FINAL REPORT

Improving Lightning Data Assimilation (LDA) in WRF

TCEQ Contract No. 582-23-45974
Work Order No. 1

Revision 2.0

Prepared by:

Jennifer Hegarty, Archana Dayalu, and Matthew Alvarado
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02421
Correspondence to: adayalu@aer.com

Prepared for:

Khalid Al-Wali
Texas Commission on Environmental Quality (TCEQ)
Air Quality Division
Building E, Room 383
Austin, Texas 78711-3087

18 June 2024

Document Change Record

Revision	Revision Date	Remarks
0.1	31 May 2024	Internal Version for Review
1.0	7 June 2024	Delivery to TCEQ
2.0	18 June 2024	Final Version to TCEQ

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

ABI – Advanced Baseline Imager
AER – Atmospheric and Environmental Research
CONUS – Contiguous United States
CTH – Cloud Top Height
CMAQ – Community Multiscale Air Quality Model
EPA – United States Environmental Protection Agency
LDA/LTGDA – Lightning Data Assimilation
LZA – Local Zenith Angle
GOES – Geostationary Operational Environmental Satellite
GLM – Geostationary Lightning Mapper
MSKF – Multiscale Kain Fritsch Scheme
NCAR – National Center for Atmospheric Research
NCEP – National Center for Environmental Prediction
RAINC – Convective Rainfall
RAINNC – Non-convective Rainfall
RAINSH – Shallow cumulus precipitation
WRF – Weather Research and Forecasting Model
WPS – WRF Preprocessing System

Executive Summary

Misplacement of deep convection can negatively affect near-surface meteorology such as temperature, humidity, winds, and boundary layer height, all of which negatively impact air quality simulations. Previous studies have shown that lighting data assimilation (LDA) improves air quality simulations through improvements in convection (Heath et al., 2016ab). Under a previous TCEQ Work Order, LDA using Geostationary Lightning Mapper (GLM) observations from the Geostationary Operational Environmental Satellite (GOES)-East satellite was incorporated into the Weather Research and Forecasting (WRF) Multi-Scale Kain-Fritsch (MSKF) cumulus parameterization scheme (WRF-MSKF-LDA) during a June 2019 case study using the TCEQ 2019 Modeling Platform (Adams-Selin et al., 2022). However, it was found that the impacts of LDA on the WRF precipitation simulations were mixed and less substantial than anticipated. The results from Adams-Selin et al. (2022) suggested contributions from insufficient parameterization of updraft strength and depth.

In this project, estimates of cloud top height (CTH) from the GOES Advanced Baseline Imager (ABI) have been used as an additional constraint for internal WRF parameterizations of convective depth. This work order uses the WRF model along with AER's previous software implementation of WRF with LDA (Adams-Selin et al., 2022). This work order adds GOES CTH assimilation into AER's previous software implementation of WRF-MSKF-LDA and investigates the impact on a re-run of the June 2019 simulations originally performed in 2022.

There were three major tasks to accomplish this project's objectives. The first of these tasks updated existing Python code from Adams-Selin et al. (2022) to download and process both GLM and CTH data into WRF-compatible format. The next task updated the WRF-MSKF-LDA scheme to incorporate the constraints provided by the GLM and CTH data. The final task evaluated the impacts on WRF estimates of convective rainfall (RAINC) and non-convective rainfall (RAINNC); surface temperature; surface humidity; and surface windspeed. Final evaluations of impacts on WRF model performance compared three versions of WRF for the June 2019 Adams-Selin et al. (2022) case study: unmodified (base) WRFv4.3.3; WRFv4.3.3-LDA and WRFv4.3.3-LDA-CTH. All versions run the Multi-Scale Kain-Fritsch cumulus parameterization scheme; the option to include LDA and/or LDA with CTH is controlled by a toggle variable in the WRF namelist.

Overall, all quality control checks indicate the updated WRFv4.3.3-LDA-CTH is working appropriately and is readily ported to newer versions of WRF. For the June 2019 case study, aggregate statistics across inner and outer modeling domains indicate no significant differences in temperature, dewpoint and wind surface variables. However, both WRFv4.3.3-LDA and WRFv4.3.3-LDA-CTH estimate small but statistically significant increases in RAINC and RAINNC aggregated across WRF domains. In addition, dates for which precipitation estimates from WRF and WRF-LDA were known to be underestimated generally had higher estimates by WRF-LDA-CTH. Finally, WRF-LDA-CTH precipitation estimates agree better with NCEP Stage IV radar-based precipitation analyses in both magnitude and spatial patterns.

While this work focuses on aggregate statistics calculated across each WRF modeling domain, regional differences for a broader range of dates might be illustrative of the relative strengths and weaknesses of the LDA and LDA-CTH approaches.

1 Introduction

1.1 Project Objectives

The purpose of this project is to improve TCEQ's simulation of ozone photochemistry by applying a proven lightning data assimilation (LDA) method to the Weather Research and Forecasting (WRF) model. The method will force deep convection in the meteorological model where lightning is observed and only allow shallow convection where it is not, which will improve the representation of clouds in the meteorological and photochemical models. This is accomplished in WRF through the Multi-Scale Kain-Fritsch (MSKF) cumulus parameterization scheme and the version of WRF with this capability is referred to here as WRF-MSKF-LDA. The original version of this software was delivered to TCEQ in 2022 under Contract Number 582-19-90498, Work Order Number 582-22-31297-009 (Adams-Selin et al., 2022). The current version has been updated to use GOES Advanced Baseline Imager (ABI) cloud top height (CTH) observations.

The objectives of this project are thus to (1) update Python code for a previous project (Multiscale Lightning Data Assimilation for Improved Weather and Air Quality Modeling) to download, preprocess, and regrid Geostationary Lightning Mapper (GLM) lightning observations to TCEQ 2019 Modeling Platform WRF Domain (including processing cloud top height from the GOES ABI); and (2) add the updated GLM-LDA to the MSKF scheme in WRF, and (3) evaluate changes in its performance relative to simulations without the updated scheme.

