Evaluating ammonia (NH_3) predictions in the NOAA National Air Quality Forecast Capability (NAQFC) using *in situ* aircraft, ground-level, and satellite measurements from the DISCOVER-AQ Colorado campaign

Authors: William H. Battye,[a,*] Casey D. Bray,[a] Viney P. Aneja,[a] Daniel Tong,[b,c,d] Pius Lee[b] and Youhua Tang[b,c]

Abstract

The U.S. National Oceanic and Atmospheric Administration (NOAA) is responsible for forecasting elevated levels of air pollution within the National Air Quality Forecast Capability (NAQFC). The current research uses measurements gathered in the DISCOVER-AQ Colorado field campaign and the concurrent Front Range Air Pollution and Photochemistry Experiment (FRAPPE) to test performance of the NAQFC CMAQ modeling framework for predicting NH_3. The DISCOVER-AQ and FRAPPE field campaigns were carried out in July and August 2014 in Northeast Colorado. Model predictions are compared with measurements of NH_3 gas concentrations and the NH_4^+ component of fine particulate matter concentrations measured directly by the aircraft in flight. We also compare CMAQ predictions with NH_3 measurements from ground-based monitors within the DISCOVER-AQ Colorado geographic domain, and from the Tropospheric Emission Spectrometer (TES) on the Aura satellite.

In situ aircraft measurements carried out in July and August of 2014 suggest that the NAQFC CMAQ model underestimated the NH_3 concentration in Northeastern Colorado by a factor of ~2.7 (NMB = –63%). Ground-level monitors also produced a similar result. Average satellite-retrieved NH_3 levels also exceeded model predictions by a factor of 1.5 to 4.2 (NMB = -33 to -76%). The underestimation of NH_3 was not accompanied by an underestimation of particulate NH_4^+, which is further controlled by factors including acid availability, removal rate, and gas-particle partition. The average measured concentration of NH_4^+ was close to the average predication (NMB = +18%).

Seasonal patterns measured at an AMoN site in the region suggest that the underestimation of NH_3 is not due to the seasonal allocation of emissions, but to the overall annual emissions estimate. The underestimation of NH_3 varied across the study domain, with the largest differences occurring in a region of intensive agriculture near Greeley, Colorado, and in the vicinity of Denver. The NAQFC modeling framework did not include a recently developed bidirectional flux algorithm for NH_3, which has shown to considerably improve NH_3 modeling in agricultural regions. The bidirectional flux algorithm, however, is not expected to obtain the

[a] North Carolina State University, 2800 Faucette Drive, Raleigh, NC 27695-8208
[b] NOAA Air Resources Laboratory, 5830 University Research Court, College Park, Maryland, MD 20740
[c] Cooperative Institute for Climate and Satellites, University of Maryland, College Park, Maryland, MD 20740
[d] Center for Spatial Information Science and Systems, George Mason University, Fairfax, Virginia, VA 22030
[*]Please address correspondence to: William Battye; Department of Marine, Earth and Atmospheric Sciences; North Carolina State University, Campus Box 8208, Raleigh, NC 27695-8208; tel. (919) 357-0425, fax (919) 484-0122, email whbattye@ncsu.edu.

magnitude of this increase sufficient to overcome the underestimation of NH_3 found in this study. Our results suggest that further improvement of the emission inventories and modeling approaches are required to reduce the bias in NAQFC NH_3 modeling predictions.

Keywords: NH3, ammonia, model evaluation, CMAQ, aircraft measurement, remote sensing

1. Introduction and background

Gaseous ammonia (NH_3) in the atmosphere contributes to the formation of ammonium (NH_4^+) compounds – including ammonium bisulfate (NH_4HSO_4), ammonium sulfate [$(NH_4)_2SO_4$], and ammonium nitrate (NH_4NO_3) – which comprise a large fraction of airborne fine particulate matter ($PM_{2.5}$) (Kwok *et al.*, 2013). Elevated levels of $PM_{2.5}$ are associated with various adverse human health impacts, including irregular heartbeat, aggravated asthma, and premature death (Pope *et al*., 2009), and can contribute to visibility impairment and regional haze (Wang *et al.*, 2012). NH_3 gas can play a role in the nucleation of new particles (Holmes, 2007), and can sometimes control nucleation events (Herb *et al.*, 2011).

Atmospheric NH_3 and NH_4^+ deposit to terrestrial and aquatic ecosystems though wet and dry deposition processes. This leads to an increase in the level of biologically available nitrogen, which can affect species diversity and can lead to eutrophication of aquatic ecosystems (Jones, 2013; Paerl, 1988; and U.S. EPA SAB, 2007). In terrestrial ecosystems, NH_3 and NH_4^+ are oxidized by soil microbes to nitrate (NO_3^-) and other oxidized nitrogen species, resulting in acidification of the soil. A portion of the NH_3 and NH_4^+ processed by soil microbes is also converted to gaseous nitrous oxide (N_2O), which reenters the atmosphere. N_2O is a long-lived absorber of infrared radiation, with a climate change potential approximately 250 times that of CO_2 (IPCC, 2013).

The U.S. National Oceanic and Atmospheric Administration (NOAA) is responsible for forecasting elevated levels of air pollution within the National Air Quality Forecast Capability (NAQFC) (Tang *et al*., 2015). NOAA uses the Community Multiscale Air Quality (CMAQ) model (Byun and Schere, 2006) to simulate atmospheric emissions and transport of NH_3, and conversion of NH_3 to $PM_{2.5}$, and deposition of NH_3 and NH_4^+ to terrestrial and aquatic ecosystems. The capability of NAQFC to predict NH_3 and NH_4^+ in $PM_{2.5}$ has not been thoroughly evaluated.

An important source of uncertainty for NH_3 modeling is the inventory of emissions used in CMAQ (Battye *et al*, 2003). Agricultural sources account for approximately 90% of atmospheric NH_3 emissions in the U.S. (Aneja *et al.,* 2009). These emissions emanate primarily from animal waste management and synthetic nitrogen fertilizer application (Battye *et al.,* 2002). NH_3 emissions estimates are calculated by applying emission factors and emission models to the agricultural census (USEPA, 2009). These emissions are allocated to different times of the year and to geographic modeling grids using temporal and spatial allocation factors, which add to the uncertainty of model emissions estimates. Validation studies of NH_3 emissions estimates in CMAQ have focused on secondary indicators such as wet deposition of NH_4^+ ions, and the concentration of NH_4^+ in $PM_{2.5}$ (Gilliland *et al*., 2006, Kelly *et al.,* 2014).

This current study evaluates NAQFC CMAQ predictions for NH_3 in Northeastern Colorado against direct measurements of NH_3 in the atmosphere. Comparisons are made using three different measurement platforms for NH_3: *in situ* sampling by aircraft, ground-level passive samplers, and satellite data retrievals. In addition, model predictions of NH_4^+ (fine-mode)

particulate matter are evaluated against *in situ* aircraft measurements. We also use long-term measurements from ground level monitors, and from the Tropospheric Emission Spectrometer (TES) on the Aura satellite to evaluate temporal patterns of atmospheric NH_3.

2. Methodology

The current research uses measurements of NH_3 and NH_4^+ collected during the DISCOVER-AQ Colorado field campaign to assess the performance of the NAQFC CMAQ modeling framework for predicting NH_3 concentrations. (DISCOVER-AQ was a program for *D*eriving *I*nformation on *S*urface *C*onditions from *CO*lumn and *VER*tically resolved observations relevant to *A*ir *Q*uality.) The DISCOVER-AQ Colorado field campaign, which was carried out from July 17 through August 10, 2014 in the Front Range of the Rocky Mountains in Northeast Colorado, included in-situ aircraft measurements, ground-based measurements, and satellite measurements. Figure 1 shows the locations of the aircraft flights, ground level monitors, and the swath of satellite measurements.

