Estimates of lightning NO\textsubscript{x} production based on high-resolution OMI NO\textsubscript{2} retrievals over the continental US

Xin Zhang\textsuperscript{1,2}, Yan Yin\textsuperscript{1,2}, Ronald van der A\textsuperscript{2,3}, Jeff L. Lapierre\textsuperscript{4}, Qian Chen\textsuperscript{1,2}, Xiang Kuang\textsuperscript{1,2}, Shuqi Yan\textsuperscript{2}, Jinghua Chen\textsuperscript{1,2}, Chuan He\textsuperscript{1,2}, and Rulin Shi\textsuperscript{1,2}

\textsuperscript{1}Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters/Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China
\textsuperscript{2}Department of Atmospheric Physics, Nanjing University of Information Science and Technology (NUIST), Nanjing 210044, China
\textsuperscript{3}Royal Netherlands Meteorological Institute (KNMI), Department of Satellite Observations, De Bilt, the Netherlands
\textsuperscript{4}Earth Networks, Germantown, Maryland, USA

Correspondence: Yan Yin (yinyan@nuist.edu.cn)

Received: 2 October 2019 – Discussion started: 27 November 2019
Revised: 8 March 2020 – Accepted: 11 March 2020 – Published: 7 April 2020

Abstract. Lightning serves as the dominant source of nitrogen oxides (NO\textsubscript{x} = NO + NO\textsubscript{2}) in the upper troposphere (UT), with a strong impact on ozone chemistry and the hydroxyl radical production. However, the production efficiency (PE) of lightning nitrogen oxides (LNO\textsubscript{x}) is still quite uncertain (32–1100 mol NO per flash). Satellite measurements are a powerful tool to estimate LNO\textsubscript{x} directly compared to conventional platforms. To apply satellite data in both clean and polluted regions, a new algorithm for calculating LNO\textsubscript{x} has been developed that uses the Berkeley High-Resolution (BEHR) v3.0B NO\textsubscript{2} retrieval algorithm and the Weather Research and Forecasting model coupled with chemistry (WRF-Chem). LNO\textsubscript{x} PE over the continental US is estimated using the NO\textsubscript{2} product of the Ozone Monitoring Instrument (OMI) data and the Earth Networks Total Lightning Network (ENTLN) data. Focusing on the summer season during 2014, we find that the lightning NO\textsubscript{2} (LNO\textsubscript{2}) PE is 32 ± 15 mol NO\textsubscript{2} per flash and 6 ± 3 mol NO\textsubscript{2} per stroke while LNO\textsubscript{x} PE is 90 ± 50 mol NO\textsubscript{2} per flash and 17 ± 10 mol NO\textsubscript{2} per stroke. Results reveal that our method reduces the sensitivity to the background NO\textsubscript{2} and includes much of the below-cloud LNO\textsubscript{2}. As the LNO\textsubscript{x} parameterization varies in studies, the sensitivity of our calculations to the setting of the amount of lightning NO (LNO) is evaluated. Careful consideration of the ratio of LNO\textsubscript{2} to NO\textsubscript{2} is also needed, given its large influence on the estimation of LNO\textsubscript{2} PE.

1 Introduction

Nitrogen oxides (NO\textsubscript{x}) near the Earth’s surface are mainly produced by soil, biomass burning, and fossil fuel combustion, while NO\textsubscript{x} in the middle and upper troposphere originates largely from lightning and aircraft emissions. NO\textsubscript{x} plays an important role in the production of ozone (O\textsubscript{3}) and the hydroxyl radical (OH). While the anthropogenic sources of NO\textsubscript{x} are largely known, lightning nitrogen oxides (LNO\textsubscript{x}) are still the source with the greatest uncertainty, though they are estimated to range between 2 and 8 Tg N yr\textsuperscript{-1} (Schumann and Huntrieser, 2007). LNO\textsubscript{x} is produced in the upper troposphere (UT) by O\textsubscript{2} and N\textsubscript{2} dissociation in the hot lightning channel as described by the Zel’dovich mechanism (Zel’dovich and Raizer, 1967). With the recent updates of UT NO\textsubscript{x} chemistry, the daytime lifetime of UT NO\textsubscript{x} is evaluated to be ~3 h near thunderstorms and ~0.5–1.5 d away from thunderstorms (Nault et al., 2016, 2017). This results in enhanced O\textsubscript{3} production in the cloud outflow of active convection (Pickering et al., 1996; Hauglustaine et al., 2001; DeCaria et al., 2005; Ott et al., 2007; Dobber et al., 2008; Allen et al., 2010; Finney et al., 2016). As O\textsubscript{3} is known as
a greenhouse gas, strong oxidant, and absorber of ultraviolet radiation (Myhre et al., 2013), the contributions of LNO\textsubscript{x} to O\textsubscript{3} production also have an effect on climate forcing. Finney et al. (2018) found different impacts on atmospheric composition and radiative forcing when simulating future lightning using a new upward cloud ice flux (IFLUX) method versus the commonly used cloud-top height (CTH) approach. While global lightning is predicted to increase by 5\%–16\% over the next century with the CTH approach (Clark et al., 2017; Banerjee et al., 2014; Krause et al., 2014), a 15\% decrease in global lightning was estimated with IFLUX in 2100 under a strong global warming scenario (Finney et al., 2018). As a result of the different effects on radiative forcing from ozone and methane, a net positive radiative forcing was found with the CTH approach while there is little net radiative forcing with the IFLUX approach (Finney et al., 2018). However, the convective available potential energy (CAPE) times the precipitation rate (P) proxy predicts a 12 \pm 5\% increase in the continental US (CONUS) lightning strike rate per kelvin of global warming (Romps et al., 2014), while the IFLUX proxy predicts the lightning will only increase 3.4 \% K\textsuperscript{−}1 over the CONUS. Recently, Romps (2019) compared the CAPE \times P proxy and IFLUX method in cloud-resolving models. They reported that higher CAPE and updraft velocities caused by global warming could lead to the large increases in tropical lightning simulated by the CAPE \times P proxy, while the IFLUX proxy predicts little change in tropical lightning because of the small changes in the mass flux of ice.

In view of the regionally dependent lifetime of NO\textsubscript{x} and the difficulty of measuring LNO\textsubscript{x} directly, a better understanding of the LNO\textsubscript{x} production is required, especially in the tropical and midlatitude regions in summer. Using its distinct spectral absorption lines in the near-ultraviolet (UV) and visible (VIS) ranges (Platt and Perner, 1983), NO\textsubscript{2} can be measured by satellite instruments like the Global Ozone Monitoring Experiment (GOME; Burrows et al., 1999; Richter et al., 2005), SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY; Boesvansmann et al., 1999), the Second Global Ozone Monitoring Experiment (GOME-2; Callies et al., 2000), and the Ozone Monitoring Instrument (OMI; Levelt et al., 2006). OMI has the highest spatial resolution, least instrument degradation, and longest record among these satellites (Krotkov et al., 2017). Satellite measurements of NO\textsubscript{2} are a powerful tool compared to conventional platforms because of their global coverage, constant instrument features, and temporal continuity.

Recent studies have determined and quantified LNO\textsubscript{x} using satellite observations. Beirle et al. (2004) constrained the LNO\textsubscript{x} production to 2.8 (0.8–14) Tg N yr\textsuperscript{−}1 by combining GOME NO\textsubscript{2} data and flash counts from the Lightning Imaging Sensor (LIS) aboard the Tropical Rainfall Measurement Mission (TRMM) over Australia. Boersma et al. (2005) estimated the global LNO\textsubscript{x} production of 1.1–6.4 Tg N yr\textsuperscript{−}1 by comparing GOME NO\textsubscript{2} with distributions of LNO\textsubscript{2} modeled by Tracer Model 3 (TM3). Martin et al. (2007) analyzed SCIAMACHY NO\textsubscript{2} columns with Goddard Earth Observing System chemistry model (GEOS-Chem) simulations to identify LNO\textsubscript{x} production amounting to 6 \pm 2 Tg N yr\textsuperscript{−}1.

As these methods focus on monthly or annual mean NO\textsubscript{2} column densities, more recent studies applied specific approaches to investigate LNO\textsubscript{x} directly over active convection. Beirle et al. (2006) estimated LNO\textsubscript{x} as 1.7 (0.6–4.7) Tg N yr\textsuperscript{−}1 based on a convective system over the Gulf of Mexico, using National Lightning Detection Network (NLDN) observations and GOME NO\textsubscript{2} column densities. However, it is assumed that all the enhanced NO\textsubscript{2} originated from lightning and did not consider the contribution of anthropogenic emissions. Beirle et al. (2010) analyzed LNO\textsubscript{x} production systematically using the global dataset of SCIAMACHY NO\textsubscript{2} observations combined with flash data from the World Wide Lightning Location Network (WWLLN). Their analysis was restricted to 30 km \times 60 km satellite pixels where the flash rate exceeded 1 flash km\textsuperscript{−}2 h\textsuperscript{−}1. But they found LNO\textsubscript{x} production to be highly variable, and correlations between flash-rate densities and LNO\textsubscript{x} production were low in some cases. Bucsel et al. (2010) estimated LNO\textsubscript{x} production as 100–250 mol NO\textsubscript{2} per flash for four cases, using the DC-8 and OMI data during NASA’s Tropical Composition, Cloud and Climate Coupling Experiment (TC4).