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

Table 1. Schedule of Deliverables for Work Order No. 1

Milestones	Planned Date
Task 1 - Work Plan	
1.1: TCEQ-approved Work Plan	February 12, 2024
1.2: TCEQ-approved QAPP	February 12, 2024
Task 2 – Progress Reports	
2.1: Monthly Progress Reports	Monthly
Task 3 – Update Python Code for a previous project (Multiscale Lightning Data Assimilation for Improved Weather and Air Quality Modeling) to download, preprocess, and regrid GLM Lightning Observations to TCEQ 2019 Modeling Platform WRF Domain (including processing cloud top height from the GOES ABI).	
3.1: Updated Python code for GLM Data and User Guide	March 1, 2024
3.2: Python code for processing cloud top height from the GOES ABI and User Guide	March 1, 2024
Task 4 – Add Lightning Data Assimilation into WRF-MSKF	
4.1: WRF Code with updated GLM-MSKF-LDA, Example namelists, and script for updating to newer WRF versions	April 1, 2024

Milestones	Planned Date
4.2: Updated LDA User Guide	April 1, 2024
Task 5 – Evaluation of WRF with updated GLM-MSKF-LDA for June 2019	
5.1: WRF model setup and input files, WRF output files, final gridded GLM observations, and a PDF technical summary of the evaluation process and results to TCEQ	May 2, 2024
Task 6 – Training of TCEQ Staff to be able to run the updated LDA scheme	
6.1: Training Session and related training instructions provided to TCEQ staff	May 30, 2024
Task 7 – Draft and Final Reports	
7.1: Draft Report	June 10, 2024
7.2: Final Report	June 24, 2024

1.2 Background

A key deficiency of many retrospective meteorological simulations is the timing and location of convective rainfall. In addition to poor simulation of rainfall itself, misplacement of deep convection can negatively affect near-surface meteorology such as temperature, humidity, winds, and boundary layer height, all of which negatively impact air quality simulations.

Heath et al. (2016a) developed a simple LDA method for improving parameterized deep convection in retrospective weather simulations. The method has a straightforward approach to force deep convection where lightning is observed and only allow shallow convection where it is not. The LDA method has been used to improve air quality simulations in the Community Multiscale Air Quality (CMAQ) model (Heath et al., 2016b) and in the modeling system used by the Environmental Protection Agency (EPA) (e.g., Kang et al., 2020 and Pleim et al., 2019). The LDA method for this project improves upon the previous work by applying the technique at higher resolutions and using publicly available satellite-derived lightning observations from the GLM (Goodman et al., 2013). In WRF, the height of the convection is determined internally by the convective parameterization. In this project the cloud top height will be estimated from the Geostationary Operational Environmental Satellite (GOES) Advanced Baseline Imager (ABI) and will be used as a constraint for the parameterized convective depth.

1.3 Report Outline

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 7.2 of the Work Plan for Work Order No. 1 under TCEQ Contract 582-23-45974:

Deliverable 7.2: Final Report

Deliverable 7.2 Due Date: June 24, 2024

2 Update Python Code to Download, Process, and Regrid GLM and CTH Data to TCEQ 2019 Modeling Platform WRF Domain

In this task, AER updated previously developed Python code (Adams-Selin et al., 2022) to download GOES ABI CTH data in addition to GLM lightning observations from the National Center for Environmental Information (NCEI). With the exception of CTH data prior to December 2019, all data are hosted on the Amazon Web Services and Google Cloud. All data was preprocessed to reduce storage, and regridded to the TCEQ 2019 Modeling Platform WRF domain. User Manuals were provided to the TCEQ with detailed instructions for data download and processing (Hegarty & Dayalu, 2024ab). The remainder of this section provides a high-level summary of those documents and the main outcomes.

2.1 Update Python Code for GLM Data

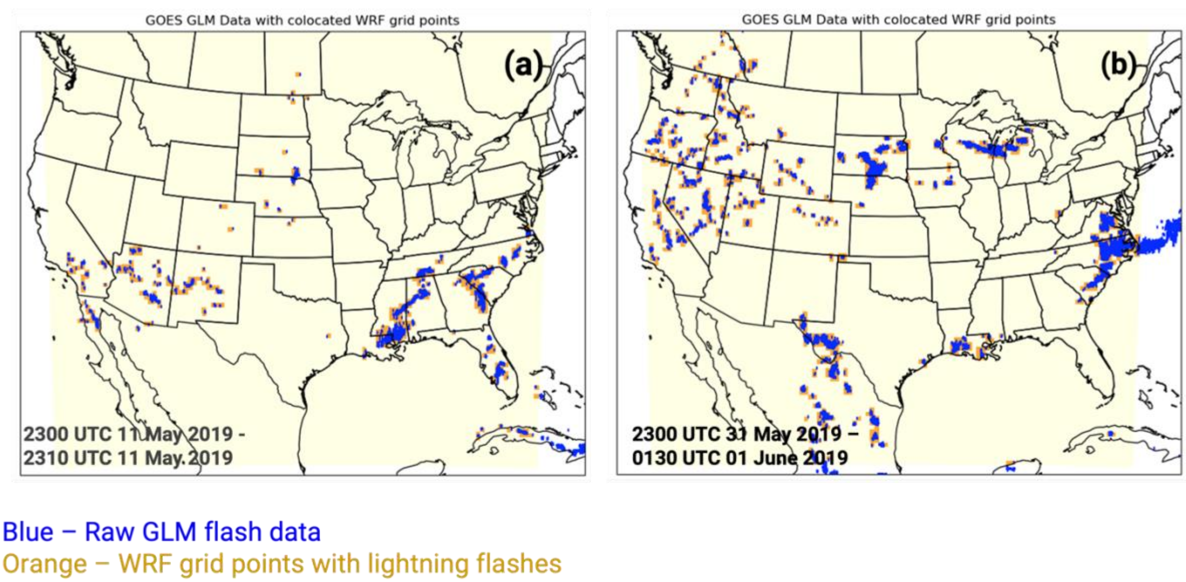

Figure 1. Raw GLM lightning flash data (blue dots) and WRF grid points (orange) corresponding with each flash. Flashes shown are from (a) 2300 UTC 11 May 2019 – 2310 UTC 11 May 2019; and (b) 2300 UTC 31 May 2019 – 0130 UTC 01 June 2019. Light yellow shows the full extent of the WRF domain.

