Final Report

Model Evaluation for Ammonia (NH_3) using Satellite Data

TCEQ Contract No. 582-23-45974
Work Order No. 4
Revision 1.0

Deliverable 5.2

Prepared by:

Archana Dayalu, Sara Vannah, Matthew Alvarado
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02421

Prepared for:

Sushil Gautam
Texas Commission on Environmental Quality
Air Quality Division
Bldg. E, Room 341S
Austin, Texas 78711-3087

June 30th, 2025

Document Change Record

Revision	Revision Date	Remarks
0.1	June 23, 2025	Internal Version for Review
1.0	June 30, 2025	Version Submitted to TCEQ

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

A – Averaging Kernel, Observation Operator
AER – Atmospheric and Environmental Research
AMoN – Ambient Ammonia Monitoring Network
AWS – Amazon Web Services
CONUS – Contiguous United States
CrIS – Cross-Track Infrared Sounder
CAMx – Comprehensive Air Quality Model with Extensions
DV – Design Value
ECCC – Environment Climate Change Canada
N – Nitrogen
NAAQS – National Ambient Air Quality Standard
NH_3 – Ammonia
$PM_{2.5}$ – Fine Particulate Matter of 2.5 microns or less
SIP – State Implementation Plan
TES – Tropospheric Emissions Spectrometer
X_{CAMx} – CAMx Ammonia Vertical Profile
X_{CAMx_xa} – CAMx Ammonia Vertical Profile with CrIS Observation Operator applied
USDA – United States Department of Agriculture

Executive Summary

Emissions of ammonia (NH_3) have become an increasing focus for air quality managers, as reducing these emissions has been proposed as a cost-effective way to reduce fine particulate matter ($PM_{2.5}$) air pollution (e.g., Liu et al., 2022). However, current emissions of NH_3 are highly uncertain (e.g., Chen et al., 2021) with errors in bottom-up inventories making it difficult to make optimal decisions on $PM_{2.5}$ control strategies.

Research has shown that current polar-orbiting satellite retrievals of NH_3 – particularly NH_3 retrievals from the Cross-Track Infrared Sounder (CrIS) – can provide qualitative and quantitative information on the geographic distribution and seasonal cycle of NH_3 emissions over the globe (e.g., Cady-Pereira et al., 2017). Due to its peak sensitivity in the boundary layer (~850-750 hPa) and, under ideal conditions, near the surface (Shepard & Cady-Pereira, 2015), CrIS data has been used in multiple air quality model evaluations and NH_3 emission inventory optimizations (e.g., Toro et al., 2024). The purpose of this project was to use CrIS data to evaluate and identify potential areas discrepancies in the TCEQ modeling platform NH_3 emissions inventory. The evaluation provides important observational information to the TCEQ modeling platform which is used for $PM_{2.5}$ modeling to support State Implementation Plan (SIP) requirements for the 2024 Annual $PM_{2.5}$ National Ambient Air Quality Standards (NAAQS).

There were two major tasks to accomplish this project's objectives. As TCEQ uses the Comprehensive Air Quality Model with Extensions (CAMx) to translate the underlying bottom-up emissions to 3D NH_3 concentrations, the first step was to develop software to directly compare CAMx output with CrIS observations. Specifically, the software was designed to (1) read in CrIS retrieval data and relevant CAMx output; (2) spatially and temporally match CrIS NH_3 columns with CAMx modeled NH_3 columns; (3) directly compare model and observations by applying the CrIS observation operator to matched CAMx NH_3 columns; and (4) calculate monthly averaged model-observation mismatch for detailed analysis for a selection of months in 2019. The second task was to conduct a detailed analysis of monthly average results; exploring discrepancies between CAMx and CrIS NH_3 estimates; compare with additional available ground-based measurements; and investigate correlations of NH_3 with $PM_{2.5}$ on days important to the annual $PM_{2.5}$ design value.

Overall, the results suggest the underlying TCEQ emissions inventory is systematically overestimating the early spring "before planting" fertilizer application in wheat, corn, and sorghum regions. In addition, the mid-to-late Spring fertilizer application and/or livestock emissions are generally somewhat underestimated by the underlying TCEQ emissions inventory. Limited ground-based data near Houston suggests that, while both CAMx and CrIS overestimate surface NH_3 in the vicinity of Houston, CrIS generally shows better agreement. Finally, an assessment of the Spring correlation between surface NH_3 and $PM_{2.5}$ in the Houston region suggests only a weak correlation in May. It is possible there are stronger correlations on days important to the Houston design values warranting additional exploration across more time periods. In addition, the importance of NH_3 to $PM_{2.5}$ design values warrants exploration in other non-attainment regions such as Dallas Fort Worth, Austin, and San Antonio.

All requirements established in the Quality Assurance Project Plan (QA Category III for Research Model Development or Application) for this Work Order were met and passed, satisfying the requirement to audit 10% of the data produced for quality.

1. Introduction

1.1. Project Objectives

The purpose of this work order is to develop a software that uses Cross-track Infrared Sounder (CrIS) retrievals of satellite data to improve estimation of ammonia (NH_3) emissions within TCEQ's modeling platform. TCEQ's modeling platform uses the Comprehensive Air Quality Model with Extensions (CAMx) to output 3D fields of NH_3. The comparison between CrIS satellite NH_3 retrievals and CAMx modeled NH_3 enable evaluation of the underlying bottom-up emissions inventory used by TCEQ for photochemical modeling.

The objectives of this project are thus to develop software to (1) read in CrIS retrieval data and relevant CAMx output; (2) spatially and temporally match CrIS NH_3 columns with CAMx modeled NH_3 columns; (3) directly compare model and observations by applying the CrIS observation operator to matched CAMx NH_3 columns; and (4) calculate monthly averaged model-observation mismatch for detailed analysis for a selection of months in 2019.

The Schedule of Deliverables for this project is given in Table 1.

Table 1. Schedule of Deliverables for TCEQ Work Order #4

Milestones	Planned Date
Task 1 - Work Plan	
1.1: TCEQ-approved Work Plan	December 19, 2024
1.2: TCEQ-approved QAPP	December 19, 2024
Task 2 – Progress Reports	
2.1: Monthly Progress Reports	Monthly with invoice
Task 3 – Apply CrIS observation operators to TCEQ CAMx Data	
3.1: Software to create monthly averages of the TCEQ CAMx data sampled to match CrIS overpasses and with CrIS observation operators applied.	April 30, 2025
Task 4 – Analyze TCEQ CAMx NH_3 Data with CrIS	
4.1: Summary report of analysis and findings	June 10, 2025
Task 5 – Draft and Final Reports	
5.1: Draft Report	June 10, 2025
5.2: Final Report	June 30, 2025
5.3: Data and Software	June 30, 2025

1.2. Background

Ammonia (NH_3) contributes to the formation of fine particles ($PM_{2.5}$) by reaction with nitric and sulfuric acid (HNO_3 and H_2SO_4), which are in turn formed by the photochemical oxidation of nitrogen oxides (NO_x = NO + NO_2) and sulfur dioxide (SO_2), respectively. Emissions of NH_3 have become an increasing focus for air quality managers, as reducing these emissions has been proposed as a cost-effective way to reduce $PM_{2.5}$ (e.g., Liu et al., 2022). However, current emissions of NH_3 are highly uncertain (e.g., Chen et al., 2021) with errors in bottom-up inventories making it difficult to make optimal decisions on $PM_{2.5}$ control strategies.

Research has shown that current polar-orbiting satellite retrievals of NH_3 can provide qualitative and quantitative information on the geographic distribution and seasonal cycle of NH_3 emissions over the globe (e.g., Cady-Pereira et al., 2017). The development of the CrIS NH_3 retrieval (Shephard and Cady-Pereira, 2015) has provided much greater spatial coverage while using essentially the same retrieval approach as was tested with Tropospheric Emission Spectrometer (TES). The CrIS retrievals have been extensively validated (e.g., Cady-Pereira et al., 2024). Due to its peak sensitivity in the boundary layer (~850-750 hPa) and, under ideal conditions, near the surface (Shepard & Cady-Pereira, 2015), CrIS data has been used in multiple air quality model evaluations and emission inventory optimizations (e.g., Toro et al., 2024). Therefore, this project uses CrIS data to evaluate the NH_3 emissions used in TCEQ's modeling platform. This evaluation will provide important observational information to the TCEQ modeling platform as it is used to conduct $PM_{2.5}$ modeling to support State Implementation Plan (SIP) requirements for the 2024 Annual $PM_{2.5}$ National Ambient Air Quality Standards (NAAQS).

However, a fair comparison of modeled NH_3 concentrations with those estimated from the CrIS satellite retrievals first requires spatial and temporal matching of modeled locations with satellite observed locations; and next, a procedure known as "applying the satellite observation operator" to the modeled quantity. As noted in Cady-Pereira et al. (2024), this enables direct comparability, ensuring that the modeled surfaces (columns) and the true surfaces (columns) are both being evaluated through the same lens – in this case, the CrIS instrument. Equation (1) below summarizes the process of "applying the observation operator" to a quantity of interest that is being compared with satellite data.

$$X_o = X_{aprior} + A(X - X_{aprior}) \qquad (1)$$

In Equation (1), X_o is a modeled vertical profile with the satellite observation operator (i.e., averaging kernel "A") applied; X_{aprior} is the initial guess of used for a final satellite retrieval "X_{retv}"; A is the averaging kernel; and X is the modeled quantity of interest interpolated to the satellite *x-y-z-t* location.

Equation (1) then enables direct comparison of a modeled quantity X_o with corresponding satellite retrievals "X_{retv}". Differences between the modeled "retrieval" X_o and a true observation retrieval "X_{retv}" can then inform subsequent emissions inventory optimizations.

1.3. Deliverable

This Final Report highlights major activities and key findings, provides pertinent analysis, describes encountered problems and associated corrective actions, and details relevant statistics including data, parameter, or model completeness, accuracy and precision. It satisfies Deliverable 5.2 of the Work Plan for Work Order No. 4 under TCEQ Contract 582-23-45974:

Deliverable 5.2: Final Report
Deliverable 5.2 Due Date: June 30, 2025

2. Task 3: Apply CrIS observation operators to TCEQ CAMx Data

In this task, AER developed software that (1) reads in CrIS retrieval data and Comprehensive Air Quality Model with Extensions (CAMx) output; (2) matches modeled NH_3 columns to the corresponding CrIS NH_3 column; and (3) applies the CrIS observation operator to the interpolated CAMx NH_3 column. In addition, diagnostic plots were created that enabled assessment of CAMx data before and after the CrIS observation operators were applied. Additional software was developed to calculate monthly averages of the model-observation mismatch for both NH_3 total column and surface concentrations. These monthly-average output files were used in the analysis in Task 4.

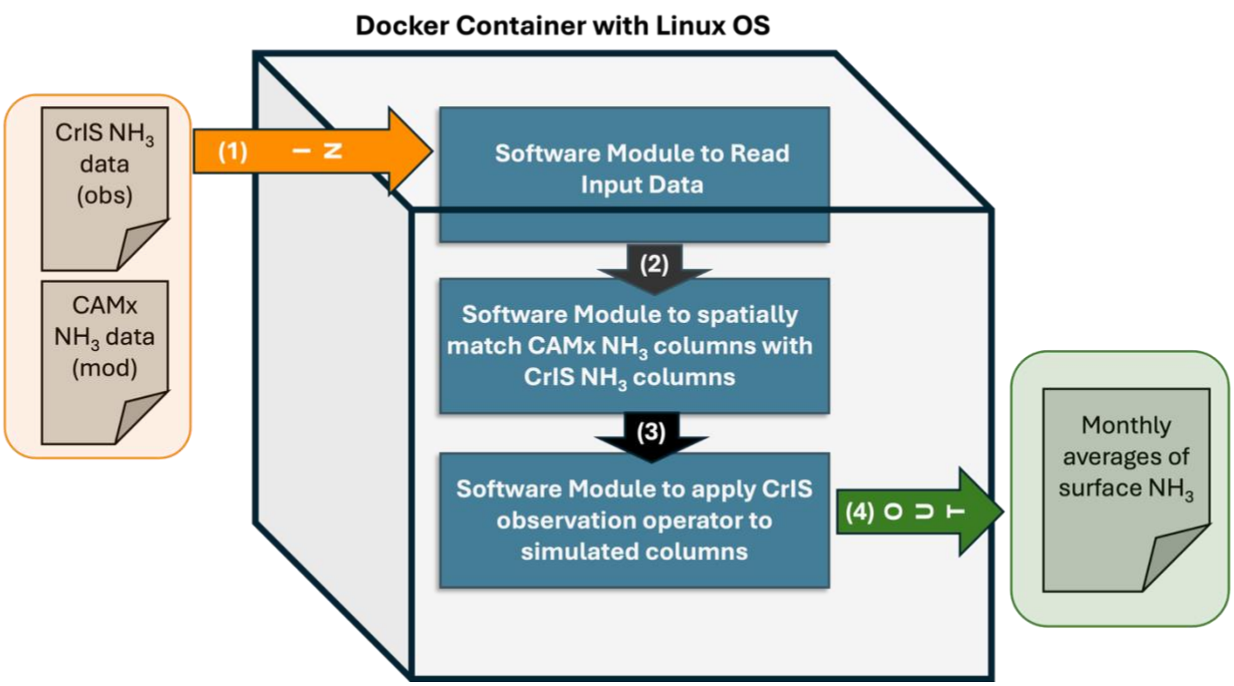


Figure 1. Overview of System Design.

