Scope of Work For

Project # 20-026
Improve Cloud Modeled by WRF using COSP and Generative Adversarial Network

Prepared for

Air Quality Research Program (AQRP)
The University of Texas at Austin

By

Principal Investigator, Zheng Lu
Department of Atmospheric Sciences, Texas A&M University

May-11-2020
Version #1

QA Requirements: Audits of Data Quality: 10% Required
Report of QA Findings: Required in Final Report

Approvals

This Scope of Work was approved electronically on **06/22/20** by Elena McDonald-Buller, The University of Texas at Austin

Elena McDonald-Buller
Project Manager, Texas Air Quality Research Program

This Scope of Work was approved electronically on **06/30/20** by Bright Dornblaser, Texas Commission on Environmental Quality

Bright Dornblaser
Project Liaison, Texas Commission on Environmental Quality

TABLE OF CONTENTS

1.0 Abstract

The cloud fields modeled by meso-scale models play an important role in the application of predicting local air quality. The cloud fields can strongly affect the formation, transportation, as well as deposition of many gaseous and particulate species, through regulating radiative transfer, influencing aqueous chemistry, and altering precipitation. However, it is very challenging to accurate predict the microphysical and macrophysical properties of cloud fields.

In this proposal, we plan to run **WRF** model with Texas in the center of model domain. Modeled cloud fields are feed into Cloud Feedback Intercomparison Project (CFMIP) Observation Simulator Package (**COSP**), so that modeled cloud can be directly compared to satellite observations. The objective is to select optimal combination of initiation state (the selection of reanalysis data) and physical packages (namely microphysics, cumulus parameterization, planetary boundary layer scheme) for the cloud simulation.

With modeled and observed cloud fields, we train a **GAN** (Generative Adversarial Network), a type of deep learning technique. We will perform super-resolution and image-to-image translation applications to modeled cloud microphysical fields over Texas, so that they can gain much detailed fine features, and become more accurate compared to observed cloud fields. Improved cloud fields will undoubtedly improve Texas air quality prediction.

2.0 Background

This proposal targets the research area as listed in Texas Air Quality Research Program (AQRP) guideline: "**Meteorological Inputs for Modeling**" with the goal "**to support scientific research related to Texas air quality, in areas of …meteorology and air quality modeling.**"

Cloud fields play an important role in regulating the formation, transportation, and lifetime of gas and particulate pollutants [e.g. Liang and Jacob, 1997; Gurciullo and Pandis, 1997; Fan et al., 2004]. The radiation field altered by cloud controls the photolysis reaction associated with ozone formation [Faust, 1994]. Photolysis rate is dependent on many factors that are capable of influencing solar actinic flux. Many previous studies have shown that the vertical distribution of aerosol and cloud, and their optical properties can have significant impacts on photolysis rate [E.g. Liao et al., 1999; Lefer et al., 2003; Tie et al., 2003; 2005; Liu et al., 2006]. In general, actinic flux and thereby photolysis rates are reduced below aerosol or cloud layer due to their extinction. On the other hand, over bright cloud the strong cloud reflection can increase the photolysis rate. The impacts on photolysis rate can in turn influences the photochemistry of ozone. Using a photochemical box model driven by airborne measurement from the TRACE-P mission, Lefer et al. [2003] showed that during the TRACE-P mission the net photochemical effect of clouds and aerosols was a large decrease in photochemical O_3 production in the boundary.

The interactions between particulate matters suspended in the air – or atmospheric aerosols with cloud fields are complicated and extremely important for climate as well as air quality application [Rosenfeld, et al., 2014, Fan et al., 2016; Seinfeld et al., 2016]. The cloud droplets must be nucleated from aerosol particles, which are referred as cloud condensation nuclei (CCN) if activated. Heavy pollution condition in metropolitan areas can enhance cloud droplet number concentration (CDNC). For fixed amount of liquid water mass, higher CDNC leads to smaller cloud droplets, which can reduce precipitation efficiency. As the precipitation falling, the raindrops can wash off the aerosol particles below the clouds. Aerosol embedding inside cloud droplets and wet removal of aerosols by raindrops, referred to as in-cloud and below cloud scavenging, represent important sink terms of atmospheric aerosols. Therefore, to a large extent, the atmospheric aerosols can control cloud microphysical and macrophysical properties, and vice versa.

Aqueous chemistry is another reason that cloud is important for air quality application. For example, SO2 mass can be efficiently transferred to sulfate aerosols via cloud processing [Wine et al., 1989; Feingold and Kreidenweis, 2002]. To sum up the aforementioned discussion, as an input to the air quality models, accurate representation of cloud fields, including their macro and microphysical properties by model is essential for the air quality prediction application.