The GLM User Guide (Hegarty & Dayalu, 2024a) describes the code provided by AER to download, process, and regrid publicly available observations from the GLM into the input format required by the WRF LDA method. Three Python scripts were provided to the TCEQ: (1) get_data.py, to retrieve the data; (2) regrid_ltg.py, to regrid the data to the WRF grid; and (3) gcdistance.py, a function called by regrid_ltg.py that checks that the regrid process worked correctly. As the LDA method only requires a yes/no for the presence of lightning to trigger convection, AER's Python code extracted the lightning flash data from the GLM to serve as that yes/no indicator for each grid cell.

Each lightning timestep includes any lightning flash occurrences between that time and 30 min into the future. The 10-min update cadence and 30-min accumulation interval was chosen to coincide with Heath et al. (2016a) and its original development of the lightning data assimilation method. The inclusion of 30 minutes into the future at each timestep allows the convective parameterization to be turned on prior to the appearance of lightning, as convection typically initiates some time before the first lightning flash. The 10-min update cadence will be tested during upcoming tasks to evaluate the sensitivity of the LDA method to that selection. In addition, the WRF namelist itself allows for specification of time in minutes between lightning data – it is currently defaulted at 10 minutes to mirror the GOES CTH data which is only available every 10 minutes. While the LDA can be run without CTH data and therefore can have a higher frequency (i.e., <10-min) update cadence, our tests have shown the 10-min time interval is adequate. Hegarty et al. (2024c) provides instructions on how to adjust the 10-min time interval for the lightning data in the absence of CTH data.

Methods for installing the libraries necessary to run the codes on TCEQ computing systems are also provided in Hegarty & Dayalu (2024a), along with quality assurance (QA) of the gridded lightning data when compared to the point lightning flash data. Figure 1 provides an example of QA for lightning data processing for two sample dates beginning on 11 May 2019 (Figure 1a) and 31 May 2019 (Figure 1b). Figure 1 shows the successful co-location of blue raw GLM points in the WRF CONUS modeling domain. Figure 1 also illustrates its value as a QA tool. For example, there are a few grid points such as in eastern Texas in Figure 1a where the lightning flash appears to occur outside the grid cell. However, further examination revealed the activated WRF grid cell was indeed the grid point closest to the raw flash location.

The ultimate output from the GLM data processing framework are WRF-compatible daily NetCDF files regridded to the specified WRF domain. Their nomenclature is ltgda_<WRFdomainID>_<yyyy-mm-dd>.nc.

2.2 Update Python Code for CTH Data

The CTH User Guide (Hegarty & Dayalu, 2024b) describes the code provided by AER to download, process, and regrid publicly available data from the GOES ABI Level 2 CTH product into the input format required by the WRF LDA method. Analogous to the GLM processing, the three Python scripts delivered to the TCEQ that accompany the User Guide retrieve the CTH data (get_data_cth.py); regrid the CTH data to WRF grid (regrid_ltg_cth.py); and the gcdistance.py QC function. As with the GLM data, methods for installing the libraries necessary to run the codes on TCEQ computing systems are also provided, along with quality assurance of the CTH regridded to the WRF 12-km domain compared to the CTH data on the native GOES ABI grid.

The CTH data product is only available at 10-min resolution.

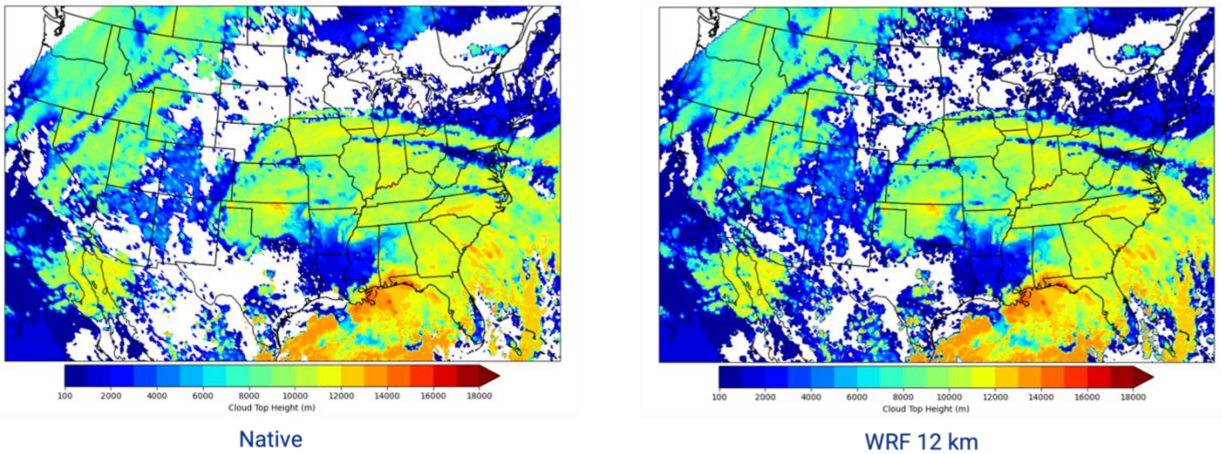

Figure 2. GOES ABI CTH Data on (left) native grid and (right) WRF 12km grid for a sample date of 6 June 2019 at 1850 UTC. Note that in the northwest corner of the CONUS there is never any valid cloud top height data for GOES East. This is because the local zenith angle (LZA) is too large to do a high-quality cloud top height retrieval.

There are a few additional notes for the CTH data that are not relevant to the lightning flash data. First, unlike the binary GLM lightning flash data which are associated with discrete latitude/longitude coordinates, CTH data is on the GOES ABI grid which requires the additional "xarray" python module and longer initial compute time to process to the WRF grid. Second, CTH data is only publicly accessible beginning in December 2019, and the automated get_data_cth.py code works only for times during or after December 2019. In this work where we restricted our analysis to the Adams-Selin et al. (2022) June 2019 case study, we had to separately register and download archived CTH data from the NOAA data archive system. See Hegarty & Dayalu (2024b) for more details.

Figure 2 illustrates the successful regridding of GOES ABI CTH data from its native grid to the WRF 12 km domain for a sample date (6 June 2019 at 1850 UTC).

The ultimate output from the CTH data processing framework is WRF-compatible daily NetCDF files regridded to the specified WRF domain. Their nomenclature is cthda_<WRFdomainID>_<yyyy-mm-dd>.nc.

