Estimates of the Lightning NOx profile in the vicinity of the North Alabama Lightning Mapping Array

William J. Koshak[1], Harold S. Peterson[2], Eugene W. McCaul[3], Arastoo Biazar[4]

[1]Earth Science Office, VP61, NASA-Marshall Space Flight Center, National Space Science & Technology Center, 320 Sparkman Drive, Huntsville, AL 35805, william.koshak@nasa.gov.

[2]NASA Post Doctoral Fellow, National Space Science & Technology Center, 320 Sparkman Drive, Huntsville, AL 35805, harold.peterson@nasa.gov.

[3]Universities Space Research Association, National Space Science & Technology Center, 320 Sparkman Drive, Huntsville, AL 35805, bill.mccaul@msfc.nasa.gov.

[4]University of Alabama in Huntsville, National Space Science & Technology Center, 320 Sparkman Drive, Huntsville, AL 35805, arastoo.biazar@nsstc.uah.edu.

ABSTRACT

The NASA Marshall Space Flight Center Lightning Nitrogen Oxides Model (LNOM) is applied to August 2006 North Alabama Lightning Mapping Array (NALMA) data to estimate the (unmixed and otherwise environmentally unmodified) vertical source profile of lightning nitrogen oxides, NOx = NO + NO_2. Data from the National Lightning Detection Network™ (NLDN) is also employed. This is part of a larger effort aimed at building a more realistic lightning NOx emissions inventory for use by the U.S. Environmental Protection Agency (EPA) Community Multiscale Air Quality (CMAQ) modeling system. Overall, special attention is given to several important lightning variables including: the frequency and geographical distribution of lightning in the vicinity of the NALMA network, lightning type (ground or cloud flash), lightning channel length, channel altitude, channel peak current, and the number of strokes per flash. Laboratory spark chamber results from the literature are used to convert 1-meter channel segments (that are located at a particular known altitude; i.e., air density) to NOx concentration. The resulting lightning NOx source profiles are discussed.

1. INTRODUCTION

The U.S. Environmental Protection Agency (EPA) Community Multiscale Air Quality (CMAQ) modeling system is used to enhance the scientific understanding and modeling capability of chemical and physical atmospheric interactions. Federal, state, local agencies and other stakeholders use CMAQ to evaluate the impact of air quality management practices for multiple pollutants at a variety of spatio-temporal scales. Consequently, CMAQ helps guide the development of air quality regulations and standards.

A modeling study conducted with funding from the NASA Applied Science Program that compared CMAQ model predictions of ozone against ozonesonde observations, found model bias in excess of 30%. Long-range transport of pollution, emissions from aircraft and lightning, and other sources are believed to account for the CMAQ errors.

To help resolve the issue, a better understanding of the lightning NOx emission inventory is desired. Presently, emissions from lightning are either omitted or are poorly represented in CMAQ. Model predictions suffer as a result, especially in the middle and upper troposphere.

Recently, Kaynak et al., (2008) estimated a 2 ppb impact on surface ozone concentration from lightning NO_x emissions. But, several simplifying assumptions were made regarding lightning modeling: (a) the ratio of the number of cloud flashes to ground flashes was held fixed at a value of 3, (b) the NO produced by each cloud flash was assumed to be constant, (c) the NO produced by each ground flash was assumed to be constant, and (d) cloud and ground flashes were assumed to produce the same amount of NO.

Unfortunately, owing to the highly variable nature of lightning, none of these assumptions hold as a general rule.

Therefore, a more sophisticated model, called the Lightning Nitrogen Oxides Model (LNOM), was introduced in Koshak et al. (2009) for the specific purpose of accounting for the variable nature of lightning. The LNOM is an IDL program that combines a detailed theory, state-of-the-art lightning measurements, and spark chamber laboratory results to substantially improve estimates of the vertical source profile (VSP) of lightning NO_x.

The VSP is the unmixed and otherwise environmentally unmodified lightning NOx profile. That is, the focus of LNOM is on the lightning NOx production in space and time, not its subsequent chemical conversion, transport (convective, advective) or removal (wet scavenging). The Pickering et al. (1998) profiles could be used to appropriately distribute the LNOM NOx source profiles.

In this study, we apply the LNOM to make the first detailed estimate of the VSP within a cylindrical volume centered over the North Alabama Lightning Mapping Array (NALMA). The NALMA provides a highly accurate spatio-temporal mapping of both ground and cloud flashes [see retrieval error analyses in Koshak et al., (2004)]. In addition to the physical constraints afforded by the NALMA data, the LNOM uses National Lightning Detection Network™ (NLDN) data to analyze ground flashes, and laboratory results derived from Wang et al. (1998).

The lightning NOx source profiles obtained in this study will serve as a baseline and a check of future NOx profiles obtained for regions lacking lightning detection coverage. In particular, results of this study will be used to check future NOx profile estimates for CMAQ grid cell regions not covered by a VHF total lightning mapping array network.

2. PRE-PROCESSING STEPS

This section describes how the VHF sources obtained from the NALMA dataset are prepared for individual flash NOx analyses.

2.1 Clustering VHF Sources to a Flash

The raw NALMA data provides VHF sources, but does not associate the sources to specific flashes. To complete the association, we utilize a standard clustering algorithm described in McCaul et al., (2009).