Modeled clouds are often too bright (high cloud brightness) and too few (low cloud fraction) compared to satellite observation [e.g. Otkin et al., 2008; Thompson et al., 2016]; however, the “general pictures” of cloud fields can be well captured by the meso-scale weather prediction models, for instance, convective frontal clouds associated with cyclonic-frontal system; or large decks of cumulus clouds over a large area when atmosphere is stable. For example, WRF model is widely used in simulating the meteorology and cloud fields that are essential for air quality prediction. WRF model parameterizations has been shown to lead to accurate simulations of southeast Texas mesoscale circulations [Ngan et al., 2013]. This indicates that over a relatively large area, the characteristic of modeled clouds is reasonable statistically and can be “adjusted” to match the observations. The direct comparison between modeled and observed cloud fields are like “apple-to-orange” comparison, because of different sampling rate. To facilitate the so-called “apple-to-apple” comparison, we must firstly use the tool called COSP [Bodas-Salcedo et al., 2011; Zhang et al., 2019].

As a new technique, the machine learning and deep learning (ML/DL) tools have not been widely used in geoscience, but have shown great potentials [Reichstein et al., 2019]. One of the advantages of ML/DL is that is data-driven – in other words, the more data we feed into the tools, the more accurate the results will be. With satellite observation, we have large amount of satellite data available for training the ML/DL tools. In this proposed work, we will use a ML/DL tool called Generative Adversarial Network (GAN) to “adjust” modeled cloud fields [Goodfellow et al., 2014].

3.0 Objectives

The two objectives of this proposal are:

(1) To conduct a series of WRF simulations as well COSP analysis to find a optimal combination of physics suite and reanalysis input for modeling clouds fields over Texas;
(2) To train a GAN model over the time series of modeled cloud fields so that the macro- and microphysical properties of modeled clouds are more accurate compared to observations.

4.0 Task Descriptions

Task 4.1. find optimal WRF model configuration for Texas

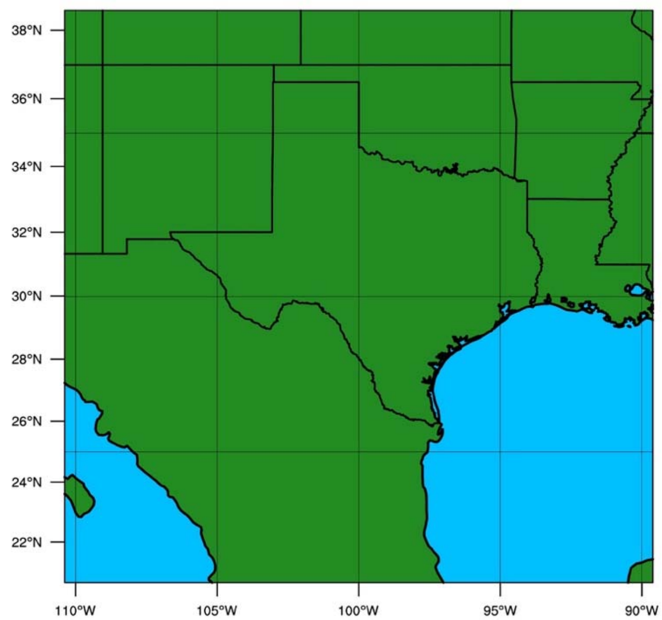

Figure 1 Potential domain setup with Texas in center

In this proposed work, we plan to use Weather Research and Forecasting model (WRF) [*Skamarock & Klemp*, 2008] to generate cloud fields, that can be used in air quality forecasting application. The domain will be set up with Texas in the center. The resolution can be flexible, varies according to the need of air quality prediction. Figure 1 shows an example of simulation domain, with 256 (west-east) by 256 (south-north) grids and 32 vertical levels. The horizontal resolution is 8 km. In the latest version of WRF model (after V3.9), a suite of physical packages is specifically recommended for simulations over CONUS (CONtinental U. S.). Namely, they are new Thompson microphysics scheme [Thompson et al., 2008], modified Tiedtke scheme for cumulus parameterization [Tiedtke, 1989]; Mellor-Yamada-Janjić TKE scheme for boundary layer scheme (PBL) [Janjić, 1994]; RRTMG radiation scheme for both shortwave and longwave radiation calculation [Iacono et al., 2008]; and unified Noah land-surface model [Koren et al., 1999]. For CONUS application, the initial and boundary conditions (IC and BC) of model is often driven by 6 - hourly 12 - km North American Mesoscale Analysis [e.g. Li et al., 2008].
However, this physics suite as well as reanalysis input may not be optimal for Texas application and/or cloud field simulations. Based on our previous research experience [Lu and Sokolik, 2013; Lu and Sokolik, 2017; Lu et al., 2018], three physics packages, namely microphysical scheme, cumulus parameterization, and PBL scheme, are most important physics packages that affect cloud simulation. The selection of re-analysis data also strongly affects large-scale dynamic and resulting cloud deck patterns. Therefore, here we propose to run several groups of one year of simulations with different combination of physics packages and reanalysis datasets,

the candidate of which are shown in Table 1. (To select which year depends on the availability of satellite products).