Based on the approach used by Bucsel et al. (2010), a special algorithm was developed by Pickering et al. (2016) to retrieve LNO\textsubscript{x} from OMI and the WWLLN. The algorithm takes the OMI tropospheric slant column density (SCD) of NO\textsubscript{2} (S\textsubscript{NO\textsubscript{2}}) as the tropospheric slant column density of LNO\textsubscript{2} (S\textsubscript{LNO\textsubscript{2}}) by using cloud radiances (CRF) greater than 0.9 to minimize or screen the lower tropospheric background. To convert the S\textsubscript{LNO\textsubscript{2}} to the tropospheric vertical column density (VCD) of LNO\textsubscript{2} (V\textsubscript{LNO\textsubscript{2}}), an air mass factor (AMF) is calculated by dividing the a priori S\textsubscript{LNO\textsubscript{2}} by the a priori V\textsubscript{LNO\textsubscript{2}}. The a priori S\textsubscript{LNO\textsubscript{2}} is calculated using a radiative transfer model and a profile of LNO\textsubscript{2} simulated by the NASA Global Modeling Initiative (GMI) chemical transport model. The a priori V\textsubscript{LNO\textsubscript{2}} is also obtained from the GMI model. Results for the Gulf of Mexico during 2007–2011 summer yield LNO\textsubscript{x} production of 80\%±45 mol NO\textsubscript{2} per flash. Since they considered NO\textsubscript{2} above the cloud to be LNO\textsubscript{x} in the algorithm due to the difficulty and uncertainty in determining the background NO\textsubscript{x}, their AMF and derived VCD of LNO\textsubscript{x} (LNO\textsubscript{2}) are named AMFLNO\textsubscript{x,Clean} (AMFLNO\textsubscript{2,Clean}) and LNO\textsubscript{2,Clean}, respectively. Note that Pickering et al. (2016) considered the two estimates of background derived from aircraft flights in the Gulf of Mexico region (3\% and 33\%) and subtracted the mean value (18\%) from the estimated mean LNO\textsubscript{x} production efficiency (PE) for the background bias. However, we use the original algorithm directly without correction to distinguish the effect of different AMFs on LNO\textsubscript{x} estimation in the remainder of this paper. Unless otherwise specified, abbreviations S and V are
respectively defined as the tropospheric SCD and VCD in this paper.

More recently Bucsela et al. (2019) obtained an average PE of 180 ± 100 mol NO$_2$ per flash over East Asia, Europe, and North America based on a modification of the method used in Pickering et al. (2016). A power function between LNO$_2$ and lightning flash rate was established, while the minimum flash-rate threshold was not applied. The tropospheric NO$_2$ background was removed by subtracting the temporal average of NO$_2$ at each box where the value was weighted by the number of OMI pixels which meet the optical cloud pressure and CRF criteria required to be considered deep convection but have one flash or fewer instead. The lofted pollution was considered to be 15% of total NO$_2$ according to the estimation from DeCaria et al. (2000, 2005), and the average chemical delay was adjusted by 15% following the 3 h LNO$_2$ lifetime in the nearby field of convection (Nault et al., 2017). However, there were negative LNO$_4$ values caused by the overestimation of the tropospheric background and stratospheric NO$_2$ at some locations.

On the other hand, Lapierre et al. (2020) constrained LNO$_2$ to 1.1 ± 0.2 mol NO$_2$ per stroke for intracloud (IC) strokes and 10.7 ± 2.5 mol NO$_2$ per stroke for cloud-to-ground (CG) strokes over the CONUS. LNO$_2$ per stroke was scaled to 24.2 mol NO$_2$ per flash using mean values of strokes per flash and the ratio of NO$_2$ to NO$_2$ in the UT. They used the regridded Berkeley High-Resolution (BEHR) v3.0A 0.05° × 0.05° “visible only” NO$_2$ VCD (V$_{vis}$) product which includes two parts of NO$_2$ that can be “seen” by the satellite. The first part is the NO$_2$ above clouds (pixels with CRF > 0.9) and the second part is the NO$_2$ detected from cloud-free areas. A threshold of 3 × 10$^{15}$ molecules cm$^{-2}$, the typical urban NO$_2$ concentration, was applied to mask the contaminated grid cells (Beirle et al., 2010; Laughner and Cohen, 2017). The main difference between Lapierre et al. (2020) and Pickering et al. (2016) is the air mass factor for lightning (AMF$_{LNO_2}$) implemented in the basic algorithm. In Lapierre et al. (2020), the air mass factor was used to convert $S_{NO_2}$ to $V_{vis}$, while in Pickering et al. (2016) it was used to convert $S_{LNO_2}$ to $V_{LNO_2}$, assuming that all $S_{NO_2}$ is generated by lightning.

To apply the approach used by Bucsela et al. (2010), Pickering et al. (2016), Bucsela et al. (2019), and Lapierre et al. (2020) without geographic restrictions, the contamination by anthropogenic emissions must be taken into account in detail. The Weather Research and Forecasting (WRF) model coupled with chemistry (WRF-Chem) has been employed to evaluate the convective transport and chemistry in many studies (Barth et al., 2012; Wong et al., 2013; Fried et al., 2016; Li et al., 2017). Meanwhile, Laughner and Cohen (2017) showed that the OMI AMF is increased by ≈ 35% for summertime when LNO$_2$ simulated by WRF-Chem is included in the a priori profiles to match aircraft observations. The simulation agrees with observed NO$_2$ profiles and

the bias of AMF related to these observations is reduced to < ±4% for OMI viewing geometries.

In this paper, we focus on the estimation of LNO$_2$ production per flash (LNO$_2$ per flash), LNO$_2$ production per flash (LNO$_2$ per flash), LNO$_2$ production per stroke (LNO$_2$ per stroke), and LNO$_2$ production per stroke (LNO$_2$ per stroke) in May–August (MJJJA) 2014 by developing an algorithm similar to that of Pickering et al. (2016) based on the BEHR NO$_2$ retrieval algorithm (Laughner et al., 2018, 2019), but it performs better over background NO$_2$ sources. Section 2 describes the satellite data, lightning data, model settings, and the algorithm in detail. Section 3 explores the suitable data criteria, compares different methods, and evaluates the effect of background NO$_2$, cloud, and LNO$_2$ parameterization on LNO$_2$ production estimation. Section 4 examines the effect of different sources of the uncertainty on the results. Conclusions are summarized in Sect. 5.

2 Data and methods

2.1 Ozone Monitoring Instrument (OMI)

OMI is carried on the Aura satellite (launched in 2004), a member of the A-train satellite group (Levelt et al., 2006, 2018). OMI passes over the Equator at ∼ 13:45 LT (ascending node) and has a swath width of 2600 km, with a nadir field-of-view resolution of 13 km × 24 km. Since the beginning of 2007, some of the measurements have become useless as a result of anomalous radiances called the “row anomaly” (Dobber et al., 2008; KNMI, 2012). For the current study, we used the NASA standard product v3.0 (Krotkov et al., 2017) as input to the LNO$_2$ retrieval algorithm.

The main steps of calculating the NO$_2$ tropospheric VCD (V$_{NO_2}$) in the NASA product include the following.

1. SCDs are determined by the OMI-optimized differential optical absorption spectroscopy (DOAS) spectral fit.

2. A corrected (“de-striped”) SCD is obtained by subtracting the cross-track bias caused by an instrument artifact from the measured slant column.

3. The AMF for stratospheric (AMF$_{strat}$) or tropospheric column (AMF$_{trop}$) is calculated from the NO$_2$ profiles integrated vertically using weighted scattering weights with the a priori profiles. These profiles are obtained from GMI monthly mean profiles using 4 years (2004–2007) of simulation.

4. The stratospheric NO$_2$ VCD (V$_{strat}$) is calculated from the subtraction of the a priori contribution from tropospheric NO$_2$ and a three-step (interpolation, filtering, and smoothing) algorithm (Bucsela et al., 2013).

5. V$_{strat}$ is converted to the slant column using AMF$_{strat}$ and subtracted from the measured SCDs to yield $S_{NO_2}$, leading to $V_{NO_2} = S_{NO_2}/AMF_{trop}$. 

www.atmos-meas-tech.net/13/1709/2020/ Atmos. Meas. Tech., 13, 1709–1734, 2020
Based on this method, we developed a new AMF_LNO_x to obtain the desired V_LNO_x (V_LNO_x = S_NO_2/AMF_LNO_x) by replacing the original step.

6. Details of this algorithm are discussed in Sect. 2.4.

2.2 The Earth Networks Total Lightning Detection Network (ENTLN)

The Earth Networks Total Lightning Network (ENTLN) operates a system of over 1500 ground-based stations around the world with more than 900 sensors installed in the CONUS (Zhu et al., 2017). Both IC and CG lightning flashes are located by the sensors with detection frequency ranging from 1 Hz to 12 MHz based on the electric field pulse polarity and wave shapes. Groups of pulses are classified as a flash if they are within 700 ms and 10 km. In the preprocessed data obtained from the ENTLN, both strokes and lightning flashes composed of one or more strokes are included.

Rudlosky (2015) compared ENTLN combined events (IC and CG) with LIS flashes and found that the relative flash detection efficiency of ENTLN over CONUS increases from 62.4 % during 2011 to 79.7 % during 2013. Lapierre et al. (2020) also compared combined ENTLN and the NLDN dataset with data from the LIS during 2014 and found the detection efficiencies of IC flashes and strokes to be 88 % and 45 %, respectively. Since we only use the ENTLN data in 2014 as Lapierre et al. (2020), and NLDN detection efficiency of IC pulses should be lower than 33 %, which is calculated by the data in 2016 (Zhu et al., 2016), only the IC flashes and strokes are divided by 0.88 and 0.45, respectively, while CG flashes and strokes are unchanged because of the high detection efficiency.

2.3 Model description

The present study uses WRF-Chem version 3.5.1 (Grell et al., 2005) with a horizontal grid size of 12 km × 12 km and 29 vertical levels (Fig. 1). The initial and boundary conditions of meteorological parameters are provided by the North American Regional Reanalysis (NARR) dataset with a 3-hourly time resolution. Based on Laughner et al. (2019), 3D wind fields, temperature, and water vapor are nudged towards the NARR data. Outputs from version 4 of the Model for Ozone and Related chemical Tracers (MOZART-4; Emmons et al., 2010) are used to generate the initial and boundary conditions of meteorological parameters are provided by the Model of Emissions of Gases and Aerosol from Nature (MEGAN; Guenther et al., 2006) is used for biogenic emissions. The domain and terrain height (m) of the WRF-Chem simulation with 350 × 290 grid cells and a horizontal resolution of 12 km.

The impact of these biases on LNO_x production is discussed and mitigated in Sect. 3.1 and 3.4. The bimodal profile modified from the standard Ott et al. (2010) profile (Laughner and Cohen, 2017) is employed as the vertical distribution of lightning NOx (LNOx) in WRF-Chem, while outputs of LNO and LNO2 profiles are defined as the difference of vertical profiles between simulations with and without lightning.