2.2 VHF Source Filtering

The cluster algorithm mentioned above considers, and outputs, only high-quality VHF source solutions (i.e., those having a reduced chi-squared of 2 or less). This represents the first filter applied to the VHF sources. The LNOM begins by reading in these high quality sources for a given flash.

Secondly, in order to minimize VHF source location and time-of-occurrence errors, and to approximate the typical (36 km x 36 km) CMAQ grid cell area, only flashes within about a 20 km range of the NALMA network centroid are analyzed. [Specifically, the CMAQ grid cell area is equated to the area of a circle, and gives radius, r_o = 20.31 km.]

Third, only flashes with at least 20 VHF sources are considered. Though this can potentially lead to underestimation in lightning NOx, events with few VHF sources are suspect. For example, the event might not be a lightning-related discharge at all, but instead a VHF noise source. Or, the event might be a weak atmospheric electrical discharge (not a lightning flash) that produces little or no NOx. In addition, with so few sources in an event, it can be difficult to define a meaningful channel and compute a meaningful channel length.

Fourth, before a channel length estimate is made, a minimum VHF source power threshold of 2 dbW is applied. Any VHF source with a power below this threshold is removed. Note that the physical mechanisms behind VHF and optical emissions from lightning differ; VHF sources only provide a rough estimate of the optical channel geometry. In some cases there are VHF emissions not associated with optical emissions, and vice versa. The primary benefit of removing the lower power sources is that it prevents construction of erroneous channel sections (see section 3) and therefore prevents overestimation of channel length. The power filter also improves the speed of computation since there are fewer sources to analyze.

Finally, since anomalous VHF sources occur in the NALMA data at the surface and at 20 km or higher, these sources are removed. Hence, the final lightning NOx profile is for flashes, *or portions of flashes*, that occur within a right circular cylinder (the “analysis cylinder”) of radius, r_o = 20.31 km, and height, h_o = 20 km.

2.3 Determining Flash-Type

The average time and location of the three lowest VHF sources in a flash is computed. If an NLDN event is within 100 ms and within 10 km of these respective average values, and if the lowest VHF source is below 7 km, the flash is deemed a ground flash. Otherwise, the flash is deemed a cloud flash. If the flash is deemed a ground flash, 5 co-located virtual VHF sources are placed at the surface directly below the lowest altitude VHF source that meets the minimum power threshold discussed above. These 5 assure that a source location will exist at the ground after spatial averaging is performed (see below).

2.4 Transforming, Averaging, and Sorting

The chi-square, range, number, power, and altitude filtering of VHF sources described above results in the final set of VHF sources optimal for further analysis by the LNOM. Three additional steps are performed.

First, we transform coordinates. The NALMA data describes the VHF source location in geodetic coordinates (latitude, longitude, height) using the WGS84 ellipsoid. This coordinate system is cumbersome for LNOM analyses, so we convert source location into the Euclidean (X,Y,Z) Earth Centered Earth Fixed (ECEF) system.

Second, spatial averaging of the sources is performed. We divide the (X,Y,Z) space into a 3-D cartesian grid, the dimensions of each grid cell is 1 km x 1 km x 1km. If there are 5 or more VHF sources in a grid cell, the average location of the sources in the cell is found, otherwise the cell and the sources it contains are ignored. This converts the VHF sources to a smaller set of average source locations. Spatial averaging prevents channel length overestimation by removing fine-structure (i.e., the high frequency interconnections between sources), and it also reduces computation time.

Third, the VHF sources are sorted by (geodetic) altitude. To do this, the spatially averaged points are first transformed back into geodetic coordinates using the non-iterative method in Heikkinen (1982), and then the altitudes are inter-compared. Sorting by altitude allows the channel length algorithm to begin constructing each channel from the highest spatially averaged source location, and therefore maintains a consistent approach for handling each flash. Starting channel construction at a different source can result in a different computed channel length (see section 3).

3. CHANNEL LENGTH

As part of this study, a new algorithm for estimating channel length from VHF sources has been developed for the LNOM; it replaces a test algorithm that was used for preliminary channel length studies.

The new channel length subroutine accepts as input the spatially averaged VHF source points for a particular flash. It then computes all the distances between each of these points. Each distance, D_{ij}, between the i^{th} and j^{th} point is stored in a matrix, **D**. Hence, **D** is both symmetric and traceless. Next, the algorithm assigns the diagonal elements to a large number (i.e., 9×10^{20} m); in addition, any element of **D** that is less than 0.1 meters is assigned this large number.

Suppose *n* points result from the spatial averaging process. By convention, the highest altitude point (called the “starting point”) is deemed to be “on the channel”, and the remaining *n*-1 points are deemed “free” (i.e., “off the channel”). The minimum distance between the starting point and all the free points is found. A line is then drawn from the starting point to the closest free point; this is the first iteration, and the line drawn is called a channel “section”.