Table 1 Physics packages and reanalysis-data used for WRF simulation

Physical parameterization scheme		**Acronym**	**Reference**
Cumulus convective	Tiedtke	Tiedtke	Tiedtke [1989]
	Grell-Fretas	GF	Grell and Freitas [2014
	Multiscale Kain-Fritsch	msKF	Zheng et al. [2016]
Microphysics	1.5-moment 6-class Thompson	Thompson	Thompson et al. [2008]
	2-moment 6 class Morrison	Morrions	Morrions et al. [2009]
PBL	Mellor-Yamada-Janjic scheme	MYJ	Janjić [1994]
	Yonsei University scheme	YSU	Hong et al. [2006]
Reanalysis input	North American Mesoscale Analysis	NAMA	Rogers et al. [2009]
	NCEP final (FNL)	FNL	NCEP [2000]

Totally 3×2×2×2=24 groups of simulations will be performed and compared to satellite observations (more details of evaluation discussed in Section 2.2). After optimal combination of physics packages is selected, we use this physics suite with reanalysis data and conduct multiple years of simulation with the same domain setup. The results will be used in training a generative adversarial network (GAN).
Modeled cloud fields that we need from simulations are **cloud water path** (sum of liquid and ice water path, CWP, in kg m^{-2}); **cloud fraction** (CF, in %); **cloud top height** (CTH, in m) and **cloud optical thickness** (COT, unitless). These four cloud fields will be compared against satellite observations.

Responsible organization: TAMU team
Expected milestones, outcomes, and deliverables: 1) select the optimal combination of physical packages and reanalysis inputs for cloud simulation and 2) conduct long-term simulation with this optimal configuration.

Task 4.2 synergize COSP and satellite observations with WRF outputs

Direct comparison between model outputs with satellite observations is challenging because of different spatiotemporal sampling of clouds; however, with the aid of COSP, the comparison becomes possible [Bodas-Salcedo et al., 2011; Zhang et al., 2019]. One big strength of COSP is to facilitate “apple-to-apple comparison of observed cloud data and model-simulated clouds” as shown in the example in Figure 2.

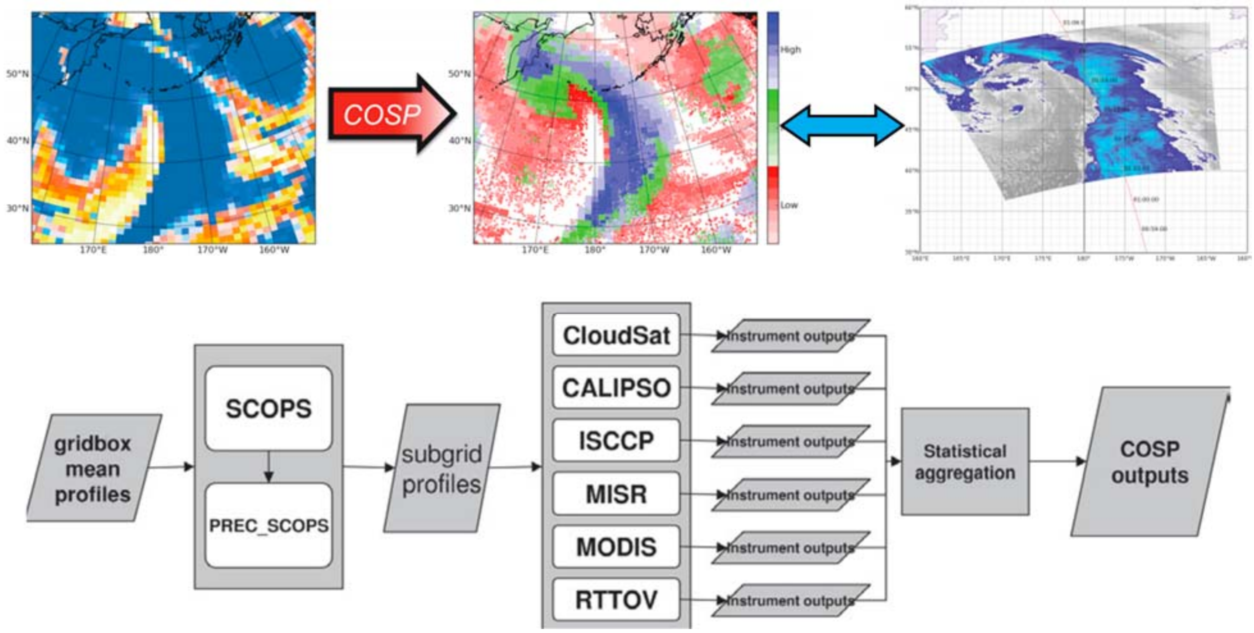

Figure 2 Upper panel: example of COSP: modeled cloud fields are converted to pseudo-satellite observations, then compared against actual satellite observations [modified from COSP webpage on https://climatedataguide.ucar.edu/]; Lower panel: flow chart of COSP, adopted from Bodas-Salcedo [2011]