3 Add LDA into WRF MSKF

In this task, AER updated the existing WRF-MSKF-LDA scheme from Adams-Selin et al. (2022) to incorporate new constraints provided by the CTH data. AER provided: (1) a detailed User Manual to the TCEQ describing changes to the WRF code; (2) sample WRF namelist files with modifications to run WRF-MSKF-LDA; and (3) script to combine WRF-compatible LDA and CTH netCDF files into a single WRF input file; and (4) script to update WRF-MSKF-LDA with newer releases of WRF. The User Manual (Hegarty et al., 2024c) provides detailed run instructions and background for all these components. The procedures described in Hegarty et al. (2024c) assume that gridded lightning and CTH data files have been created as outlined in Section 2.

There have been several changes to the Hegarty et al. (2024c) WRF-MSKF-LDA User Guide since Revision 1 (as Adams-Selin et al., 2022) was delivered to the TCEQ in

FY2022. In addition to minor corrections and clarifications being made to the text where needed, the following major revisions have been made:

- ***A provision has been added to combine the lightning data with the CTH data into a single WRF-readable input file.*** A python script to combine the LDA and CTH NetCDF files into a file consistent with the WRF namelist file were provided to the TCEQ to facilitate a run of WRF-MSKF-LDA. The output of this script is a NetCDF file of nomenclature ltgda_cthda_<WRFdomainID>_<yyyy-mm-dd>.nc. The option for including this file has been added to the *&time_control* section of the WRF *namelist.input* file. For reference, an example combined gridded LDA and CTH file was provided to the TCEQ as part of the Task 4 Deliverables.
- ***A flag for turning on and off the use of CTH in the WRF-MSKF-LDA has been added into the WRF-MSKF-LDA namelist.*** The CTH flag has been added to the *&physics* section of the WRF *namelist.input* file with three options for convection suppression (*suppress_opt*) in the event lightning is not present. In all cases where LDA is turned on, the presence of lightning forces deep convection with an upper constraint provided by CTH when the CTH flag is turned on. Where lightning is not present, however, a *suppress_opt* = 0 tells WRF to run the MSKF scheme as normal; a *suppress_opt* = 1 tells WRF to skip the MSKF scheme entirely; a *suppress_opt* = 2 tells WRF to run only the shallow part of the MSKF scheme. *suppress_opt* = 2 is the recommended option for most cases. The *suppress_opt* = 1 where MSKF is skipped entirely for non-lightning cases is generally not recommended. Figure 3 summarizes basic QC results along the impacts of the two recommended suppression options (0, 2) on WRF estimated convective precipitation (RAINC).

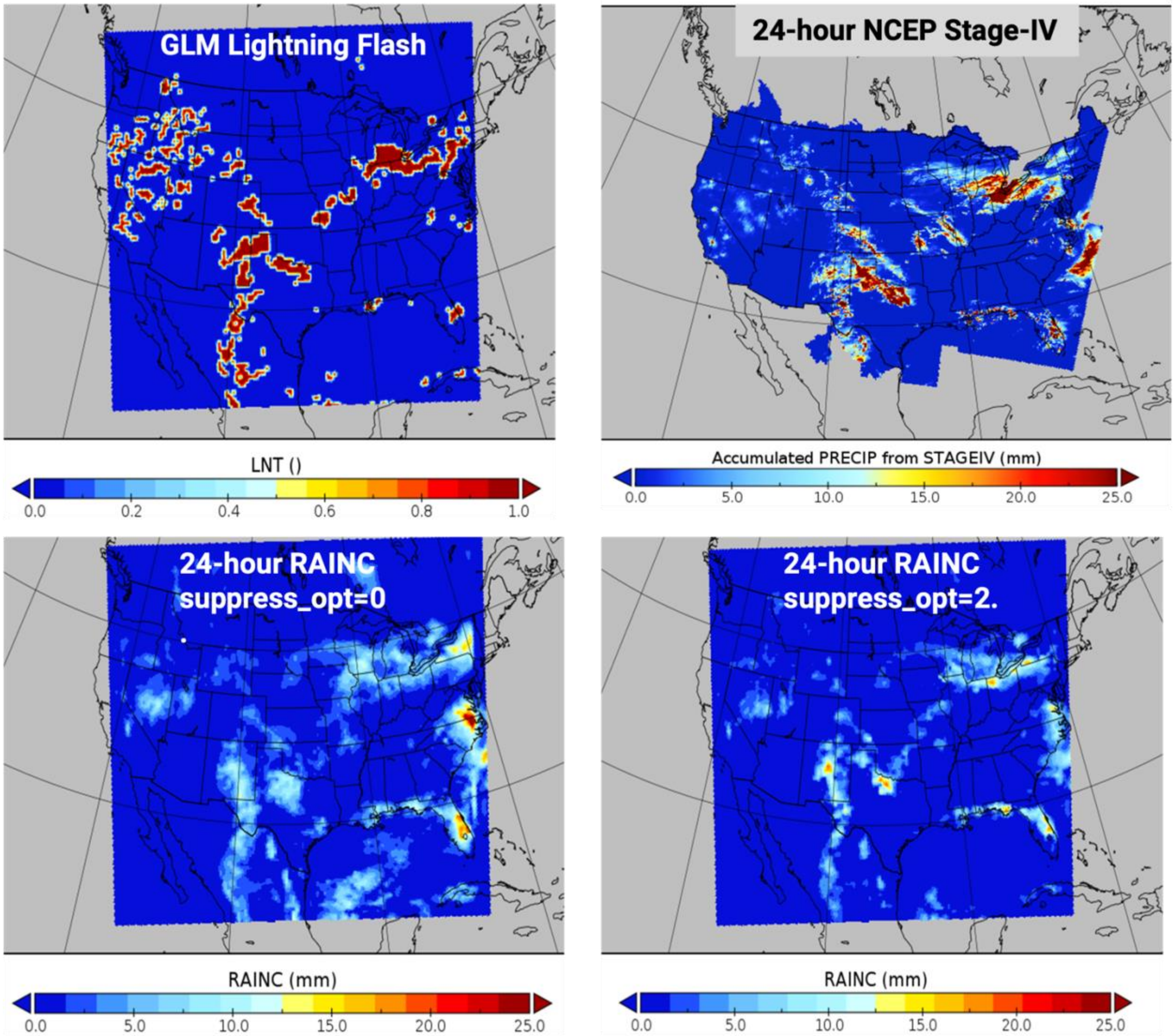

Figure 3. Impact of convection suppression options on Accumulated 24-hour WRF-MSKF-LDA-CTH RAINC precipitation 24-hour WRF-MSKF-LDA simulation beginning 0000 UTC 1 June 2019. (Top Left): GLM lightning flash data on WRF grid. (Top Right): Accumulated 24-hour NCEP Stage IV precipitation for reference. (Bottom Left): RAINC precipitation estimated using *suppress_opt*=0. (Bottom Right): RAINC precipitation estimated using *suppress_opt=2*.

