LNOx results will be sensitive to uncertainty in aircraft emissions, which should be acknowledged.

The sentence has been rewritten as:
"The total NOx emission by aircraft is obtained from EDGAR as 0.55 TgN yr-1, which is similar to a more recent estimate of 0.49 TgN yr-1 for 2004 (Wilkerson et al., 2010)."

The following sentence has been added in Section 6.1.2:
"Although the aircraft NOx emissions likely have relatively small uncertainties (e.g., Wilkerson et al., 2010), the LNOx source estimates might be influenced by errors in the aircraft emissions especially along the major flight routes in the northern extratropics."

Section 3.1.2. This section could use clarification, particularly for readers not familiar with data assimilation and/or EnKF. It would be helpful to include a sentence or two that qualitatively describe how the EnKF works. Does the error covariance matrix take into account errors in the observations (e.g., those discussed in Section 6.1.1), or does EnKF blindly treat all the satellite products as truth, even in instances where we know the observations to be poor or highly uncertain? What averaging kernels are used in H(x), assumably those from each satellite product? What is the value of k?

Section 3.1.2 has been expanded and reformulated. The following sentences have been added:
"The EnKF uses an ensemble forecast to estimate the background error covariance matrix. The advantage of the EnKF over 4D-VAR is its easy implementation for complicated systems; i.e., it does not require the development of an adjoint code. The EnKF data assimilation technique employed is local ensemble transform Kalman filter (LETKF, Hunt et al., 2007). The LETKF scheme, which is based on the ensemble square root filter (SRF) method (e.g., Whitaker and Hamill, 2002), generates an analysis ensemble mean and covariance that satisfy the Kalman filter equations for linear models. The LETKF has conceptual and computational advantages over the original EnKF. The analysis performed locally in space and time reduces sampling errors caused by limited ensemble size, which also enable us to perform parallel computation. The computational advantages are important for this study because of the large state vector size."

"The spatial interpolation operator (S) is first applied to the model fields x in order to interpolate to the horizontal location of each observation and the height of each of the vertical layers. The averaging kernel (A) is then applied to define the sensitivity of the estimated state to changes to the true state. Because of the operator, the a priori profile (xapriori) does not, or only weakly, influence the model-observation difference in the data assimilation. The averaging kernel (A) and the a priori profile (xapriori) information provided for each retrieval is used in the data assimilation."

"In conclusion, the data assimilation updates model variables (the concentrations and the emission multiplication factors) for every grid point. This analysis is based on the observational information (i.e., the satellite retrievals) and the background error covariance estimated from the ensemble forecast with 48 members in our case. The estimated concentrations and emissions are used as initial conditions in the next step of ensemble model simulations and updated at every analysis step (i.e., 100 min.)."

p29213 l22-23, p29231 l9-10 neglect to acknowledge the existence of ground-based networks with global coverage, e.g., the World Wide Lightning Location Network (WWLLN; Abarca et al., 2010, http://dx.doi.org/10.1029/2009JD013411) or Vaisala's GLD360.

  The sentence has been rewritten as:
"The ground-based operational lightning detection networks (e.g., the World Wide Lightning Location Network (WWLLN)) provide lightning maps but they have low detection efficiencies (Abarca et al., 2010), whereas satellite instruments provide limited coverage on a daily basis."

p29214 l6-10 - why was a global scaling factor chosen to give 41.2 flashes s-1, rather than one to match the climatological value from satellites? Also, the more recent climatology using the combined LIS and OTD instruments (46 flashes s-1; Cecil et al., 2012, http://doi.dx.org/10.1016/j.atmosres.2012.06.028) should be referenced, instead of the old OTD-only reference.

  The scaling factor was used to obtain a realistic flash estimate based on a comparison with an older observation data. The sentences have been rewritten as:
"A globally and annually constant tuning factor is applied for the total flash frequency in CHASER simulations to obtain a realistic estimate of the global total flash occurrence, whereas the spatial distribution of the flash frequency is determined by the model parameterization."

"The simulated average global flash rate for 2007 is 41.2 flashes s-1, which is comparable to the climatological estimates of 44+-5 flashes s-1 derived from the Optical Transient Detector (OTD) measurements (Christian et al., 2003) and 46 flashes s-1 derived from the Lightning Imaging Sensor (LIS) and OTD measurements (Cecil et al., 2014). The difference between the model simulation and the

observations is partly attributed to interannual variations in flash activity; the annual total flash rate for the latitude band 35S-35N in 2007 observed from the LIS measurement is about 3 % lower than those from the climatology. Because only LIS measurements are available in 2007 and because the global coverage was not obtained, this study uses the climatological observations obtained from a combination of LIS and OTD measurements to validate the global flash rate."

p29214 l20-24 - z is not the IC/CG ratio as stated by the authors, but the CG proportion of total flashes. (Otherwise, setting z to zero makes no sense). Also, the coefficients for z given here are those from Price and Rind (1993, http://dx.doi.org/10.1029/93GL00226), not those in Price et al. (1997).