The monthly averaging process and the software usage are summarized in Figure 1 and detailed in Appendix A. The software developed in Task 3 was delivered to the TCEQ on May 1, 2025, as part of Deliverable 3.1. Deliverable 3.1 included a User Guide (reproduced in Appendix A); python scripts (as both a standard and interactive jupyter notebook); and an AWS cloudfront link to a tar gzipped file that included the code, User's Guide, and the CrIS and CAMx data for the sample month of April 2019. The software was used to calculate monthly averages for April through October 2019 as these were the most

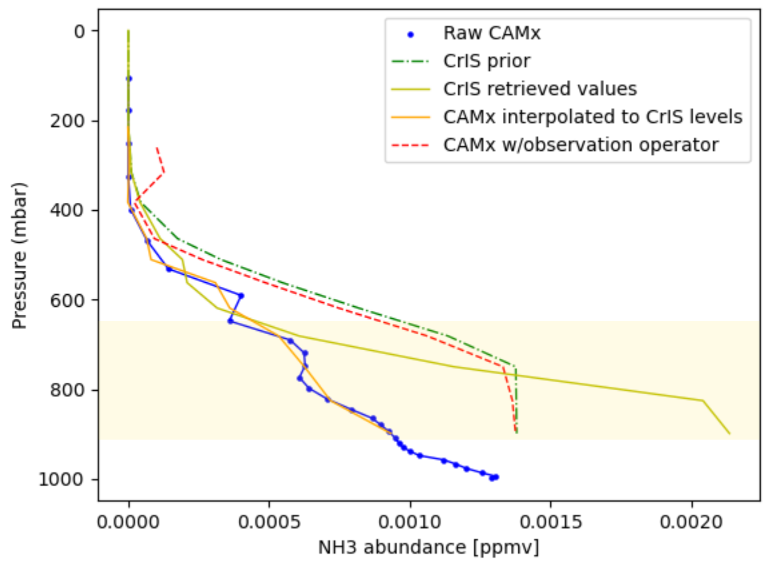


Figure 2. Sample output for 17 April 2019 at row 0 column 17 of intermediate processed data. Highlighted yellow region indicates theoretical levels of peak sensitivity of CrIS instrument between ~750-900 mbar (hPa). Note that interpolation truncates at lowest CrIS pressure level.

recent time periods available with 3D CAMx NH_3 fields. However, given the necessary CAMx and CrIS input files (Appendix A), the software can be readily used to output additional months/years.

Equation (2) below replaces the general variables in Equation (1) with variables specific for this project. Here, 3D modeled NH_3 fields from CAMx are being compared with CrIS NH_3 retrievals.

$$X_{CAMx\_xa} = X_{aprior} + A(X_{CAM_X} - X_{aprior}) \qquad (2)$$

In Equation (2), X_{CAMx_xa} is the CAMx modeled vertical profile with the CrIS observation operator (i.e., "averaging kernel") applied; X_{aprior} is the initial guess of NH_3 from the CrIS retrieval; A is the averaging kernel; and X_{CAMx} is the CAMx modeled column interpolated to the CrIS *x-y-z-t* location. At levels where CrIS has highest sensitivity (boundary layer and/or near-surface), A is sufficiently non-zero such that X_{CAMx_xa} retains some amount of information from the interpolated X_{CAMx}. Conversely, where A has least sensitivity (all other altitudes greater than the boundary layer), X_{CAMx_xa} will be the same as X_{aprior}. Figure 2 provides sample diagnostic output illustrating the processing in Equation (2).

All output files and processing were focused on the CAMx eastern Texas 4 km domain (Figure 3). The software outputs monthly averages of both CAMx and CrIS NH_3 as surface and total column concentrations. At each CAMx gridcell, the concentrations are reported if there is a collocated CrIS sounding with satisfactory quality (Appendix A). For ease of analysis, AER explored regridding the point-by-point results to averages over 10x, 15x, and 20x resolutions. Figure 3 displays the regridding results, using CAMx_xa surface concentrations for April 2019 as an example. In coordination with the TCEQ Project Manager, the ideal resolution for analysis was selected as a 10x regrid, compromising between the original 4 km CAMx domain and a useful signal for evaluation of the underlying TCEQ NH_3 emissions inventory. The monthly averages were used in the Task 4 analysis (Section 3). In the final evaluation, total column and/or surface NH_3 values are extracted from X_{CAMx_xa} and compared directly to corresponding CrIS retrievals. Differences between the CrIS retrieval of true surface emissions and the "retrieved" surface emissions simulated as CAMx_xa are used to inform future TCEQ NH_3

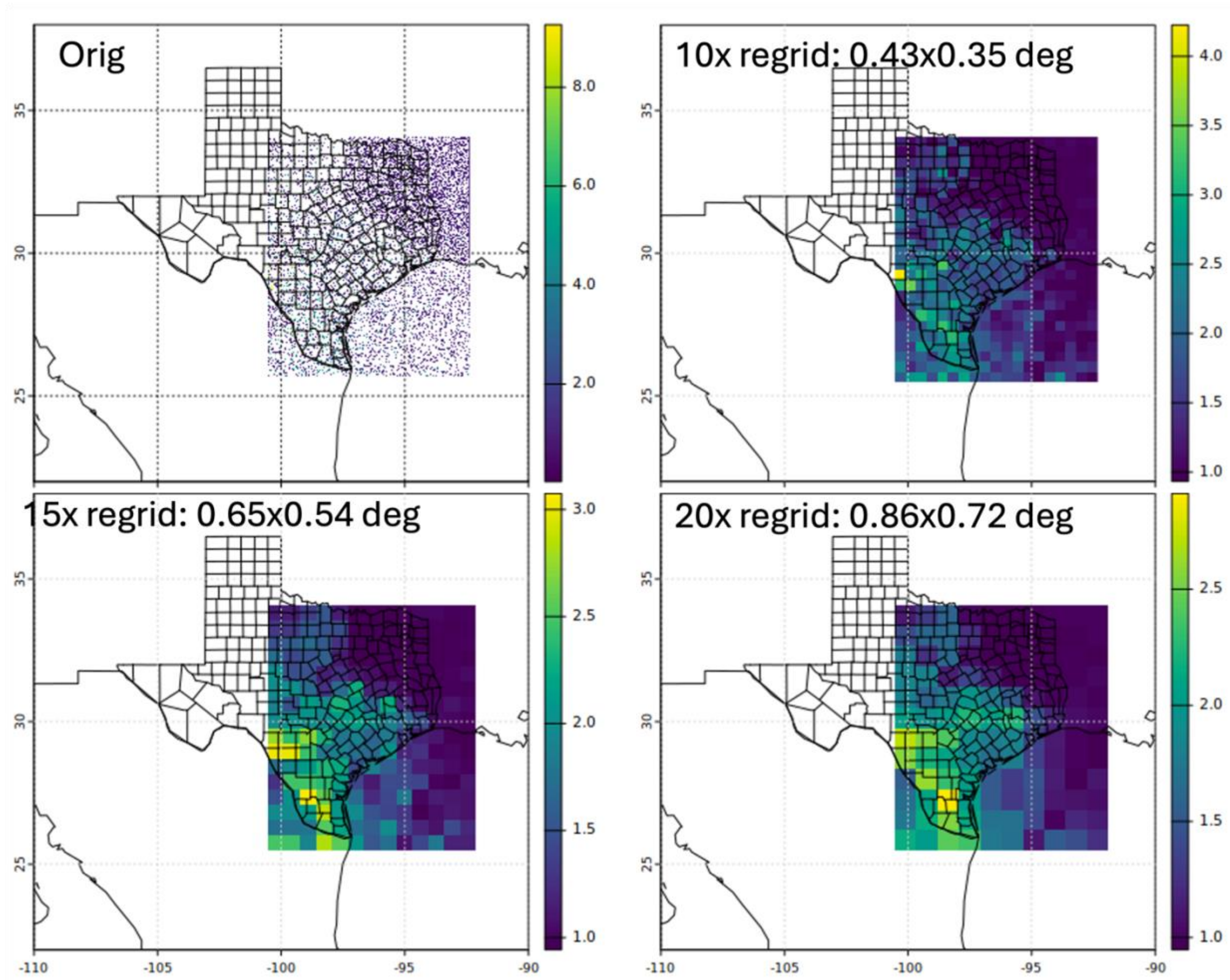


Figure 3. Sample CAMx_xa surface NH3 monthly averages (μg/m3) for April 2019 co-located and interpolated spatially and temporally to CrIS sounding locations. Focus region is CAMx 4 km Eastern TX modeling domain. Clockwise from top, the panels show (1) CAMx_xa point locations co-located with CrIS; and regridded by factors of (2) 10x, (3) 15x, and (4) 20x. Texas county borders are provided for reference.

emissions inventory optimizations. Comparison to independent ground-based Ammonia Observation Network (AMoN) NH_3 concentrations was also conducted to the extent possible.

3. Task 4: Analyze TCEQ CAMx NH_3 Data with CrIS

The output files generated by the software developed in Task 3 were used to evaluate the emissions used in the TCEQ CAMx modeling for April through October 2019. The focus of analysis was on monthly average NH_3 concentrations at CrIS overpass time (~13:30 Local Standard Time), as monthly averaging reduces the impact of clouds and missing data on the analysis.

In addition, day-to-day variability was examined for days important to the 2024 annual $PM_{2.5}$ design values and was limited to the Houston area (Figure 4). Houston was selected as it was the only non-attainment region in the modeling domain that was located near an Ambient Ammonia Monitoring Network (AMoN; NADP, 2022). The only other relevant AMoN site is located outside of the modeling domain in the Texas Panhandle, and was not included in this analysis. Note that CrIS retrievals have already been extensively validated with AMoN network data for multiple regions in the United States (e.g., Cady-Pereira et al., 2024).