- ***An evaluation of the impact on convective and non-convective precipitation of including the CTH data in the LDA has been added.*** Generally, the inclusion of CTH data into the WRF-MSKF-LDA leads to an increase in RAINC and RAINNC across CONUS. Figures 4 and 5 summarize results of this analysis for a sample 24-hour period beginning 6 June 2019 0700 UTC. This date was selected as it was cited in Adams-Selin et al. (2022) for under simulating RAINC and so it is encouraging that the inclusion of CTH increases, albeit slightly, the simulated RAINC, and provides confidence that the CTH code added to the MSKF module in

WRF is working properly (Figure 4). In addition, the banded structure of the RAINNC in Alabama and Mississippi for the case with the CTH seems to be more consistent with NCEP Stage IV radar-based precipitation estimates (Figure 5).

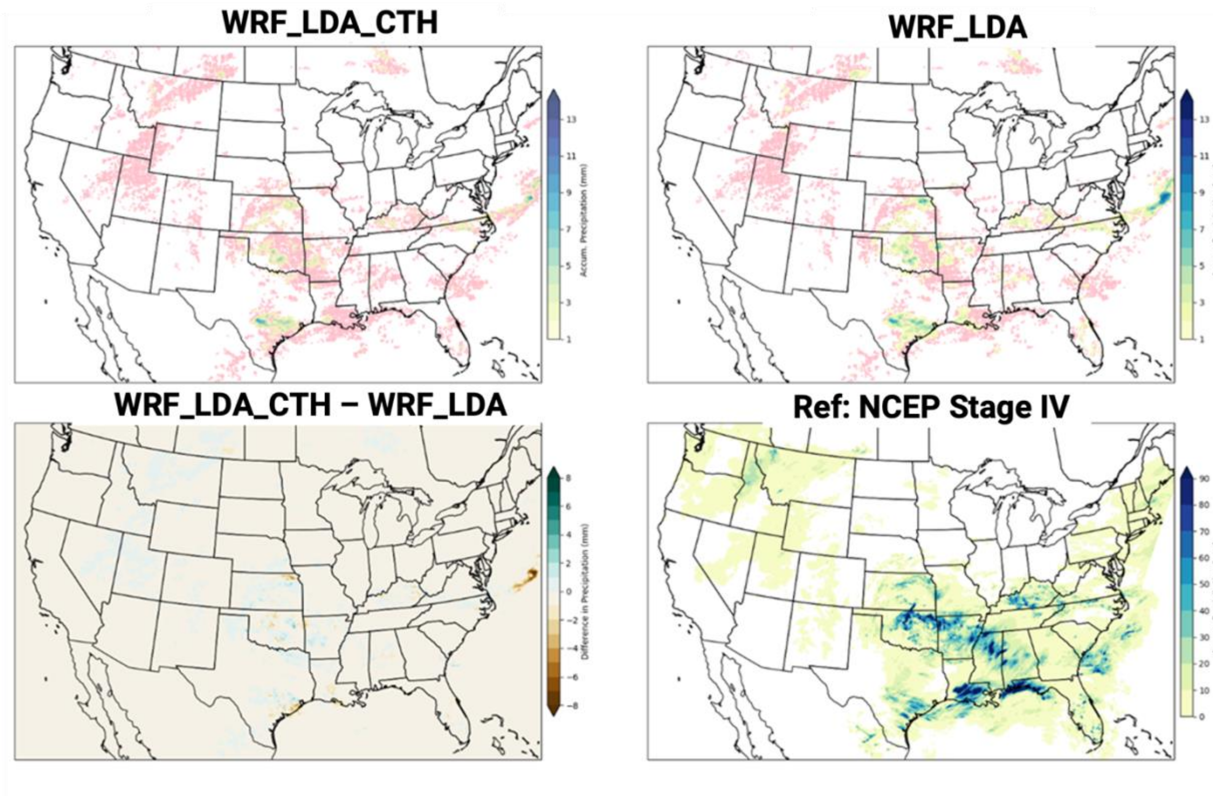

Figure 4. WRF 24 hour accumulated RAINC (mm) for the period ending 0600 UTC 7 June 2019 over the CONUS portion of the 12 km domain of the TCEQ 2019 WRF Modeling Platform for WRF LDA with CTH (upper left), WRF LDA without CTH (upper right), Difference between RAINC in LDA with CTH – LDA without CTH (lower left). For reference, the NCEP Stage IV precipitation analysis is shown in the lower right.