Now there are two points on the channel, and *n*-2 free points. On the next iteration, the minimum distance between the first channel point and the free points is found. The minimum distance between the 2nd channel point and the free points is also found. The minimum of the two minimums defines what line is drawn. *That is, the algorithm always draws a line between the free point and the channel point that results in the smallest distance.* The process continues until there are no more free points (i.e., until all free

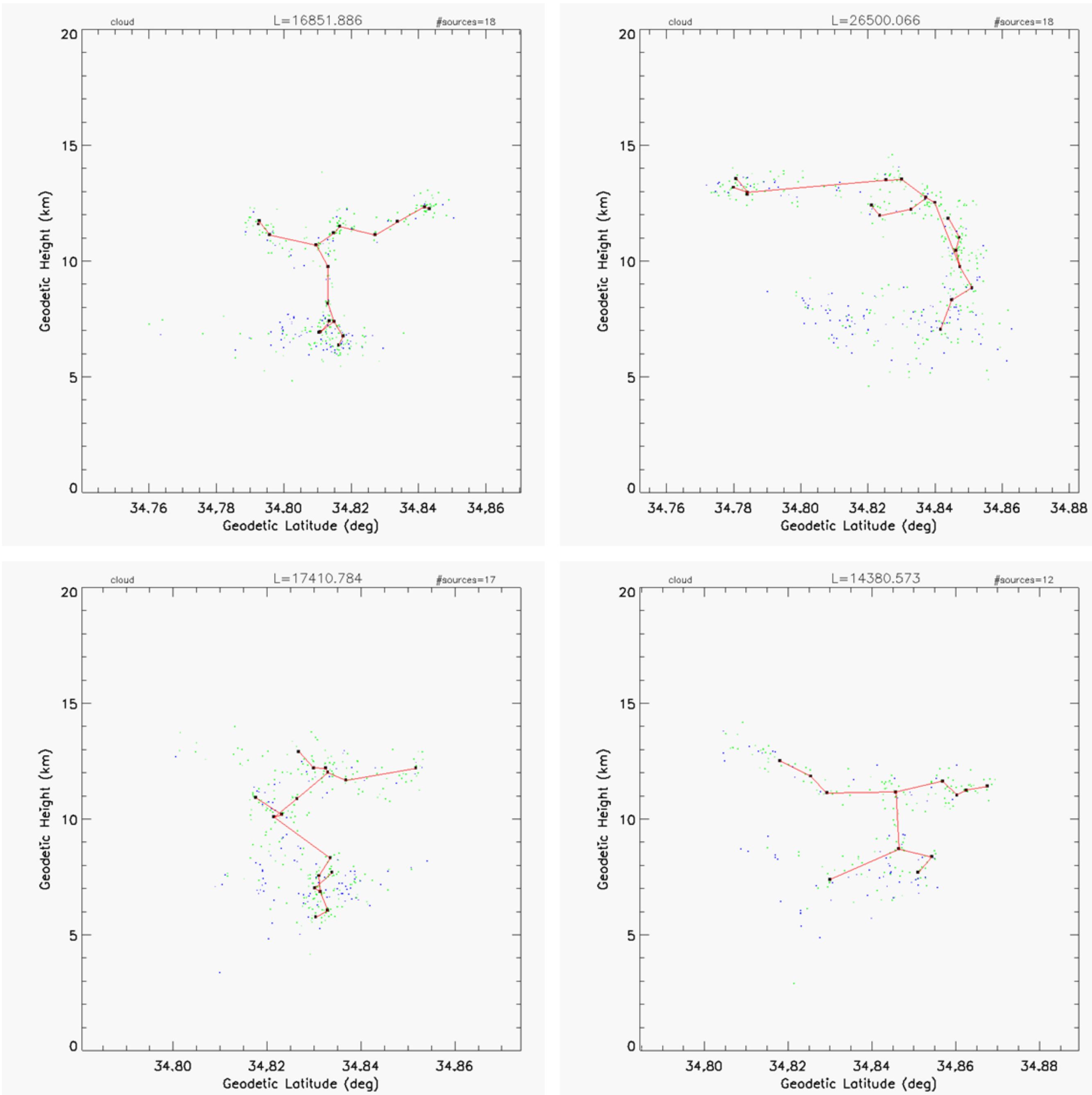

Figure 1. Sample channel length, L, computed for 4 cloud flashes. Number of sources shown in upper right is the number of spatially averaged points (black dots). Green dots are the VHF sources that meet the 2 dBW power threshold, and the blue dots are those that do not.