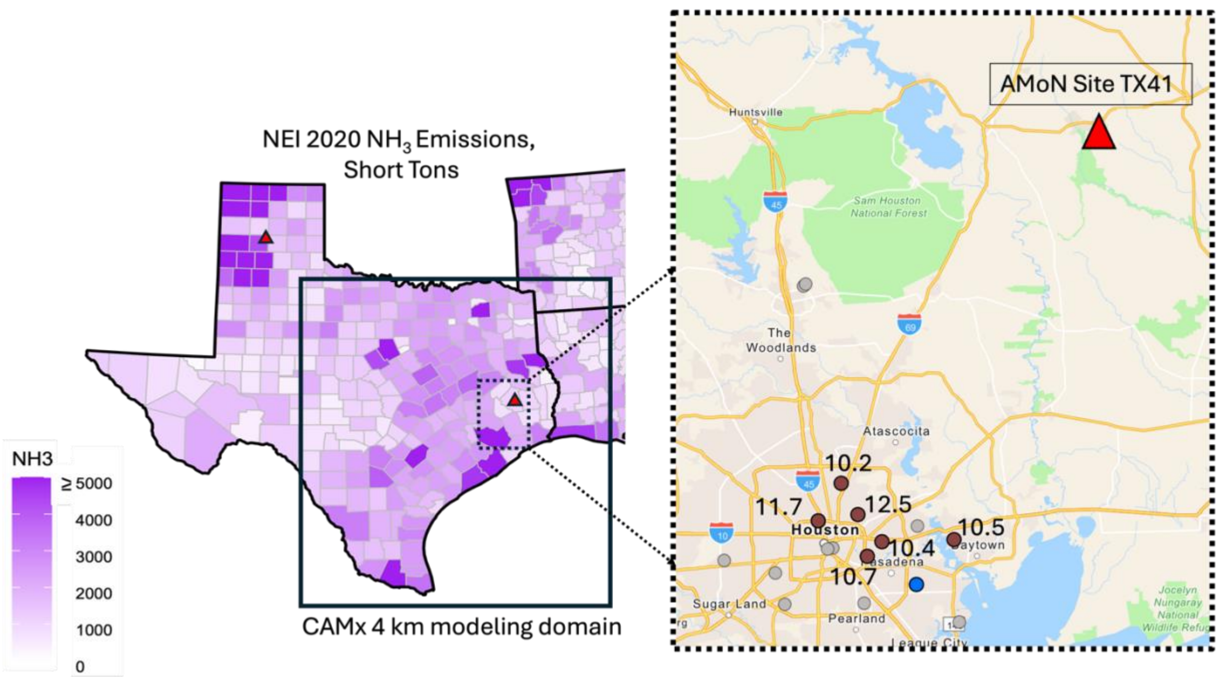


Figure 4. CAMx 4 km modeling domain (solid black rectangle) overlaid on NEI 2020 NH3 Emissions (Short Tons) map. Monthly averages are evaluated within CAMx modeling domain. Red triangles show AMoN network sites in Texas. Zoom region, used in AMoN cross-validation and daily variability analysis, shows Houston and area including AMoN site TX41 . Colored circles in zoom region show monitoring sites, with sites exceeding current annual 9µg/m^3 $PM_{2.5}$ standards colored in brown. $PM_{2.5}$ Design Values (µg/m^3) for non-attainment sites are noted. Blue circles are sites in attainment; grey circles are sites not valid for use annual attainment determinations. AMoN data from NADP, 2022. $PM_{2.5}$ DV data from EPA (2025).

3.1. Monthly Average Results

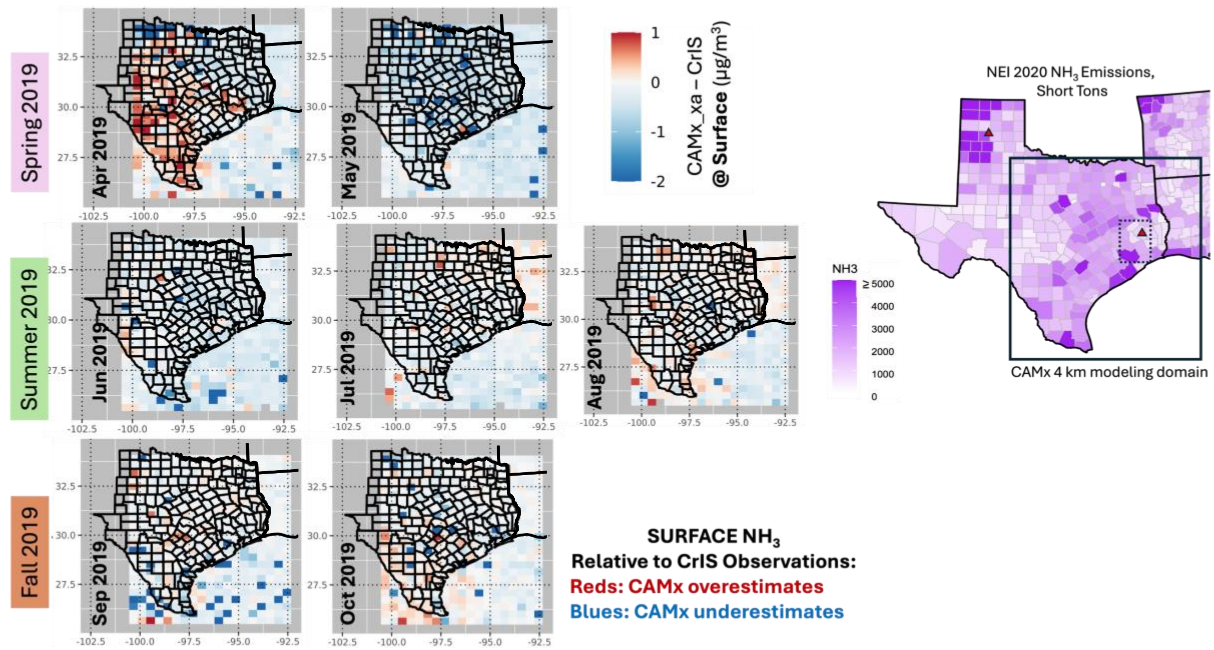


Figure 5. Monthly Average CAMx_xa-CrIS differences at the surface (μg/m³), grouped by season, in the CAMx 4 km modeling domain. For reference, the NEI 2020 NH3 emissions are also provided. Relative to CAMx collocated with CrIS points, regridding is 10x resolution (providing average values over 0.43x0.35 degree gridcell).

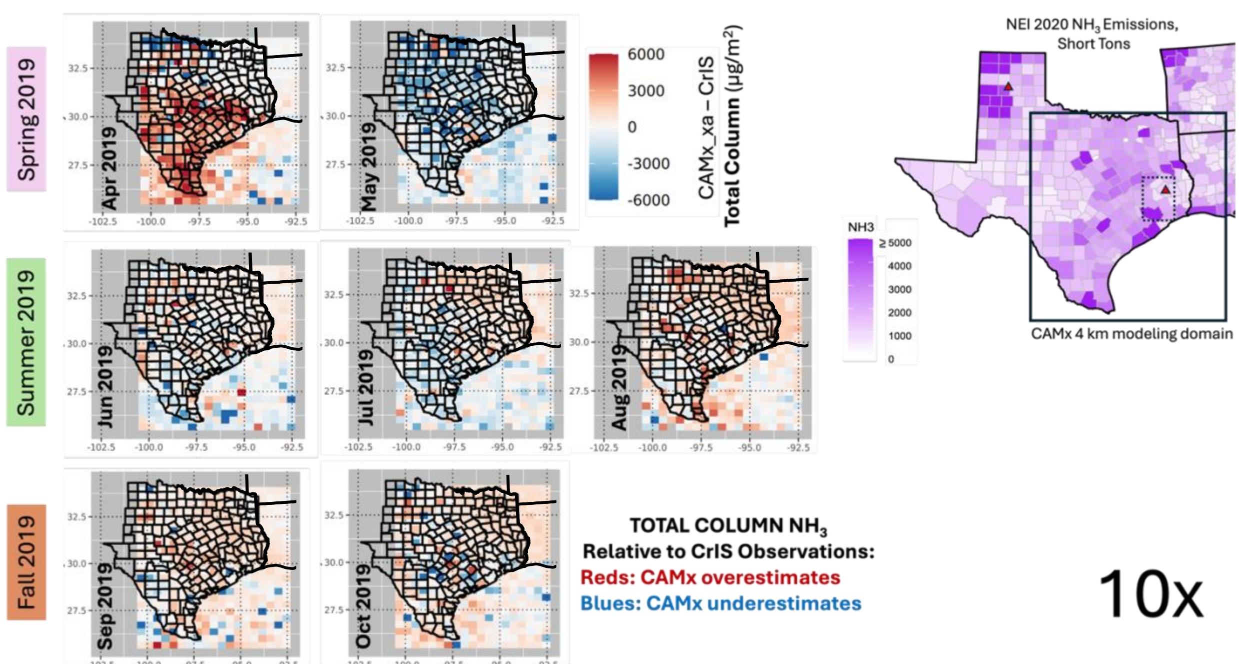


Figure 6. Same as Figure 5, but as monthly average CAMx_xa-CrIS differences as total columns (μg/m²).

Monthly averages of surface (μg/m3) and total column (μg/m2) NH_3 from CAMx_xa and CrIS were obtained for April through October 2019. The monthly average differences between CAMx_xa and CrIS were calculated at the surface (Figure 5) and for total columns (Figure 6). In the Spring, and April in particular, CAMx_xa can significantly differ from CrIS estimates in both surface and total column concentrations. There is generally better agreement in the Summer and Fall. More specifically, the underlying CAMx_xa emissions inventory appears to be *overestimating* NH_3 in *April* and *somewhat underestimating* NH_3 in *May* in both surface and total column concentrations.

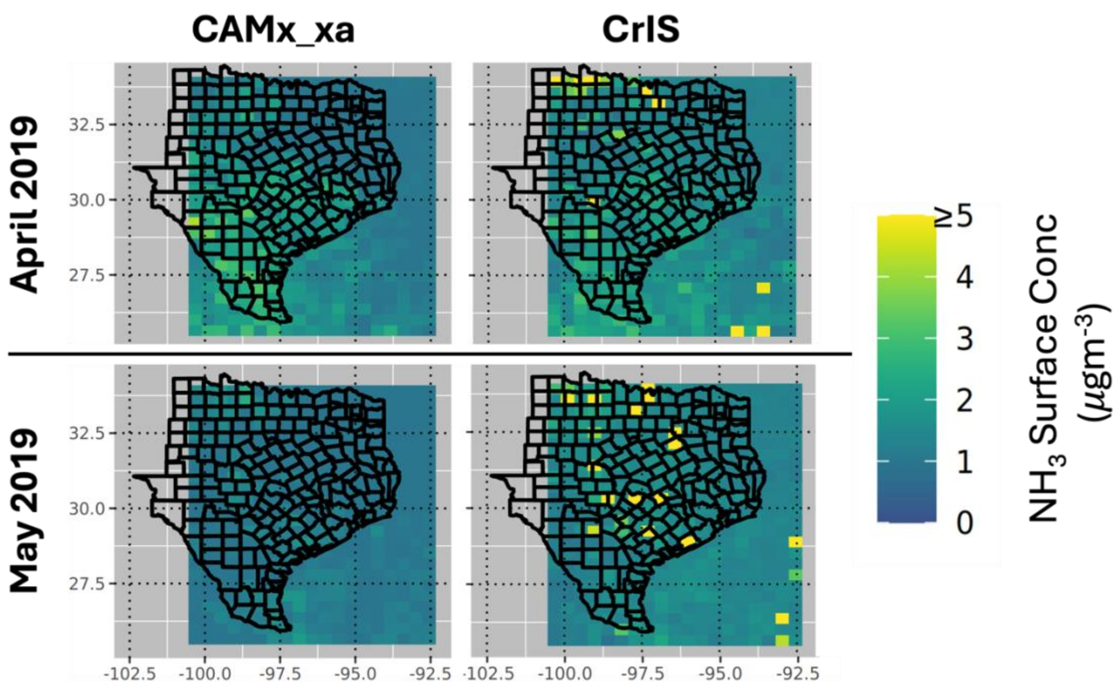


Figure 7. CAMx_xa and CrIS surface NH_3 (μg m-3) concentrations in Spring 2019. Top row: April 2019 surface concentrations for (left) CAMx_xa and (right) CrIS. Bottom row: May 2019 surface concentrations for (left) CAMx_xa and (right) CrIS.

While total column results are largely consistent with the surface results, there is less confidence in the greater spatial spread of model-observation mismatch in the Spring total column concentrations. Noting that CrIS retrievals generally introduce more uncertainty in the total column calculations due to lower averaging kernel values at altitudes greater than 750 hPa (e.g., Figure 2), the CrIS total column observational constraints are less reliable. At the surface, however, CrIS has the highest sensitivity and is more reliable as an observational constraint**.** Therefore, the CAMx_xa and CrIS differences in the Spring were further investigated through an assessment of their raw surface concentration estimates (Figure 7); in combination with the difference plots (Figure 5), the evaluation suggests the following:

- ***<u>April (Figure 5).</u>***

CAMx_xa tends to significantly overestimate surface NH_3 concentrations throughout the western half of the CAMx 4 km modeling domain. The exception is the northwesternmost section of the domain, where CAMx_xa significantly underestimates surface NH_3 concentrations. CAMx_xa generally agrees with CrIS in the eastern half of the domain.