Model predictions are compared with measurements of NH_3 gas concentrations and NH_4^+ fine particulate matter concentrations measured directly by the aircraft during flight. We also compare CMAQ predictions with NH_3 measurements from ground-based monitors within the DISCOVER-AQ Colorado geographic domain, and from TES.

2.1. Air Quality Model

Within the NAQFC framework, CMAQ model version 5.0.2 was used to predict air pollutant concentrations for the continental U.S. during the summer of 2014 (CMAS 2016). Meteorological predictions to drive the CMAQ model were generated using the Weather Research and Forecasting Advance Research WRF (WRF-ARW) regional meteorological model. The horizontal resolution of both models is 12 km, with 42 vertical layers with a domain top at 50 hPa. More vertical layers are used below 1 km. The height of the lowest vertical layer was 8 meters above the ground within the DISCOVER-AQ domain. The configuration of the CMAQ and WRF-ARW models within the NAQFC is described in more detail in Tang *et al*. (2015).

Air pollutant emissions for the NAQFC are derived from the U.S. National Emissions Inventory (NEI). At the time of the Colorado field study, the 2005 NEI was being used, with several major updates as described in Tong *et al* (2015). For NH_3, the NEI provides county-level estimates of annual emissions. These annual emissions estimates are allocated the 12-km model grid and to hourly values using the Sparse Matrix Operator Kernel Emissions (SMOKE) system (Vukovich and Pierce, 2002). Aerosol chemistry is based on the AERO5 module of CMAQ version 4.7.1 (Binkowski and Shankar, 1995), and dry deposition computed for NH_3 is based on the M3Dry module (Mathur *et al*., 2005).

It must be noted that the NAQFC modeling framework at the time of the 2014 field study did not account for the potential bidirectional flux of NH_3 between the bottom layer of the model and the surface. A bidirectional surface exchange model for NH_3 has recently been developed and implemented in CMAQ (Cooter *et al*. 2012; Bash *et al*. 2013; Pleim *et al*. 2013). This model replaces the unidirectional deposition flux algorithm for NH_3 and adds a term for the potential evaporation of NH_3 to the air from vegetated landscapes. This upward flux of NH_3 offsets the deposition flux, resulting in higher atmospheric concentrations of NH_3. Testing of the

bidirectional flux model has predicted NH_3 concentrations 10% higher, on average, than previous predictions with the unidirectional deposition flux approach (Cooter *et al*. 2012; Bash *et al*. 2013).

2.2. Aircraft measurements

We compared CMAQ model predictions of gaseous NH_3 with measurements made in flight by a Lockheed P3B Orion aircraft operated by the National Aeronautics and Space Administration (NASA). The rate of conversion of gaseous NH_3 particulate NH_4^+ is a potential source of discrepancy between the modeled and measured NH_3 concentrations. Therefore, we also compared modeled and measured values for the sum of gaseous NH_3 and particulate NH_4^+, NH_X. The aircraft measurements were made at elevations ranging from ground level to 5 km above ground level, and included upward spirals, downward spirals, and transect flights in the Front Range of the Rocky Mountains, around Denver, Boulder, Fort Collins, and Greeley, Colorado.

The measured values of NH_3 and NH_4^+ were obtained from the DISCOVER-AQ Colorado field campaign archive. P3B aircraft measurements of NH_3 and NH_4^+ are described in detail in Müller *et al*. (2014) and Sun *et al*. (2015). Ambient air was directed to an array of instruments located on-board the aircraft. NH_3 was measured using a proton transfer reaction time-of-flight mass spectrometer (PTR-MS). NH_3 concentrations were measured every 10 s; and 1-minute averages were also computed. The 1-minute averages were used for model-to-measurement comparisons.

The PTR-MS measurement system for NH_3 was evaluated in a previous DISCOVER-AQ campaign in the San Joaquin Valley of California (Sun *et al.,* 2015). The PTR-MS system was found to have a measurement accuracy of ±35% and a 1σ measurement precision of 5.5–6.5 ppbv at 1 s time resolution, or 0.75 ppbv for a 1-minute average. This variability results in a low signal-to-noise ratio, especially for NH_3 in the free troposphere, where concentrations are below 1 ppbv. In order to reduce the impact of this high value for measurement precision, our comparisons of aircraft data with model predictions focus on measurements made at altitudes below 1,000 m above ground level, as measured by radar.

Concentrations of NH_4^+ aerosol, and other soluble aerosols were measured by a Particle-into-Liquid-Sampler followed by ion chromatography (PILS-IC). The NH_4^+ concentration was recorded every minute. In side-by-side comparisons, the NASA PILS-IC system showed good correlation with filter measurements, giving a slope of ~0.93, intercept of ~0.24 $\mu g\ m^{-3}$, and r-value of 0.94. Precision for calculated at ~0.4 $\mu g\ m^{-3}$. (Orsini *et al*, 2003).

As air pollutant concentrations were recorded, the location, altitude, speed, bearing, and angle of ascent or descent were recorded using data from the aircraft navigation system and global positioning system (GPS). The height above ground level was also measured using radar. CMAQ model predictions of NH_3 and NH_4^+ were extracted for comparison with for each 1-minute average aircraft measurement. The CMAQ prediction at a given measurement location and time is computed by 4-dimensional interpolation across space and time, using the model grid cells surrounding the measurement point at the appropriate model layer height.

2.3. Ground-level measurements

Ground level measurements of NH_3 were obtained from 3 monitoring sites of the Ammonia Monitoring Network (AMoN) located within the DISCOVER-AQ Colorado domain, for the period 2007 through 2014. The AMoN network is operated under the National Atmospheric Deposition Program (NADP) to provide a consistent, long-term record of NH_3 gas concentrations across the U.S. (NADP 2014). AMoN monitors use passive diffusion collectors which are changed every two weeks. The detection limit of the AMoN passive sampler is approximately 1.5 ppbv for samples collected over a 24 hour period, or 100 pptv for samples collected over a two-week period (Sigma Aldrich). The accuracy is estimated at $\pm$6%. NH_3 measurements were also obtained for 12 passive samplers in the study domain operated by Colorado State University (CSU) during the DISCOVER-AQ campaign timeframe (Benedict, 2015). The CSU measurement network used Radiello passive samplers, changed every week. Methods used by CSU are described in more detail in Day, *et al* (2012).

CMAQ NH_3 predictions were compared with these passive sampler measurements. NH_3 concentration results were extracted for the grid cells surrounding each monitor location, in the lowermost model layer. The model grid cell results were interpolated to the monitor location sites and averaged for the passive sampler measurement periods.

2.4. Satellite measurements

CMAQ predictions were also compared with NH_3 concentrations retrieved from infrared spectra gathered by the TES instrument on the Aura satellite. TES performed 5 transect measurements over the DISCOVER-AQ study domain between July 29 and August 14, 2014. These were all daytime passes, between 1:00 and 1:30 PM local standard time.

The NH_3 retrievals rely on the change in intensity of infrared radiation across a number of specific wavelength bands which are chosen to cover a sharp feature in the NH_3 infrared absorption spectrum (940–970 cm^{-1}). A forward radiative transfer model (RTM) is used to compute the expected intensity of radiation in the selected bands at the top of the atmosphere. The RTM requires input information on the atmospheric density, relative humidity and concentrations of other trace gases, as well as an a priori assumption on the concentration of NH_3. The retrieval for NH_3 is carried out after retrievals for temperature and other trace gases. The assumed concentration profile of NH_3 is varied to minimize the error between the spectrum predicted by the RTM and the spectrum actually measured by the satellite. This results in an estimate of the concentration of NH_3 for the region sensed by the satellite. (Shephard *et al*, 2012). In the current study, only those measurements which passed TES quality assurance checks were used (Species Retrieval Quality = 1).

The estimated concentration of NH_3 is affected by and may tend to be biased toward the a priori assumption made for NH_3. In addition, the satellite is seeing an absorption by the entire atmospheric column. Although the retrieval algorithm is used to estimate the vertical distribution of NH_3, this vertical distribution is also subject to uncertainties and is affected by the a priori assumption.