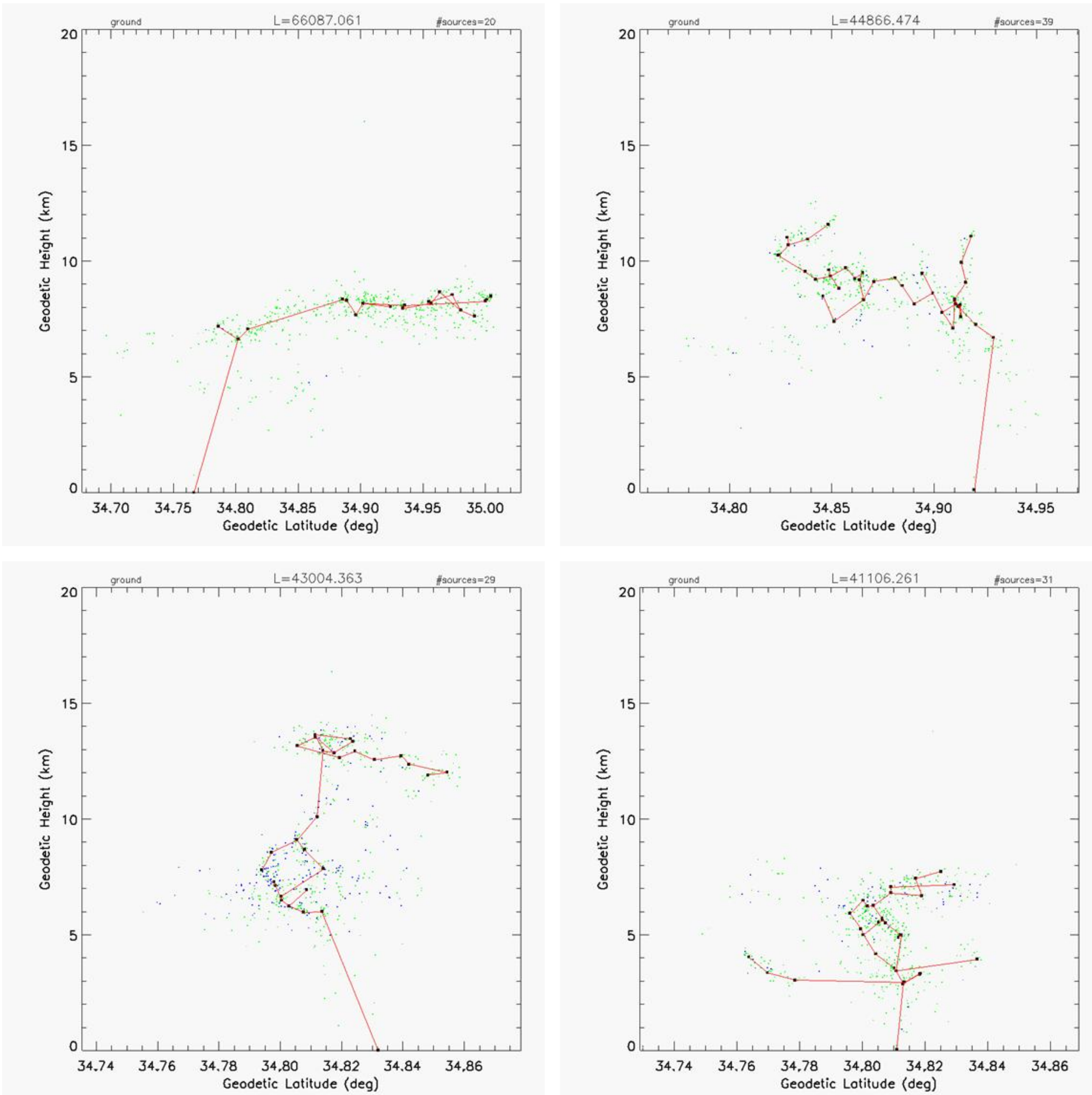

Figure 2. Sample channel length computations for 4 ground flashes.

points are connected up to the channel). The length of a channel section is stored at each iteration, and the sum of all the channel sections defines the channel length estimate.

Figures 1 and 2 show samples of the channel length analyses for both cloud and ground flashes, respectively. Note that the algorithm avoids drawing channel sections into regions containing low power sources (blue dots) or into regions having a low number density of sources that meet the power threshold (green dots).

4. SEGMENT ALTITUDE DISTRIBUTION

In the channel length algorithm described above, the (X, Y, Z) endpoints of each channel section is stored. The LNOM takes the two endpoints and creates a unit vector that points from one endpoint to the other. It uses the unit vector to move in 1-meter steps along the section. In this way, and with the section length rounded to the nearest 1-meter, the section is broken up into 1-meter "segments". Repeating this process for the rest of the sections in the channel, the (X, Y, Z) location of the midpoint of each segment in the channel is obtained.

The (X, Y, Z) coordinates of each segment are then transformed back into geodetic coordinates by again using the non-iterative Heikkinen (1982) method. This provides the geodetic (latitude, longitude, height) location of each segment in the channel. Next, the 1-m segments that are contained within the analysis cylinder (section 2.2) are identified; so, for example, a single flash that crosses the walls of the analysis cylinder will be appropriately split.

The heights of the segments within the analysis cylinder are then tallied for each 500 meter layer, from the surface to an altitude of 20 km; this is the segment altitude distribution within the analysis cylinder contributed by a single flash. The process is repeated for all flashes to obtain the total segment altitude distribution for a given time period.

The segment altitude distribution within the analysis cylinder for the entire month of August, 2006 is provided in Figure 3. A total of 5994 flashes contributed segments to the analysis cylinder. This is in good agreement with the Lightning Imaging Sensor (LIS) climatology for the month of August (i.e., ~45 flashes/km^2/yr, or roughly 5000 flashes for the analysis cylinder). Based on the NLDN flash-typing approach applied here, 5128 of the 5994 flashes were cloud flashes and 866 were ground flashes. This implies a cloud flash to ground flash ratio of 5.9. The daily profiles, not shown for brevity, have also been obtained and obviously vary from day-to-day. Note in Figure 3(c) that the peaks in the distribution are consistent with the expected upper positive and lower negative charge centers of a thundercloud. Interestingly, Figure 3(b) shows that many of the ground flashes are actually hybrid flashes (i.e., contain an in-cloud component) as evident by the peak in the distribution just below 10 km.

