TECHNICAL MEMO

Development and Implementation of a WRF Planetary Boundary Layer Height Operator

Task 7 .2 – Final Report and Docker container with other electronic data

TCEQ Contract No. 582-19-90498
Work Order No. 582-23-42428-017

Deliverable 7.2
Revision 1.0

Prepared by:

John Henderson and Hannah Vagasky
Atmospheric and Environmental Research, Inc. (AER)
131 Hartwell Ave.
Lexington, MA 02421
Correspondence to: jhenders@aer.com

Prepared for:

Bryce Kuchan
Texas Commission on Environmental Quality
Air Quality Division
Bldg. E, Room 383
Austin, Texas 78711-3087

23 June 2023

Document Change Record

Revision	Revision Date	Remarks
0.1	**22 June 2023**	**Internal Version**
1.0	**23 June 2023**	**Report Delivery to TCEQ**

TABLE OF CONTENTS

TABLE OF FIGURES

1 Preamble

The overall goal of this project is to compare planetary boundary layer height (PBLH) estimates from the Weather Research and Forecasting (WRF) model with radar-estimated PBLH values computed by the technique outlined by Banghoff et al. (2018). To achieve the project goal, AER first developed an observation operator to extract the WRF PBLH values from grid boxes collocated with the radar pixels used by the radar QVP radar algorithm to estimate PBLH. The software was then applied to TCEQ-generated WRF simulations from select months in 2019 and compared against radar estimates. Analysis and recommendations for future study are provided. Finally, TCEQ staff was trained to use the software. This effort adds to the set of methods available to evaluate TCEQ's meteorological simulations, further ensuring that model performance is adequate to use in air quality modeling for state implementation plan (SIP) purposes.

This Final Report is to be delivered to the TCEQ Project Manager electronically (i.e., via file transfer protocol (FTP) or e-mail) in Microsoft Word and PDF formats no later than the deliverable due date shown below. This Final Report provides a comprehensive overview of activities undertaken and any data collected and analyzed. The Final Report highlights major activities and key findings, provides pertinent analysis, describes encountered problems, and associated corrective actions, and/or detailed relevant statistics, including data, parameter, or model completeness, accuracy, and precision.

Deliverable 7.2: Final Report and Docker container with other electronic data

Deliverable Due Date: 23 June 2023

2 Executive Summary

This project refined and extended the capabilities of software that AER developed in 2022 for TCEQ that retrieved PBLH from National Weather Service (NWS) radars based on a method developed by Banghoff et al. (2018). The retrieved NWS PBLH field has the potential to validate model simulated PBLH under certain weather conditions.

The current project developed software that first obtains an average PBLH at appropriate grid points from the Weather Research and Forecasting (WRF) model fields generated by TCEQ and then compares the simulated PBLH to NWS radar-retrieved PBLH. AER ran this software on the TCEQ machine 'eira' in order to conduct an 8 – month comparison between NWS and WRF PBLH for simulations over Texas in 2019. The software has been scientifically and technically inspected within the constraiints of the project to ensure results are valid.

NWS and WRF PBLH estimates were compared using the Houston, Dallas, and El Paso radars from March – October 2019 for each daytime period (1200 to 0000 UTC). An additional layer of quality control applied here - beyond what is contained in the radar approach - used WRF precipitation fields to prevent comparisons at times and locations where the WRF PBLH may have been contaminated by the presence of rain. Hourly comparisons of WRF – radar height estimates were aggregated monthly and for the entire period . Graphics were genenerated to aid in quality control and analysis purposes.

Results from the analysis indicate PBLH characteristics in WRF that are broadly consistent with the radar estimates and also are expected features of the diurnal growth and decay of the PBL. The seasonal and diurnal variability in the WRF PBLH field is appropriate across different geographical locations. A pronounced reduction in PLBH depth over water bodies, such as Galveston Bay, compared to inland locations is consistent with expectations of the PBL in marine environments. While the simulated WRF PBLH estimates often differ from the NWS retrieved PBLH, the WRF PBLH are not unreasonable. Based on our analysis, we do not have concerns about the overall ability of the WRF model to well represent the top of the boundary layer for the three metroplitian areas in Texas during the spring, summer, and fall of 2019. It is expected that these findings can be extended to all regions of Texas.