Modeled vertical profiles of temperature, humidity, hydrometeor mixing ratios, cloud optical thickness and emissivity (a function of cloud water content and particle size), as well surface temperature at satellite overpassing time are feed into COSP. Firstly, the vertical profiles of model grids are broken into sub-columns to commensurate satellite pixels. Next, vertical profiles of sub-columns are passed to several instrument simulators, which apply models to simulate the radiance signals received by each sensor. Finally, statistical modules gather output from all instrument simulators, and build pseudo-cloud fields that can be directly compared to observations. Out of many products retrieved from several instruments (e.g. as shown in Figure 2 or Table 1 in Bodas-Salcedo et al. [2011]), we select total **COT** retrieved from **MODIS** (Moderate Resolution Imaging Spectroradiometer) [Levy et al., 2009] from both Aqua and Terra satellites and **CTH** from **CALIPSO** (Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations) [Winker et al., 2010] as criteria when evaluating the performance of different WRF simulations discussed in Section 2.1. This is because not many assumptions are made in the retrievals of COT and CTH. In addition, COT and CTH directly impact the radiation fields that controls photolysis reaction associated with ozone formation. Potentially, the successor of MODIS, VIIRS (Visible Infrared Imaging Radiometer Suite) can provide the same cloud products as MODIS, which will be used in our study too.

Usually for each day, MODIS will generate two snapshots (granules) of 2D cloud fields over Texas (10:30 and 13:30 local time), while CALIPSO observation will generate two swaths (cross-sections) of cloud profiles (daytime and night time) over Texas. Compared to these four "pictures", we rank the performance of all 24 groups of WRF simulations each day. The goal is to find the WRF simulation that achieve the highest score for a relative long period (one year or one season). The physics packages and reanalysis data used in this WRF simulation will be considered as the optimal configuration for this period (one year or one specific season).

Responsible organization: TAMU team
Expected milestones, outcomes, and deliverables: 1) download long-term satellite observations; 2) create COSP outputs; 3) perform the evaluation of WRF model simulation.

Task 4.3 train a GAN to improve cloud simulation over Texas

Generative adversarial networks (GANs) are a type of deep learning technique [Goodfellow et al., 2014] that is commonly used in many areas (e.g. super-resolution application that can enhance the details of images). A GAN contains two neutral networks (NN), a generator and a discriminator. The purpose of the generator is to generate fake samples of data/image and tries to "fool" the discriminator. The discriminator on the other hand tries to distinguish the real and fake samples — in other words, two NNs try to compete each other and play zero-sum game. The GANs are formulated as a mini-max game, where the discriminator is trying to minimize its reward V:

$$min_G max_D \mathrm{V(D, G)} = E_{x\sim p_{data}}[\log D(x)] + E_{z\sim p_z}[\log(1 - D(G(z)))]$$

,where x is satellite observed images of CWP, CF, or CTH, and COT, z is COSP simulation outputs of CWP, CF, CTH, and COT.

We consider the 2D cloud properties (CWP, CF, and COT) as different layers of one "image" and apply only one GAN model training. We will train one GAN model for Calipso CTH separately using the similar approach. To prepare input data and target data for GAN training, we run multiple years of WRF simulations with the optimal configuration, feed vertical profiles of variables into COSP, which generate pseudo-observed CF, CWP, and COT, as input data for the generator to generate fake cloud fields. Target or real fields is simply the corresponding observed MODIS CF, CWP, and COT fields.

Figure 3 shows the workflow of GAN training, which contains two parts. In the first part, only discriminator is trained as the network is only forward propagated. The discriminator is trained on target data (observed cloud fields) for n epochs and see if it can correctly predict them as real. Also, in this part, the discriminator is also trained on the fake generated cloud fields from the generator and see if it can correctly predict them as fake. In the second part, the generator is trained while the discriminator is idle. After the discriminator is trained by the generated fake cloud fields of the generator, we can get its predictions and use the results for training the generator and get better from the previous state to try and fool the discriminator. The above method is repeated for a few epochs and then manually check the fake cloud fields how it seems compared to target cloud fields.

Figure 3 also shows the architecture of two deep NNs. The generator has this "encoder-decoder" structure. The encoder part of the model is comprised of convolutional layers that use a 2×2 stride to downsample the input source "image" down to a bottleneck layer. The decoder part of the model reads the bottleneck output and uses transpose convolutional layers to upsample to the required output image size. Both encoder and decoder use ReLU activation function. The Adam optimizer will be used in training [Kingma & Ba, 2014].

A well-trained GAN is expected to 1) adjust large-scale cloud distributions, especially over Texas; 2) generate the fine features associated with modeled cloud decks (CF); 3) improve the accuracy of modeled cloud so that COT, CWP, as well as CTH become much closer to the observations, in terms of magnitudes and location. With GAN-generated COT, CWP, CF, and CTH, we are also able to revise 3D field of clouds accordingly, for example, increase or decrease cloud water content proportionally to GAN-generated CWP; enhance or reduce cloud top height; revise cloud brightness by modifying cloud hydrometeor size based on GAN-generated COT. As discussed in abstract, improved cloud fields over Texas are expected to increase the accuracy of air quality prediction.