2.4 Method for deriving AMF

The V_LNO_x near convection is calculated according to

\[ V_{LNO_x} = \frac{S_{NO_2}}{AMF_{LNO_x}} \]

where S_NO_2 is the OMI-measured tropospheric slant column NO2, and AMF_LNO_x is a customized lightning air mass factor. The concept of AMF_LNO_x was also used in Beirle et al. (2009) to investigate the sensitivity of satellite instruments to freshly produced lightning NOx. In order to estimate LNO_x, we define the AMF_LNO_x as the ratio of the “visible” modeled NO2 slant column to the total modeled tropospheric LNO_x vertical column (derived from the a priori NO and NO2 profiles, scattering weights, and cloud radiance fraction):

\[ AMF_{LNO_x} = \frac{(1 - f_c)p_{\text{surf}}w_{\text{clear}}(p)\text{NO}_2(p) \, dp + f_c p_{\text{cloud}} w_{\text{cloudy}}(p)\text{NO}_2(p) \, dp}{\int p_{\text{surf}} LNO_x(p) \, dp} \]

where f_c is the cloud radiance fraction (CRF), p_surf is the surface pressure, p_TP is the tropopause pressure, p_cloud is the cloud optical pressure (CP), w_clear and w_cloudy are respectively the pressure-dependent scattering weights from
the TOMRAD lookup table (Bucsela et al., 2013) for clear and cloudy parts, and NO$_2$($p$) is the modeled NO$_2$ vertical profile. Details of these standard parameters and calculation methods are given in Laughner et al. (2018). LNO$_x$($p$) is the LNO$_x$ vertical profile calculated by the difference of vertical profiles between WRF-Chem simulations with and without lightning.

Please note that the CP is a reflectance-weighted pressure retrieved by the collision-induced O$_2$–O$_2$ absorption band near 477 nm (Acarreta et al., 2004; Sneep et al., 2008; Stammes et al., 2008). For a deep convective cloud with lightning, the CP lies below the geometrical cloud top, which is much closer to that detected by thermal infrared sensors, such as CloudSat and the Aqua Moderate Resolution Imaging Spectrometer (MODIS) (Vasilkov et al., 2008; Joiner et al., 2012). Hence, much of the tropospheric NO$_2$ measured by OMI lies inside the cloud rather than above the cloud top.

In the following, “above cloud” or “below cloud” is relative to the cloud pressure detected by OMI. The sensitivity study of Beirle et al. (2009) compared the chemical composition from the cloud bottom to that of the cloud top and revealed that a significant fraction of the NO$_2$ within the cloud originating from lightning can be detected by the satellite. This valuable cloud pressure concept has been applied not only in the LNO$_x$ research but also in the cloud slicing method of deriving the UT O$_3$ and NO$_x$ (Ziemke et al., 2009, 2017; Choi et al., 2014; Strode et al., 2017; Marais et al., 2018). As discussed in Pickering et al. (2016), the ratio of $V_{LNO_x}$ seen by OMI to $V_{LNO_x}$ is partly influenced by $p_{cloud}$. The effects of LNO$_2$ below the cloud will be discussed in Sect. 3.4.

To compare our results with those of Pickering et al. (2016) and Lapierre et al. (2020), we calculate their

\[
\text{AMF}_{LNO_x,\text{Clean}} \quad \text{and} \quad \text{AMF}_{NO_2 Vis},
\]

respectively:

\[
\text{AMF}_{LNO_x,\text{Clean}} = \frac{(1 - f_1) \int_{p_{\text{part}}}^{p_{\text{cloud}}} w_{\text{clear}}(p) LNO_2(p) \, dp + f_1 \int_{p_{\text{cloud}}}^{p_{\text{cloud}}} w_{\text{cloudy}}(p) LNO_2(p) \, dp}{\int_{p_{\text{part}}}^{p_{\text{cloud}}} LNO_2(p) \, dp},
\]

(3)

\[
\text{AMF}_{NO_2 Vis} = \frac{(1 - f_2) \int_{p_{\text{part}}}^{p_{\text{cloud}}} w_{\text{clear}}(p) NO_2(p) \, dp + f_2 \int_{p_{\text{cloud}}}^{p_{\text{cloud}}} w_{\text{cloudy}}(p) NO_2(p) \, dp}{\int_{p_{\text{part}}}^{p_{\text{cloud}}} NO_2(p) \, dp},
\]

(4)

where $f_g$ is the geometric cloud fraction and LNO$_2$($p$) is the modeled LNO$_2$ vertical profile. Besides these AMFs, another AMF called AMF$_{LNO_2 Vis}$ is developed for later comparison.

\[
\text{AMF}_{LNO_2 Vis} = \frac{(1 - f_3) \int_{p_{\text{part}}}^{p_{\text{cloud}}} w_{\text{clear}}(p) NO_2(p) \, dp + f_3 \int_{p_{\text{cloud}}}^{p_{\text{cloud}}} w_{\text{cloudy}}(p) NO_2(p) \, dp}{\int_{p_{\text{part}}}^{p_{\text{cloud}}} NO_2(p) \, dp}
\]

(5)

A full definition list of the used AMFs is shown in Appendix A.

2.5 Procedures for deriving LNO$_x$

$V_{LNO_x}$ is re-gridded to 0.05° × 0.05° grids using the constant value method (Kuhlmann et al., 2014). Then, it is analyzed in 1° × 1° grid boxes with a minimum of 50 valid 0.05° × 0.05° grids to minimize the noise. The main procedures of deriving LNO$_x$ are as follows.

CRFs (CRF ≥ 70 %, CRF ≥ 90 %, and CRF = 100 %) and CP ≤ 650 hPa are various criteria of deep convective clouds for OMI pixels (Ziemke et al., 2009; Choi et al., 2014; Pickering et al., 2016). The effect of different CRFs on the retrieved LNO$_x$ is explored in Sect. 3.2. Furthermore, another criterion of cloud fraction (CF) is applied to the WRF-Chem results for the successful simulation of convection. The CF is defined as the maximum cloud fraction calculated by the Xu–Randall method between 350 and 400 hPa (Xu and Randall, 1996; Strode et al., 2017). This atmospheric layer (between 350 and 400 hPa) avoids any biases in the simulation of high clouds. We choose CF ≥ 40 % suggested by Strode
et al. (2017) to determine cloudy or clear for each simulation grid.

Besides cloud properties, a time period and sufficient flashes (or strokes) are required for fresh LNO$_2$ to be detected by OMI. The time window ($t_{\text{window}}$) is the hours prior to the OMI overpass time. $t_{\text{window}}$ is limited to 2.4 h by the mean wind speed at pressure levels 500–100 hPa during OMI overpass time and the square root of the $1^\circ \times 1^\circ$ box over the CONUS (Lapierre et al., 2020). Meanwhile, 2400 flashes per box and 8160 strokes per box per 2.4 h time window are chosen as sufficient for detecting LNO$_2$ (Lapierre et al., 2020). These criteria will result in a low bias in the PE results, as Bucse et al. (2019) found that the PE is larger at small flash rates, which are discarded here. Since our study focuses on developing a new AMF and compares results with other works using similar lightning thresholds (Lapierre et al., 2020; Pickering et al., 2016), we will only discuss results based on the strict criteria in the main text. For comparisons between the criterion of 2400 flashes per box and that of one flash per box, scatter diagrams using different lightning criteria are presented in Appendix B.

To ensure that lightning flashes are simulated successfully by WRF-Chem, the threshold of simulated total lightning flashes (TL) per box is set to 1000, which is fewer than that used by the ENTLN lightning observation, considering the uncertainty of lightning parameterization. In view of other NO$_2$ sources in addition to LNO$_2$, the ratio of modeled lightning NO$_2$ above cloud (LNO$_2$ Vis) to modeled NO$_2$ above cloud (NO$_2$ Vis) is defined to check whether enough LNO$_2$ can be detected by OMI. The ratio $\geq 50\%$ indicates that more than half of the NO$_2$ above the cloud must have an LNO$_2$ source.

Finally, the NO$_2$ lifetime due to oxidation should be taken into account. As estimated by Nault et al. (2017), the lifetime ($\tau$) of NO$_2$ in the near field of convections is $\sim 3$ h. The initial value of NO$_2$ is solved by Eq. (6) as

$$\text{NO}_2(0) = \text{NO}_2(\text{OMI}) \times e^{0.5t/\tau},$$

where NO$_2(0)$ is the moles of NO$_2$ emitted at time $t = 0$, NO$_2$(OMI) is the moles of NO$_2$ measured at the OMI overpass time, and $0.5t$ is the half cross grid time, which is 1.2 h, assuming that lightning occurred at the center of each $1^\circ \times 1^\circ$ box. For each grid box, the mean LNO$_3$ vertical column is obtained by averaging $V_{\text{LNO}_3}$ values from all regridded 0.05$^\circ \times 0.05^\circ$ pixels in the box. This mean value is converted to moles of LNO$_3$ using the dimensions of the grid box. Two methods are applied to estimate the seasonal mean LNO$_2$ per flash, LNO$_3$ per flash, LNO$_2$ per stroke, and LNO$_3$ per stroke:

1. summation method, dividing the sum of LNO$_3$ by the sum of flashes (or strokes) in each $1^\circ \times 1^\circ$ box in MJJA 2014;
2. linear regression method, applying the linear regression to daily mean values of LNO$_3$ and flashes (or strokes).

3 Results

3.1 Criteria determination

To determine the suitable criteria from the conditions defined in Sect. 2.5, six different combinations are defined (Table 1) and applied to the original data with a linear regression method (Table 2).