There are two important considerations about the radar approach that should be noted. First, further refinements guided by additional validation with independent datasets are needed to better interpret the radar profiles. This would, for instance, minimize identification of elevated minima in differential reflectivity in the morning and afternoon shoulder periods. The second is the need to limit application during time periods that have poorly developed PBLs. Use of additional WRF fields could improve quality control approaches to restrict comparison to times that have relatively uncontaminated PBLs in both the radar and WRF fields.

3 Introduction

The TCEQ is required under the federal Clean Air Act (FCAA) to perform air quality modeling for attainment demonstration purposes. Evaluating model performance is a crucial step in the air quality modeling process. These evaluations involve comparing simulated values of a certain parameter to estimated or measured values of the same parameter and quantifying the difference using various statistical measures.

The planetary boundary layer (PBLH) is the region of the lower troposphere where meteorological conditions are largely impacted by the Earth's surface and is important for a variety of air quality applications. Of considerable importance is that the depth of this layer strongly influences air quality in a given region by determining the volume of air into which pollutants are mixed. Observations of PBLH, however, are spatially and temporally limited.

Banghoff et al. (2018) presented a new method to potentially obtain more observations of PBLH by using the differential reflectivity field from National Weather Service (NWS) radars to estimate PBLH under certain weather conditions. Differential reflectivity became available from NWS radars following an upgrade in recent years to dual polarization. Work done by AER for the TCEQ in 2022 developed software to compute PBLH from WSR-88D weather radars. Please see Task 8.1 deliverable report named P2304.12_Task8.1-deliverable-v1.2-20220630.pdf for a full description of the previously developed software.

The purpose of this project is to develop software to obtain an average PBLH from output of the Weather Research and Forecasting (WRF) model used by the TCEQ and compare it with radar-estimated PBLH values. This will add to the set of methods already used to evaluate meteorological simulations, further ensuring that model performance is adequate to use in air quality modeling for state implementation plan (SIP) purposes.

4 Project Output Data

Here we describe the data generated by this project. We focus primarily on changes to the 2022 code that result in additional capabilities and output datasets. Additional detail of the software mentioned in this section follows in Section 5.

4.1 New output from NWS radar retrieval

The majority of the output from the software delivered in 2022 is unchanged. Please refer to the Task 8.1 deliverable report named P2304.12_Task8.1-deliverable-v1.2-20220630.pdf for a full description of the 2022 output products.

A new output product is now generated by time-vs-height-radar-data.py in netcdf format. It contains the latitude and longitude of each radar point used in calculating the NWS radar retrieved PBLH. There are four fields in this dataset: Dates, PBLH, Lat, Lon. Each point used to calculate the PBLH is in a separate row. Multiple latitude – longitude pairs have the same estimated PBLH and date since multiple points via the azimuthal averaging are used to calculate the final estimate of PBLH. This file will be used in the forward operator to verify PBLH estimates in WRF. NetCDF format was chosen for efficiency since Python can more rapidly read and write netCDF files than CSV files.

Example filename and size:
`HGX-20190801-20190802-pblh_latlon-5.0976562deg_azimuths_270-360.nc [2.4 MB]`

4.2 Daily NWS radar and WRF comparisons

The new script wrf_validation.py generates two types of figures and a text file for each day that is analyzed. The figures demonstrate algorithm performance, as well as aid in quality control efforts. The text file provides summary statistics that can be used for validation analyses over longer time periods.

The first type of figure is a time series from 1200-0000 UTC and is created to show radar-retrieved PBLH and the mean Point WRF and Regional WRF PBLH estimates. Additionally, +/- one standard deviation is also shown for the Point WRF and Regional WRF PBLH estimates. An example is Figure 1 shown below in Section 5.3.3.

Example filename and size:
`HGX-20190801-mean-stdev-PBLH-timeseries.png [74 KB]`

A second type of figure contains a two-dimentional plot of the WRF PBLH field. The WRF grid points used in the Point estimate, and the outline of the 30-km x 30-km box used in the Regional estimate, are shown. An example is Figure 2 shown below in Section 5.3.3.