5. LIGHTNING NOx PRODUCTION

The LNOM uses the geodetic segment altitude, an estimate of the peak lightning current *I* (in kiloamps) flowing through a segment, and the laboratory results of Wang et al. (1998) to estimate the NO production by a segment. By combining equations (6) and (9) in Wang et al. (1998), the NO production, *Q* (in x10^{21} NO molecules) from a

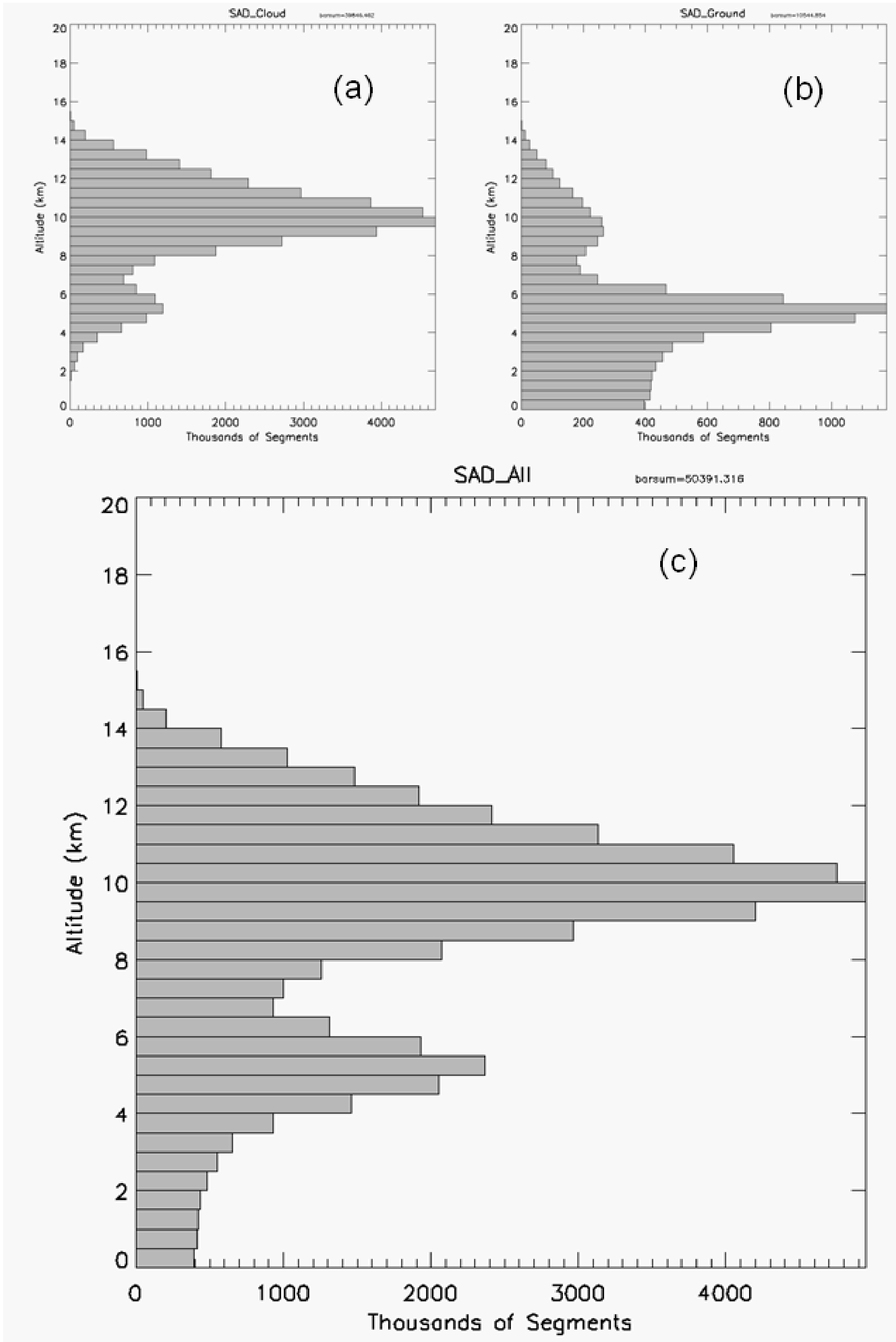

Figure 3. Segment Altitude Distribution (SAD) for (a) cloud, (b) ground, and (c) all flashes.