A daily search of the NO$_2$ product for coincident ENTLN flash (stroke) data results in 99 (102) valid days under the CRF90_Entln condition. Taking the flash-type ENTLN data as an example, the number of valid days decreases from 99 to 81 under the CRF90_Entln_TL100_ratio50 condition, while LNO$_3$ per flash increases from 52.1 $\pm$ 51.1 to 54.5 $\pm$ 48.1 mol per flash. The result is almost the same as that under the CRF90_Entln_TL100 condition, which is without the condition of more than half of the above-cloud NO$_x$ having an LNO$_x$ source. Although this indicates the criterion of TL works well, it is better to include the ratio criterion in case there are some exceptions in the different AMF methods. Since $\text{CF} \geq 40\%$ leads to a sharp loss of valid numbers and production, it is not a suitable criterion. Instead the CRF criteria are used. Finally, coincident ENTLN data, TL $\geq 1000$, and ratio $\geq 50\%$ are chosen as the thresholds to explore the effects of three different CRF conditions (CRF $\geq 70\%$, CRF $\geq 90\%$, and CRF $\geq 100\%$) on LNO$_3$ production (Table 3). Apart from the fewer valid days under higher CRF conditions (CRF $\geq 90\%$ and CRF $\geq 100\%$), LNO$_3$ per flash increases from 35.7 $\pm$ 36.8 to 54.5 $\pm$ 48.1 mol per flash and decreases again to 20.8 $\pm$ 37.4 mol per flash while LNO$_3$ per stroke enhances from 4.1 $\pm$ 3.9 to 7.0 $\pm$ 4.8 mol per stroke and drops again to 2.6 $\pm$ 4.0 mol per stroke (Table 3), as the CRF criterion increases from 70% to 90% and to 100%. When the CRF increases from 90% to 100%, the LNO$_3$ PE decreases because of the higher lightning density with less LNO$_3$ (not shown). The increment of LNO$_3$ PE caused by the CRF increase from 70% to 90% is opposite to the result of Pickering et al. (2016). This is an effect of the consideration of NO$_3$ contamination transported from the boundary layer in our method. Although enhanced NO$_3$ is often observed in regions with CRF $> 70\%$ (Pickering et al., 2016), the following analysis will be based on the criterion of CRF $\geq 90\%$ considering the contamination by low and midlevel NO$_2$ and comparisons with the results of Pickering et al. (2016) and Lapierre et al. (2020).

3.2 Comparison of LNO$_3$ production based on different AMFs

Lapierre et al. (2020) derived LNO$_2$ production based on the BEHR NO$_2$ product. In order for our results to be comparable with those of Pickering et al. (2016) and Lapierre et al. (2020), we choose NO$_2$ instead of NO$_x$ to derive production per flash (production efficiency, PE). In Fig. 3, time series of NO$_2$ Vis, LNO$_2$ Vis, LNO$_2$, and LNO$_2$Clean production per
are significantly overestimated. In other words, NO\textsubscript{2} larger than 200\%), PEs based on NO\textsubscript{2} is highly polluted (\textgreater\ 17\%\ days over CONUS are plotted for MJJA 2014 with the criterion of CRF \textgeq\ 90\% and a flash threshold of 2400 flashes per 2.4 h. LNO\textsubscript{2}PEs are mostly in the range from 20 to 80 mol per flash. LNO\textsubscript{2}Vis PEs are smaller than LNO\textsubscript{2}PEs, which contain LNO\textsubscript{2} below clouds. The simulation of GMI in Pickering et al. (2016) indicated that 25\%–30\% of the LNO\textsubscript{x} column lies below the CP, while the ratio in our WRF-Chem simulation is 56\%±20\%. The effect of cloud properties on LNO\textsubscript{2}PE will be discussed in more detail in Sect. 3.4. Generally, the order of estimated daily PEs is LNO\textsubscript{2}Clean > LNO\textsubscript{2} > NO\textsubscript{2}Vis > LNO\textsubscript{2}Vis. The percent difference in the estimated PE (\Delta PEs) between NO\textsubscript{2}Vis and LNO\textsubscript{2}Vis indicates a certain amount of background NO\textsubscript{x} exists above clouds. Overall, the tendency of that \Delta PEs is consistent with another \Delta PEs between NO\textsubscript{2}Vis and LNO\textsubscript{2}Clean. When the region is highly polluted (\Delta PEs between NO\textsubscript{2}Vis and LNO\textsubscript{2}Vis is larger than 200\%), PEs based on NO\textsubscript{2}Vis and LNO\textsubscript{2}Clean are significantly overestimated. In other words, NO\textsubscript{2}Vis and LNO\textsubscript{2}Clean are more sensitive to background NO\textsubscript{2}. The extent of the overestimation of NO\textsubscript{2}Vis is larger than that of LNO\textsubscript{2}Clean in highly polluted regions, while it is usually opposite in most regions.

Figure 4 shows the linear regression for ENTLN data versus NO\textsubscript{2}Vis, LNO\textsubscript{2}Vis, LNO\textsubscript{2}, and LNO\textsubscript{2}Clean with the same criteria as shown in Fig. 3. LNO\textsubscript{2}Clean PE (the largest slope) is 25.2±22.3 mol NO\textsubscript{2} per flash with a correlation of 0.25 and 2.3±2.1 mol NO\textsubscript{2} per stroke with a correlation of 0.22. As shown in Fig. 3, positive percent differences between NO\textsubscript{2}Vis PE and LNO\textsubscript{2}Clean PE occur much less often than negative differences. As a result, NO\textsubscript{2}Vis PE (17.1±17.2 mol NO\textsubscript{2} per flash and 0.4±1.0 mol NO\textsubscript{2} per stroke) is smaller than LNO\textsubscript{2}Clean PE using the linear regression method.

In order to compare our result with that of Lapierre et al. (2020), we tried to remove the CP \textleq\ 650 hPa, TL \textgeq\ 1000, and ratio \textgeq\ 50\% conditions from criteria. But, our result

### Table 1. Definitions of the abbreviations for the criteria used in this study.

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Full form (source)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRF</td>
<td>Cloud radiance fraction (OMI)</td>
</tr>
<tr>
<td>CP</td>
<td>Cloud optical pressure (OMI)</td>
</tr>
<tr>
<td>CF</td>
<td>Cloud fraction (WRF-Chem)</td>
</tr>
<tr>
<td>TL</td>
<td>Total lightning flashes (WRF-Chem)</td>
</tr>
<tr>
<td>Ratio</td>
<td>Modeled LNO\textsubscript{2}Vis / modeled NO\textsubscript{2}Vis (WRF-Chem)</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_ENTLN</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) (ENTLN)</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_CF40_ENTLN</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) + CF \textgeq\ 40%</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_ENTLN_TL1000</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) + TL \textgeq\ 1000</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_CF40_ENTLN_TL1000</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) + CF \textgeq\ 40% + TL \textgeq\ 1000</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_ENTLN_TL1000_ratio50</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) + TL \textgeq\ 1000 + ratio \textgeq\ 50%</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_CF40_ENTLN_TL1000_ratio50</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 2400 (8160) + CF \textgeq\ 40% + TL \textgeq\ 1000 + ratio \textgeq\ 50%</td>
</tr>
<tr>
<td>CRF\textsubscript{\alpha}_ENTLN(3.4)_TL1_ratio50</td>
<td>CRF \textgeq\ \alpha + ENTLN flashes (strokes) \textgeq\ 1 (3.4) + TL \textgeq\ 1 + ratio \textgeq\ 50%</td>
</tr>
</tbody>
</table>

\textalpha\ has three options: 70\%, 90\%, or 100\%.

### Table 2. LNO\textsubscript{x} production efficiencies for different combinations of criteria defined in Table 1.

<table>
<thead>
<tr>
<th>Condition\textsuperscript{1}</th>
<th>ENTLN data type\textsuperscript{2}</th>
<th>LNO\textsubscript{x} per flash or stroke</th>
<th>R value</th>
<th>Intercept (10\textsuperscript{6} mol)</th>
<th>Days\textsuperscript{3}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRF90_ENTLN</td>
<td>Flash</td>
<td>52.1 ± 51.1</td>
<td>0.20</td>
<td>0.21</td>
<td>99</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN</td>
<td>Flash</td>
<td>84.2 ± 31.5</td>
<td>0.54</td>
<td>-0.04</td>
<td>70</td>
</tr>
<tr>
<td>CRF90_ENTLN_TL1000</td>
<td>Flash</td>
<td>61.9 ± 49.1</td>
<td>0.27</td>
<td>0.33</td>
<td>83</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN_TL1000</td>
<td>Flash</td>
<td>63.4 ± 52.9</td>
<td>0.38</td>
<td>0.26</td>
<td>38</td>
</tr>
<tr>
<td>CRF90_ENTLN_TL1000_ratio50</td>
<td>Flash</td>
<td>54.5 ± 48.1</td>
<td>0.25</td>
<td>0.39</td>
<td>81</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN_TL1000_ratio50</td>
<td>Flash</td>
<td>90.0 ± 65.0</td>
<td>0.46</td>
<td>0.15</td>
<td>32</td>
</tr>
<tr>
<td>CRF90_ENTLN</td>
<td>Stroke</td>
<td>6.7 ± 4.1</td>
<td>0.31</td>
<td>0.23</td>
<td>102</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN</td>
<td>Stroke</td>
<td>10.3 ± 3.6</td>
<td>0.55</td>
<td>0.08</td>
<td>79</td>
</tr>
<tr>
<td>CRF90_ENTLN_TL1000</td>
<td>Stroke</td>
<td>7.5 ± 5.1</td>
<td>0.29</td>
<td>0.38</td>
<td>94</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN_TL1000</td>
<td>Stroke</td>
<td>8.6 ± 6.2</td>
<td>0.39</td>
<td>0.27</td>
<td>46</td>
</tr>
<tr>
<td>CRF90_ENTLN_TL1000_ratio50</td>
<td>Stroke</td>
<td>7.0 ± 4.8</td>
<td>0.29</td>
<td>0.42</td>
<td>93</td>
</tr>
<tr>
<td>CRF90_CF40_ENTLN_TL1000_ratio50</td>
<td>Stroke</td>
<td>8.9 ± 7.0</td>
<td>0.39</td>
<td>0.31</td>
<td>40</td>
</tr>
</tbody>
</table>

\textsuperscript{1} These conditions are defined in Table 1. \textsuperscript{2} The thresholds of ENTLN data are 2400 flashes per box and 8160 strokes per box during the period of 2.4 h before OMI overpass time. \textsuperscript{3} The number of valid days with specific criteria in MJJA 2014.
Table 3. LNO$_x$ production efficiencies for different thresholds of CRF with coincident ENTLN data, TL $\geq$ 1000, and ratio $\geq$ 50%.