Example filename and size:
`HGX-20190801-1800-wrfPBLH-map.png [200 KB]`

Finally a summary csv file is output. This comma delimited text file lists the following data: date and time, radar PBLH, mean WRF PBLH using the Point method, standard deviation of the WRF PBLH using the Point method, mean WRF PBLH using the Regional method, and standard deviation of the WRF PBLH using the Regional method. Each WRF output time between 1200-0000 UTC corresponds to one line in the csv file.

Example filename and size:
`HGX-20190801-NEXRAD_WRF_pblh_estimates.csv [4 KB]`

4.3 Multi-day NWS radar and WRF comparisons

The aggregate_plot_comparison.py script aggregates these csv files generated by wrf_validation.py so that multi – day statistics can be generated. Two figures are generated from this code. One figure shows a time series of the NWS and WRF means as a line and the +/- one standard deviation is shaded. The second figure is a spaghetti plot showing a time series of each hourly NWS and WRF PBLH trace used in the analysis. An example of these figures is shown in Figure 3.

Example filename and size:
`HGX-20190801-20190802-mean-stdev-PBLH-timeseries.png [4 MB]`
`HGX-20190801-20190802-spaghetti-PBLH-timeseries.png [2.4 MB]`

5 Software

The software and data needed for the example case study used during training was placed on the TCEQ machine 'eira' in /eirun/Training/AER_WRF.14jun2023:

-rw-r--r--@ 1 jhenders green 2468352 Jun 14 10:15 wrf_validation.tar

The necessary software run-time environment, including all software dependencies, is managed by Docker and contained in the Dockerfile. The code underlying the forward operator and the subsequent radar – WRF comparisons are based on publicly – available Python packages that are also contained within the Dockerfile.

The final version of the project software was attached via a file named "P2304.17_Task7.2-software-deliverable-v1.0-20230623.tar" to the email that delivered this Final Report to TCEQ on 23 June 2023.

This section briefly reviews the software that was developed to support this project, provides details about the execution of that software, and how users can modify the software for future use.

5.1 Updates to pre-existing software

For the current project, we modified software AER delivered in 2022 that retrieved PBLH from NWS radars called time-vs-height-radar-data.py, plus the underlying Dockerfile.

The Dockerfile was updated as follows:

- "FROM continuumio/miniconda3" was changed to "FROM continuumio/miniconda3:22.11.1"

Modifications that were made to time-vs-height-radar-data.py include:

- Instead of listing each date that will be processed in "datstr" only the first date and last date need to be provided in the "datestart" and "dateend" variables, respectively.
- Radar data are now automatically accessed through Amazon's registry of open data (https://registry.opendata.aws/noaa-nexrad/). Users no longer need to manually download radar data.
- Level 2 NWS radar data are now used instead of Level 3 data. The benefits of Level 2 NEXRAD data include additional elevation angles, a single file per time level that contains all radar variables, a common grid for all radar variables and a longer data archive on Amazon's servers.
- The minimum height of the retrieved PBLH was reduced from 250 m to 100 m.
- Data are quality controlled in a manner that is more consistent with Banghoff et al. (2018). The quality control procedure (apply_qc_qpv.py) was also moved into a separate subroutine called by time-vs-height-radar-data.py.
- A subroutine called retrieve_ring_latlon.py was added to time-vs-height-radar-data.py to output the latitude and longitude of the radar bins used to calculate the PBLH. These data are output to a new netcdf file called {Radar Identifier}-{Start Date}-{End Date}_pblh_latlon.nc. There are four vectors in this dataset: Dates, PBLH, Lat, Lon. Each point used to calculate the PBLH is separate row.

From a user's perspective, the most important modifications to the 2022 scripts are that radar data no longer has to be manually downloaded and that the method of specifying dates has been slightly changed. Please refer to the Task 8.1 deliverable report named P2304.12_Task8.1-deliverable-v1.2-20220630.pdf for a full description of the 2022 software and how to execute the software.