2.5. Model to measurement comparisons

Prediction accuracy for the NAQFC CMAQ model was quantified by computing the normalized mean bias (NMB), and the ratio of the average measured concentration to the average model prediction ($R_{o/m}$):

$$NMB = \frac{1}{N}\frac{\sum_{i=1}^{N}[C_{mod}(i) - C_{obs}(i)]}{\sum_{i=1}^{N} C_{obs}(i)}$$

and:

$$R_{o/m} = \frac{\sum_{i=1}^{N} C_{obs}(i)}{\sum_{i=1}^{N} C_{mod}(i)}$$

Where $C_{mod}(i)$ and $C_{obs}(i)$ are, respectively, the model prediction and the observed concentration at a given location and time, and N is the number of observations. $R_{o/m}$ and NMB are related to one another as follows:

$$NMB = \frac{1}{R_{o/m}} - 1$$

The Pearson correlation coefficient (r) and the concordance correlation coefficient (ρ_c) were used to evaluate correlation of the measured concentrations with predicted concentrations. The concordance correlation coefficient is also known as the reproducibility index, and gives a more rigorous test of whether modeled values predict observed values (Lin, 1989, 1992).

3. Results and discussion

3.1. Comparison of model predictions with in situ aircraft measurements

Table 1 summarizes the results of the comparison of *in situ* aircraft measurements with model predictions for of NH_3, NH_4^+ in $PM_{2.5}$ and NH_X. For each measurement location, the corresponding CMAQ prediction was interpolated based on the surrounding grid cells at the appropriate model layer heights. The concentration pairs were then compared directly, without any adjustment for altitude. The aircraft measurements were carried out during the day, and our comparisons were restricted to measurements taken below 1000 m in altitude. Therefore, these measurements are generally within the well-mixed planetary boundary layer (Arya, 1999). The average measured NH_3 concentration was 6.1 ppbv (3.9 µg/m^3), with a standard deviation of 6.9 ppbv (4.2 µg/m^3) and a maximum measured value of 90.0 ppbv (53.1 µg/m^3). In comparison, the average model prediction at the locations and times corresponding to these measurements was 2.2 ppbv (1.4 µg/m^3). The standard deviation of the model prediction was 1.6 ppbv (1.4 µg/m^3) and the maximum model prediction was 15.3 ppbv NH_3 (9.1 µg/m^3). The average measured concentration of NH_3 was a factor of 2.7 higher than the average of model predictions at the sample locations. This corresponds to a normalized mean bias for NH_3 of −63%.

The average measured concentration of particulate NH_4^+ was 0.29 µg/m^3, which reflects an average conversion of 7% of NH_3 to NH_4^+. The average model prediction was 0.34 µg/m^3, corresponding to an average conversion of 23%. The ratio of the average measured concentration on particulate NH_4^+ to the average model prediction was 0.85, corresponding to a normalized

mean bias of +18%. Thus, the underestimation of gaseous NH_3 was not accompanied by an underestimation of particulate NH_4^+. However, the relative magnitude of predicted NH_3 gas and particulate NH_4^+ suggests that the formation of NH_4^+ was not limited by availability of NH_3.

The results of a comparison for NH_X (the combination of NH_3 vapor and particulate NH_4^+) are similar to the results for gaseous NH_3 alone. The average measured concentration of NH_X is a factor of 2.5 higher than the average of corresponding model predictions, and the normalized mean bias is an under-prediction of 60%. These values are slightly lower than the values for NH_3 vapor alone.

Figure 2a plots the measured concentrations of NH_3, on the y-axis, against model predictions on the x-axis. Figures 2b and 2c provide similar plots for NH_4^+ and NH_X, respectively. Each measurement is plotted as a point. Two lines are also included in each plot. The dotted lines show a 1:1 slope, where points would have fallen if the measurements and model predictions were in complete agreement (measured = modeled). The dashed lines show the 1:1 slope displaced by the NMB.

The graphs in Figure 2 show substantial scatter for all three pollutants. In all three cases, high measured values can occur where model predictions are low, and vice versa. For both NH_3 and NH_X, the majority of measurements fall above the prediction line (measured = modeled). For NH_4^+, the measurements fall evenly on both sides of the line.

Figure 3 compares a histogram of the measured NH_3 concentration with a histogram of the modeled NH_3 concentration. The figure illustrates that the distribution of model predictions falls off much more swiftly than the distribution of measured concentrations. However, the structure of the two profiles is similar. Figure 3 shows that the model does not produce the full range of values found in the measured data set at the high end. The 98th percentile of measured values was 23 ppbv while the 98th percentile level of corresponding modeled values was 6 ppbv. However, Figure 2a shows that the underestimation in not restricted to the high end, but affects the full range of NH_3 concentrations.

In order to identify spatial patterns in the model prediction error, NMB and $R_{o/m}$ were computed using the *in situ* aircraft measurements within each 12 km modeling grid. Figure 4 presents the results of this analysis. In the figure, a background raster (in blue) shows the average CMAQ prediction during the DISCOVER-AQ campaign. Round icons indicate the ratio of the average measured concentration to the average model prediction ($R_{o/m}$). The largest differences between modeled and measured NH_3 were around Greeley, in Weld County.

Over 1,300 cattle operations are located in Weld County (USDA, 2014), including two feedlots which are among the largest in the U.S. (CSU, 2016). The inventory of cattle in Weld County is over 500,000, the 3rd largest cattle population of any U.S. county (USDA, 2014). The concentration of cattle operations in the Greeley area resulted in model predictions of NH_3 which were higher than those in the rest of the modeling region. Measured NH_3 in the Greeley area were a factor of 3 to 4.3 higher than the model predictions. Similar ratios of measured-to-modeled NH_3 were found near Denver; however the magnitude of both modeled and measured NH_3 concentrations were lower than in the area around Greeley.

Each of the icons for $R_{o/m}$ in Figure 4 represents multiple measurements (85 on average), with the icon at the Northeastern of the loop near Greeley representing 113 measurements. Nevertheless, these measurements are localized along the path of the aircraft. Thus, it is possible that the measurements are affected by local hotspots of NH_3, so that the large values of $R_{o/m}$ may apply to only a fraction of the modeling grid.

In summary, the average concentration of NH_3 measured by in situ aircraft sampling was a factor of 2.7 higher than the average of model predictions at the sample locations. However, the underestimation of gaseous NH_3 was not accompanied by an underestimation of particulate NH_4^+. The under prediction of NH_3 was more pronounced in an area around Greeley with high NH_3 emissions. In addition, the highest concentrations of NH_3 predicted by the model were considerably lower than the highest measurements.

3.2. Model predictions compared with ground level passive measurements

Table 2 and Figure 5 summarize the results of the comparison of measured concentrations with model predictions for 3 passive NH_3 samplers operated under the AMoN network and 12 passive samplers operated by CSU. One of the AMoN sites is located in Fort Collins, Colorado, with the Rocky Mountains to the west and an agricultural region to the east. The remaining three AMoN sites are in remote areas, including two in the Rocky Mountain National Park. Most of the CSU sampling sites are in areas of intensive agriculture around the city of Greeley. In each comparison between the model and a ground-level measurement, we computed the average model prediction for the entire duration of the ground-level measurement (14 days for AMoN and 7 days for CSU). Thus, the measurement and the model prediction were compared on the same basis, from the standpoint of averaging time.

The CSU monitoring results are high in comparison with the AMoN results. However, as noted above, these monitors are located in an area of intensive agriculture. The results for the CSU monitors are comparable to the results of in situ aircraft measurements made near Greeley. In addition, the NMB for the CSU monitors is comparable to the NMB for the AMoN monitors.