<table>
<thead>
<tr>
<th>CRF (%)</th>
<th>ENTLN data type$^1$</th>
<th>LNO$_x$ per flash or LNO$_x$ per stroke</th>
<th>$R$ value</th>
<th>Intercept (10$^5$ mol)</th>
<th>Days$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td>Flash</td>
<td>35.7 $\pm$ 36.8</td>
<td>0.21</td>
<td>4.91</td>
<td>85</td>
</tr>
<tr>
<td>90</td>
<td>Flash</td>
<td>54.5 $\pm$ 48.1</td>
<td>0.25</td>
<td>3.90</td>
<td>81</td>
</tr>
<tr>
<td>100</td>
<td>Flash</td>
<td>20.8 $\pm$ 37.4</td>
<td>0.13</td>
<td>5.67</td>
<td>71</td>
</tr>
<tr>
<td>70</td>
<td>Stroke</td>
<td>4.1 $\pm$ 3.9</td>
<td>0.21</td>
<td>5.16</td>
<td>96</td>
</tr>
<tr>
<td>90</td>
<td>Stroke</td>
<td>7.0 $\pm$ 4.8</td>
<td>0.29</td>
<td>4.16</td>
<td>93</td>
</tr>
<tr>
<td>100</td>
<td>Stroke</td>
<td>2.6 $\pm$ 4.0</td>
<td>0.14</td>
<td>5.41</td>
<td>82</td>
</tr>
</tbody>
</table>

$^1$ The thresholds of ENTLN data are 2400 flashes per box and 8160 strokes per box during the period of 2.4 h before OMI overpass time. $^2$ The number of valid days with specific criteria in MJJA 2014.

Figure 3. (a) Time series of NO$_2$ Vis, LNO$_2$ Vis, LNO$_2$, and LNO$_2$Clean production per day over the CONUS for MJJA 2014 with CRF $\geq$ 90% and a flash threshold of 2400 flashes per 2.4 h. (b) Time series of the percent differences between NO$_2$ Vis and LNO$_2$ Vis and the percent differences between NO$_2$ Vis and LNO$_2$Clean with CRF $\geq$ 90%. The value of the black dot on 23 August (not shown) is 1958%.

3. The surface pressure is now calculated according to Zhou et al. (2009).

The detailed log of changes is available at https://github.com/CohenBerkeleyLab/BEHR-core/blob/master/Documentation/Changelog.txt (last access: 20 March 2020). Note that Lapiere et al. (2020) used the monthly NO$_2$ profile. While the daily profile is used in our study and the interval of our outputs from WRF-Chem is 30 min, which is more frequent than 1 h in the BEHR daily product, the AMF could be affected by different NO$_2$ profiles. In view of these factors, we compare different methods based on our data to minimize these effects.

Meanwhile, LNO$_2$ PE (18.7 $\pm$ 18.1 mol per flash and 2.1 $\pm$ 1.8 mol per stroke) is between LNO$_2$Clean PE and NO$_2$Vis...
PE, which coincides with the daily results in Fig. 3. Furthermore, the LNO$_2$ PE based on the linear regression of daily summed values, the same method used in Pickering et al. (2016), is $114.8 \pm 18.2$ mol per flash (or $17.8 \pm 2.9$ mol per stroke), which is larger than $91$ mol per flash in Pickering et al. (2016), possibly due to the differences in geographic location, lightning data, and chemistry model.

The mean and standard deviation of LNO$_2$ PE under CRF $\geq 90\%$ using the summation method is $46.2 \pm 35.1$ mol per flash and $9.9 \pm 8.1$ mol per stroke, while LNO$_3$ PE is $125.6 \pm 95.9$ mol per flash and $26.7 \pm 21.6$ mol per stroke (Fig. 5). The LNO$_2$ PE and LNO$_3$ PE are both higher in the southeast US (denoted by the red box in Fig. 5, 25–37°N, 75–95°W), consistent with Lapierre et al. (2020) and Bucsel et al. (2019). Compared with Fig. 3, Fig. 6a and b present some large differences between NO$_2$ Vis PE and LNO$_2$ Vis PE, which are consistent with what we expect for polluted regions. Meanwhile, the differences between LNO$_2$ PE and NO$_2$ Vis PE depend on background NO$_2$, the strength of updraft, and the profile. The negative differences are caused by background NO$_2$ carried by the updraft while parts of the below-cloud LNO$_2$ result in LNO$_2$ PE higher than NO$_2$ Vis PE (Fig. 6c). Figure 6d shows that the ratio of LNO$_2$ Vis to LNO$_2$ ranges from 10 % to 80 %. This may be caused by the height of the clouds and the profile of LNO$_2$. If the CP is near 300 hPa, the ratio should be smaller because of the coverage of clouds. While peaks of the LNO$_2$ profile are below the CP, the ratio would also be smaller. Therefore, a better understanding of the LNO$_2$ profile and LNO$_3$ below clouds is required.

### 3.3 Effects of tropospheric background on LNO$_x$ production

With respect to the LNO$_2$ production, the patterns in Fig. 6 indicate the improvement of our approach is different in polluted and clean regions. To simplify the quantification, we select six grids with similar NO$_2$ profiles ($\sim 100$ pptv) above the cloud with CRF $= 100\%$. These grid boxes contain the polluted and clean cities denoted by stars and triangles in Fig. 6a, respectively. Then, the differences between AMFs are dependent on fewer parameters.

$$\text{AMF}_{\text{LNO}_2} = \frac{\int_{p_{\text{cloud}}} \text{w}_{\text{cloudy}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{cloud}}} \text{w}(p) \text{NO}_2(p) \, dp} \quad (7)$$

$$\text{AMF}_{\text{NO}_2\text{Vis}} = \frac{\int_{p_{\text{cloud}}} \text{w}_{\text{cloudy}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{cloud}}} \text{w}(p) \text{NO}_2(p) \, dp} \quad (8)$$

$$\text{AMF}_{\text{LNO}_2\text{Clean}} = \frac{\int_{p_{\text{cloud}}} \text{w}_{\text{cloudy}}(p) \text{LNO}_2(p) \, dp}{\int_{p_{\text{cloud}}} \text{w}(p) \text{LNO}_2(p) \, dp} \quad (9)$$

Figure 7 compares the mean profiles of NO$_2$, background NO$_2$ and background NO$_2$ ratio in polluted and clean grids. Generally, the profiles of the ratio of background NO$_2$ to total NO$_2$ are C shaped because UT LNO$_2$ concentrations are higher than UT background NO$_2$ concentrations. However, the ratio profile in Fig. 7e has one peak between the cloud pressure and tropopause as background NO$_2$ increases and LNO$_2$ decreases. Besides, the percentage of UT background NO$_2$ in polluted regions is steady and higher than that in clean regions.

Table 4 presents the relative changes among three methods in six cities. The difference between AMF$_{\text{LNO}_2}$ (Eq. 7) and AMF$_{\text{NO}_2\text{Vis}}$ (Eq. 8) is the numerator: $\int_{p_{\text{cloud}}} \text{w}_{\text{cloudy}}(p) \text{NO}_2(p) \, dp - \int_{p_{\text{cloud}}} \text{w}(p) \text{NO}_2(p) \, dp$. When the ratio of LNO$_2$ is higher or the region is cleaner, the relative difference is smaller (e.g. 5.0%–12.0%, Fig. 7d–f). The largest relative difference (46.3%) occurs when the ratio of background NO$_2$ is continuously high in the UT (Fig. 7e). As a result, our approach is less sensitive to background NO$_2$ and more suitable for convective cases over polluted locations.

In contrast, production estimated by our method is larger than that based on NO$_2$ Vis due to the LNO$_2$ below the cloud. When the cloud is higher, in particular the peak of the LNO profile is lower than the cloud (Fig. 7b). The relative difference is larger (121.2%) because more LNO$_2$ can not be included in the NO$_2$ Vis, which has been discussed in Sect. 3.2. The relative change between AMF$_{\text{NO}_2\text{Clean}}$ (Eq. 9) and AMF$_{\text{NO}_2\text{Vis}}$ (Eq. 8) depends on $\int_{p_{\text{cloud}}} \text{w}_{\text{cloudy}}(p) \text{LNO}_2(p) \, dp \big/ \int_{p_{\text{cloud}}} \text{w}(p) \text{LNO}_2(p) \, dp$, which is also affected by cloud, not the background NO$_2$. The largest relative change (153.8%) occurs at New Orleans, which has the lowest cloud pressure and consequently the smallest visible column.

### 3.4 Effects of cloud and LNO$_x$ parameterization on LNO$_x$ production

Figure 8a presents the daily distribution of CP and the ratio of NO$_2$ Vis to LNO$_2$ during MJJA 2014 with the criteria defined in Sect. 3.1 under CRF $\geq 90\%$. Since the ratio of NO$_2$ Vis to LNO$_2$ decreases from 0.8 to 0.2 as the cloud pressure decreases from 600 to 300 hPa, NO$_2$ Vis PE is smaller than LNO$_2$ PE in relatively clean areas as shown in Fig. 4. Apart from NO$_2$ Vis, the LNO$_2$ PE is also affected by CP. For LNO$_2$ PEs larger than 30 mol per stroke, the CPs are all smaller than 550 hPa (Fig. 8b). However, smaller LNO$_2$ PEs (< 30 mol per stroke) occur on all levels between 650 and 200 hPa. Because of the limited number of large LNO$_2$ PEs and lightning data, we cannot derive the relationship between LNO$_2$ PE and cloud pressure or other lightning properties at this stage. Because the CP only represents the development of clouds, the vertical structure of flashes can not be derived from the CP values only. As discussed in several previous studies, the flash channel length varies and depends on the environmental conditions (Carey et al., 2016; Mecikalski and Carey, 2017; Fuchs and Rutledge, 2018). Davis et al. (2019) compared two kinds of flash: normal flashes and anomalous flashes. Because updrafts are stronger and flash...
rates are higher in anomalous storms, UT LNOx concentrations are larger in anomalous than normal polarity storms. In general, normal flashes are coupled with an upper-level positive charge region and a midlevel negative charge region, while anomalous flashes are opposite (Williams, 1989). It is not straightforward to estimate the error resulting from the vertical distribution of LNOx. There are mainly two methods of distributing LNOx in models: LNOx profiles (postconvection) in which LNOx has already been redistributed by convective transport and LNOx production profiles (preconvection) made before the redistribution of convective transport (Allen et al., 2012; Luo et al., 2017). However, given the similarity of results compared to other LNOx studies, we believe that our 1° × 1° results based on postconvective LNOx profiles are sufficient for estimating average LNOx production.