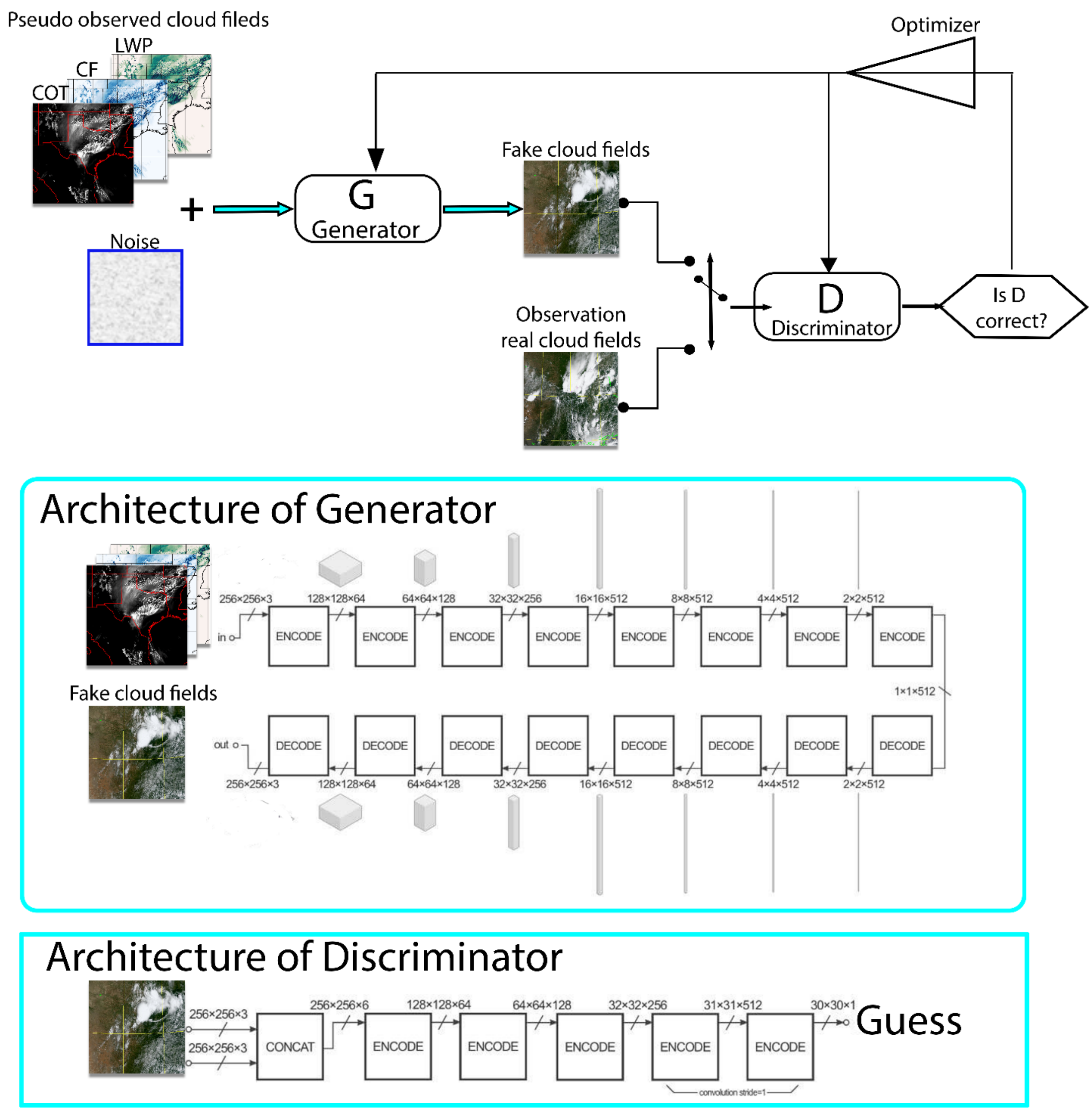

Figure 3. workflow of GAN training and the architecture of generator

Responsible organization: TAMU team
Expected milestones, outcomes, and deliverables: A well-trained GAN neutral network.

Task 4.4. Project Reporting and Presentation

As specified in Section 7.0 “Deliverables” of this Scope of Work, AQRP requires the regular and timely submission of monthly technical, monthly financial status and quarterly reports as well as an abstract at project initiation and, near the end of the project, submission of the draft final and final reports. Additionally, at least one member of the project team will attend and present at the AQRP data workshop. For each reporting deliverable, one report per project will be submitted (collaborators will not submit separate reports), with the exception of the Financial Status Reports (FSRs). The lead PI (or their designee) will electronically submit each report to both the AQRP and TCEQ liaisons and will follow the State of Texas accessibility requirements as set forth by the Texas State Department of Information Resources. The report templates and accessibility guidelines found on the AQRP website at http://aqrp.ceer.utexas.edu/ will be followed. ****Draft copies of any planned presentations (such as at technical conferences) or manuscripts to be submitted for publication resulting from this project will be provided to both the AQRP and TCEQ liaisons per the Publication/Publicity Guidelines included in Attachment G of the subaward.**** Finally, our team will prepare and submit our final project data and associated metadata to the AQRP archive.