The average measured NH_3 concentration for all ground-level passive monitors was 16.0 ppbv (9.5 $\mu g/m^3$), with a standard deviation of 19.8 ppbv (11.7 $\mu g/m^3$) and a maximum measured value of 116.3 ppbv (68.7 $\mu g/m^3$). In comparison, the average model prediction at the locations and times corresponding to these measurements was 6.0 ppbv (3.5 $\mu g/m^3$). The standard deviation of the model prediction was 3.7 ppbv (2.2 $\mu g/m^3$) and the maximum model prediction was 12.8 ppbv NH_3 (7.6 $\mu g/m^3$). The average measured concentration of NH_3 was a factor of 2.7 higher than the average of the corresponding model predictions. The normalized mean bias (NMB) for NH_3 was an under-prediction of 63%. This confirms the result for in situ aircraft measurements, discussed above.

3.3. Model predictions compared with satellite retrievals

Table 3 and Figure 6 compare CMAQ model predictions with NH_3 concentration estimates retrieved from TES satellite spectroscopic measurements. Three separate comparisons were made: one using the estimated total atmospheric column loading, the second using the estimated concentration in the lowest layer of the atmosphere, and the third using the estimated concentration at an altitude of 1740 m above ground level (AGL). This is the altitude where the averaging kernel indicates that the retrieved concentration from the satellite measurement is most sensitive to the actual atmospheric concentration.

The NMB for the model prediction of total column loading (−76%, $R_{o/m} = 4.2$) is somewhat more negative than the NMB for the comparisons with aircraft data and ground level

monitoring data. The model prediction for the lowest layer of the atmosphere has a less negative NMB (−33%, $R_{o/m}$ = 1.5) than the prediction for total column loading, or than the comparisons with aircraft and ground level monitor data. The average TES retrieval for the lowest layer of the atmosphere is also lower than concentrations measured in the same region by aircraft and by the CSU monitors (Tables 1 and 2). The NMB of the model prediction at 1740 m AGL (−53%, $R_{o/m}$ = 2.1) is midway between the results for the total column loading and the ground level concentration. Model predictions for this altitude also give a better correlation with the satellite retrieval (r = 0.52) than the ground level concentration (r = 0.09) or the total column loading (r = 0.11).

The NMB from the satellite data analysis is subject to considerable uncertainty, as highlighted by the variability among the different satellite metrics for NH_3 (Table 3). However, the satellite results for NH_3 are in agreement with the aircraft and ground-level results discussed above.

3.4. Satellite retrievals compared with in situ aircraft measurements

The TES satellite swath was not aligned with aircraft spiral measurements; however, a number of aircraft flight paths crossed the satellite swath close to the time of satellite passage. We identified 46 *in situ* observations which occurred within an hour of a TES satellite pass, and within 15 km of the center of the satellite swath. These *in situ* measurements were compared with the TES NH_3 retrievals for the atmospheric layer corresponding to the aircraft elevation. Table 4 and Figure 7 summarize the results of this comparison. The average of aircraft measurements overlapping the TES track was 2.9 ppbv, with a standard deviation of 2.4 ppbv, and a maximum value of 8.1 ppbv. The average of TES retrievals corresponding to these measurement locations was 2.8 ppbv, with a standard deviation of 2.5 ppbv, and a maximum value of 6.6 ppbv. Thus, the normalized mean bias of the TES retrieval with respect to the *in situ* measurement was only −1%. The correlation coefficient (r) and concordance correlation coefficient between the TES retrieval and the aircraft measurement are both 0.78. Thus, the TES results exhibit good correlation with the aircraft measurements.

3.5. Analysis of model bias in relation to previous studies and the NH_3 emissions inventory

Gilliland *et al* (2006) performed inverse modeling in order to evaluate the emissions inventory for NH_3. Measurements of NH_4^+ in precipitation were used with a 2001 CMAQ simulation for the continental U.S. Annual emissions estimates were found to be reasonable on average, but inverse modeling results indicated that the NH_3 emissions inventory was too high in winter and too low in summer. On a domain-wide basis, the posterior NH_3 emissions inventory for the July-August timeframe was 17% higher than the prior inventory. Smaller-scale analyses of the data suggested that the error may have been higher in the western U.S., however these results were unstable due to low precipitation rates.

Butler *et al* (2014) evaluated CMAQ predictions in Susquehanna River Watershed of New York and Pennsylvania using ambient concentration measurements conducted in 2008 and 2009. The model estimates were lower than measured values by 8% to 60%.

Kelly *et al* (2014) evaluated CMAQ predictions in the San Joaquin Valley of California using measurements from the measurement campaign for “California Research at the Nexus of

Air Quality and Climate Change" (CalNex) in May and June of 2010. The study analyzed multiple pollutants, including NH_4^+ and NH_3. The model performed well for NH_4^+. NH_3 was over-predicted in some urban areas; however, this was attributed to errors in prediction of the mixing layer behavior. The model under-predicted NH_3 in agricultural regions. In addition, model predictions did not capture the large variations in measured NH_3.

Zhu *et al* (2013) performed inverse modeling of ambient NH_3 in the Continental U.S. using TES satellite data in conjunction with the GEOS-Chem model. TES data were assimilated for April, July, and October of 2006 through 2009. AMoN data were used to evaluate the inverse modeling results. The study found that the initial NH_3 emissions inventory appeared to be an underestimate, especially in the Western U.S.

The current study found that the NAQFC CMAQ model underestimated the NH_3 concentration in Northeastern Colorado in July and August of 2014 by a factor of ~2.7 (NMB = −63%). This difference is larger than the differences found by Gilliland *et al* (2006) and Butler *et al* (2014). However, these studies differed from the current study in important ways. The Gilliland study used deposition measurements to evaluate CMAQ predictions; and the Butler study focused on a region of low NH_3 concentration. The findings of the current study are comparable to the findings of Kelly *et al* (2014) for an agricultural region in California. Both the current study and the Kelly study included regions with intensive agriculture. A European study using CMAQ as part of the CALIOPE-EU modeling system also found that NH_3 concentrations were underestimated in the summer months (Pay *et al*, 2012).

Measured and modeled concentrations of NH_4^+ were much lower than the measured concentration of NH_3. Therefore, any differences in the conversion of NH_3 to NH_4^+ would be too small to account for the underestimation of NH_3. Rather, the model error for NH3 is believed to result from either the NH_3 emissions inventory, or to the rate of NH_3 deposition. As noted in Section 2.1, the NAQFC modeling framework used in the current study did not include a recently-developed bidirectional flux algorithm for NH_3 between the bottom layer of the model and the surface. (Cooter *et al*. 2012; Bash *et al*. 2013; Pleim *et al*. 2013). Testing of the bidirectional flux model has predicted NH_3 concentrations 10% higher, on average, than previous predictions with the unidirectional deposition flux approach (Cooter *et al*. 2012; Bash *et al*. 2013). Thus, we would not expect the incorporation of bidirectional flux, by itself, to correct the underestimation of NH_3 for the DISCOVER-AQ domain.

As illustrated in Figure 4, the model bias varies across the DISCOVER-AQ domain, with larger differences in the neighborhood of Greeley and Denver. The Greeley area is a region of intensive agriculture, with high levels of NH_3 emissions in the 2005 NEI. Thus, CMAQ predictions of NH3 in this area are higher than the surrounding region. However, results of the model-to-measurement comparison indicate that emissions in the Greeley region may have been still higher than the levels reflected in the inventory.

The current study also uses NH_3 emissions estimates from the 2005 NEI, which have recently been updated in the 2011 NEI. However, the change in estimated NH_3 emissions from the 2005 NEI to the 2011 NEI was only an increase of 10% within the DISCOVER-AQ Colorado domain (USEPA, 2009 and 2015). Long term NH_3 monitoring trends at the Fort Collins AMoN site also do not show an increase in measured NH_3 concentrations over this period. Figure 8 shows that measured concentrations in 2014 at Fort Collins fall within the range of concentrations measured for the preceding 7 years.