The LNO production settings in WRF-Chem varied in different studies. Zhao et al. (2009) set a NOx production rate of 250 mol NO per flash in a regional-scale model, while Bela et al. (2016) chose 330 mol NO per flash used by Barth et al. (2012). Wang et al. (2015) assumed approximately 500 mol NO per flash, which was derived by a cloud-scale chemical transport model and in-cloud aircraft observations (Ott et al., 2010). To illustrate the impact of LNOx parameterization on LNOx estimation, we apply another WRF-Chem NO2 profile setting (2× base flash rate, 500 mol NO per flash, hereinafter referred to as “2 × 500 mol NO per flash”) to a priori profiles and evaluate the changes in AMF\sub{LNO2}, AMF\sub{LNOx}, LNO2 PE, and LNOx PE. For the linear regression method (Fig. 9), LNO2 PE is 29.8 ± 20.5 mol per flash, which is 59.4 % larger than the basic one (18.7 ± 18.1 mol per flash). Meanwhile, LNOx PE (increasing from 54.5 ± 48.1 mol per flash to 88.5 ± 61.1 mol per flash) also depends on the configuration of LNO production in WRF-Chem. The comparison between Figs. 4 and 9 shows that LNO2Clean PE and LNO2 PE are more similar while LNOx PE and NOxPE Vis present the same tendency. It remains unclear as to whether the NO–NO2–O3 cycle or other LNOx reservoirs account for the increment of LNOx PE. This would need detailed source analysis in WRF-Chem and is beyond the scope of this study.

Figure 10 shows the average percentage changes in AMF\sub{LNO2}, AMF\sub{LNOx}, LNO2, and LNOx between retrievals using profiles based on 1 × 200 and 2 × 500 mol NO per flash. These results were obtained by averaging data over MJJA 2014 based on the method described in Sect. 2.5 with the criterion of CRF ≥ 90 %. The effects on LNO2 and LNOx...
retrieval from increasing LNO profile values show mostly the same tendency: smaller AMF\textsuperscript{LNO\textsubscript{2}} and AMF\textsuperscript{LNO\textsubscript{x}} lead to larger LNO\textsubscript{2} and LNO\textsubscript{x}, but the changes are regionally dependent. This is caused by the nonlinear calculation of AMF\textsubscript{LNO\textsubscript{2}} and AMF\textsubscript{LNO\textsubscript{x}}. As the contribution of LNO\textsubscript{2} increases, both the numerator and denominator of Eq. (2) increase. Note that the LNO\textsubscript{2} accounts for a fraction of NO\textsubscript{2} above the clouds. The magnitude of an increasing denominator could be different than that of an increasing numerator, resulting in a different effect on the AMF\textsubscript{LNO\textsubscript{2}} and AMF\textsubscript{LNO\textsubscript{x}}. As mentioned in Zhu et al. (2019), the lightning densities in the southeast US might be overestimated using the 2 × 500 mol NO per flash setting and the same lightning parameterization as ours. Fortunately, the AMFs and esti-
mated LNO$_2$ change little in that region. Because the southeast US has the highest flash density (Fig. 2), the NO$_2$ in the numerator of AMP is dominated by LNO$_2$. Both the SCD and VCD will increase when the model uses higher LNO$_2$. In other words, the sensitivity to the LNO setting decreases and the relative distribution of LNO$_2$ matters.

Figure 11 shows the comparison of the mean LNO and LNO$_2$ profiles in two specific regions where the 2 × 500 mol NO per flash setting leads to lower and higher LNO$_2$ PEs, respectively. The first one (Fig. 11a) is the region (36°–37° N, 89°–90° W) containing the minimal negative percent change in LNO$_2$ (Fig. 10c). The second one (31°–32° N, 97°–98° W), Fig. 11b, has the largest positive percent change in LNO$_2$ (Fig. 10c). Although the relative distributions of mean LNO and LNO$_2$ profiles are similar in both regions, the magnitude differs by a factor of 10. This phenomenon implies that the performance of lightning parameterization in WRF-Chem is regionally dependent, and an unrealistic profile could appear in the UT. Although this sensitivity analysis is false in some regions, it allows the calculation of an upper limit on the NO$_2$ due to LNO and LNO$_2$ profiles. As discussed in Laughner and Cohen (2017), the scattering weights are uniform under cloudy conditions and the sensitivity of NO$_2$ is nearly constant with different pressure levels because of the high albedo. However, the relative distribution of LNO$_2$ within the UT should be taken carefully into consideration. If the LNO$_2$/NO$_2$ above the cloud is large enough (Fig. 11a), the AMF$_{LNO_2}$ is largely determined by the ratio of LNO$_2$Vis to LNO$_2$, which is related to the relative distribution. When the condition of high LNO$_2$/NO$_2$ is not met, both relative distribution and ratio are important (Fig. 11b).

To clarify this, we applied the same sensitivity test of different simulating LNO amounts for all four methods mentioned in Sect. 2.4: LNO$_2$, LNO$_2$Vis, LNO$_2$Clean, and NO$_2$Vis (Fig. 12). Note that the threshold for CRF is set to 100% to simplify Eq. (2) to Eq. (7). The overall differences of LNO$_2$Clean and NO$_2$Vis are smaller than those of LNO$_2$ and LNO$_2$Vis. Comparing the numerator and denom-
Figure 7. Comparisons of mean WRF-Chem NO\textsubscript{x} and background NO\textsubscript{2} profiles in six grids with CRF \( \geq 100\% \) on specific days during MJJA 2014. The (a), (b), and (c) data are selected from polluted regions (stars in Fig. 6a) while the (d), (e), and (f) data are from clean regions (triangles in Fig. 6a). The green dashed lines are the mean ratio profiles of background NO\textsubscript{2} to total NO\textsubscript{x}. The zoomed figures show the profiles from the cloud pressure to the tropopause. The titles present the mean productions based on three different methods mentioned in Sect. 2.4.

Figure 8. Kernel density estimation of the (a) daily ratio of LNO\textsubscript{2}Vis to LNO\textsubscript{2} and (b) daily LNO\textsubscript{2} production efficiency versus the daily cloud pressure measured by OMI with CRF \( \geq 90\% \) for MJJA 2014. The kernel density estimation was generated by kdeplot in the Python package named seaborn.

In the equations, it is clear why the impact of different simulating LNO amounts is smaller in Fig. 12c and d. For LNO\textsubscript{2}Clean and NO\textsubscript{2}Vis, both the SCD and VCD will increase (decrease) when more (less) LNO\textsubscript{2} or NO\textsubscript{2} presents. The difference between Fig. 12a and b is the denominator: the total tropospheric LNO\textsubscript{2} vertical column and visible LNO\textsubscript{2} vertical column, respectively. As a result, the negative values in Fig. 12a are caused by the part of LNO\textsubscript{2} below the cloud. The uncertainty of retrieved LNO\textsubscript{2} and LNO\textsubscript{x} PEs is driven by this error, and we conservatively estimate this to be \( \pm 13\% \) and \( \pm 25\% \), respectively.
4 Uncertainty analysis

The uncertainties of the LNO$_2$ and LNO$_x$ PEs are estimated following Pickering et al. (2016), Allen et al. (2019), Bucsela et al. (2019), Laughner et al. (2019) and Lapierre et al. (2020). We determine the uncertainty due to BEHR tropopause pressure, cloud radiance fraction, cloud pressure, surface pressure, surface reflectivity, profile shape, profile location, $V_{\text{strat}}$, the detection efficiency of lightning, $t_{\text{window}}$, and LNO$_2$ lifetime numerically by perturbing each parameter in turn and re-retrieving the LNO$_2$ and LNO$_x$ with the perturbed values (Table 5).

The GEOS-5 monthly tropopause pressure, which is consistent with the NASA standard product, is applied instead of the variable WRF tropopause height to evaluate the uncertainty (6 % for LNO$_2$ PE and 4 % for LNO$_x$ PE) caused by the BEHR tropopause pressure. The cloud pressure bias is given as a function of cloud pressure and fraction by Acarreta et al. (2004), implying an uncertainty of 32 %, the most likely uncertainty in the production analysis, for LNO$_2$ PE and 34 % for LNO$_x$ PE. The resolution of GLOBE terrain height data is much higher than the OMI pixel, and a fixed scale height is assumed in the BEHR algorithm. As a result, Laughner et al. (2019) compared the average WRF surface pressures to the GLOBE surface pressures and arrived at the largest bias of 1.5 %. Based on the largest bias, we vary the surface pressure (limited to less than 1020 hPa), and the uncertainty can be neglected.

The error in cloud radiance fraction is transformed from cloud fraction using

$$\sigma = 0.05 \cdot \frac{\partial f_r}{\partial f_g} \bigg|_{f_g,\text{pix}},$$

where $f_r$ is the cloud radiance fraction, $f_g$ is the cloud fraction, and $f_{g,\text{pix}}$ is the cloud fraction of a specific pixel. We calculate $\partial f_r/\partial f_g$ under $f_{g,\text{pix}}$ by the relationship between all binned $f_r$ and $f_g$ with the increment of 0.05 for each specific OMI orbit. Considering the relationship, the error in cloud fraction is converted to an error in cloud radiance fraction of 2 % for the LNO$_2$ and LNO$_x$ PEs.

The accuracy of the 500 m MODIS albedo product is usually within 5 % of albedo observations at the validation sites, and those exceptions with low-quality flags have been found to be primarily within 10 % of the field data (Schaaf et al., 2011). Since we use the bidirectional reflectance distribution function (BRDF) data directly, rather than including a radiative transfer model, 14 % Lambertian equivalent reflectivity
Figure 10. Average percent differences in (a) AMF_{LNO_2}, (b) AMF_{LNO_x}, (c) LNO_2, and (d) LNO_x with CRF ≥ 90% over MJJA 2014. Differences between profiles are generated by 2 × 500 and 1 × 200 mol NO per flash.

Table 5. Uncertainties for the estimation of LNO_2 per flash, LNO_x per flash, LNO_2 per stroke, and LNO_x per stroke.