Deliverables: Abstract, monthly technical reports, monthly financial status reports, quarterly reports, draft final report, final report, attendance and presentation at AQRP data workshop, submissions of presentations and manuscripts, project data and associated metadata

Schedule: The schedule for Task 4.4 Deliverables are shown in Section 7.

5.0 Project Participants and Responsibilities

- Dr. Zheng Lu and a graduate student are responsible for all the tasks.

6.0 Timeline

PI (Dr. Lu) and a graduate student majored in Atmospheric Science promise to deliver the following results:

	09/20	10/20	11/20	12/20	01/21	02/21	03/21	04/21	05/21	06/21	07/21	08/21
1. Find optimal WRF model configuration												

2. Synergize COSP with WRF outputs												
3. Train a GAN to improve cloud simulation												
4. Final report												

The proposed work will be done within one year, starting from Sept. 1, 2020. The task 1 and task 2 will be conducted in parallel. After we find the optimal configuration, we will conduct long-term simulation. We plan to spend six months performing GAN training and improving the cloud fields.

7.0 Deliverables

AQRP requires certain reports to be submitted on a timely basis and at regular intervals. A description of the specific reports to be submitted and their due dates are outlined below. One report per project will be submitted (collaborators will not submit separate reports), with the exception of the Financial Status Reports (FSRs). The lead PI will submit the reports, unless that responsibility is otherwise delegated with the approval of the Project Manager. All reports will be written in third person and will follow the State of Texas accessibility requirements as set forth by the Texas State Department of Information Resources. Report templates and accessibility guidelines found on the AQRP website at http://aqrp.ceer.utexas.edu/ will be followed.

Abstract: At the beginning of the project, an Abstract will be submitted to the Project Manager for use on the AQRP website. The Abstract will provide a brief description of the planned project activities, and will be written for a non-technical audience.

Abstract Due Date: Friday, Aug 28, 2020

Quarterly Reports: Each Quarterly Report will provide a summary of the project status for each reporting period. It will be submitted to the Project Manager as a Microsoft Word file. It will not exceed 2 pages and will be text only. No cover page is required. This document will be inserted into an AQRP compiled report to the TCEQ.

Quarterly Report Due Dates:

Report	Period Covered	Due Date
Quarterly Report #1	September, October, November, 2020	Friday, November 27, 2020
Quarterly Report #2	December 2020, January February 2021	Friday, February 26, 2021
Quarterly Report #3	March, April, May 2021	Friday, May 28, 2021
Quarterly Report #4	June, July August 2021	Friday, Aug 27, 2021

Monthly Technical Reports (MTRs): Technical Reports will be submitted monthly to the Project Manager and TCEQ Liaison in Microsoft Word format using the AQRP FY20-21 MTR Template found on the AQRP website.

MTR Due Dates:

Report	Period Covered	Due Date
Technical Report #1	September 1 - 30 2020	Thursday, September 10, 2020
Technical Report #2	October 1 - 31, 2020	Friday, October 9, 2020
Technical Report #3	November 1 - 30, 2020	Tuesday, November 10, 2020
Technical Report #4	December 1 - 31, 2020	Thursday, December 10, 2020
Technical Report #5	January 1 - 31, 2021	Friday, January 8, 2021
Technical Report #6	February 1 - 28, 2021	Wednesday, February 10, 2021
Technical Report #7	March 1 - 31, 2021	Wednesday, March 10, 2021
Technical Report #8	April 1 - 30, 2021	Friday, April 9, 2021
Technical Report #9	May 1 - 31, 2021	Monday, May 10, 2021
Technical Report #10	June 1 - 30, 2021	Thursday, June 10, 2021

Technical Report #11	July 1 - 31, 2021	Friday, July 9, 2021

DUE TO PROJECT MANAGER

Financial Status Reports (FSRs): Financial Status Reports will be submitted monthly to the AQRP Grant Manager (RoseAnna Goewey) by each institution on the project using the AQRP 20-21 FSR Template found on the AQRP website.

FSR Due Dates:

Report	Period Covered	Due Date
FSR #1	September 1 - 30 2020	Thursday, October 15, 2020
FSR #2	October 1 - 31, 2020	Friday, November 13, 2020
FSR #3	November 1 - 31, 2020	Tuesday, December 15, 2020
FSR #4	December 1 - 31, 2020	Friday, January 15, 2021
FSR #5	January 1 - 31, 2021	Monday, February 15, 2021
FSR #6	February 1 - 28, 2021	Monday, March 15, 2021
FSR #7	March 1 - 31, 2021	Thursday, April 15, 2021
FSR #8	April 1 - 30, 2021	Friday, May 14, 2021
FSR #9	May 1 - 31, 2021	Tuesday, June 15, 2021
FSR #10	June 1 - 30, 2021	Thursday, July 15, 2021
FSR #11	July 1 - 31, 2021	Friday, August 13, 2021
FSR #12	August 1 - 31, 2021	Wednesday, September 14, 2021
FSR #13	Final FSR	Friday, October 15, 2021

DUE TO GRANT MANAGER

Draft Final Report: A Draft Final Report will be submitted to the Project Manager and the TCEQ Liaison. It will include an Executive Summary. It will be written in third person and will follow the State of Texas accessibility requirements as set forth by the Texas State Department of Information Resources. It will also include a report of the QA findings.