- ***May (Figure 5).***

CAMx_xa tends to slightly underestimate surface NH_3 in the western half of the domain, where it was previously overestimating in April. CAMx_xa continues to significantly underestimate surface NH_3 concentrations in the northwesternmost section of the domain. CAMx_xa continues to generally agree with CrIS in the eastern half of the domain. (Figure 5).

- ***Raw surface concentrations (Figure 7).***

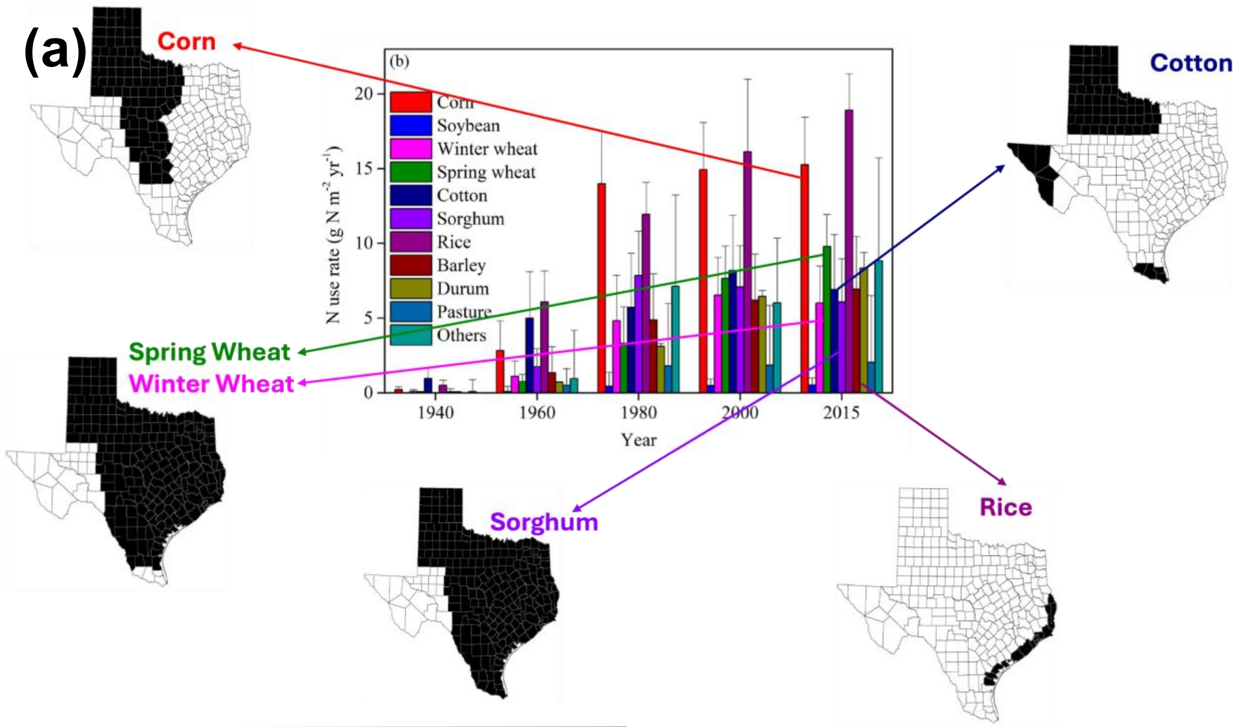
CrIS generally estimates higher NH_3 in April relative to May in the southwestern section of the domain. This suggests that while CAMx_xa is qualitatively capturing increased emissions it is overestimating how much is being emitted by the associated processes. In contrast, both CAMx_xa and CrIS agree in their estimates of lower surface concentrations in May relative to their April estimates. This suggests both are quantitatively capturing lower source strengths in May in the southwest of the domain.

Over the whole domain, CAMx_xa estimates lower source strengths in May relative to its April estimates. In May, the CAMx_xa has better agreement with CrIS than in April. Based on the differences, however, CAMx_xa has a small but systematic underestimate of surface concentrations (Figure 5).

In the northwesternmost section of the domain, CAMx_xa significantly underestimates NH_3 in both April and May (Figure 5). Based on the raw concentrations, CAMx_xa estimates similar source strengths in that region in both April and May. Note that this discrepancy is not apparent in the summer, but possibly begins again in the Fall (Figure 5). Appendix B provides raw surface concentrations for all months, indicating peak surface concentrations estimated by CAMx_xa and CrIS are in April and May.

To explore the possible sources of the Spring discrepancy, AER analyzed existing literature on Nitrogen (N) fertilizer application intensity categorized by major crop types in Texas (Figure 8a); NH_3 emissions based on livestock presence (Figure 8b); Nitrogen (N) fertilizer application schedules (Figure 9); and comparisons of N fertilizer usage rates across time (Figure 10) and multiple studies (Figure 11).

Figure 8 shows expected hotspots of NH_3 emissions resulting from cropland fertilizer use and livestock (cattle) presence. Overall, as of 2015, the most N intensive crops in Texas are rice, corn, wheat, cotton, and sorghum. Rice and corn are the most N intensive but have less spatial coverage in the CAMx modeling domain than wheat and sorghum. In particular, rice is the most N-intensive but is limited to the eastern section of the domain, along the Gulf coast. Cotton is also N intensive, but again has limited spatial coverage in the CAMx domain. In terms of livestock (cattle) emissions, the Texas panhandle (outside the modeling domain) has the highest population density of cattle, with a few secondarily dense regions in the northern and eastern portions of the modeling domain.



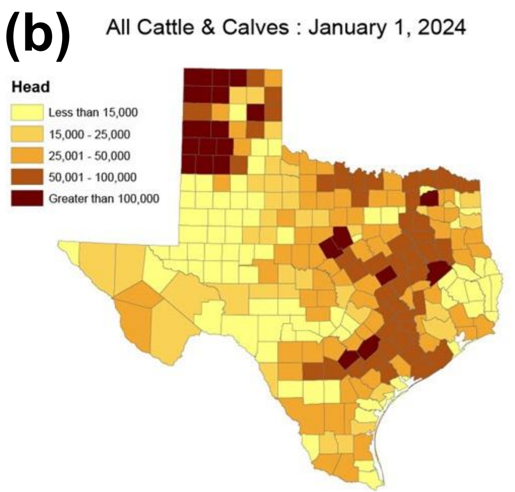


Figure 8. Major agricultural contributors of NH3 emissions. (a) CONUS cropland N application rate from 1940-2015, categorized by crop type. Maps link most intensive N use crop types to planting locations in Texas; and (b) livestock populations as a proxy for NH3 emissions, focusing on cattle. N application figure from Cao et al. (2018); crop maps from Texas Department of Agriculture (TXDA, 2025); cattle map from United States Department of Agriculture (USDA, 2025).

Figure 9 shows the typical N fertilizer application schedule for CONUS by season (Cao et al., 2018). Overall, the Cao et al. (2018) CONUS-wide study showed intensive cropland N applications primarily in the “before planting” season in regions outside of Texas (especially the Midwest and the Northern Great Plains); in contrast, there was relatively low N application estimated in the CAMx modeling domain at that time (Figure 9b).

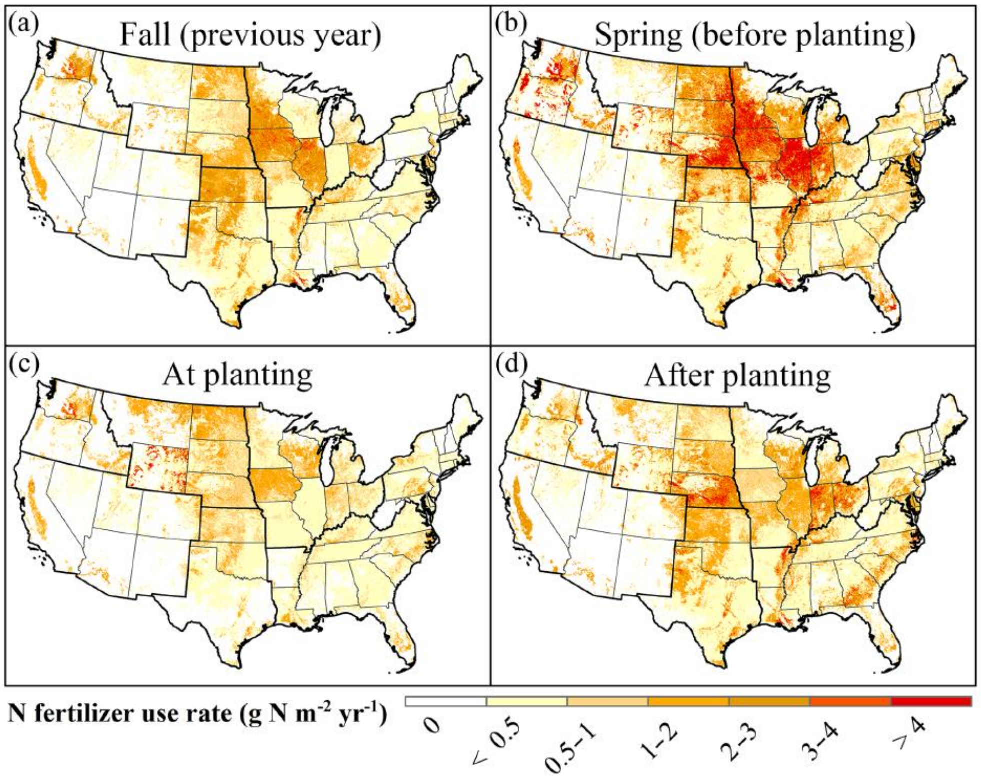


Figure 9. Nitrogen (N) Fertilizer Application Schedules in 2015 for CONUS during (a) Fall 2014, (b-d) before, at, and after planting. Figure from Cao et al. (2018).

Figure 10 compares N fertilizer use rates between 2000 and 2015 (Cao et al., 2018). Notably, N fertilizer usage rates decreased overall in the modeling domain. The exceptions are the hotspots corresponding to usage rates ≥7 g N m^{-2} yr. These modeling domain "hotspots" largely remained consistent between 2000 and 2015 (and, likely, 2019) and correspond to corn and rice growing regions in the central and eastern section of the domain, as well as cotton and sorghum regions in the south.

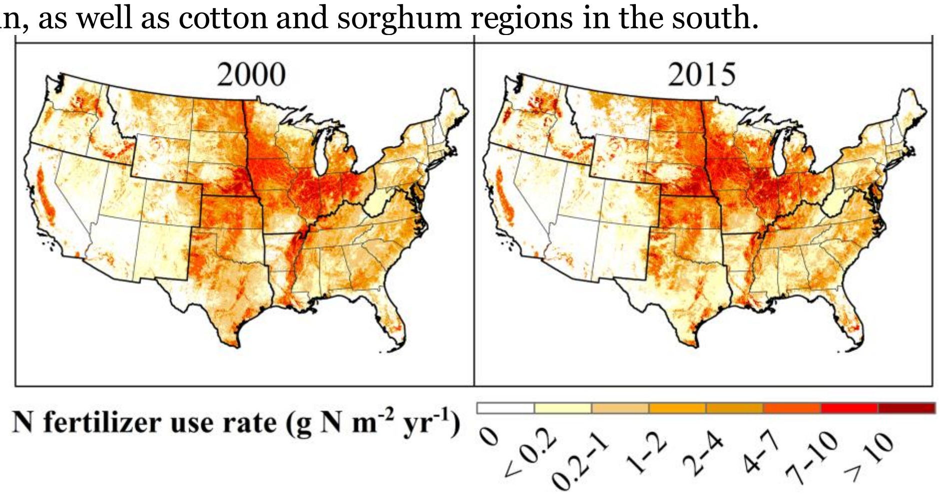


Figure 10. Comparison of N fertilizer use across CONUS between 2000 and 2015. Figure adapted from Cao et al. (2018).

Figure 11 assesses the state of science for the CONUS N fertilizer use across the 2010-2019 decade. The four different studies snapshot annual estimates for 2015 (Cao et al., 2018; IPNI/NuGIS); 2013 (Lu & Tian, 2017); and 2010 (Nishina et al., 2017). The studies all used different datasets and downscaling methods, resulting in key differences evident in the CAMx modeling domain. Namely, the USDA data-based studies (Cao et al., 2018; IPNI/NuGIS) suggest significantly lower overall N fertilizer use in the CAMx modeling domain than the other studies that incorporate international data sets from the Food and Agriculture Organization and/or the International Fertilizer Association).