<table>
<thead>
<tr>
<th>Type</th>
<th>Perturbation</th>
<th>LNO_2</th>
<th>LNO_x</th>
<th>LNO_2</th>
<th>LNO_x</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>per flash</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>per stroke</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BEHR tropopause pressure^1</td>
<td>NASA product tropopause</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Cloud radiance fraction^1</td>
<td>±5%</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Cloud pressure^2</td>
<td>Variable</td>
<td>32</td>
<td>34</td>
<td>32</td>
<td>34</td>
</tr>
<tr>
<td>Surface pressure^1</td>
<td>±1.5%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Surface reflectivity^1</td>
<td>±17%</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LNO_2 profile^1</td>
<td>2 × 500 mol NO per flash</td>
<td>13</td>
<td>25</td>
<td>13</td>
<td>25</td>
</tr>
<tr>
<td>Profile location^1</td>
<td>Quasi-Monte Carlo</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Lightning detection efficiency^3</td>
<td>IC: ±16%, CG: ±5%</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>t^window^3</td>
<td>2–4 h</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>LNO_x lifetime^3</td>
<td>2–12 h</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>V^strat^4</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Systematic errors in slant column^4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Tropospheric background^4</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>NO/NO_2^4</td>
<td>20% ± 15%</td>
<td>0</td>
<td>15</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Net</td>
<td>49</td>
<td>56</td>
<td>48</td>
<td>56</td>
<td>56</td>
</tr>
</tbody>
</table>

PE^uncertainty = (Error\_rising\_perturbed\_value + Error\_lowering\_perturbed\_value)/2, where Error\_perturbed\_value = (PE\_perturbed\_value - PE\_original\_value)/PE\_original\_value.

Figure 11. LNO and LNO$_2$ profiles with different LNO settings in (a) the region containing the minimal negative percent change in LNO$_2$ and (b) the region containing the largest positive percent change in LNO$_2$ when the LNO setting is changed from 1 × 200 to 2 × 500 mol NO per flash, averaged over MJJA 2014. The profiles using 1 × 200 (2 × 500) mol NO per flash are shown in blue (red) lines. Solid (dashed) green lines are the mean ratio of LNO$_2$ to NO$_2$ with 1 × 200 (2 × 500) mol NO per flash. (LER) error and 10 % uncertainty are combined to get a perturbation of 17 % (Laughner et al., 2019). The uncertainty due to surface reflectivity can be neglected with the 17 % perturbation.

As discussed at the end of Sect. 3.4, another setting of LNO$_2$ (2 × 500 mol NO per flash) is applied to determine the uncertainty of the lightning parameterization and the vertical distribution of LNO in WRF-Chem. Differences between the two profiles lead to an uncertainty of 13 % and 25 % in the resulting PEs of LNO$_2$ and LNO$_x$. Another sensitivity test allows each pixel to shift by −0.2, 0, or +0.2 ° in the directions of longitude and latitude, taking advantage of the high-resolution profile location in WRF-Chem. The resulting uncertainty of LNO$_x$ PE is 1 %, including the error of transport and chemistry by shifting pixels.

Compared to the NASA standard product v2, Krotkov et al. (2017) demonstrated that the noise in V$_{strat}$ is 1 × 10$^{14}$ cm$^{-2}$. Errors in polluted regions can be slightly larger than this value, while errors in the cleanest areas are typically significantly smaller (Bucsela et al., 2013). We estimated the uncertainty of the V$_{strat}$ component and the slant column errors to be 10 % and 5 %, respectively, following Allen et al. (2019).

Based on the standard deviation of the detection efficiency estimation over the CONUS relative to LIS, ENTLN detection efficiency uncertainties are ± 16 % for total and IC flashes and strokes. Due to the high detection efficiency of CG over the CONUS, the uncertainty is estimated to be ± 5 % (Lapierre et al., 2020). It is found that the resulting uncertainty of detection efficiency is 15 % in the production analysis. We have used the $t_{\text{window}}$ of 2.4 h for counting ENTLN flashes and strokes to analyze LNO$_2$ and LNO$_x$ production. Because $t_{\text{window}}$ derived from the ERA5 reanalysis can not represent the variable wind speeds, a sensitivity test is performed which yields an uncertainty of 10 % for production per flash and 8 % for production per stroke using $t_{\text{window}}$ of 2 and 4 h. Meanwhile, the lifetime of UT NO$_x$ ranges from 2 to 12 h depending on the convective location, the methyl peroxy nitrate and alkyl, and multifunctional nitrates (Nault et al., 2017). The lifetime (τ) of NO$_2$ in Eq. (6) is replaced by 2 and 12 h to determine the uncertainty as 24 % due to lifetime. This is comparable with the uncertainty (25 %) caused by lightning parameterization for the LNO$_x$ type.

Recent studies revealed that the modeled NO/NO$_2$ ratio departs from the data in the SEAC4RS aircraft campaign (Travis et al., 2016; Silvern et al., 2018). Silvern et al. (2018) attributed this to the positive interference on the NO$_2$ measurement or errors in the cold-temperature NO−NO$_2$−O$_3$ photochemical reaction rate. We assign a 20 % bias with ±15 % uncertainty to this error considering the possible positive NO$_2$ measurement interferences (Allen et al., 2019; Bucsela et al., 2019) and estimate the uncertainty to be 15 % for LNO$_x$ PE.

In addition, the estimation of LNO$_x$ PE also depends on the tropospheric background NO$_2$. In our method, the main factors affecting this factor are the emissions inventory and
the amount of transported NO$_2$. For the emissions inventory, the sources of uncertainty are assumptions, methods, input data, and calculation errors. As a result, the uncertainties for different species or pollutants related to NO$_2$ are different and the U.S. EPA also does not publish the quantified uncertainty measures because the parties that submit emissions estimates to the U.S. EPA are not asked to include quantitative uncertainty measurements or estimates (EPA, 2015). For the simulated convective transport, Li et al. (2018) compared the cloud-resolving simulations with these based on convective parameterization and pointed out that the convective transport was weaker in the parameterization. But, we believe that the ratio condition (LNO$_2$Vis/NO$_2$Vis $\geq$ 50 %) should reduce these two kinds of uncertainty, and we assume an uncertainty of 10 %, which is less than the 20 % assigned in Allen et al. (2019) and Bucsela et al. (2019).

The overall uncertainty is estimated as the square root of the sum of the squares of all individual uncertainties in Table 5. The net uncertainty is 48 % and 56 % for the LNO$_2$ type and LNO$_x$ type, respectively. The mean LNO$_2$ per flash, LNO$_x$ per flash, LNO$_2$ per stroke, and LNO$_x$ per stroke based on the linear regression and summation method are 32 mol per flash, 90 mol per flash, 6 mol per stroke, and 17 mol per stroke. Applying the corresponding uncertainty to these mean values, we arrive at $32 \pm 15$ mol LNO$_2$ per flash, $90 \pm 50$ mol LNO$_x$ per flash, $6 \pm 3$ mol LNO$_2$ per stroke, and $17 \pm 10$ mol LNO$_x$ per stroke. This is in the range of current literature estimates ranging from 33 to 500 mol LNO$_x$ per flash (Schumann and Huntrieser, 2007; Beirle et al., 2010; Bucsela et al., 2010). Bucsela et al. (2010) estimated LNO$_x$ PE of 100–250 mol per flash, which is higher than but overlaps with our estimate. Pickering et al. (2016) estimated LNO$_x$ PE to be $80 \pm 45$ mol per flash for the Gulf of Mexico. This is 50 % smaller than our flash-based results over the CONUS, if we use the same linear regression method which is based on the daily summed values instead of daily mean values. Note that the criteria defined in Sect. 3.1 lead to many missing data over the Gulf of Mexico; thus it is actually a comparison between different regions. For the stroke-based results, Lapierre et al. (2020) found lower LNO$_2$ PE of $1.6 \pm 0.1$ mol per stroke, and the difference is caused by the different version of BEHR algorithm and several settings as mentioned in Sect. 3.2. Bucsela et al. (2019) inferred an average value of $200 \pm 110$ mol (122 % larger than our results) of
LNO\textsubscript{x} produced per flash over the North America, this is related to the different algorithm, lightning data, and lightning thresholds.

5 Conclusions

In this study, a new algorithm for retrieving LNO\textsubscript{2} (LNO\textsubscript{x}) from OMI, including LNO\textsubscript{2} (LNO\textsubscript{x}) below cloud, has been developed for application over active convection. It works in both clean and polluted regions because of the consideration of tropospheric background pollution in the definition of AMFs. It uses specific criteria combined with several other conditions (sufficient CRF, coincident ENTLN data, TL ≥ 1000, and ratio ≥ 50 %) to ensure that the electrically active regions are detected by OMI and simulated by WRF-Chem successfully. We conducted an analysis on 1° × 1° daily boxes in MJJA 2014 and obtained the seasonal mean LNO\textsubscript{2} and LNO\textsubscript{x} production efficiencies over the CONUS. Considering all the uncertainties (Table 5) and applying the summation and regression methods, the final mean production efficiencies are estimated to be 32 ± 15 mol LNO\textsubscript{2} per flash, 90 ± 50 mol LNO\textsubscript{x} per flash, 6 ± 3 mol LNO\textsubscript{2} per stroke, and 17 ± 10 mol LNO\textsubscript{x} per stroke.

Compared with Lapierre et al. (2020), we find that the LNO\textsubscript{2} production could be larger when the below-cloud LNO\textsubscript{2} is taken into account, especially for the high clouds. Meanwhile, if the method of Pickering et al. (2016) is applied without the background NO\textsubscript{2} correction, the derived LNO\textsubscript{x} production efficiency is similar to ours in clean regions or regions with a high LNO\textsubscript{2} concentration above the cloud, but it could be overestimated by more than 18 % in polluted regions. Finally, implementing profiles generated with different model settings of lightning (1 × 200 and 2 × 500 mol NO\textsubscript{2} per flash), we find that the larger LNO production setting leads to 62 % larger retrieval of LNO\textsubscript{x} on average despite some regionally dependent effects caused by the nonlinear calculation of AMF. Both the ratio of the tropospheric LNO\textsubscript{2} above the cloud to the total tropospheric LNO\textsubscript{2} and the ratio of LNO\textsubscript{2} to NO\textsubscript{2} cause different comprehensive effects due to the nonlinear calculation of AMFNNO\textsubscript{2} and AMFNNO\textsubscript{x}.