Corrected.

p29214 l26 – p29215 l1, p29232 l18-19 - The difference in yields between IC and CG flashes is still very uncertain. Comparison of what is used here with the literature should be given. Most recent work suggests the CG/IC production ratio should be closer to unity, cf. Table 19 of Schumann and Huntrieser (2007), although not all (e.g., Koshak et al., 2013; http://dx.doi.org/10.1016/j.atmosres.2012.12.015).

The sentences have been rewritten as:
"Second, following Price et al. (1997), the LNOx source amounts are calculated on the basis of a lightning NO production of 1100 moles per CG flash and 110 moles per IC flash, with a mean energy per CG flash of 6.7×10^9 J flash-1."
"However, it has been suggested that the ratio should be closer to 1 than to 10 (Gallardo and Cooray, 1996; Fehr et al., 2004; DeCaria et al., 2005), although a more recent estimate by Koshaz et al (2014) showed the ratio to be closer to 10."

p29215 l1-5 - Were the Pickering et al. (1998) profiles scaled to local cloud top height, or were fixed altitudes used? Why were the Pickering et al. (1998) profiles used instead of the Ott et al. (2010) profiles?

The Pickering et al. (1998) profiles were scaled to local cloud top height. CHASER uses the Pickering et al. (2008) profiles because the Ott et al. (2010) profiles were not available when CHASER was developed (and the model has not yet been updated).

p29216 l4 - "lighting" should be "lightning."

Corrected.

p29216 l5-6 - The Cooper et al. (2007) and Hudman et al. (2007) studies examined North America, not the tropical upper troposphere. Better references for comparison would be Sauvage et al. (2007, http://dx.doi.org/10.5194/acp-7-815-2007) or Murray et al. (2012), who examined the influence of lightning in the tropics.

Corrected.

p29221 l3-7 - The assimilated changes in mean OH could be independently evaluated by comparison to the methyl chloroform and methane lifetimes, available from observational constraints (cf. John et al., 2012, and references therein; http://dx.doi.org/10.5194/acp-12-12021-2012). In addition to OH, I would also expect a major benefit of the multiple-species to be in its ability to constrain ozone production efficiencies (OPE, which may be approximated as PO3/PHNO3, cf. Cooper et al., 2010, http://dx.doi.org/10.1029/2010JD015056), which are non-linearly dependent on NOx, and would be important for inversely determining LNOx emissions from ozone observations.

Evaluations of the analysed OH fields and the OPE are very interesting topics. However, these are beyond the main scope of this study and should be discussed in a separate paper. To note the importance, the following sentence has been added in Section 4.2:

"The simultaneous assimilation also has the ability to constrain ozone production efficiencies (OPE) through the NOx-CO-OH-O3 set of chemical reactions, which may improve the LNOx source estimation with the assimilation of TES O3 data. Detailed analyses are required to measure the impact of the simultaneous assimilation on OPE."

p29222 Section 5.1 - The authors might consider showing Ascension instead of Irene, given the expected strong influence of lightning on the South Atlantic ozone maximum, the dominant mode of seasonal variability in tropical ozone (e.g., Sauvage et al., JGR, 2007, http://dx.doi.org/10.1029/2006JD008008).

Added.

p29223 l4-5 - Convection and lightning are heavily parameterized everywhere in the model. Please cut, or give an objective argument as to why tropical W Pacific is expected to have worse convection or lightning than elsewhere in the model.

The sentence has been rewritten as:

"Large uncertainties in the LNOx sources are expected over the tropical western Pacific because of errors in the tropical Pacific ITCZ cumulus clouds simulated by the AGCM (Emori et al., 2005)."

p29223 l15-17 - Please clarify what is being compared in these sentences.

The following sentence has been added:

"The analysed NO2 and O3 concentrations show better agreements with the observations (Fig. 11) because of the simultaneous data assimilation."

p29224 l26-27 - Please justify why large uncertainties in cumulus cloud and biomass burning activity are "expected" in this region

The sentence has been rewritten as follow:

"In this region, large uncertainties in the simulated cumulus cloud and biomass burning activity are expected, as suggested by Emori et al. (2005) and Stroppiana et al. (2010), respectively."

p29228 l18 - "tests is" should be "tests are"

Corrected.

p29228 l25-28 - Please clarify what is meant by the phrases "mean analysis spread" and "spin-up period for the assimilation" (I thought Kalman filters only require the previous state?). Also, "week" should be plural.

The sentences have been rewritten as

"The mean analysis spread, as estimated by transforming the background ensemble in the data assimilation (c.f., Eq. (4)), is about 0.9 TgNyr-1 for the annual global source strength"

"The LNOx analysis is obtained from information of roughly two weeks of measurements, as demonstrated by the spin-up period of the assimilation (i.e., the spin-up period was required to obtain a converged solution in the analysis)"

p29231 l16-19, p29237 l2-3 - I find this conclusion weak unless more is done to objectively evaluate the flash rate distribution in the model. It could easily be due to a systematic low bias in the a priori NOx production per flash over the ocean. Whether or not this is primarily due to underestimation of (1) the

flash rate, or (2) NOx yields per flash over marine regions could be determined by comparison of the simulated flash rates with the LIS/OTD climatology.