On the timescale of the summer measurement campaign, errors in the emissions inventory can arise not only from the overall emission factors, but also from the seasonal

allocation of emissions. However, the increase in the measured NH_3 concentration at Fort Collins is less than the increase in NH_3 emissions in the modeling domain, based on the seasonal factors used in the NEI. The measured NH_3 concentration during the monitoring campaign was 1.44 times the annual average concentration at the Fort Collins site in the calendar year 2014. Based on seasonal allocation factors used in the NEI for NH_3, emissions used in July and August are 1.8 times the annual average. Thus, the underestimation in NH_3 for the campaign is not believed to result from errors in seasonal allocation.

Summary and Conclusions

This paper describes an evaluation of the NOAA NAQFC predictions of NH_3 and NH_4^+ using a number of different data sources. The primary data source is a large set of aircraft-based *in situ* measurements from the DISCOVER-AQ Colorado campaign. In addition, data were obtained from the ground-based AMoN network, a ground-based study carried out by CSU in concert with the DISCOVER-AQ campaign, and satellite-based TES instrument. The NAQFC model underestimated Northeastern Colorado NH_3 concentrations during the July and August of 2014 by a factor of ~2.7 when compared to aircraft emissions measurements. Similar results were observed for the AMoN, CSU, and TES datasets, with the model underestimating NH_3 by 1.5 to 4.2 times. However, the underestimation of gaseous NH_3 was not accompanied by an underestimation of particulate NH_4^+.

The model error for NH3 is believed to result from either the NH_3 emissions inventory, or to the rate of NH_3 deposition. The NAQFC modeling framework did not include a recently-developed bidirectional flux algorithm for NH_3. Although the bidirectional flux algorithm could be expected to raise NH3 concentrations in the summer months; however, the magnitude of this increase is not believed to be sufficient to overcome the underestimation of NH3 which was found in this study.

The underestimation of NH3 varied across the study domain, with the highest errors occurring in a region of intensive agriculture near Greeley, and in the vicinity of Denver. Seasonal patterns measured at an AMoN site in the region suggest that the underestimation of NH3 is not due to the seasonal allocation of emissions, but to the overall annual emissions estimate.

Acknowledgements

Support for this research was provided by the NASA Earth and Space Science Fellowship (NESSF) program, grant No. NNX15AN15H. We thank Dr. Katie Benedict, Colorado State University, for providing us with the passive sampler data. We also acknowledge the NOAA National Air Quality Forecast Program, the DISCOVER-AQ and FRAPPE measurement campaigns, and the North Carolina State University Air Quality Research Group.

References

Aneja, V.P., W.H. Schlesinger, and J.W. Erisman. 2009. Effects of Agriculture upon the Air Quality and Climate: Research, Policy and Regulations. *Environmental Science and Technology.* 43(12): 4234-4240.

Arya, S.P. 1999. Air pollution meteorology and dispersion. Oxford University Press, New York, NY.

Bash, J.O, E.J. Cooter, R.L. Dennis, J.T. Walker, and J.E. Pleim. 2013. Evaluation of a regional air-quality model with bidirectional NH_3 exchange coupled to an agroecosystem model. *Biogeosciences*. 10: 1635-1645.

Battye, W., V.P. Aneja, P.A. Roelle. 2003. Evaluation and improvement of ammonia emissions inventories. *Atmospheric Environment.* 37(27): 3873-3883.

Benedict, K. 2015. Results of passive NH3 sampling by Colorado State University during the FRAPPE campaign. Personal communication. May 18.

Binkowski, F.S., and U. Shankar. 1995. The regional particulate model 1. Model description and preliminary results. Journal of Geophysical Research. 100 (D12): 26,191-26,209. DOI:10.1029/95JD02093.

Butler, T., R. Marino, D. Schwede, R. Howarth, J. Sparks, and K. Sparks. 2014. Atmospheric ammonia measurements at low concentration sites in the northeastern USA: implications for total nitrogen deposition and comparison with CMAQ estimates. *Biogeochemistry.* DOI: 10.1007/s10533-014-0036-5.

Byun, D., and K.L. Schere. 2006. Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Review*. 59 (1-6), 51-77.

Cooter, E.J., J.O. Bash, V. Benson, and L. Ran. 2012. Linking agricultural crop management and air quality models for regional to national-scale nitrogen assessments. *Biogeosciences*. 9: 4023-4035.

CMAS. 2016. Community Modeling and Analysis System. https://www.cmascenter.org/download.cfm

CSU. 2016. Agriculture – Weld County. Colorado State University, Cooperative Extension, Fort Collins, CO. http://coopext.colostate.edu/WELD/agriculture.htm

Day, E.E. X. Chen, K.A. Gebhart, C.M. Carrico, F.M. Schwandner, K.B. Benedict, B.A. Schichtel, J.L. Collett Jr.. 2012. Spatial and temporal variability of ammonia and other inorganic aerosol species. *Atmospheric Environment*, 61: 490-498, doi.org/10.1016/j.atmosenv.2012.06.045.

Gilliland, A.B., K.W. Appel, R.W. Pinder, and R.L. Dennis. 2006. Seasonal NH_3 emissions for the continental united states: Inverse model estimation and evaluation. *Atmospheric Environment*. 40: 4986–4998.

Herb, J., A.B. Nadykto, and F. Yu. 2011. Large ternary hydrogen-bonded pre-nucleation clusters in the Earth's atmosphere. *Chemical Physics Letters*. 518: 7–14.

Holmes, N.S. 2007. A review of particle formation events and growth in the atmosphere in the various environments and discussion of mechanistic implications. Atmospheric Environment. 41: 2183–2201.

Intergovernmental Panel on Climate Change (IPCC). 2014. Climate Change 2013: The Physical Science Basis, Intergovernmental Panel on Climate Change. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jones, L., M.S. Nizam, B. Reynolds, S. Bareham, and E.R.B. Oxley. 2013. Upwind impacts of ammonia from an intensive poultry unit. *Environmental Pollution*. 180: 221-228.

Kelly, J.T., K.R. Baker, J.B. Nowak, J.G. Murphy, M.Z. Markovic, T.C. VandenBoer, R.A. Ellis, J.A. Neuman, R.J. Weber, J.M. Roberts, P.R. Veres, J.A. de Gouw, M.R. Beaver, S. Newman, and C. Misenis. 2014. Fine-scale simulation of ammonium and nitrate over the South Coast Air Basin and San Joaquin Valley of California during CalNex-2010. *Journal of Geophysical Research Atmospheres*. 119: 3600–3614. DOI:10.1002/2013JD021290.

Kwok, R.H.F., Napelenok, S.L., and Baker, K.R. 2013. Implementation and evaluation of $PM_{2.5}$ source contribution analysis in a photochemical model. *Atmospheric Environment*. 80: 398–407.

Lin, L.I-K. 1989. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 45:255-268.

Lin, L.I-K. 1992. Assay validation using the concordance correlation coefficient. *Biometrics*. 48:599-604.

Mathur, R., J. Pleim, K. Schere, G. Pouliot, J. Young, and T. Otte. 2005. The community Multiscale Air Quality (CMAQ) model: Model configuration 475 and enhancements for 2006 air quality forecasting. Air Quality Forecaster Focus Group Meeting, September 16, 2005, Washington, DC.

National Atmospheric Deposition Program (NADP). 2014. Ambient Ammonia Monitoring Network (AMoN). http://nadp.sws.uiuc.edu/AMoN/AMoNfactsheet.pdf

Orsini, D.A., Y. Ma, A. Sullivan, B. Sierau, K. Baumann, R.J. Weber. 2003. Refinements to the particle-into-liquid sampler (PILS) for ground and airborne measurements of water soluble aerosol composition. *Atmospheric Environment* 37: 1243–1259.

Paerl, H.W. 1988. Nuisance Phytoplankton Blooms in Coastal, Estuarine, and Inland Waters, *Limnology and Oceanography.* 33: 823-847.