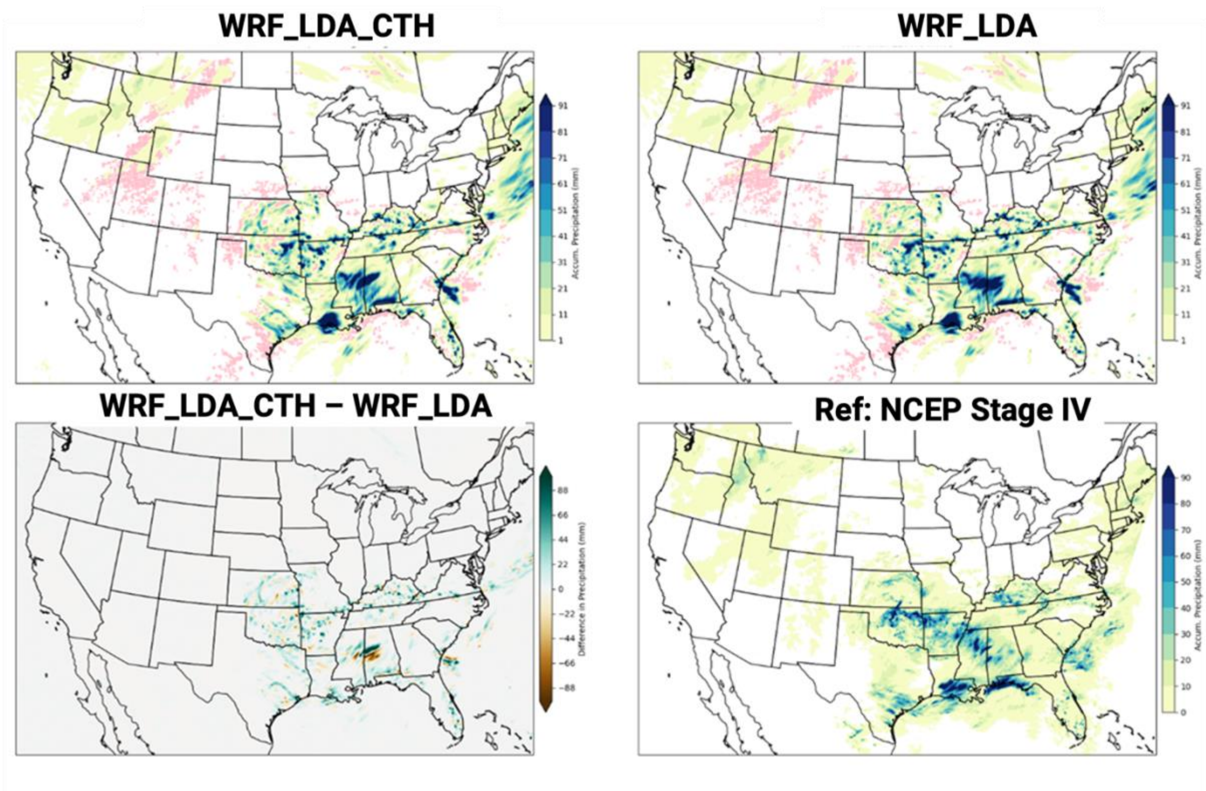

Figure 5. WRF 24 hour accumulated RAINNC (mm) for the period ending 0600 UTC 7 June 2019 over the CONUS portion of the 12 km domain of the TCEQ 2019 WRF Modeling Platform for WRF LDA with CTH (upper left), WRF LDA without CTH (upper right), Difference between RAINC in LDA with CTH – LDA without CTH (lower left). For reference, the NCEP Stage IV precipitation analysis is shown in the lower right.

- ***An option to update newer versions of WRF with the LDA/LDA-CTH scheme has been added.*** Hegarty et al. (2024c) provides instructions on updating newer versions of WRF to include the MSKF-LDA option developed in this work. The update option uses the shell script called "update_to_new_WRF.sh" and assumes a compute environment that is set up with GitHub. Using the shell script requires: 1) a base version without the MSKF-LDA modifications (e.g., in this work it would have been WRFv4.3.3), 2) a version with the MSKF-LDA modifications that is the same WRF version as the base version (e.g., WRFv4.3.3_MSKF-LDA, provided via the Hegarty et al. 2024c deliverable) and 3) a new WRF version without the WRF-LDA modifications (e.g., WRFv4.4).

4 Evaluation of WRF with updated GLM-MSKF-LDA for June 2019

In this task, AER compared precipitation, temperature, and wind speed forecasts from the WRF-MSKF-LDA configuration (both with and without CTH observations) to the unchanged (standard) WRF-noLDA configuration (Hegarty et al., 2024d). The MSKF convective parameterization is used for all runs. Changes since Revision 1 (as Adams-Selin et al., 2022) include (1) use of WRFv4.3.3 instead of WRFv4.2.2 and

WRFv4.3.1; and (2) the addition of the WRF-LDA-CTH runs in the comparison with WRF and WRF-LDA.

Overall, no statistically significant differences in verification statistics between the WRF-LDA-CTH and WRF-LDA runs were found for any of the variables. However, statistically significant differences were found in precipitation between the WRF-noLDA case and the other two cases for both domains. For 24-hour precipitation in Domain 1, the mean error shifted from -0.21 mm for WRF-noLDA to 0.17 mm for both WRF-LDA and WRF-LDA-CTH. In Domain 2 the mean error reduced, shifting from -0.87 mm for the WRF-noLDA to 0.37 mm and 0.41 mm for the WRF-LDA and WRF-LDA-CTH respectively. For 6-hour precipitation in Domain 1 the mean error shifted from -0.05 mm underprediction for the WRF-noLDA to ~0.05 mm in the other two cases. In addition, the WRF-noLDA case had a statistically significant lower error standard deviation than the other two cases. In Domain 2 the 6-hour precipitation mean error shifted from -0.22 mm to 0.09 mm and 0.10 mm overpredictions for the WRF-LDA and WRF-LDA-CTH respectively. Sections 4.1 through 4.6 present the results for each of the temperature, wind, and precipitation variables considered in this work. Lower and upper confidence limits are calculated assuming a bivariate normal distribution of the two fields (forecast and observations) with an alpha (a) value of 0.05. The value selected for a indicates that if the test were repeated many times, 100 ×(1- a) of the skill scores (here, 95%) would fall within the confidence interval. Statistical significance was evaluated with a two-tailed T-Test with a confidence interval of p=0.05. Results where there is a statistically significant difference between the skill scores are shown in ***bold italics*** font.

We note that for the June 2019 case study, the comparisons are provided as aggregate differences for each of WRF domains 1 and 2. While the spatial and temporally aggregated differences were small and/or insignificant across all variables, they could be regionally significant, including when comparing specific events. For example, referring back to Figures 4 and 5 there are pockets of potentially significant precipitation differences in southeast Texas.