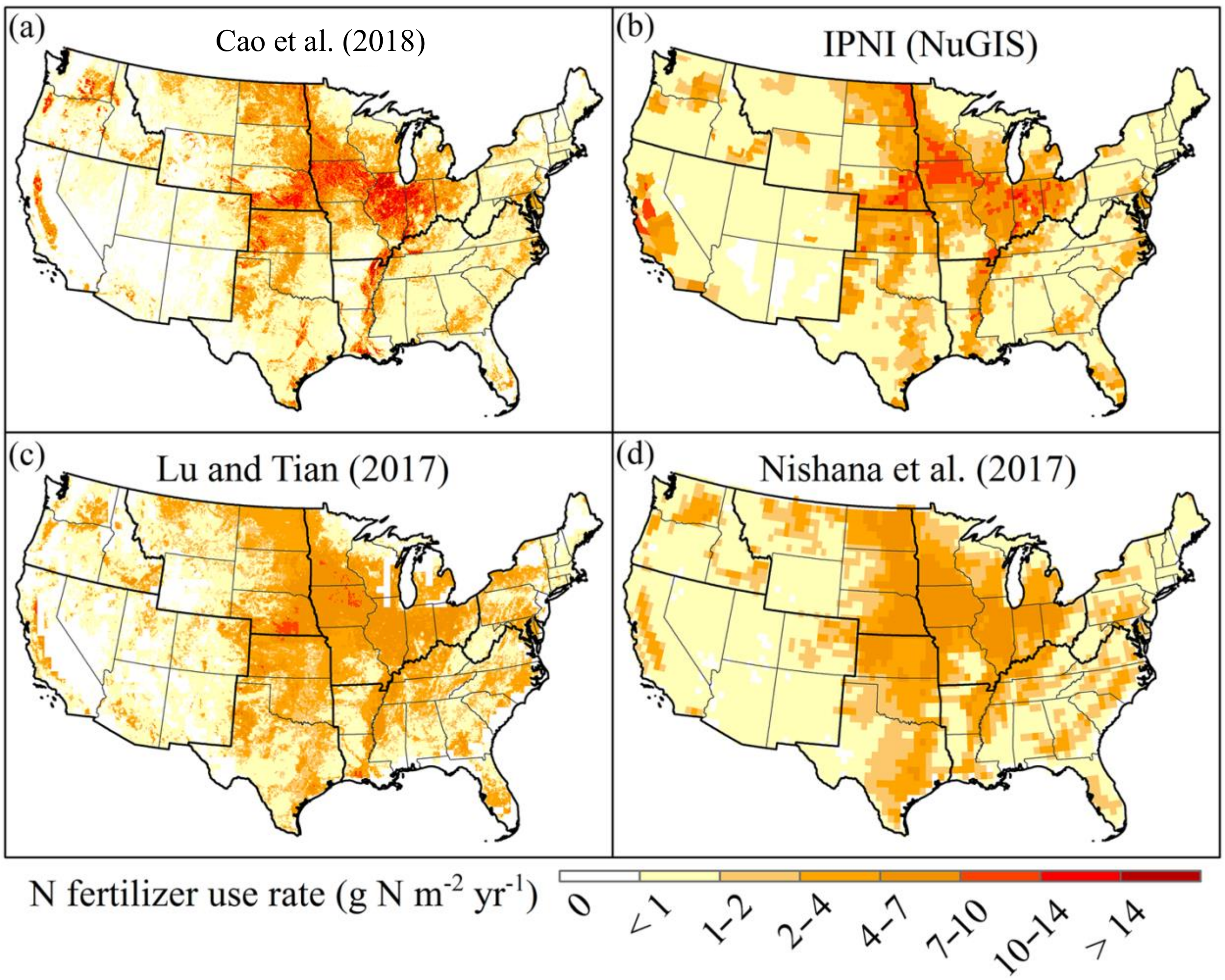


Figure 11. N fertilizer use rates across CONUS circa 2015 estimated by four different studies. Figure adapted from Cao et al. (2018).

The information synthesized from Figures 8-11, suggests the following about the CAMx_xa and CrIS spring time surface NH_3 discrepancies:

Overall, N fertilizer use rates have decreased in the modeling domain between 2000 and 2015 (Figure 10). Relative to the Texas panhandle and Gulf coast regions, the western half of the domain has lower N fertilizer usage rates (Figure 9).

- Generally, the April overestimate by CAMx relative to CrIS coincides with the Spring “before planting” (Figure 9b) Nitrogen (N) fertilizer application schedules.

The CAMx_xa April "before planting" systematic underestimate of surface NH_3 concentrations is consistently associated with wheat, corn, and sorghum regions (Figures 7-9). Notably, in April, the N-intensive rice corridor along the Gulf coast generally agrees with CrIS. ***Overall, this suggests the underlying TCEQ emissions inventory is systematically overestimating April's "before planting" fertilizer application in wheat, corn, and sorghum regions.***

- The small but systematic underestimate by CAMx relative to CrIS in the northwest sector of the domain in May suggests that ***late Spring fertilizer application and/or livestock emissions are being underestimated by the underlying TCEQ emissions inventory.***
- In both April and May, CAMx_xa systematically underestimates NH_3 surface concentrations in the northwesternmost section of the domain, corresponding to regions associated with higher densities of livestock and all crops except rice (Figure 8). Figures 10 and 11 indicate that these regions border N fertilizer application hotspots in Oklahoma, suggesting that surface NH_3 in these regions is potentially influenced by neighboring state fertilizer application. However, as these hotspots do not feature significantly in the "before planting" fertilizer use schedule (Figure 9b), it is unlikely to be the major cause.

3.2. AMoN Network Cross-Validation

To further investigate results from Section 3.1, the CrIS and CAMx_xa surface NH_3 at the nearest co-located CrIS overpass location was compared to AMoN TX 41 surface NH_3. Figure 4 shows the location of the AMoN site. Note that AMoN measurements are limited – sample collection is approximately every two weeks, such that only a few days per month are available for comparison.

The focus is the April and May 2019 periods of maximum discrepancy between CAMx and CrIS. Table 2 provides the output of the AMoN comparison to the nearest CrIS and CAMx point. Figure 12 provides summary statistics over the comparison days on the dates in April and May 2019 for which AMoN data exists. , and AMoN NH_3 at or near the CrIS 13:30 LST overpass in the vicinity of the AMoN TX 41 site (Figure 12).

Overall, the limited data suggests that CAMx_xa and CrIS both overestimate NH_3 relative to the AMoN TX 41 ground station. However, on the few comparison dates, CrIS generally shows better agreement. Note that depending on location of the nearest point, the influence of higher surface NH_3 from Houston could impact CrIS and CAMx results (Figure 4).

Table 2. Comparison of CrIS and CAMx_xa surface NH_3 with bi-weekly AMoN measurements. AMoN site located at approximately 30.7015N, -94.674E.

Nearest Lat	Nearest Lon	Measurement Date	AMoN NH_3 ($\mu g\ m^{-3}$)	CAMX_xa NH_3 ($\mu g\ m^{-3}$)	CrIS NH_3 ($\mu g\ m^{-3}$)
31.1059	-94.174	2019/04/02	0.45	0.96	0.50
30.3278	-94.575	2019/04/16	0.50	0.96	0.48
30.3278	-94.575	2019/04/16	0.06	0.96	0.48
31.1615	-93.627	2019/05/01	0.40	0.96	1.9
30.6845	-94.523	2019/05/14	0.44	4.0	1.7
30.6882	-94.689	2019/05/28	0.45	0.96	1.5

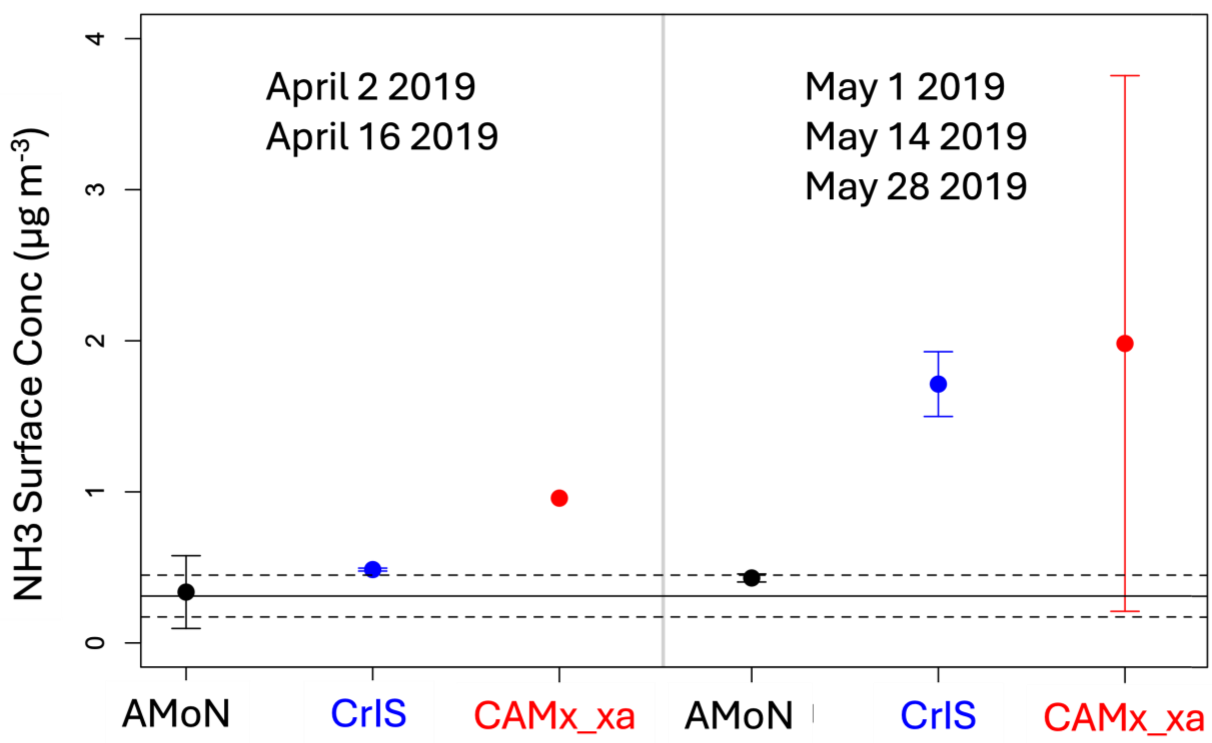


Figure 12. AMoN, CrIS, and CAMx_xa surface NH3 averaged over AMoN measurement dates in April and May. Uncertainties are based on standard deviation from the limited data points.

3.3. Assessment of Daily variability in the context of Houston Area PM2.5

Daily averaged surface PM2.5 across the Houston region for April and May 2019 is shown in Figure 13. The sites used for the averaging are shown in Figure 4.

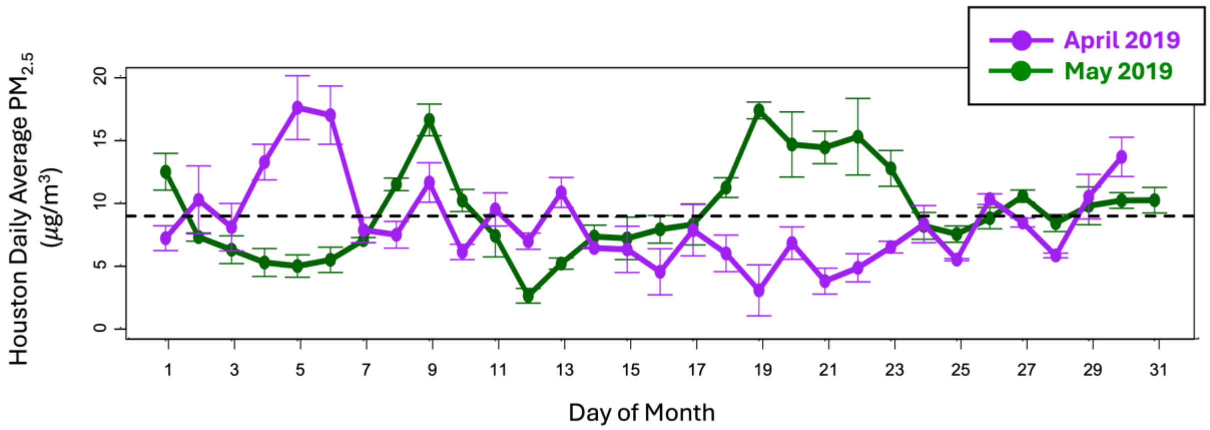


Figure 13. Daily Averages of PM2.5 (*μg m⁻³*) across Houston area measurement sites for April and May 2019. The 2024 PM2.5 Annual Standard of 9.0 *μg m⁻³* (dashed black horizontal line) is shown for reference

Anomalies (relative to the average over April and May 2019) were calculated for each of PM2.5, CrIS surface NH_3, and CAMx_xa surface NH_3. Daily averages of surface NH_3 were calculated for the region encompassing the Houston PM2.5 measurement sites shown in Figure 4. Figure 14 displays the daily variation in PM2.5 anomalies along with the daily variations in surface NH_3 for April (Figure 14a) and May (Figure 14b). Overall, there is not a strong correlation in April, while there is only a weak correlation in May for CrIS (R^2=0.25). It is possible there are stronger correlations on days important to the Houston design values warranting additional exploration across more time periods. In addition, the importance of NH_3 to PM2.5 design values warrants exploration in other non-attainment regions such as Dallas Fort Worth, Austin, and San Antonio.