Pay, M.T., P.Jiminez-Guerrero, and J.M. Baldasano. 2012. Assessing sensitivity regimes of secondary inorganic aerosol formation in Europe with the CALIOPE-EU modeling system. *Atmospheric Environment*. 51: 146-164.

Pope III, C., M. Ezzati, and D.W. Dockery. 2009. Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine.* 360 (4): 376–386.

Shephard, M.W., K. Cady-Pereira, M. Luo, D.K. Henze, R.W. Pinder, J.T. Walker, C.P. Rinsland, J.O. Bash, L. Zhu, V. H. Payne, and L. Clarisse. 2011. TES ammonia retrieval strategy and global observations of the spatial and seasonal variability of ammonia. *Atmospheric Chemistry and Physics*, 11(20), 10743.

Sigma Aldrich. www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Supelco/Application_Notes/radiello_i1_i2.pdf

Sun, K., K. Cady-Pereira, D.J. Miller, L. Tao, M.A. Zondlo, J.B. Nowak, J.A. Neuman, T. Mikoviny, M. Müller, A. Wisthaler, A.J. Scarino, and C.A. Hostetler. 2015. Validation of TES ammonia observations at the single pixel scale in the San Joaquin Valley during DISCOVER-AQ. *Journal of Geophysical Research: Atmospheres*. 120 (10): 5140-5154. DOI:10.1002/2014JD022846.

Tang, Y., T. Chai, L. Pan, P. Lee, D. Tong, H. Kim, and W. Chen. 2015. Using optimal interpolation to assimilate surface measurements and satellite AOD for ozone and $PM_{2.5}$: A case study for July 2011. *Journal of the Air & Waste Management Association.* 65 (10): 1206-1216. DOI: 10.1080/10962247.2015.1062439.

USDA. 2014. 2012 Census of Agriculture. U.S. Department of Agriculture, National Agricultural Statistical Service, Washington, DC.

U.S. Environmental Protection Agency Science Advisory Board. 2007. Hypoxia in the Northern Gulf of Mexico: An Update by the EPA Science Advisory Board. EPA-SAB-08-003, U.S. Environmental Protection Agency Science Advisory Board, Washington, DC.

U.S. Environmental Protection Agency. 2009. 2005 National Emissions Inventory Data & Documentation. www.epa.gov/ttn/chief/net/2005inventory.html

U.S. Environmental Protection Agency. 2015. 2011 National Emissions Inventory Data & Documentation. www.epa.gov/ttn/chief/net/2011inventory.html

Vukovich, J., and T. Pierce. 2002. The Implementation of BEIS-3 within the SMOKE modeling framework. Presented at "Emissions Inventories—Partnering for the Future," the 11th Emissions Inventory Conference of the U.S. Environmental Protection Agency, Research Triangle Park, NC.

Wang, X., W. Wang, L. Yang, X. Gao, W. Nie, Y. Yu, P. Xu, Y. Zhou, and Z. Wang. 2012. Secondary formation of inorganic aerosols in the droplet mode through heterogeneous aqueous reactions under haze conditions. *Atmospheric Environment.* 63: 68–76.

Zhu, L., D.K. Henze, K.E. Cady-Pereira, M.W. Shephard, M. Luo, R.W. Pinder, J.O. Bash, and G.-R. Jeong. 2013. Constraining U.S> ammonia emissions using TES remote sensing observations and the GEOS-Chem adjoint model. *Journal of Geophysical Research: Atmospheres*. 118 (8): 3355-3368. DOI: 10.1002/jdrd.50166.

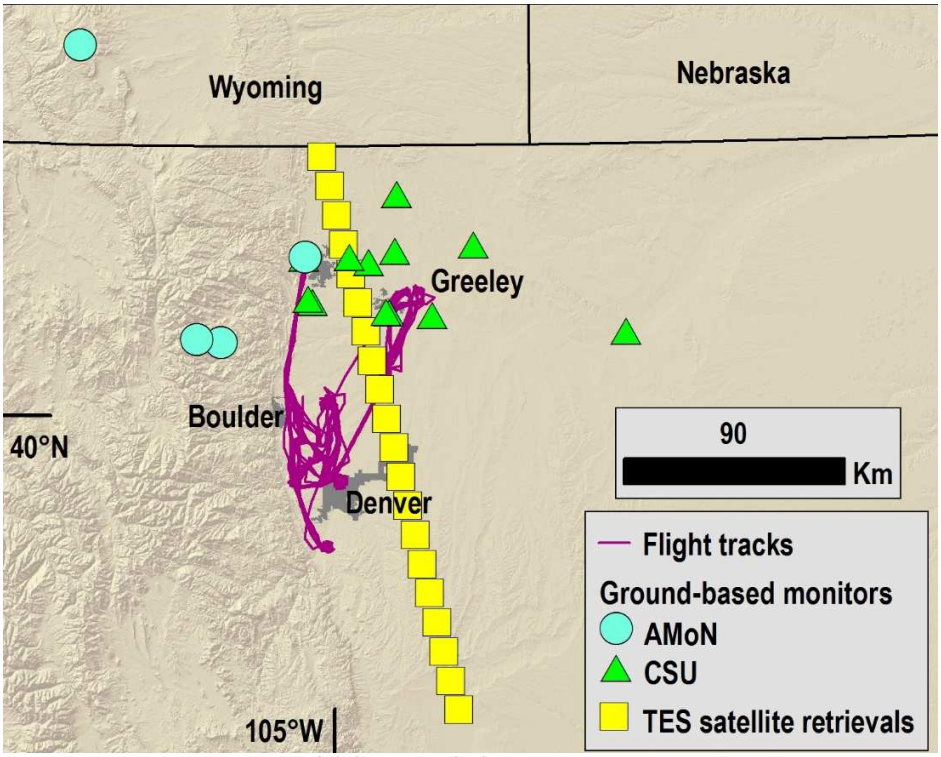

Figure 1. Domain of the DISCOVER-AQ Colorado measurement campaign, showing flight paths for low level in situ aircraft measurements (<1km AGL), locations of ground level monitors, and the path for TES satellite measurements. [1.5 column image]