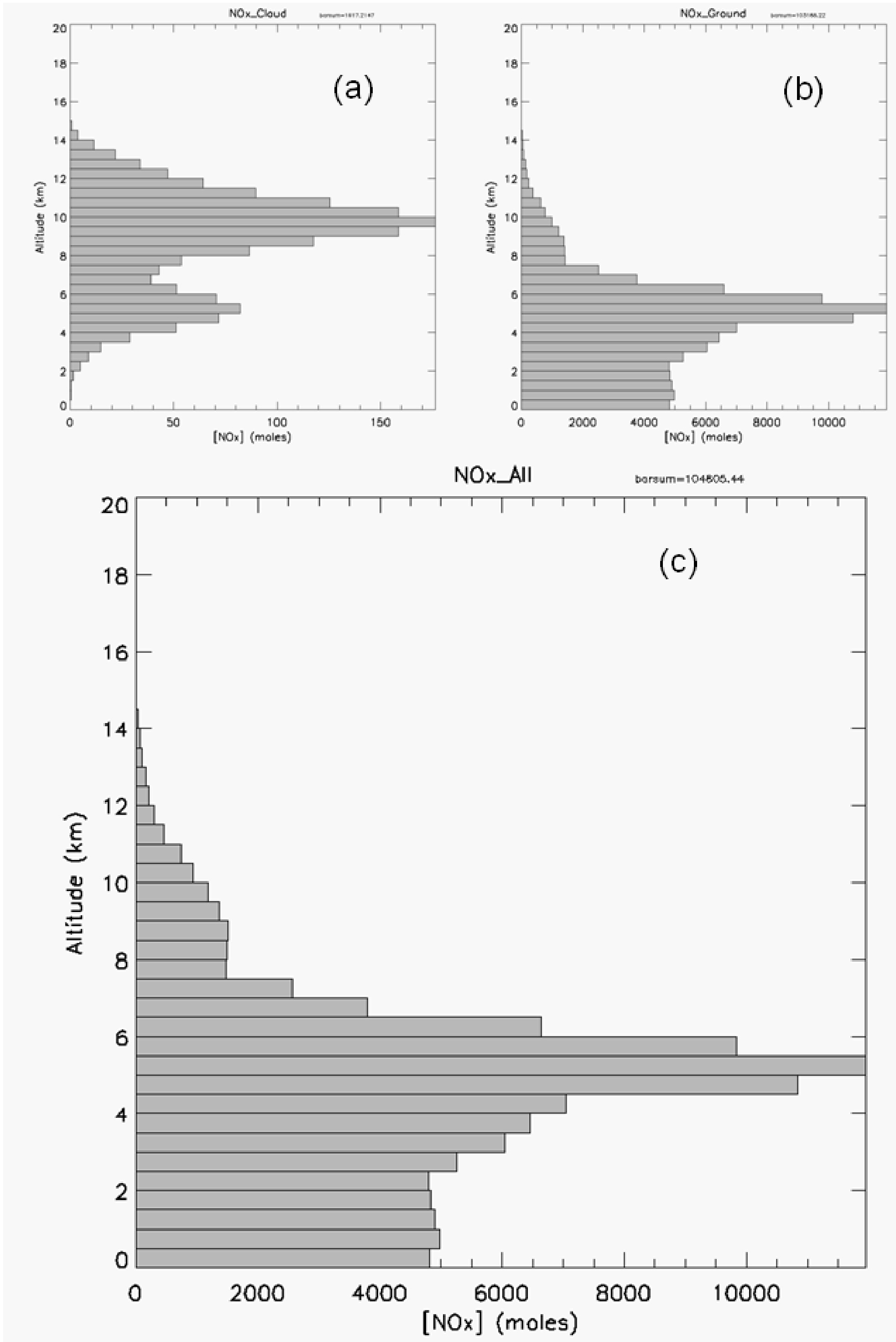

Figure 4. Lightning NOx Vertical Source Profile (VSP) for (a) cloud, (b) ground, and (c) all flashes. Cloud flash current is 4kA.