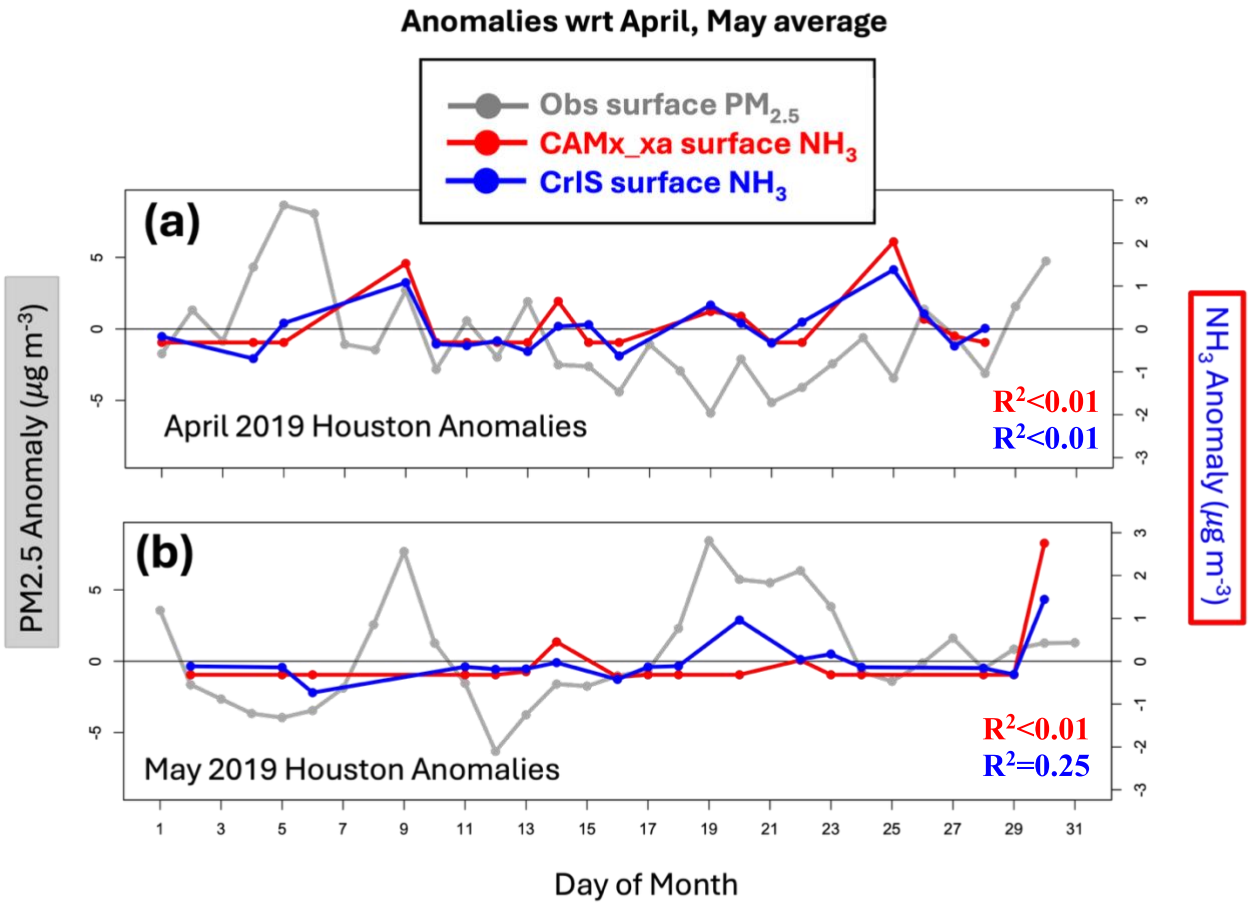


Figure 14. Anomalies in daily average surface PM2.5 and NH3 in the Houston area for (a) April and (b) May. Anomalies are calculated relative to the mean across April and May for each of PM2.5, CAMx_xa NH3, and CrIS NH3. R^2 refers to correlation between NH3 and PM2.5 anomalies.

4. Issues and Resolutions

- ***Issue:*** CAMx 3D NH3 columns are only available for April through October 2019. ***Resolution:*** In coordination with TCEQ PM, scope of project was limited to April – October 2019.
- ***Issue:*** We obtained a sample month (April 2019) of CAMx data, however of the two CrIS instruments that were collecting measurements in that time, only the CrIS instrument on the NOAA20 satellite was operational. Due to an instrument issue, CrIS on the Suomi NPP (SNPP) platform was offline from 2019/03/26 to 2019/08/12.
  Resolution: While CrIS on NOAA20 has NH3 data in 2019, only ECCC has the data readily available. The NASA earth data portal only provides access to NOAA20/JPSS-1 CrIS data beginning 2021. No issue is anticipated between usage

of the ECCC retrieval algorithm over the NASA retrieval algorithm; however, as data structures are different depending on ECCC vs. NASA sources, we note that the code will be built on assumption that ECCC CrIS data will be used in future. As these data are publicly available, continued access for TCEQ in the future should not be an issue.

- ***Issue:*** The global CrIS files that contain the averaging kernel/observation operator information are very large, with a single day of tar.gz files ~1.5GB. ***Resolution:*** We have circumvented that issue by selecting the files relevant to the Texas domain. The code has been written such that it downloads a given date as requested to a temporary location, selectively extracts the relevant files from the tar.gz global archive, and deletes the tar.gz archive from the temporary location.
- ***Issue:*** Challenges with Latitude-Longitude CAMx to CrIS matching. CAMx variables have dimensions XY, while CrIS observations give latitude and longitude. CAMx files provide latitude and longitude as a function of X and Y but not the inverse. Several versions of computations of Lambert conformal projection did not give expected results.
  ***Resolution*:** We used a WRF geogrid file on the same grid, an open-source Python package (“salem”) to read the grid, and another Python package (“pyproj”) to generate the projection from the grid.
- ***Issue:*** Challenges with vertical interpolation CAMx to CrIS matching. The CAMx files are on pressure levels, but the vertical coordinate given for the pressure levels is altitude. To interpolate onto CrIS pressure levels, we needed the pressure at each CAMx level.
  Resolution: The TCEQ PM provided us with external CAMx met3D files that do have pressure levels. We then matched up to the correct date, latitude, and longitude for each point in the CAMx grid. (The met3d files were obtained via sftp to the TCEQ servers: /TXO3/camx/bexar_dfw_hgb_fy23_s15/input/met/2019_wrf415_noah_ysu_txe_lyr45t30)

5. Quality Assurance

The processing and analysis scripts used in this project were inspected by a team member not involved in their creation for accuracy. All automated calculations and at least 10% of manual calculations were inspected for correctness. This meets the requirement of Level III QAPPs that 10% of the data must be inspected.

As the quality of the information, including secondary data was not evaluated by EPA, the below disclaimer applies to all project deliverables:
Disclaimer: The information contained in this report or deliverable has not been evaluated by EPA for this specific application

6. Conclusions

- Generally, the April overestimate by CAMx relative to CrIS coincides with the Spring “before planting” (Figure 9b) Nitrogen (N) fertilizer application schedules.

The CAMx_xa April “before planting” systematic underestimate of surface NH_3 concentrations is consistently associated with wheat, corn, and sorghum regions (Figures 7-9). Notably, in April, the N-intensive rice corridor along the Gulf coast generally agrees with CrIS. ***Overall, this suggests the underlying TCEQ emissions inventory is systematically overestimating April’s “before planting” fertilizer application in wheat, corn, and sorghum regions.***

- The small but systematic underestimate by CAMx relative to CrIS in the northwest sector of the domain in May suggests that ***late Spring fertilizer application and/or livestock emissions are being underestimated by the underlying TCEQ emissions inventory.***
- In both April and May, CAMx_xa systematically underestimates NH_3 surface concentrations in the northwesternmost section of the domain, corresponding to regions associated with higher densities of livestock and all crops except rice (Figure 8). Figures 10 and 11 indicate that these regions border N fertilizer application hotspots in Oklahoma, suggesting that surface NH_3 in these regions is potentially influenced by neighboring state fertilizer application. However, as these hotspots do not feature significantly in the “before planting” fertilizer use schedule (Figure 9b), it is unlikely to be the major cause.
- Comparisons with AMoN ground data: Overall, the limited data suggests that CAMx_xa and CrIS both overestimate NH_3 relative to the AMoN TX 41 ground station. However, on the few comparison dates, CrIS generally shows better agreement. Note that depending on location of the nearest point, the influence of higher surface NH_3 from Houston could impact CrIS and CAMx results (Figure 4).
- Correlation with $PM_{2.5}$ in the Houston area during April/May: Overall, there is not a strong correlation in April, while there is only a weak correlation in May for CrIS (R^2=0.25). It is possible there are stronger correlations on days important to the Houston design values warranting additional exploration across more years, dates, and/or seasons. In addition, the importance of NH_3 to PM2.5 design values warrants exploration in other non-attainment regions such as Dallas Fort Worth, Austin, and San Antonio.

7. Recommendations for Future Work

Based on the results, AER recommends:

- Further evaluation of correlation between $PM_{2.5}$ anomalies and NH_3 for DFW, ARR, SAN, HGB regions; including speciated $PM_{2.5}$ data and PMF analysis of speciated $PM_{2.5}$ data will be especially useful for source attribution across different regions;
- Further examination of the NH_3 underestimate in the northwesternmost section of domain will provide information on potential missing sources. Extension of modeling domain to include Texas Panhandle and AMoN TX43 site can potentially help with this analysis.
- Further examination of underlying NH_3 emissions inventory by crop type and timing of fertilizer application;

- Extension beyond 2019 test period upon availability of 3D CAMx fields, and associated comparison with available AMoN data.

8. References

Cao, P., Lu, C., and Yu, Z. (2018): Historical nitrogen fertilizer use in agricultural ecosystems of the contiguous United States during 1850–2015: application rate, timing, and fertilizer types, Earth Syst. Sci. Data, 10, 969–984, https://doi.org/10.5194/essd-10-969-2018 .

IPNI (International Plant Nutrition Institute): A Nutrient Use Information System (NuGIS) for the U.S. Norcross, GA, 1 November 2011, available at: www.ipni.net/nugis, last access: 1 March 2018. Note: link no longer accessible, but possibly available here: https://nugis.tfi.org/about

Cady-Pereira, K. E., V. H. Payne, J. L. Neu, K. W. Bowman, K. Miyazaki, E. A. Marais, S. Kulawik, Z. A. Tzompa-Sosa, and J. D. Hegarty (2017), Seasonal and spatial changes in trace gases over megacities from AURA TES observations, Atmos. Chem. Phys., 17, 9379–9398.

Cady-Pereira, K. E., Guo, X., Wang, R., Leytem, A. B., Calkins, C., Berry, E., Sun, K., Müller, M., Wisthaler, A., Payne, V. H., Shephard, M. W., Zondlo, M. A., and Kantchev, V. (2024): Validation of MUSES NH_3 observations from AIRS and CrIS against aircraft measurements from DISCOVER-AQ and a surface network in the Magic Valley, Atmos. Meas. Tech., 17, 15–36, https://doi.org/10.5194/amt-17-15-2024 .