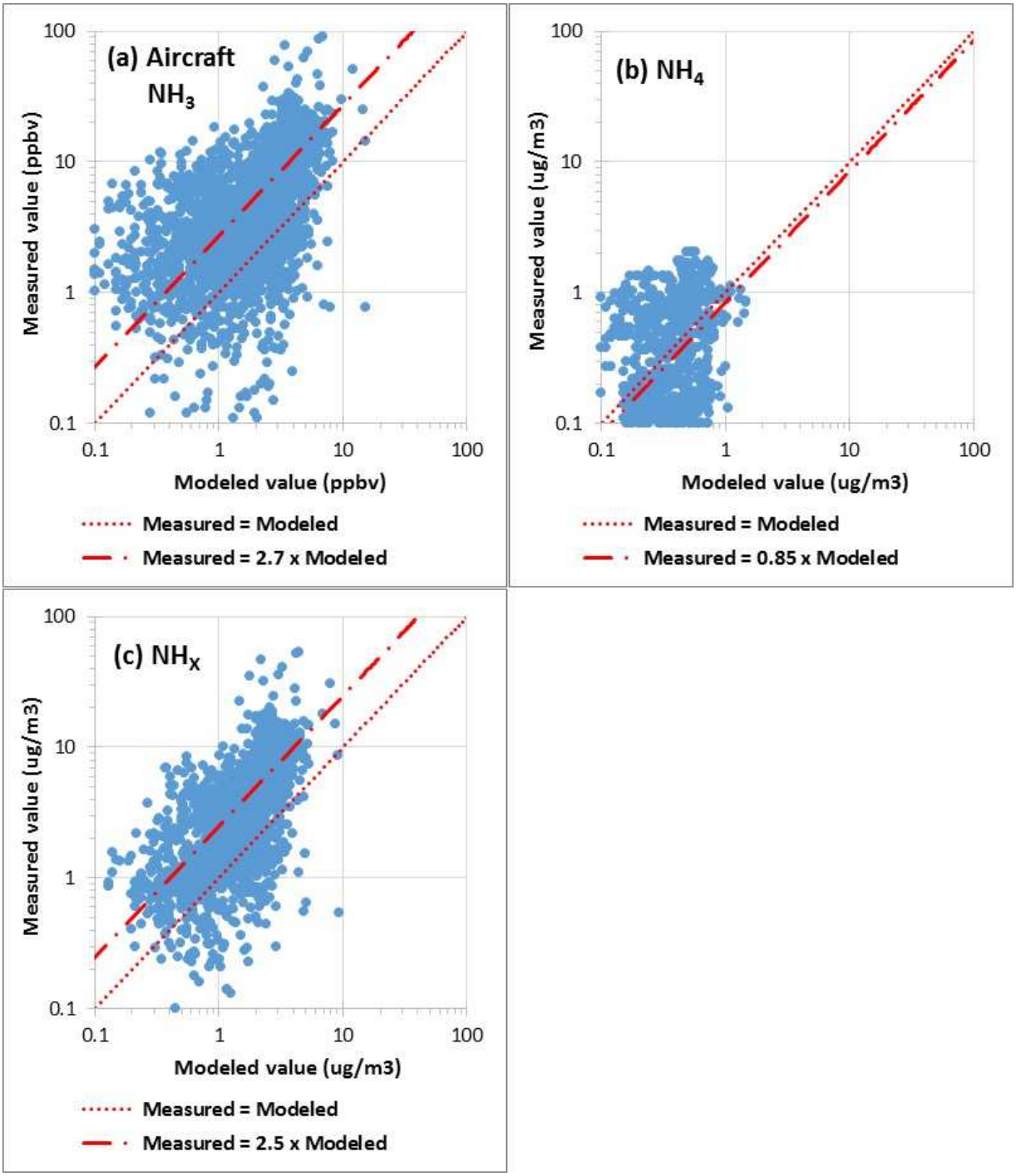

Figure 2. Aircraft in situ measurements of NH_3 (a), NH_4^+ (b), and NH_X (c), plotted against model predictions. Each measurement is plotted as a point. Dotted lines show a 1:1 slope, where points would have fallen if the measurements and model predictions were in complete agreement (measured = modeled). Dashed lines show the 1:1 slope displaced by the NMB.

[2-column image]

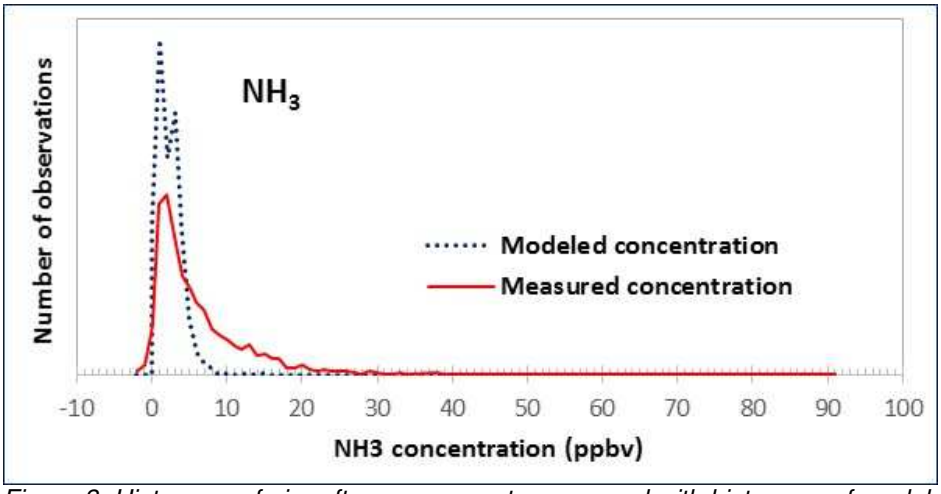

Figure 3. Histogram of aircraft measurements compared with histogram of model predictions at the corresponding times and locations.
[1.5 column image]

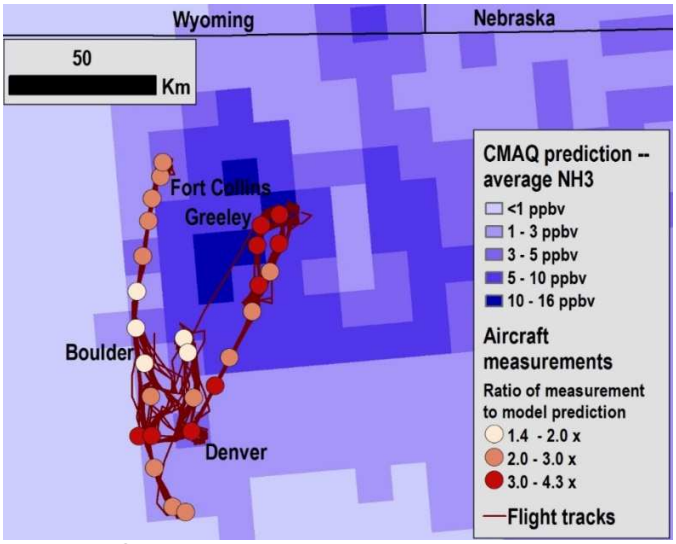

Figure 4. Spatial variation of model prediction error from in situ aircraft measurements. [1-column image]

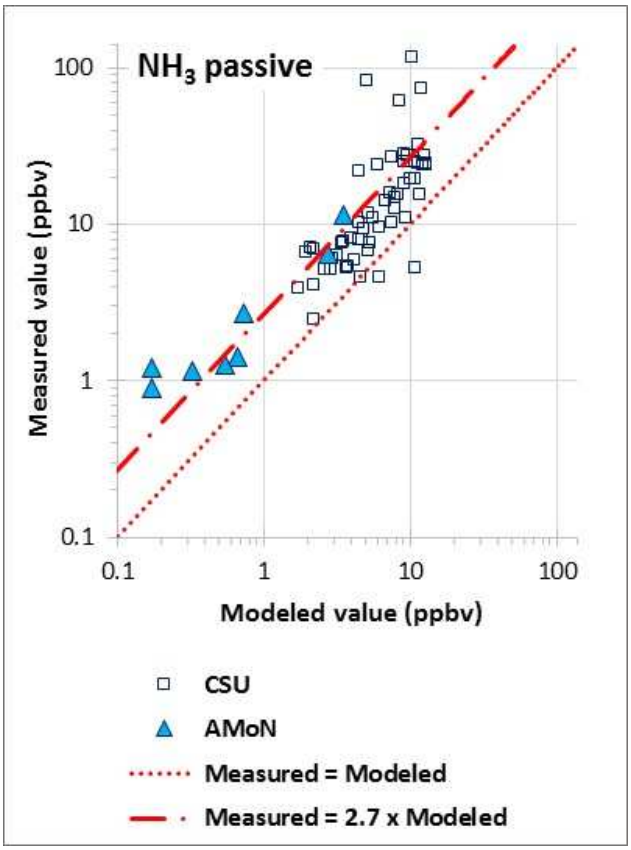

Figure 5. Ground-level measurements of NH_3 plotted against model predictions. Dotted line shows where the measured points should have fallen if the model predictions were exactly correct (measured = modeled). Dashed line shows the actual measured trend lines based on the ratio of the average measured concentration to the average model prediction. [1-column image]