4.1 Temperature Statistics

Table 2. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for surface temperature field aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA-CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA-CTH	WRF-noLDA
Pearson Correlation Coeff.	0.8956	0.8949	0.8949
lower confidence limit	0.8868	0.8860	0.8860
upper confidence limit	0.9038	0.9031	0.9031
Mean Error (K)	-0.0792	-0.0795	-0.0640
lower confidence limit (K)	-0.1849	-0.1857	-0.1699
upper confidence limit (K)	0.0266	0.0266	0.0418
Error Standard Deviation (K)	2.451	2.459	2.453
lower confidence limit (K)	2.378	2.386	2.380
upper confidence limit (K)	2.528	2.537	2.530

Table 3. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for surface temperature field aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.6677	0.6658	0.6596
lower confidence limit	0.5923	0.5901	0.5833
upper confidence limit	0.7321	0.7306	0.7251
Mean Error (K)	0.2321	0.2246	0.4026
lower confidence limit (K)	-0.0189	-0.0274	0.1520
upper confidence limit (K)	0.4831	0.4767	0.6533
Error Standard Deviation (K)	1.855	1.862	1.852
lower confidence limit (K)	1.693	1.700	1.690
upper confidence limit (K)	2.052	2.060	2.048

4.2 Dewpoint Temperature

Table 4. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for dewpoint temperature field aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.9282	0.9281	0.9286
lower confidence limit	0.9220	0.9218	0.9225
upper confidence limit	0.9339	0.9338	0.9343
Mean Error (K)	-0.0604	-0.0727	0.0592
lower confidence limit (K)	-0.1769	-0.1892	-0.0578
upper confidence limit (K)	0.0561	0.0438	0.1762
Error Standard Deviation (K)	2.696	2.696	2.708
lower confidence limit (K)	2.616	2.616	2.627
upper confidence limit (K)	2.781	2.781	2.793

Table 5. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for dewpoint temperature field aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.7289	0.7294	0.7265
lower confidence limit	0.6608	0.6612	0.6581
upper confidence limit	0.7854	0.7859	0.7834
Mean Error (K)	0.2907	0.2788	0.4413
lower confidence limit (K)	0.0435	0.0309	0.1968
upper confidence limit (K)	0.5378	0.5267	0.6859
Error Standard Deviation (K)	1.827	1.832	1.807
lower confidence limit (K)	1.667	1.672	1.649
upper confidence limit (K)	2.020	2.026	1.999

4.3 U Wind Speed

Table 6. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for U surface wind speed component aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.6324	0.6304	0.6427
lower confidence limit	0.6057	0.6035	0.6166
upper confidence limit	0.6577	0.6558	0.6675
Mean Error (m/s)	0.1499	0.1512	0.1322
lower confidence limit (m/s)	0.0541	0.0551	0.0392
upper confidence limit (m/s)	0.2457	0.2474	0.2252
Error Standard Deviation (m/s)	2.181	2.189	2.118
lower confidence limit (m/s)	2.115	2.123	2.054
upper confidence limit (m/s)	2.251	2.259	2.186

Table 7. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for U surface wind speed component aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.4428	0.4396	0.4637
lower confidence limit	0.3301	0.3265	0.3533
upper confidence limit	0.5434	0.5406	0.5616
Mean Error (m/s)	0.0941	0.0922	0.0721
lower confidence limit (m/s)	-0.1652	-0.1676	-0.1739
upper confidence limit (m/s)	0.3534	0.3520	0.3181
Error Standard Deviation (m/s)	1.893	1.896	1.795
lower confidence limit (m/s)	1.725	1.728	1.636
upper confidence limit (m/s)	2.096	2.100	1.988

4.4 V Wind Speed

Table 8. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for V surface wind speed component aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.6541	0.6523	0.6645
lower confidence limit	0.6287	0.6268	0.6397
upper confidence limit	0.6781	0.6764	0.6879
Mean Error (m/s)	0.1511	0.1494	0.1645
lower confidence limit (m/s)	0.0503	0.0482	0.0664
upper confidence limit (m/s)	0.2520	0.2505	0.2655
Error Standard Deviation (m/s)	2.296	2.302	2.233
lower confidence limit (m/s)	2.227	2.233	2.166
upper confidence limit (m/s)	2.370	2.376	2.305

Table 9. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for V surface wind speed component aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.4626	0.4590	0.4778
lower confidence limit	0.3540	0.3502	0.3714
upper confidence limit	0.5594	0.5562	0.5824
Mean Error (m/s)	0.3530	0.3459	0.4086
lower confidence limit (m/s)	0.0558	0.0472	0.1243
upper confidence limit (m/s)	0.6501	0.6445	0.6930
Error Standard Deviation (m/s)	2.169	2.180	2.076
lower confidence limit (m/s)	1.978	1.988	1.892
upper confidence limit (m/s)	2.402	2.415	2.299

4.5 24-hour Precipitation

24-h accumulated precipitation is the sum of all rain including convective, non-convective, and shallow cumulus precipitation (i.e., RAINC+RAINNC+RAINSH).

Table 10. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for 24-hour accumulated precipitation aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.4061	0.4003	0.4103
lower confidence limit	0.3993	0.3935	0.4035
upper confidence limit	0.4129	0.4071	0.4170
Mean Error (mm)	0.1758	0.1726	***-0.2076***
lower confidence limit (mm)	0.0896	0.0856	-0.2825
upper confidence limit (mm)	0.2620	0.2596	-0.1326
Error Standard Deviation (mm)	10.51	10.61	9.138
lower confidence limit (mm)	10.45	10.55	9.085
upper confidence limit (mm)	10.57	10.67	9.191

Table 11. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for 24-hour accumulated precipitation aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.2934	0.2936	0.2675
lower confidence limit	0.2855	0.2858	0.2594
upper confidence limit	0.3012	0.3014	0.2754
Mean Error (mm)	0.3732	0.4130	***-0.8746***
lower confidence limit (mm)	0.2548	0.2954	-0.9850
upper confidence limit (mm)	0.4916	0.5306	-0.7642
Error Standard Deviation (mm)	13.35	13.26	12.45
lower confidence limit (mm)	13.27	13.18	12.37
upper confidence limit (mm)	13.43	13.34	12.53

4.6 6-hour Precipitation

As with 24-h accumulated precipitation, 6-h accumulated precipitation is the sum of all rain including convective, non-convective, and shallow cumulus precipitation (i.e., RAINC+RAINNC+RAINSH).