Since other regions, like China and India, have much more NO\textsubscript{2} pollution than the CONUS, it is necessary to consider the background NO\textsubscript{2} in detail. These analyses will be complemented by the recently launched satellite instrument (TROPOspheric Monitoring Instrument, TROPOMI) (Veefkind et al., 2012; Boersma et al., 2018; Griffin et al., 2019) and Lightning Mapping Imager (LMI) on the next-generation Chinese geostationary meteorological satellite Fengyun-4 (Min et al., 2017; Yang et al., 2017; Zhang et al., 2019). Future work investigating the flash channel length and more detailed lightning parameterization in WRF-Chem would greatly benefit LNO\textsubscript{x} estimation. Applying current methods in future studies may enhance the accuracy of LNO\textsubscript{x} production at both local and global scales.
Appendix A: AMF definitions used in this study

\[
\text{AMF}_{\text{LNO}_2} = \frac{(1-f_i)\int_{p_{\text{surf}}} \omega_{\text{clear}}(p) \text{NO}_2(p) \, dp + \int_{p_{\text{cloud}}} \omega_{\text{cloud}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{surf}}} \text{LNO}_2(p) \, dp}
\]  
(A1)

\[
\text{AMF}_{\text{LNO}_2 \text{Clean}} = \frac{(1-f_i)\int_{p_{\text{surf}}} \omega_{\text{clear}}(p) \text{NO}_2(p) \, dp + \int_{p_{\text{cloud}}} \omega_{\text{cloud}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{surf}}} \text{LNO}_2(p) \, dp}
\]  
(A3)

\[
\text{AMF}_{\text{LNO}_2 \text{Vis}} = \frac{(1-f_i)\int_{p_{\text{surf}}} \omega_{\text{clear}}(p) \text{NO}_2(p) \, dp + \int_{p_{\text{cloud}}} \omega_{\text{cloud}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{surf}}} \text{NO}_2(p) \, dp + \int_{p_{\text{cloud}}} \text{NO}_2(p) \, dp}
\]  
(A4)

\[
\text{AMF}_{\text{NO}_x \text{Vis}} = \frac{(1-f_i)\int_{p_{\text{surf}}} \omega_{\text{clear}}(p) \text{NO}_x(p) \, dp + \int_{p_{\text{cloud}}} \omega_{\text{cloud}}(p) \text{NO}_x(p) \, dp}{\int_{p_{\text{surf}}} \text{NO}_x(p) \, dp + \int_{p_{\text{cloud}}} \text{NO}_x(p) \, dp}
\]  
(A5)

\[
\text{AMF}_{\text{NO}_x \text{Vis}} = \frac{(1-f_i)\int_{p_{\text{surf}}} \omega_{\text{clear}}(p) \text{NO}_2(p) \, dp + \int_{p_{\text{cloud}}} \omega_{\text{cloud}}(p) \text{NO}_2(p) \, dp}{\int_{p_{\text{surf}}} \text{LNO}_2(p) \, dp + \int_{p_{\text{cloud}}} \text{LNO}_2(p) \, dp}
\]  
(A6)

Here \( f_i \) is the cloud radiance fraction, \( p_{\text{surf}} \) is the surface pressure, \( p_{\text{tp}} \) is the tropopause pressure, \( p_{\text{cloud}} \) is the cloud optical pressure (CP), \( \omega_{\text{clear}} \) and \( \omega_{\text{cloud}} \) are respectively the pressure-dependent scattering weights from the TOMRAD lookup table (Bucsela et al., 2013) for clear and cloudy parts, and \( \text{NO}_2(p) \) is the modeled \( \text{NO}_2 \) vertical profile. \( \text{LNO}_2 \) and \( \text{NO}_x(p) \) are respectively the \( \text{LNO}_2 \) and \( \text{NO}_x \) vertical profiles calculated by the difference of vertical profiles between WRF-Chem simulations with and without lightning.

Appendix B: LNO\textsubscript{\text{x}} production based on lower lightning thresholds

While we used 2400 flashes per box and 8160 strokes per box per 2.4 h time window for detecting LNO\textsubscript{\text{x}}, here we show results obtained when using one flash per box and 3.4 strokes per box in the same time window. We note that the WRF total lightning threshold is also reduced to one flash per box, but we keep the ratio condition unchanged. Briefly, the condition is CRF90\_ENTLN1(3.4)_TL1\_ratio50 as shown in Table 1.

Similarly, the order of estimated daily PEs is LNO\textsubscript{\text{Clean}} > LNO\textsubscript{2} > NO\textsubscript{2}Vis > LNO\textsubscript{2}Vis (Fig. B1). Compared with Fig. 4, the LNO\textsubscript{2} per flash and LNO\textsubscript{\text{x}} per flash are larger while PEs based on stroke data are smaller. Considering the additional boxes of fewer lightning counts, differences in the daily mean flashes and NO\textsubscript{\text{x}} result in different PEs, and the relationship presents more like the power function as mentioned in Bucsela et al. (2019).

Instead of using the nonlinear regression of power function

\[
y = ax^b,
\]  
(B1)

where \( x \) is flashes or strokes and \( y \) is NO\textsubscript{2} or NO\textsubscript{\text{x}}, we take the logarithm of both sides and apply the linear regression to data:

\[
\log_{10} y = \log_{10} a + \beta \log_{10} x.
\]  
(B2)

As expected, the linear regression based on logarithmized data performs better in this situation and yields \( \alpha = 38 \text{ kmol} \) and \( \beta = 0.3 \) for LNO\textsubscript{2} per flash (Fig. B2). Since we use the unbinned data (flashes not divided into many groups), we compare our results with Bucsela et al. (2019) based on the same kind of data (\( \alpha = 10.3 \text{ kmol} \) and \( \beta = 0.42 \)). The large difference of \( \alpha \) is related to the method of estimating LNO\textsubscript{\text{x}}, different lightning data (WWLLN and ENTLN), and different regions (northern midlatitudes and CONUS). Note that the resolution (13 km x 24 km) of OMI could weaken the signal of LNO\textsubscript{\text{x}}. We believe the phenomenon of higher production efficiency as flash rate decreases (Fig. B3) could be explored in much more detail with higher-resolution data like the TROPOMI data.

Table A1. Simple forms of abbreviations for AMFs.

<table>
<thead>
<tr>
<th>Abbreviations</th>
<th>Numerator\textsuperscript{1}</th>
<th>Denominator\textsuperscript{2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMFLNO\textsubscript{2}</td>
<td>S\textsubscript{NO}_2</td>
<td>\text{V}_\text{LNO}_2</td>
</tr>
<tr>
<td>AMFLNO\textsubscript{2}Vis</td>
<td>S\textsubscript{NO}_2</td>
<td>\text{V}_\text{LNO}_2 \text{Vis}</td>
</tr>
<tr>
<td>AMFLNO\textsubscript{2}Clean</td>
<td>S\textsubscript{LNO}_2</td>
<td>\text{V}_\text{LNO}_2</td>
</tr>
<tr>
<td>AMFNO\textsubscript{2}Vis</td>
<td>S\textsubscript{NO}_2</td>
<td>\text{V}_\text{NO}_2 \text{Vis}</td>
</tr>
<tr>
<td>AMFLNO\textsubscript{\text{x}}</td>
<td>S\textsubscript{NO}_2</td>
<td>\text{V}<em>\text{LNO}</em>\text{x}</td>
</tr>
<tr>
<td>AMFNO\textsubscript{\text{x}}Vis</td>
<td>S\textsubscript{NO}_2</td>
<td>\text{V}<em>\text{NO}</em>\text{Vis}</td>
</tr>
</tbody>
</table>

\textsuperscript{1} The part of simulated VCD seen by OMI. \textsuperscript{2} The simulated VCD.
Figure B1. Linear regressions with CRF $\geq 90\%$ and a flash threshold of one flash per box or 3.4 strokes per box per 2.4 h. (a) Daily NO$_2$ Vis, LNO$_2$ Vis, LNO$_2$, and LNO$_2$ Clean versus ENTLN total flash data. (b) Same as (a) but for strokes. (c) Daily LNO$_x$ Vis and LNO$_x$ versus total flashes. (d) Same as (c) but for strokes.
Figure B2. Same as Fig. B1 but using log-log axes.

Figure B3. (a) Daily \( \text{LNO}_x \) production efficiencies versus ENTLN total flash data, with CRF \( \geq \) 90% and a flash threshold of one flash per box. (b) Same as (a) but for strokes.
Author contributions. YY directed the research and RvdA, XZ, and YY designed the research with feedback from the other co-authors; RvdA and XZ developed the algorithm; JLL provided guidance and supporting data on the ENTLN data; XZ performed simulations and analysis with the help of YY, RvdA, QC, XK, SY, JC, CH, and RS; YY, RvdA, JLL, and XZ interpreted the data and discussed the results. XZ drafted the manuscript with comments from the co-authors; JLL, RvdA, and YY edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We acknowledge use of the computational resources provided by the National Supercomputer Centre in Guangzhou (NSCC-GZ). We thank the University of California Berkeley Satellite Group for the basic BEHR algorithm. We also thank Earth Networks Company for providing the Earth Networks Total Lightning Network (ENTLN) dataset. We appreciate the discussions with Joshua L. Laughner for BEHR codes and Mary Barth for the WRF-Chem lightning NO\textsubscript{x} module. The authors would also like to thank all anonymous reviewers as well as Kenneth E. Pickering, Eric J. Bucsela, and Dale J. Allen for detailed comments which greatly improved this paper. Finally, we thank all contributors of Python packages used in this paper (Met Office, 2010–2015; Hoyer and Hamman, 2017; Hunter, 2007; Zhuang et al., 2019; McKinney, 2011; Inc., 2015; Seabold and Perktold, 2010; van der Walt et al., 2011; Waskom et al., 2017).

Financial support. This research has been supported by the National Natural Science Foundation of China (grant nos. 91644224 and 41705118).

Review statement. This paper was edited by Steffen Beirle and reviewed by three anonymous referees.

References


X. Zhang et al.: Estimates of lightning NO$_2$ production


X. Zhang et al.: Estimates of lightning NO$_x$ production