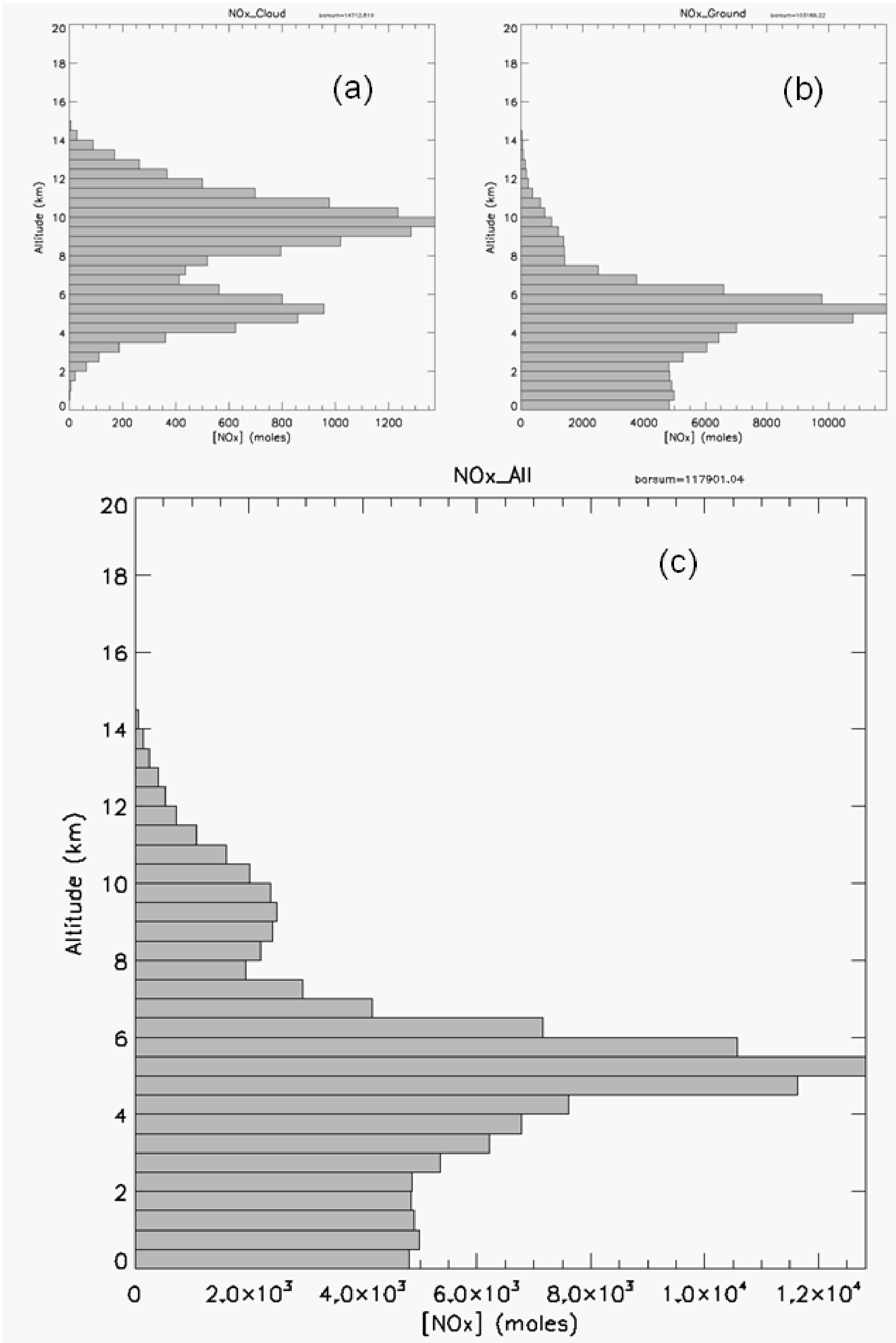

Figure 5. Lightning NOx Vertical Source Profile (VSP) for (a) cloud, (b) ground, and (c) all flashes. Cloud flash current is 15kA.

1-meter channel segment can be estimated as

$$Q(I,h) = m\left[a + b\left|I\right| + cI^2 - B(p_o - p(h)) \right]. \quad (1)$$

Here, *h* is geodetic segment altitude, and *p*(*h*) is the variation of pressure with height appropriate for MM5 (Grell et al., 1994), the dynamical model used in conjunction with CMAQ. The variable *m* is multiplicity (the number of strokes in a flash). The constant p_o is surface pressure, and (*a, b, c, B*) are positive empirical laboratory constants provided in Wang et al. (1998). As expected, equation (1) shows that the NO production increases for a channel segment that has a larger peak current and a lower altitude.

The actual computation of *Q* in the LNOM is more complicated than indicated in (1). In particular, the value of *Q* in (1) can go negative for high altitude low peak current segments (i.e., the second term in (1) can exceed the first term). In addition, equation (9) in Wang et al. (1998) gives a positive value of *Q* at *p* = 0 atm. To mitigate each of these problems, the LNOM employs 2 linear functions to model depletion of *Q* with decreasing pressure.

For ground flashes, the NLDN provides both the peak current and the multiplicity. However, subsequent strokes typically involve smaller peak currents than the 1st return stroke. Also, the peak current flowing through a segment that is part of a channel branch situated off the main channel is likely to be smaller than the peak current of the 1st return stroke. Hence, in both of these cases, NO production is likely overestimated by (1). Follow-on studies will attempt to improve this estimation process.

For cloud flashes, values of *I* are more difficult to obtain but are typically smaller than for ground flashes; e.g., Uman (1969) indicates cloud flash peak current of about 4 kA. We perform two LNOM runs, one with a cloud flash peak current of 4 kA, and another with a cloud flash peak current of 15 kA. The multiplicity for cloud flashes is taken as unity.

With these assumptions in mind, the LNOM computes the value of *Q* for each segment inside the analysis cylinder, as contributed by all flashes in a given time period. The values of *Q* are converted to units of moles, and are added up for each 500 m layer in the analysis cylinder; this comprises the VSP of lightning NOx for the analysis cylinder in the given time period. [Since the study by Wang et al. (1998) reported that 90-95% of the NOx produced from laboratory sparks was in the form of NO, the NO and NOx profiles are used interchangeably here.]

The final VSP results associated with the segment altitude distributions shown in Figure 3 (i.e., for the entire month of August, 2006) are provided in Figure 4 (4 kA cloud flash peak current) and Figure 5 (15 kA cloud flash peak current).

6. DISCUSSION

We have applied the NASA-MSFC Lightning Nitrogen Oxides Model (LNOM) to estimate the environmentally unmodified lightning NOx profile in a cylindrical volume centered over the NALMA network. The dimensions of the cylindrical volume were chosen to approximate a CMAQ grid volume. Since the NALMA and NLDN data utilized in this study provided detailed information about several key variables (i.e., the number of ground and cloud flashes that fall within the analysis cylinder, the channel geometry/location, and the peak currents and multiplicity associated with ground flashes), we have obtained direct estimates of the lightning NOx source profile. Laboratory spark chamber results were used to convert 1-m channel segments of a known geodetic altitude into NOx production.

The lightning NOx profiles obtained in this study will serve as both a reference for, and a "sanity check" of, future LNOM analyses that are conducted for regions lacking lightning observations. For example, LNOM will be applied to estimate the lightning NOx profiles in CMAQ grid volumes that are not covered by a lightning mapping array network. In this case, the number of cloud flashes in a CMAQ grid volume will be roughly estimated using the number of NLDN ground flash detections and an estimate of the "Z ratio" (the number of cloud flashes to ground flashes). Specifically, Z could be estimated using the climatological values offered in Boccippio et al. (2001), or derived by the empirical modeling in Price and Rind (1993).

Since we use the NLDN peak current of the 1st return stroke in a ground flash as an estimate of the peak current flowing through *all* 1-meter segments (i.e., even those segments that occur in subsequent return strokes, or in channel sections off the main channel, where peak currents are expected to be less), we would expect our results to overestimate the amount of lightning NOx produced in ground flashes.