Chen, Y., Shen, H., Kaiser, J., Hu, Y., Capps, S.L., Zhao, S., Hakami, A., Shih, J.S., Pavur, G.K., Turner, M.D. and Henze, D.K. (2021): High-resolution hybrid inversion of IASI ammonia columns to constrain US ammonia emissions using the CMAQ adjoint model. Atmospheric Chemistry and Physics, 21(3), pp.2067-2082.

EPA (2025): US Environmental Protection Agency Air Quality Design Values for Criteria Pollutants, PM2.5 Annual (2024). https://epa.maps.arcgis.com/apps/MapSeries/index.html?appid=bc6f3a961ea14013afb2e0d0e450b0d1. Accessed 23 June 2025.

Liu, L., Xu, W., Lu, X., Zhong, B., Guo, Y., Lu, X., Zhao, Y., He, W., Wang, S., Zhang, X. and Liu, X (2022): Exploring global changes in agricultural ammonia emissions and their contribution to nitrogen deposition since 1980. Proceedings of the National Academy of Sciences, 119(14), p.e2121998119.

Lu, C. and Tian, H. (2017): Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: shifted hotspots and nutrient imbalance, Earth Syst. Sci. Data, 9, 181–192, https://doi.org/10.5194/essd-9-181-2017.

NADP, (2022): National Atmospheric Deposition Program Ammonia Observation Network site TX41. https://nadp.slh.wisc.edu/sites/amon-TX41/ . Accessed 30 May 2025.

Nishina, K., Ito, A., Hanasaki, N., and Hayashi, S. (2017): Reconstruction of spatially detailed global map of NH4+ and NO3- application in synthetic nitrogen fertilizer, Earth Syst. Sci. Data, 9, 149–162, https://doi.org/10.5194/essd-9-149-2017.

Shephard, M.W. & Cady-Pereira, K.E. (2015): Cross-track Infrared Sounder (CrIS) satellite observations of tropospheric ammonia, Atmos. Meas. Tech., 8, 1323-1336, doi:10.5194/amt-8-1323-2015.

Sitwell, M., Shephard, M. W., Rochon, Y., Cady-Pereira, K., and Dammers, E. (2022): An ensemble-variational inversion system for the estimation of ammonia emissions using CrIS satellite ammonia retrievals, Atmos. Chem. Phys., 22, 6595–6624, https://doi.org/10.5194/acp-22-6595-2022 .

Toro, C., Sonntag, D., Bash, J., Burke, G., Murphy, B.N., Seltzer, K.M., Simon, H., Shephard, M.W. and Cady-Pereira, K.E. (2024): Sensitivity of air quality to vehicle ammonia emissions in the United States. Atmospheric Environment, 327, p.120484.

TXDA (2025). Crop Sheets. https://texasagriculture.gov/Portals/0/Publications/PEST/Crop_Sheets.pdf?ver=bVYahXf9oN3AcbWddT5HQ%3d%3d. Accessed 4 June 2025.

USDA (2025). County Estimate Map – Cattle. https://www.nass.usda.gov/Statistics_by_State/Texas/Publications/County_Estimates/ce_maps/ce_catt.php Accessed 4 June 2025.

9. Appendix A: Software User Guide

9.1. Overview

This User Guide accompanies the deliverable for TCEQ Work Order No. 582-23-45974, "Model Evaluation for Ammonia (NH_3) using Satellite Data". This project improves estimations of ammonia emissions from CAMx data by applying the CrIS observation operator to CAMx simulations. We have also produced diagnostics to compare both surface abundance and total column densities of NH3 approximated by CAMx, CAMx with the CrIS observation operator applied, CrIS alone, and AMON. This guide will enable the user to reproduce the work done, as well as extend the analysis to additional time periods as needed.

9.2. Directory Structure

***This guide and all scripts are in the parent directory*, *DELIVERABLE_5_3/*.** The contents of the parent directory include data directories for raw and processed data; and a figures directory. These are described below. Note that due to file sizes, only the relevant files required to run a sample month (April 2019) are provided. Additional months can be run assuming access to relevant CAMx files (see below) and CrIS files (automatically downloaded via Docker container).

CAMx_2019_Apr_to_Oct/CAMx_<YYYYMM>/

3D NH3 abundance simulations (netCDF). Note that daily files for a given month must all be located in unzipped format in their dedicated CAMx_<YYYYMM> directory. Example files for CAMx_201904 are provided in the DELIVERABLE_5_3 directory. Required file type is shown below, using a sample date of 1 April, 2019:

NH3.camx720_cb6r5CF2_avrg.20190401.txo3.bc19_19apr.v2dP_v2b1P1.2019_wrf415_noah_ysu_lyr45t30.txs_4km.nc

CAMx_MET/

A CAMx file with pressure levels (netCDF). This is a static file and shouldn't need to be updated:

camx7_met3d.20191012.2019_wrf415_noah_ysu_txe_lyr45t30.txs_4km.v51.nc

ECCC_CrIS_NH3/CrIS-<YYYY-MM-DD>

CrIS NH3 observations (netCDF). When running the Docker container, daily files will be automatically downloaded for the requested month(s) for the regions relevant to the Texas domain. Example:

Combined_NH3_n100_0_n080_0_p015_0_p030_0_20190401.nc

WRF_GRID/

Files for translating between lat-lon and projection coordinates. These are static files:

geo_em.d01.nc
geo_em.d02.nc

diagnostics/

If requested, diagnostic plots of CrIS and CAMx before and after interpolation are output, with and without observation operator applied. (e.g., Figure A6)

monthly-averages/

Monthly average outputs of CrIS and CAMx surface abundance and column density, on WRF grid, provided as files of name format:

avgs_<YYYYMM>.nc

Ground_Obs_2019/PM2.5_2019_Ground_Obs/

PM2.5 HGB region data used for Task 4 analysis as a daily average file.

Ground_Obs_2019/AMoN Obs/

AMoN README and 2019 data for TX41 site used in Task 4 analysis.

9.3. Data

The CrIS observations are provided in the deliverable under ECCC_CrIS_NH3. These files were accessed from https://hpfx.collab.science.gc.ca/~mas001/satellite_ext/cris/snpp/nh3/v1_6_4 (note this data requires an account to access, and credentials are automatically supplied when running via the Docker container). Please review the ECCC CrIS data usage guide at https://hpfx.collab.science.gc.ca/~mas001/satellite_ext/cris/snpp/nh3/v1_6_4/CrIS_NH3_data_usage_statement.pdf prior to use.

The CrIS files describe a 20×15° lon×lat bounding box. The regional coverage of the file can be parsed from the filename as follows:

```
<n100_0>_<n080_0> = -100.0 to -80.0 Lon #n = negative
<n120_0>_<n100_0> = -120.0 to -100.0 Lon

<p015_0>_<p030_0> = +15.0 to +30.0 Lat #p = positive
<p030_0>_<p045_0> = +30.0 to +45.0 Lat
```

In the deliverable, the ECCC CrIS files have been isolated to include only four sets of latitude-longitude boundaries which cover the state of Texas, as shown in the figure below. Each color in the figure indicates a different segment of ECCC CrIS data.

Figure A1. *ECCC CrIS file coverage over Texas. Each file's coverage region is indicated by a different color (orange, red, purple, green).*

The CAMx ammonia observations cover Texas and the file names have the form

```
NH3.camx720_cb6r5CF2_avrg.<YYYY><MM><DD>.txo3.bc19_<YY><month>v2dP_v2b1P1.<YY
>_wrf415_noah_ysu_lyer<max latitude>t<min latitude>txs_<grid size>.nc
```

Accessory data defining the WRF grid and CAMx pressure levels have also been provided in the WRF_GRID and CAMX_MET directories, respectively.

9.4. Processing and Analyzing the Data

a) With Docker

The script `get_monthly_averages.sh` builds a Docker container, runs the script to compute monthly averages (`monthly_averages.py`), then exits and removes the container. If the CrIS files for particular month are not downloaded in ECCC_CrIS, the program will download them automatically.

The monthly_averages.py script takes several arguments: `months`, `year`, `surface index`, and `maindir`. These are set in at the top of the `get_monthly_averages.sh`, then passed to the Python script. Only these four parameters should be edited by the user in a regular run. Note that year may only be one year, while months may be a list of zero-padded months, separated by spaces – e.g.

```
months='06 07 08'
```

or

```
months='10'
```

The surface index must be an integer, and refers to the CrIS pressure level to consider the "surface" when computing the CAMx abundance. We suggest a default of –3, as the third pressure level from the bottom is the lowest pressure level that is above the surface (non-NaN in both CrIS and CAMx) for nearly all of Texas. The main directory should usually be set to "./" (the current directory), as the data directories for CAMx and CrIS are in the same directory as `get_monthly_averages.sh`, However, if the directory containing these directories is changed, this parameter should be updated to give the full path to that data directory. This is also the directory where the monthly averages output will be stored.

```
###########################################################
#                    USER-SET PARAMETERS
#   All parameters you may wish to set are in this block.
###########################################################
# Note if you run multiple months, this argument must be in quotes (e.g. --months '06 07').
months='04'
year='2019'
# Pressure level to consider the "surface" in surface abundance.
# -3 gives the 3rd layer from the bottom.
surfaceindex=-3
# Directory where data is stored.
maindir="./"
###########################################################
# Text below this line should not be edited.
###########################################################
```

Figure A2. *User-set parameters section of the get_monthly_averages.sh script that launches the Docker container and produces monthly average files.*

To execute the code, simply run `./get_monthly_averages.sh` in a terminal window from the deliverable directory.

Note: The Docker image is a Linux system and uses a Linux Miniconda installer. The Miniconda installer is dependent on the system architecture of your local machine. The Dockerfile assumes an x86 architecture, which will be accurate for most systems. If your local machine uses a different architecture (e.g. aarch on an M1/M2 chip Mac), the Miniconda installation will fail when building the Docker image. If this is the case, select the appropriate Linux Miniconda 3.9 installer from https://repo.anaconda.com/miniconda/ and substitute the correct link into line 15 in the Dockerfile.

```
11  #  Install miniconda 3.9.
12
13  ENV MINICONDA="/app/miniconda"
14
15  ADD https://repo.anaconda.com/miniconda/Miniconda3-py39_24.7.1-0-Linux-x86_64.sh miniconda.sh
16
17  RUN sh miniconda.sh -b -p /opt/miniconda && rm -fr miniconda.sh
18  ENV PATH="/opt/miniconda/bin:$PATH"
```

Figure A3. *Screenshot of where to substitute the relevant non-default miniconda installer link.*

b) With Jupyter notebook

The script `monthly-avgs_jupyter.ipynb` computes a monthly average of both surface abundance (in units of ppmv and in µg/m³) and column density (in units of molec/cm² and in µg/m²), for both the CrIS instrument and CAMx with the CrIS observation operator applied. The top portion of the code (the second cell) contains all the parameters which could need to be adjusted by the user.

```
[ ]:
###################################################################################
#         User Input Variables.
###################################################################################

month_strs = ["04"]
year = "2019"
makeplot = False #Output diagnostic plots?
surface_index = -3
maindir = "~/DELIVERABLE_5_3/"

# Interpret month strings as a list of months.
months = month_strs.split(' ')
print(f"Computing monthly averages for {months} of {year}...")

################# Locations of data files ####################
cris_path     = maindir+"ECCC_CrIS_NH3/"
# Note: Prefixes of TX relevant CrIS files:
#"Combined_NH3_n100_0_n080_0_p015_0_p030",
#"Combined_NH3_n100_0_n080_0_p030_0_p045",
#"Combined_NH3_n120_0_n100_0_p015_0_p030",
#"Combined_NH3_n120_0_n100_0_p030_0_p045"
camx_path     = maindir+"CAMx_2019_Apr_to_Oct/"
wrf_grid      = maindir+"WRF_GRID/" #Needed for interpolation
# Grab a CAMx file that includes pressure levels.
camx_p_path   = maindir+"CAMx_MET/camx7_met3d.20191012.2019_wrf415_noah_ysu_txe_lyr45t30.txs_4km.v51.nc"
output_dir    = maindir
```

Figure A4. *Screenshot of Jupyter Notebook (ipynb) file's User Input section.*

These include the following parameters:

- `month_strs` – This is a list of months to compute monthly averages for. These must be zero-padded strings. For example [“04”, “05”, “10”] would compute the monthly means for April, May, and October.
- `year` – The year to compute monthly averages for. This must be a string.