5.2 Software developed for the project

5.2.1 WRF forward operator

The wrf_validation.py script contains the forward operator, which is the code that interpolates the WRF simulated PBLH values to the locations of the radar observations, and facilitates comparison between the radar and WRF PBLH values. The script first reads a netcdf file that contains the lat/lon coordinates of the radar cells used in the computation of the radar estimated PBLH. Then PBLH data is read from WRF netcdf output files. The WRF data are temporally and spatially matched to the radar times and locations, and statistics (mean and standard deviation) of the WRF PBLH values are calculated for each (hourly) model output time. Quality control to avoid WRF grid points with precipitation has been implemented. The forward operator can be run for multiple days of WRF and radar data. Each day processed by wrf_validation.py has its own set of time series and map figures and its own summary csv file.

Two independent estimates of WRF PBLH are created. These are based on two different approaches to extracting and averaging the WRF data. One estimate uses the same locations as the radar retrieval (referred to as the "Point" method) from the prior AER code deliverable in 2022. The second estimate defines a fixed rectangular box that is 0.6° degrees wide and uses all WRF values within that box (referred to as the "Regional" method). This second method, including the size of the box, is based on Eure et al. (2023).

5.2.2 Multi-day comparisons

The aggregate_plot_comparison.py script is used to compare the NWS radar observed and WRF simulated PBLH over multiple days. The script reads in NWS radar derived PBLH, the mean WRF point PBLH, and the mean WRF region PBLH from the summary csv files created by wrf_validation.py for the time period of interest. Data is aggregated by hour and, for each hour between 12 and 00 UTC, a mean and standard deviation of NWS radar and WRF data is calcaluated across all days. Two png files are created. One figure shows a timeseries of the NWS and WRF mean as a line and +/- one standard deviation in shading. The second figure is a spaghetti plot showing a time series of every NWS and WRF PBLH used in the analysis.

5.3 Execution of software

In this section we document the steps required to install and run the scientific code developed through this project for a case study. The case demonstrated in the 2023 training analyzed August 1st and 2nd, 2019 in Houston, Texas. The radar-estimated PBLH values retrieved using data from the northwest quadrant of Houston's NWS radar (HGX).

5.3.1 Build Docker

Docker is used to install the software and their required dependencies, including Python. AER has placed a tarball containing all the necessary software and data on eira:

```
[jhenderson@eira /home/jhenderson/]$ tar xf wrf_validation.tar
```

Once the tarball has been expanded, the directory named "docker" is created. This directory contains two additional directories. The directory named "main" contains the Dockerfile and time-vs-height-radar-data.py. The directory named "step2" contains a netcdf file, and six other python codes. These python codes include retrieve_qvp.py, apply_qc_qvp.py, retrieve_ring_latlon.py, wrf_validation.py, pblh_validation.py, and aggregate_plot_comparison.py.The WRF forward operator and the accompanying daily comparison with the NWS radar data is contained in wrf_validation.py.

Ensure that the docker/ directory and all files within it are fully available by using the chmod command with the '777' option.

```
[jhenderson@eira /home/jhenderson/]$ chmod -R 777 docker
```

The creation of the Docker image should only need to be performed once. Use of the -–no-cache argument forces building from scratch; otherwise, Docker will attempt to only build changes to the Dockerfile.

To build the Docker image, move into the docker/main directory and use the following command:

```
[jhenderson@eira /home/jhenderson/docker/main]$ docker build --no-
cache -t time-vs-height-radar-data . >&! Build.log
```

The output will be a Docker image named time-vs-height-radar-data:

```
[jhenderson@eira /home/jhenderson/docker/main]$ docker image ls
REPOSITORY                 TAG     IMAGE ID      CREATED     SIZE
time-vs-height-radar-data  latest  9988937d5688  1 day ago  3.94 GB
```

The scripts and data in docker/step2 are originally isolated from the docker/main directory so that the Docker image builds faster. Now that the image is built we need to move all items within the step2/ directory to the main directory before we can run the code.

```
[jhenderson@eira /home/jhenderson/docker/main]$ mv ../step2/* .
```

5.3.2 Run updated NWS radar retrieval code

The version of time-vs-height-radar-data.py provided in the tarball is already set up to run the 1 August 2019 case using Houston radar data. How to configure time-vs-height-radar-data.py is provided in the next section.

The software is run at the command line without arguments. The run command is as follows:

```
docker run --rm --name time-vs-height-radar-data -v `pwd`:/opt/src -v
`pwd`/data:/opt/data time-vs-height-radar-data ./time-vs-height-radar-
data.py &
```

5.3.3 Run WRF forward operator

The version of wrf_validation.py provided in the tarball is configured to run the 1 – 2 August 2019 case using data from the Houston radar. An explanation of how to configure wrf_validation.py is provided in the next section.