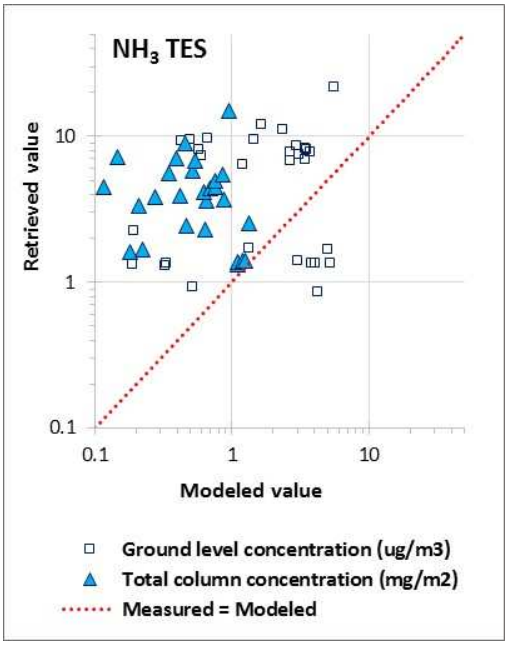

Figure 6. TES NH_3 retrievals plotted against model predictions. Each measurement is plotted as a point. Dotted line shows where the measured points should have fallen if the model predictions were exactly correct (measured = modeled).
[1-column image]

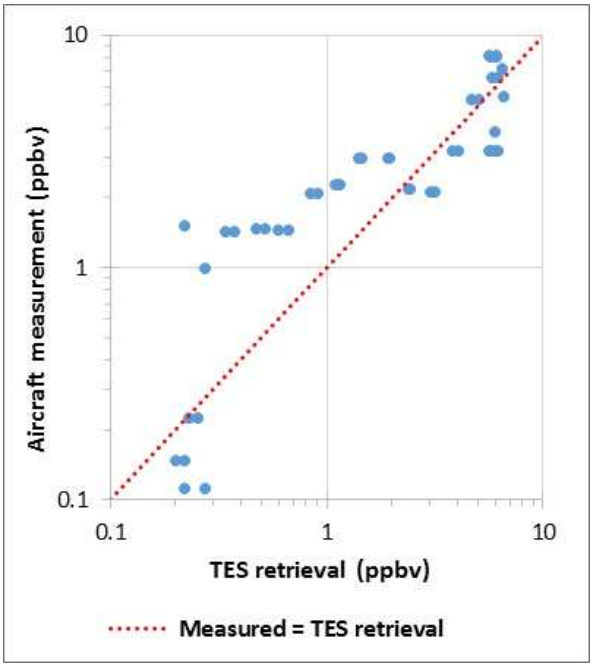

Figure 7. Aircraft in situ measurements of NH_3 plotted against TES satellite retrievals. Each measurement is plotted as a point. Dotted line shows where the measured points should have fallen if the satellite retrievals were exactly correct (measured = TES retrieval).
[1-column image]

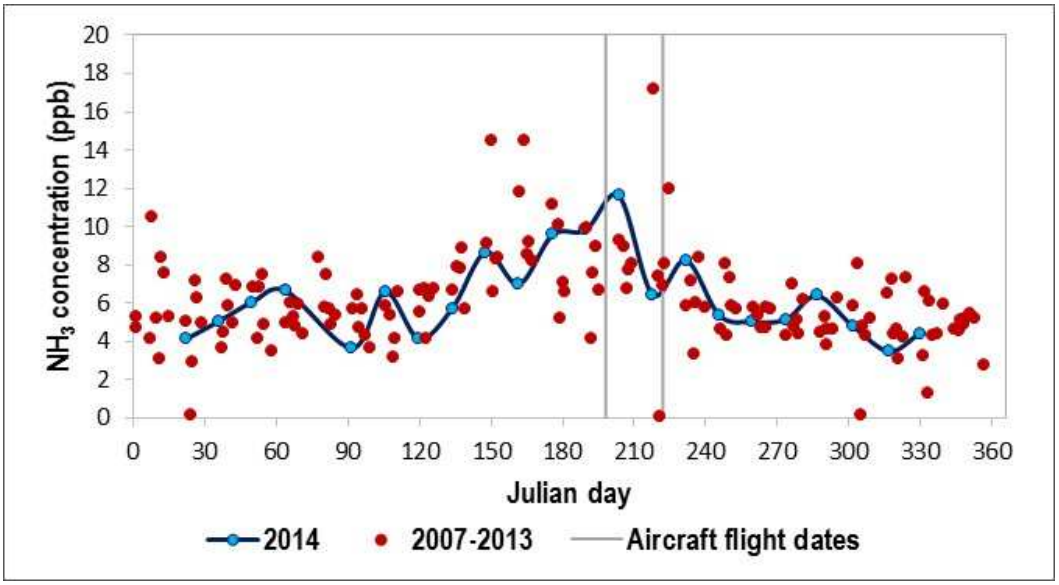

Figure 8. Seasonal pattern of NH_3 vapor at the Fort Collins AMoN site in 2014 compared with NH_3 vapor in previous years.
[2-column image]

Table 1. Comparison of *in situ* aircraft measurements with model predictions for NH_3, NH_4^+, and NH_X.

	NH_3 (ppbv)	NH_3 (μg/m^3)	NH_4 (μg/m^3)	NH_X (μg/m^3)
Measured concentrations				
Average	6.1	3.9	0.29	4.2
Standard deviation	6.9	4.2	0.38	4.6
Maximum	90.0	53.1	2.05	53.3
Model predictions				
Average	2.2	1.4	0.34	1.7
Standard deviation	1.6	1.0	0.20	1.1
Maximum	15.3	9.1	1.46	9.2
Comparison statistics				
Normalized mean bias		-63%	18%	-60%
Ratio of average measured value to average modeled value		2.7	0.85	2.5
Correlation coefficient (r)		0.52	0.37	0.54
Concordance correl. coeff. (ρ_c)		0.16	0.29	0.17
Number of observations		2,372	1,700	1,637

681

682

Table 2. Comparison of ground-based measurements with model predictions for NH_3.

	AMoN monitors		CSU monitors		All passive monitors	
	ppbv	µg/m³	ppbv	µg/m³	ppbv	µg/m³
Measured concentrations						
Average	3.3	2.0	17.8	10.5	16.0	9.5
Standard deviation	3.8	2.3	20.5	12.1	19.8	11.7
Maximum	11.6	6.8	116.4	68.7	116.3	68.7
Model predictions						
Average	1.1	0.7	6.7	3.9	6.0	3.5
Standard deviation	1.3	0.8	3.4	2.0	3.7	2.2
Maximum	3.5	2.1	12.8	7.7	12.8	7.6
Comparison statistics						
Normalized mean bias	-67%		-63%		-63%	
Ratio of average measured value to average modeled value	3.0		2.7		2.7	
Correlation coefficient (r)	0.97		0.47		0.52	
Concordance correl. coeff. (ρ_c)	0.45		0.10		0.14	
Number of observations	8		58		66	

683

684

Table 3. Comparison of TES retrievals with model predictions for NH_3.

	Total atmospheric column loading (mg/m^2)	Concentration in the lowest atmospheric layer		Concentration at the regional averaging kernel peak (ppbv)
		ppbv	$\mu g/m^3$	
Measured concentrations				
Average	2.0	3.0	1.8	0.83
Standard deviation	2.9	4.5	2.7	1.1
Maximum	14.9	21.8	12.9	4.5
Model predictions				
Average	0.5	2.0	1.2	0.39
Standard deviation	0.4	2.1	1.3	0.54
Maximum	1.5	9.2	5.4	2.6
Comparison statistics				
Normalized mean bias	-76%	-33%		-53%
Ratio of average measured value to average modeled value	4.2	1.5		2.1
Correlation coefficient (r)	0.11	0.09		0.52
Concordance correl. coeff. (ρ_c)	0.02	0.07		0.39
Number of observations	65	65		65

685

686

Table 4. Comparison of *in situ* aircraft measurements with TES retrievals for NH_3

	NH_3 (ppbv)
In situ aircraft measurements	
Average	2.9
Standard deviation	2.4
Maximum	8.1
TES retrievals	
Average	2.8
Standard deviation	2.5
Maximum	6.6
Comparison statistics	
Normalized mean bias of TES retrieval	-1%
Ratio of average measured value to average TES retrieval	1.01
Correlation coefficient (r)	0.78
Concordance correl. coeff. (ρ_c)	0.78
Number of observations	46

687