However, from Figure 4 we obtained about 104,805 moles of NOx for the entire month, and from Figure 3 there was a total channel length of 50,391 km enclosed by the analysis cylinder during the month. Dividing the numbers results in a NOx production of 2.08 moles/km. Hence, *defining* a 'typical' flash as having a 10 km length, the typical NOx produced per typical flash would be (2.08 moles/km)(10km/flash) = 20.8 moles/flash. This value is reasonable given the spread of values in Table 1 of Labrador et al. (2004), and is actually on the low-end of the table. Part of the reduction may be attributable to those "flashes" that are removed for having fewer than 20 VHF sources. Using the 10^7J/km energy value provided in Hill (1979), we infer an average NOx chemical yield of (2.08 moles$\times N_A$/km)/(10^7J/km) = 1.25×10^{17} NO/J, where N_A is Avogadro's constant.

By similar means, for ground flashes we get: 103,188 moles/10,545 km = 9.8 moles/km, whereas for cloud flashes it is only: 1617 moles/39,846 km = 0.04 moles/km.

For the 15 kA cloud flash peak current run (Figure 5) we obtain the following results:

- All flashes: 2.34 moles/km, and an average yield of 1.41×10^{17} NO/J.
- Ground flashes: 9.8 moles/km.
- Cloud flashes: 0.37 moles/km.

Hence, by increasing the cloud flash peak current from 4 kA to 15 kA, the NOx production (per kilometer of cloud flash channel) increased by almost an order of magnitude.

So our preliminary results show that when channel length, altitude, peak current, and multiplicity are accounted for, ground flashes produce substantially more NOx than cloud flashes. Of course, this conclusion is linked to all of the assumptions we have made, including, for lack of measurements, an assumed constant peak current for all cloud flashes. Since the Wang et al. (1998) laboratory results for NOx production are quadratic in peak current, increases in the assumed cloud flash peak current will have a progressively stronger effect on the cloud flash NOx production.

Future studies will involve varying the adjustable parameters within the LNOM across reasonable ranges and examining how the output results change; in particular, even larger values of cloud flash peak current will be studied.

7. REFERENCES

Boccippio, D. J., K. L. Cummins, H. J. Christian, and S. J. Goodman, Combined satellite and surface-based estimation of the intracloud:cloud--to-ground lightning ratio over the continental United States, *Mon. Weather Rev.*, 129, 108-122, 2001.

Grell, G. A., J. Dudhia, D. R. Stauffer, A description of the fifth generation Penn State/NCAR Mesoscale Model, NCAR Technical Note (NCAR/TN-398+STR), December, 1994.

Heikkinen, M., Geschlossene formeln zur berechnung raumlicher geodatischer koordinaten aus rechtwinkligen koordinaten, *Z. Vermess.*, **107**, 207-211, 1982.

Hill, R. D., A survey of lightning energy estimates, *Rev. of Geophys. and Space Phys.*, **17**, 155-164, 1979.

Kaynak, B., Y. Hu, R. V. Martin, A. G. Russell, Y. Choi, and Y. Wang, The effect of lightning NO_x production on surface ozone in the continental United States, *Atmos. Chem. Phys. Discuss.*, **8**, 5061-5089, 2008.

Koshak, W. J., M. N. Khan, A. P. Biazar, M. Newchurch, R. T. McNider, A NASA model for improving the lightning NOx emission inventory for CMAQ, Joint Session: 4th Conference on the Meteorological Applications of Lightning Data and and the 11th Conference on Atmospheric Chemistry; 89th Annual AMS Conference, Phoenix, AZ, January 11-15, 2009.

Koshak, W. J., R. J. Solakiewicz, R. J. Blakeslee, S. J. Goodman, , H. J. Christian, J. M. Hall, J. C. Bailey, E. P. Krider, M. G. Bateman, D. J. Boccippio, D. M. Mach, E. W. McCaul, M. F. Stewart, D. E. Buechler, W. A. Petersen, D. J. Cecil, North Alabama Lightning Mapping Array (LMA): VHF source retrieval algorithm and error analyses, J. Atmos. Oceanic Technol., **21**, 543-558, 2004.

Labrador, L. J., R. von Kuhlmann, M. G. Lawrence, The effects of lightning-produced NOx and its vertical distribution on atmospheric chemistry: sensitivity simulations with MATCH-MPIC, Atmos. Chem. Phys. Discuss., **4**, 6239-6281, 2004.

McCaul, E. W., S. J. Goodman, K. M. LaCasse, D. J. Cecil, Forecasting lightning threat using cloud resolving model simulations, *Wea. Forecasting*, **24**, 709-729, 2009.

Pickering, K.E., Y.S. Wang, W.K. Tao, C. Price, and J.F. Muller, Vertical distributions of lightning NO_x for use in regional and global chemical transport models, *J. Geophys. Res.*, **103** (D23), 31203-31216, 1998.

Price, C. and D. Rind, What determines the cloud-to-ground lightning fraction in thunderstorms?, *Geophys. Res. Lett.*, **20**, 463-466, 1993.

Uman, Lightning, New York: McGraw-Hill, 1969.

Wang, Y., A. W. DeSilva, and G. C. Goldenbaum, Nitric oxide production by simulated lightning: dependence on current, energy, and pressure, *J. Geophys. Res.*, **103**, 19149-19159, 1998b.