The software is run at the command line without arguments. The run command is as follows:

```
docker run --rm --name time-vs-height-radar-data -v `pwd`:/opt/src -v
/metrun/run1/bcd/WRF-4.1.5/run-b/wd2.era-
int.hvc.noah.ysu.2019.07.31:/opt/data/wrf time-vs-height-radar-data
./wrf_validation.py &
```

Figures 1 and 2 provide examples of figures output from this script.

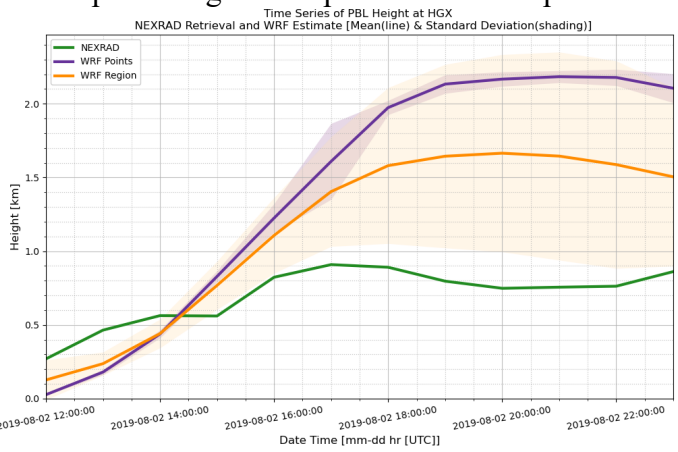

Figure 1 - Time series of the mean (solid line) and +/- one standard deviation (shading) of PBLH from the radar (green), Point WRF (purple), and Regional WRF (orange) estimates from 1200 – 0000 UTC 2 August 2019 in Houston, TX.

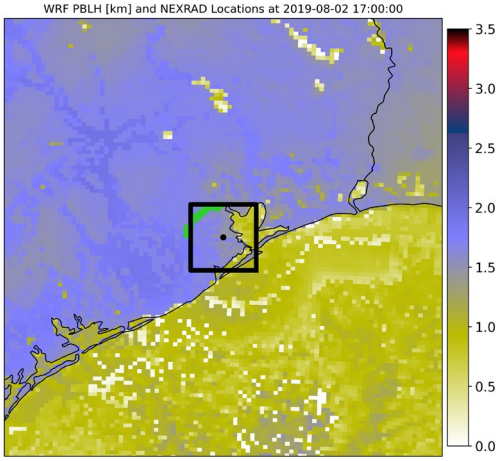

Figure 2 - Map of WRF PBLH field (km, shaded) at 1700 UTC on 2 August 2019 around Houston, TX. The green dots indicate the location of the radar PBLH estimates used in the WRF Point estimates. The black square shows the fixed 30 x 30-km box used in the Regional WRF estimates. The black dot is the location of the radar.

5.3.4 Run multi-day Comparison

Multiple day comparisons are created using aggregate_plot_comparison.py. The version of aggregate_plot_comparison.py provided in the tarball is already set up to run the 1 – 2 August 2019 case using Houston radar data. How to configure aggregate_plot_comparison.py is provided in the next section.

The software is run at the command line without arguments. The run command is as follows:

```
docker run --rm --name time-vs-height-radar-data -v
`pwd`:/opt/src -v `pwd`/data:/opt/data time-vs-height-radar-data
./aggregate_plot_comparison.py &
```

Figure 3 provides examples of figures output from this script.