- `makeplot` – If set to `True`, this creates a diagnostic plot for each CrIS observation showing the raw CAMx NH_3 abundance CrIS prior, CrIS retrived abundances, CAMx interpolated onto the CrIS pressure levels, and CAMx abundances after interpolation with the CrIS obvservation operator applied. These plots will be saved with the row and column number of the point in the CAMx grid that they correspond to. They will be saved to a subfolder called "diagnostics" in the current working directory. If set to `False`, the diagnostic plot building will be skipped and the monthly averages will be computed more quickly.
- `surface_index` – The CrIS pressure levels nominally go down to roughly sea level. However, in some parts of Texas, the land is actually higher than this pressure level. Therefore, to choose one consistent level for all CrIS and CAMx data, we select an index in the pressure level list to be considered the surface value. A `surface_index` of –3 means that the third highest pressure value is considered the surface. -3 is a good choice because this is the lowest level (closest to the actual surface) where most locations in Texas have non-nan ammonia concentrations.
- `maindir` – The local path to `DELIVERABLE_5_3`

The following parameters set paths to data locations. Since this data is provided in `DELIVERABLE_5_3`, these should not need to be adjusted (they are set relative to the deliverable path).

- `cris_path` – the location where ECCC CrIS NH_3 retrieval data is stored.
- `camx_path` – the path to CAMx NetCDF files containing ammonia abundances at different heights.
- `camx_p_path` – The path to a CAMx file that contains pressure level information.

A schematic of the input data, outputs, and processing steps is shown below. Note that WRF grid files are also used to generate the conformal projection to convert between latitude and longitude and the WRF grid parameters (X and Y).

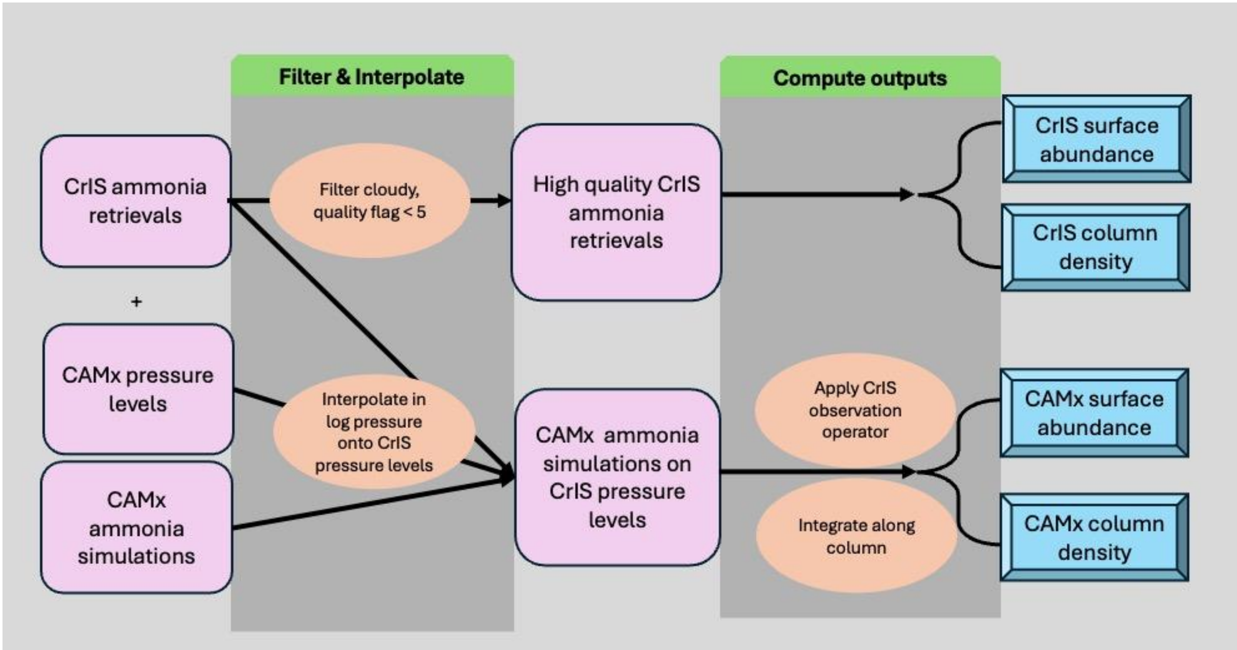


Figure A5. *Schematic illustrating all steps in data processing and interpolation.*

Optional diagnostic plots can be created during the first processing step for each CrIS observation by setting makeplot=True. Two sample diagnostic plots are shown below. Raw CAMx data (blue) contains more pressure levels, especially near the surface, than raw CrIS data (green, solid). As a result, CAMx interpolated onto CrIS pressure levels (yellow, solid) is only a rough approximation of NH_3 structure close to the surface. These are saved to `diagnostics/` with the naming convention `interp_diagnostic-r<row in WRF grid>-c<col in WRF grid>_<YYYY><MM><DD>.png.`

While producing diagnostic plots for each CrIS observation causes the analysis to process significantly slower, the plots can be useful for identifying the cause of differences in surface abundance or column density. For example, the rougher interpolation near the surface, where NH_3 abundances are typically several orders of magnitude higher than they are above ~600 mbar, can lead to significant biases when computing the CAMx column density as the integral of the abundance at each pressure level.

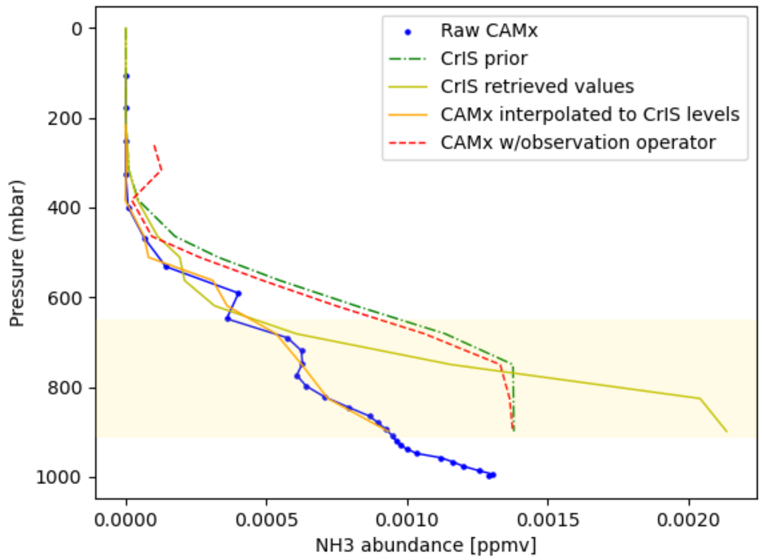


Figure A6. *Sample diagnostic output for 17 April 2019 at row 0 column 17 of intermediate processed data. Highlighted yellow region indicates theoretical levels of peak sensitivity of CrIS instrument between ~750-900 mbar (hPa). Note that interpolation truncates at lowest CrIS pressure level.*

The CrIS observation operator is then applied to the interpolated CAMx data at each CrIS location as

$$X_o = X_a + A(X_{camx}-X_{aprior})$$

where:
X_o is the CAMx NH_3 profile with CrIS observation operator applied;
X_a is the initial guess NH_3 profile from the CrIS retrieval ("xa" in the CrIS netCDF file);
A is the averaging kernel/observation operator ("avg_kernel" in the CrIS netCDF file) ;

X_{camx} is the CAMx modeled mixing ratios at each height, interpolated to the CrIS heights and matched to CrIS positions and times.

The column density, N_{NH3}, is then computed for CAMx as

$$N_{NH3} = \Sigma(X_o dz)$$

for dz the computed distance between the pressure levels. This is reported in both μg/m² and in molecules/cm². CrIS observations provide column density values directly; these are also reported in both μg/m² and in molecules/cm².

The surface abundance for both CrIS and CAMx (with the observation operator applied) is then selected as the mixing ratio at the `surface_index` pressure level in. This is reported in both μg/m³ and in ppmv.

The monthly averages in each WRF grid cell for CrIS and CAMx column densities and surface abundances are then computed and saved as a netCDF file in `outputs` with the file name `avgs_<YYYY><MM>.png`.

9.5. Troubleshooting

For issues, contact Archana Dayalu (adayalu@aer.com) and Sara Vannah (svannah@aer.com)

Appendix B: Supplementary Figures

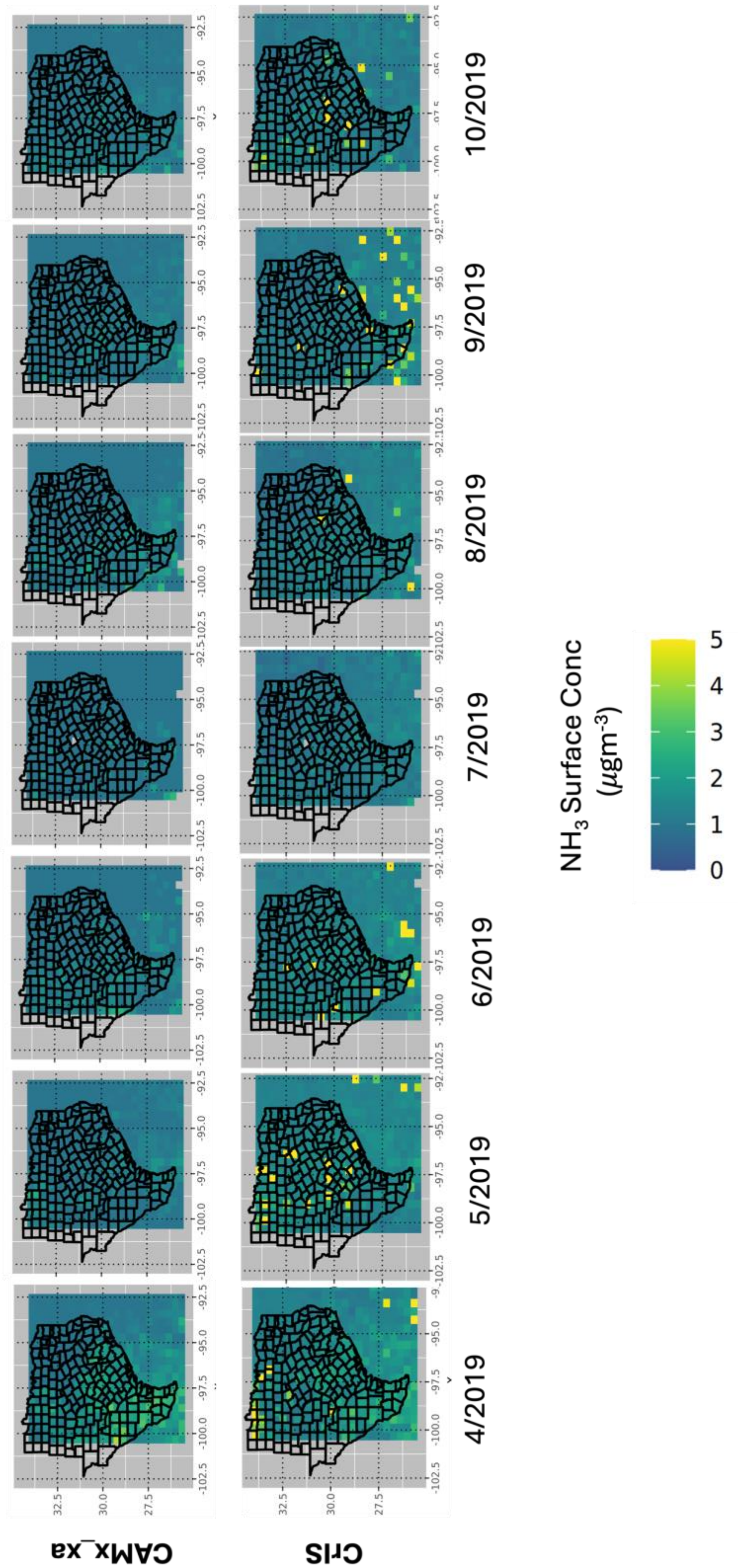


Figure B 1. Monthly average NH3 concentrations for CAMx_xa and CrIS from April to October 2019.