Table 12. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for 6-hour accumulated precipitation aggregated across domain d01 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.2947	0.2847	0.2875
lower confidence limit	0.2873	0.2772	0.2801
upper confidence limit	0.3021	0.2921	0.2948
Mean Error (mm)	0.0490	0.0476	***-0.0472***
lower confidence limit (mm)	0.0088	0.0070	-0.0824
upper confidence limit (mm)	0.0891	0.0883	-0.0121
Error Standard Deviation (mm)	4.902	4.958	***4.285***
lower confidence limit (mm)	4.873	4.929	4.260
upper confidence limit (mm)	4.930	4.987	4.310

Table 13. Comparison of results among WRF-LDA, WRF-LDA_CTH, and WRF-noLDA for 6-hour accumulated precipitation aggregated across domain d02 for June 2019. Instances where WRF-LDA and/or WRF-noLDA differ significantly from WRF-LDA_CTH are indicated by ***bold italic*** font.

Statistic	WRF-LDA	WRF-LDA_CTH	WRF-noLDA
Pearson Correlation Coeff.	0.1554	0.1543	0.1541
lower confidence limit	0.1470	0.1459	0.1457
upper confidence limit	0.1638	0.1627	0.1625
Mean Error (mm)	0.0932	0.1035	***-0.2214***
lower confidence limit (mm)	0.0411	0.0512	-0.2685
upper confidence limit (mm)	0.1453	0.1557	-0.1742
Error Standard Deviation (mm)	5.876	5.889	5.313
lower confidence limit (mm)	5.839	5.852	5.280
upper confidence limit (mm)	5.913	5.926	5.346

5 Quality Assurance

Analysis of all plots and statistics in this report provided a subjective assessment that algorithms were correctly implemented. 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

Here we summarize the conclusions of our project, with reference to the corresponding report section.

- Overall, all quality control checks indicate the updated WRFv4.3.3-LDA-CTH is working appropriately and is readily ported to newer versions of WRF.
- For the June 2019 case study, no statistically significant differences were found in temperature, dewpoint and wind surface variables. (Aggregated, examined for both domains).
- WRF-LDA-CTH generally increases RAINC relative to WRF-LDA by a few millimeters across CONUS.
    - For the June 2019 case study, the aggregate differences were not significant, but could be regionally significant.
    - The June 6/7 2019 case was previously identified by Adams-Selin et al. (2022) as underestimating RAINC. It was encouraging that CTH tends to increase (albeit slightly) RAINC relative to WRF-LDA and WRF-noLDA.
    - WRF-LDA-CTH spatial patterns tend to have higher agreement with NCEP Stage IV than WRF-LDA or WRF-noLDA.

7 Recommendations for Future Work

Based on the results of this work, we make the following recommendations for further study:

- Expanding beyond the June 2019 case study to other seasons and specific events with a regional focus will provide important information on the extent to which incorporating LDA and/or CTH adds value.

8 References

Adams-Selin, R., Hegarty, J., Pernak, R., Fanny, E., and Alvarado, M. (2022), Multiscale Lightning Data Assimilation for Improved Weather and Air Quality Modeling, Final Report to Texas Commission on Environmental Quality (TCEQ) for Work Order No. 582-22-31297-009 under TCEQ Contract No. 582-19-90498, June 30.

Goodman, S. J., and Coauthors (2013), The GOES-R Geostationary Lightning Mapper (GLM). Atmos. Res., 125–126, 34–49. https://doi.org/10.1016/j.atmosres.2013.01.006.

Heath, N. K., Pleim, J. E., Gilliam, R. C., and Kang, D. (2016a), A simple lightning assimilation technique for improving retrospective WRF simulations, J. Adv. Model. Earth Syst., 8, 1806– 1824, doi:10.1002/2016MS000735.

Heath N.K., Pleim, J.E., Gilliam, R.C., Kang, D., Woody, M., Foley, K., and Appel, W. (2016b), Impacts of WRF lightning assimilation on offline CMAQ simulations. Paper presented at 15th Annual CMAS Conference, Chapel Hill, NC.

Hegarty, J. D., R. R. Draxler, A. F. Stein, J. Brioude, M. Mountain, J. Eluszkiewicz, T. Nehrkorn, F. Ngan, and A. E. Andrews (2013), Evaluation of Lagrangian particle dispersion models with measurements from controlled tracer releases, J. Appl. Meteor. Climatol., 52(12), 2623–2637, doi:10.1175/JAMC-D-13-0125.1.

Hegarty, J., A. McVey, J. Henderson, and M. Alvarado (2018), Evaluation of High Resolution WRF Modeling for El Paso, Texas, Final Report to Texas Commission on Environmental Quality (TCEQ) for Work Order No. 582-18-81899-09 under TCEQ Contract No. 582-15-50414, July 30.

Hegarty, J., Dayalu, A. (2024a). Geostationary Lightning Mapper (GLM) Processing for Lightning Data Assimilation (LDA) in WRF, User Guide provided to the Texas Commission on Environmental Quality (TCEQ) for Work Order No. 1 under TCEQ Contract No. 582-23-45974, 1 March.

Hegarty, J., Dayalu, A. (2024b). GOES Cloud Top Height (CTH) Processing for Improved Lightning Data Assimilation (LDA) in WRF, User Guide provided to the Texas Commission on Environmental Quality (TCEQ) for Work Order No. 1 under TCEQ Contract No. 582-23-45974, 1 March.

Hegarty, J., Dayalu, A. (2024c). WRF Multi-Scale Kain Fritsch (MSKF) Lightning Data Assimilation, User Guide provided to the Texas Commission on Environmental Quality (TCEQ) for Work Order No. 1 under TCEQ Contract No. 582-23-45974, 1 April.

Hegarty, J., Adams-Selin, R., Dayalu, A. (2024d). WRF Multi-Scale Kain Fritsch (MSKF) Lightning Data Assimilation Evaluation, Technical Memo provided to the Texas Commission on Environmental Quality (TCEQ) for Work Order No. 1 under TCEQ Contract No. 582-23-45974, 2 May.

Kang, D., Mathur, R., Pouliot, G.A. et al. (2020), Significant ground-level ozone attributed to lightning-induced nitrogen oxides during summertime over the Mountain West States. npj Clim. Atmos. Sci., 3, 6 (2020). https://doi.org/10.1038/s41612-020-0108-2.

Pleim, J. E., Ran, L., Appel, W., Shephard, M. W., and Cady-Pereira, K. (2019), New bidirectional ammonia flux model in an air quality model coupled with an agricultural model. J. Adv. Model. Earth Sys., 11, 2934– 2957. https://doi.org/10.1029/2019MS001728.