Draft Final Report Due Date: Monday, August 2, 2021

Final Report: A Final Report incorporating comments from the AQRP and TCEQ review of the Draft Final Report will be submitted to the Project Manager and the TCEQ Liaison. It will be written in third person and will follow the State of Texas accessibility requirements as set forth by the Texas State Department of Information Resources.

Final Report Due Date: Tuesday, August 31, 2021

Project Data: All project data including but not limited to QA/QC measurement data, metadata, databases, modeling inputs and outputs, etc., will be submitted to the AQRP Project Manager within 30 days of project completion (September 20, 2021). The data will be submitted in a format that will allow AQRP or TCEQ or other outside parties to utilize the information. It will also include a report of the QA findings.

AQRP Workshop: A representative from the project will present at the AQRP Workshop in the first half of August 2021.

Presentations and Publications/Posters: All data and other information developed under this project which is included in **published papers, symposia, presentations, press releases, websites and/or other publications** shall be submitted to the AQRP Project Manager and the TCEQ Liaison per the Publication/Publicity Guidelines included in Attachment G of the Subaward.

8.0 References

Bodas-Salcedo, A., et al. (2011), COSP: Satellite simulation software for model assessment,

Bull. Amer. Meteor. Soc., 92, 1023– 1043.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., & Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. https://doi.org/10.1002/qj.828

Fan, S. M., Horowitz, L. W., Levy, H., & Moxim, W. J. (2004). Impact of air pollution on wet deposition of mineral dust aerosols. Geophysical research letters, 31(2).

Fan, J., Wang, Y., Rosenfeld, D., & Liu, X. (2016). Review of aerosol–cloud interactions: Mechanisms, significance, and challenges. Journal of the Atmospheric Sciences, 73(11), 4221-4252.

Feingold, G., & Kreidenweis, S. M. (2002). Cloud processing of aerosol as modeled by a large eddy simulation with coupled microphysics and aqueous chemistry. Journal of Geophysical Research: Atmospheres, 107(D23), AAC-6.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative Adversarial Networks. Retrieved from https://arxiv.org/abs/1406.2661v1

Grell, G.A. and S. R. Freitas (2014). A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos. Chem. Phys., 14, 5233–5250, https://doi.org/ 10.5194/acp-14-5233-2014.

Gurciullo, C. S., & Pandis, S. N. (1997). Effect of composition variations in cloud droplet populations on aqueous-phase chemistry. Journal of Geophysical Research: Atmospheres, 102(D8), 9375-9385.

Hong S,-Y, Y. Noh, and J. Dudhia (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. Mon. Wea. Rev., 134, 2318–2341, https://doi.org/10.1175/ MWR3199.1.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. Journal of Geophysical Research, 113, D13103. https://doi.org/10.1029/2008JD009944

Janjić, Z. I. (1994). The Step-Mountain Eta Coordinate Model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. Monthly Weather Review, 122(5), 927–945. https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. ArXiv:1412.6980[Cs]. Retrieved from http://arxiv.org/abs/1412.6980

Koren, V., Schaake, J., Mitchell, K., Duan, Q.-Y., Chen, F., & Baker, J. M. (1999). A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. Journal of Geophysical Research, 104(D16), 19,569–19,585. https://doi.org/10.1029/1999JD900232

Lefer, B. L., R. E. Shetter, S. R. Hall, J. H. Crawford, and J. R. Olson (2003), Impact of clouds and aerosols on photolysis frequencies and photochemistry during TRACE-P: 1. Analysis using radiative transfer and photochemical box models, J. Geophys. Res., 108(D21), 1125, doi:10.1029/2002JD003171.

Levy, R. C., Remer, L. A., Tanre, D., Mattoo, S. and Kaufman, Y. J. (2009). Algorithm for remote sensing of tropospheric aerosol over dark targets from MODIS: Collections 005 and 051: Revision 2, MODIS Algorithm Theoretical Basis Document for the MOD04_L2 Product.

Li, Y., Pickering, K. E., Barth, M. C., Bela, M. M., Cummings, K. A., & Allen, D. J. (2018). Evaluation of parameterized convective transport of trace gases in simulation of storms observed during the DC3 field campaign. Journal of Geophysical Research: Atmospheres, 123, 11,238–11,261. https://doi.org/10.1029/2018JD028779

Liang, J., & Jacob, D. J. (1997). Effect of aqueous phase cloud chemistry on tropospheric ozone. Journal of Geophysical Research: Atmospheres, 102(D5), 5993-6001.