Table 7 and the following discussions have been added in Section 6.2.1:
"We note that comparisons against the LIS/OTD observations consistently reveal a larger underestimation in the parameterised global flash rate over the oceans (about 27 %) than over land (about 5 %). On the other hand, over the tropical oceans (Pacific, Atlantic, and Indian Oceans), the difference between the observed and the parameterised flash rate is relatively small, as summarised in Table 7. This suggests that errors in the NOx production efficiency rather than those in the flash rate could be responsible for the large increase in the LNOx sources over the tropical oceans. This will be further discussed in Sect. 6.2.2."

The relevant discussions in Section 6.2.2 have been rewritten as follow:
"The annual global LNOx source from our estimates corresponds to a mean NO production of about 350 mol flash-1 based on the parameterized flash rate, as summarized in Table 7. Because errors in the parameterized flash rate influence this estimation, we also use the LIS/OTD climatological observations; a global mean NO production of about 310 mol flash-1 is estimated using the flash observations. Both these values are within the range of most other recent estimates."

"Our analysis for July consistently reveals a large production per flash of 430 and 350 mol of NO in the NH (20-90N) compared to 360 and 240 mol of NO in the tropics (20S-20N) based on the parameterised flash rate and the LIS/OTD observations, respectively. There are also obvious regional differences; e.g., a large production per flash of about 440 and 570 are estimated for the northern Eurasia continent based on the parameterised flash rate and the LIS/OTD observations, respectively, as summarised in Table 7 and shown in Fig. 14. The detailed spatial structures in the production efficiencies estimated from the analysed LNOx sources and the observed and the parameterised flash rates (Fig. 14) may reflect not only variations in flash characteristics but also noises and errors in the assimilated and flash measurements (c.f., Section 6.1.2). Note that the production efficiency estimated using the observed flash rate becomes unrealistically large locally where the observed flash rate is much smaller than the model flash rate."

The following paragraph has been also added in Section 6.2.2:
"The NO production efficiencies estimated using the simulated total LNOx sources and the simulated flash rate by the model parameterization (without any assimilation) are about 20 % lower over land and about 11 % lower over the oceans, compared with those estimated using the analysed LNOx sources and the LIS/OTD observations. The obtained results imply general underestimations in the NOx production efficiency simulated by the model, although there are obvious regional differences in the estimates (Table

7). The underestimation could be attributed to errors either in the parameterised IC/CG flash ratio (c.f., Eq. (5)) or in the assumptions on the production efficiency of IC and CG flashes."

The following sentence has added in Section 7:
"It is also suggested that the model parameterisation may underestimate the annual and global mean NO production efficiency by about 10 % over land and 20 % over the oceans."

p29231 l23-24 - 6.3 Tg N yr-1 using a global mean flash rate of 46 flashes s-1 from the LIS/OTD climatology corresponds to 310 mol per flash. Do you expect your 41.2 flashes s-1 for 2006 could be explained by interannual variability in the global mean lightning flash rate?

Please see my reply above.

p29232 l10-14 - An extremely useful figure for the community would be a map of the average NOx yield per flash, calculated by using the assimilated LNOx emissions divided by the flash rate distribution from (1) the model parameterization, and (2) the LIS/OTD climatology. This would be helpful for informing CTMs/CCMs as how to implement differential LNOx yields per flash, which are typically done in arbitrary manner, but necessary for matching global ozone distributions. To me, this is the most useful and unique scientific contribution enabled by this work. The greatest uncertainty global models face at present in reproducing the lightning NOx source is in NOx yields per flash, since most constrain the flash rate magnitude and distribution to the LIS/OTD climatology.

Thank you for your advice. In the revised paper, Table 7 and Fig. 14 demonstrate the NO production efficiencies estimated from the model flash rate and the LIS/OTD observations. Table 8 compares our estimates with other estimates.

The following sentence has been added in Section 6.3:
"The annual global LNOx source from our estimates corresponds to a mean NO production of about 310 mol flash-1 based on the LIS/OTD climatological observations. This value is also within most of the recent estimates (c.f., Table 8)."

Please also see my reply above.

p29237 l7-8 - Please rephrase to make it clear that this is because errors in simulated flash rates are small in this study. Many CTM studies find it necessary to constrain the lightning flash rates for their

ozone simulations (e.g., Martin et al., 2007; Sauvage et al., 2007; Jourdain et al., 2010; Allen et al., 2010; Murray et al., 2012).

The sentence has been rewritten as:

"First, errors in flash rates can explain only a small fraction of the uncertainty in LNOx estimates, as the main observed features of the annual global flash rate are generally reproduced by the model, except for the large low bias over central Africa."

Fig. 9 - superfluous axis labels and titles could be removed to increase panel box sizes

Removed.

Fig. 12 caption should clearly state which difference is taken (I assume with minus without the cloud-covered observations)?

Corrected.