Liao, H., Y. L. Yung, and J. H. Seinfeld (1999), Effects of aerosols on tropospheric photolysis rates in clear and cloudy atmospheres, J. Geophys. Res., 104(D19), 23697–23707, doi:10.1029/1999JD900409.

Liu, H., J. Crawford, R. Pierce, P. Norris, S. Platnick, G. Chen, J. Logan, R. Yantosca, M. Evans, and C. Kittaka (2006), Radiative effect of clouds on tropospheric chemistry in a global three-dimensional chemical transport model, J Geophys Res, 111(20,303).

Lu Z., X. Liu, Z. Zhang, C. Z., K. Meyer., C. Rajapakshe., C. Wu., Z. Yang., J. E. Penner. (2018). Biomass smoke from southern Africa can significantly enhance the 4 brightness of stratocumulus over the southeastern Atlantic Ocean. Proceedings of the National Academy of Sciences of the United States of America. 115 (12) 2924-2929. DOI:10.1073/pnas.1713703115.

Lu Z., and I.N. Sokolik. (2017). Examining the impact of smoke on clouds and precipitation under different fire regimes: A case study of Yakutsk wildfires of 2002, J. Geophys. Res. Atmos., 122, 12,765–12,785. DOI: 10.1002/2017JD027001. (Cover page paper)

Lu, Z., and I. N. Sokolik. (2013). The effect of smoke emission amount on changes in cloud properties and precipitation: A case study of Canadian boreal wildfires of 2007. J. Geophys. Res. Atmos., 118. DOI:10.1002/2013JD019860.

Morrison, H., G. Thompson, and V. Tatarskii (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of oneand two-moment schemes. Mon. Wea. Rev., 137, 991–1007, https://doi.org/10.1175/2008MWR2556.1.

NCEP (National Centers for Environmental Prediction). (2000. NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, https://doi.org/10.5065/d6m043c6.

Otkin, J. A., & Greenwald, T. J. (2008). Comparison of WRF model-simulated and MODIS-derived cloud data. Monthly Weather Review, 136(6), 1957-1970.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743), 195-204.

Rogers, E., and Coauthors (2009). The NCEP North American Mesoscale Modeling System: Recent changes and future plans. Preprints, 23rd Conf. on Weather Analysis and Forecasting/19th Conf. on Numerical Weather Prediction, Omaha, NE, Amer. Meteor. Soc., 2A.4.

Rosenfeld, D., Sherwood, S., Wood, R., & Donner, L. (2014). Climate effects of aerosol-cloud interactions. Science, 343(6169), 379-380.

Seinfeld, J. H., Bretherton, C., Carslaw, K. S., Coe, H., DeMott, P. J., Dunlea, E. J., ... & Kraucunas, I. (2016). Improving our fundamental understanding of the role of aerosol−cloud interactions in the climate system. Proceedings of the National Academy of Sciences, 113(21), 5781-5790.

Skamarock, W. C., & Klemp, J. B. (2008). A time-split nonhydrostatic atmospheric model for Weather Research and Forecasting applications. Journal of Computational Physics, 227(7), 3465–3485. https://doi.org/10.1016/j.jcp.2007.01.037

Thompson, G., Field P.R., Rasmussen R.M., and Hall W.D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. Mon. Wea. Rev., 136, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Thompson, G., Tewari, M., Ikeda, K., Tessendorf, S., Weeks, C., Otkin, J., & Kong, F. (2016). Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. Atmospheric Research, 168, 92-104.

Tie, X., S. Madronich, S. Walters, R. Zhang, P. Rasch, and W. Collins (2003), Effect of clouds on photolysis and oxidants in the troposphere, J. Geophys. Res., 108(D20), 4642, doi:10.1029/2003JD003659.

Tie, X., S. Madronich, S. Walters, D. P. Edwards, P. Ginoux, N. Mahowald, R. Zhang, C. Lou, and G. Brasseur (2005), Assessment of the global impact of aerosols on tropospheric oxidants, J. Geophys. Res., 110(D3), 13,791, doi:10.1029/2004JD005359.

Tiedtke, M., (1989). A Comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon. Wea. Rev., 117, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117,1779: ACMFSF.2.0.CO;

Winker, D. M., and Coauthors, 2010: The CALIPSO mission: A global 3D view of aerosols and clouds. Bull. Amer. Meteor. Soc., 91, 1211–1229.

Zhang, Y., Xie, S., Lin, W., Klein, S. A., Zelinka, M., Ma, P.-L., Rasch, P. J., Qian, Y., Tang, Q., & Ma, H.-Y. (2019). Evaluation of clouds in version 1 of the E3SM atmosphere model with satellite simulators. Journal of Advances in Modeling Earth Systems, 11, 1253– 1268. https://doi.org/10.1029/2018MS001562

Zheng, Y., K. and co-authors (2016). Improving high-resolution weather forecasts using the Weather Research and Forecasting (WRF) Model with an updated Kain–Fritsch scheme. Mon. Wea. Rev., 144, 833–860, https://doi.org/10.1175/MWR-D-15-0005.1.