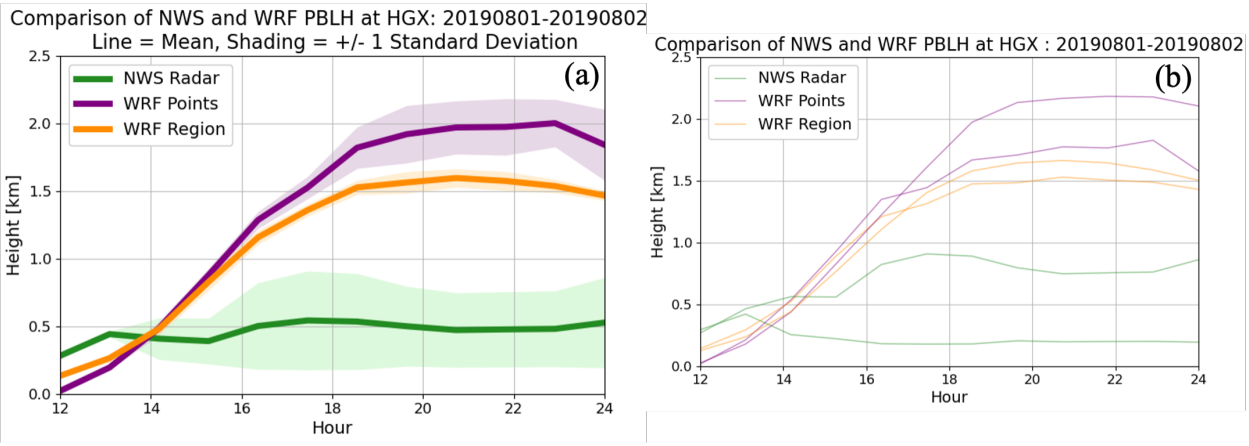

Figure 3 - (a) Time series of the mean (lines) and +/- one standard deviation (shading) of the NWS radar retrieved and WRF simulated PBLH and (b) Spaghetti plot of time series of all NWS radar retrieved and WRF simulated PBLH on August 1st and 2nd, 2019 in Houston, TX. NWS radar retrieved PBLH in green, point – based WRF PBLH estimates in purple, and region– based WRF estimates in orange.

5.4 User defined options

These options determine how the code functions and must be set prior to execution of the Python scripts.

5.4.1 time-vs-height-radar-data.py

The majority of these options have not changed from 2022. The only difference in that 'datestr' has been removed and replaced with 'datestart' and 'dateend'. Datestart lists the first day that will be processed. Dateend lists the last day that will be processed. Both of these variables have YYYYMMDD format. Please refer to the Task 8.1 deliverable report named P2304.12_Task8.1-deliverable-v1.2-20220630.pdf for description of the other user defined options.

5.4.2 wrf_validation.py

The location of the WRF data files on TCEQ's computers is provided in the Docker run command. The portion of the call that does this task is highlighted below:

```
docker run --rm --name time-vs-height-radar-data -v `pwd`:/opt/src -v
/metrun/run1/bcd/WRF-4.1.5/run-b/wd2.era-
int.hvc.noah.ysu.2019.07.31:/opt/data/wrf time-vs-height-radar-data
./wrf_validation.py &
```

The directory that is currently used has WRF data from August 2019. A different month of WRF data can be used by changing the path highlighted above.

All other user defined options are contained within the first 40 lines of code in wrf_validation.py. A list of these options, the values used for the training session, and a brief description of the option, is listed below.

```
#---------------- USER OPTIONS SECTION ------------------------
datestart = '20190801' # first date to be processed in YYYYMMDD format
dateend = '20190824' # last date to be processed in YYYYMMDD format
sitestr = 'HGX' # Radar site to be processed
radarlev = 'l2' # Type of radar data that was used generate NWS PBLH
data
radarloc = {'HGX':[-95.0833,29.4667,5]} # Longitude, latitude,
elevation of radar
workdir = '/opt/src/' # home directory
radar_file = 'HGX-20190801-20190831-pblh_latlon-
5.0976562deg_azimuths_270-360.csv' # name of csv file from NEXRAD
retrieval
radar_indir = '/opt/src/' # Location of NEXRAD csv file relative to
the Docker structure
wrf_indir = '/opt/data/wrf/' # Location of WRF output relative to the
Docker structure
outdir = '/opt/src/' # Location where output will be saved relative to
the Docker structure
```

5.4.3 aggregate_plot_comparison.py

All user defined options are contained within the first 35 lines of code in aggregate_plot_comparison.py. A list of these options, the values used for the training session, and a brief description of the option, is listed below.

```
#---------------- USER OPTIONS SECTION ------------------------
radarlev = 'l2' # Type of radar data that was used generate NWS PBLH
data
datestart = '20190801' # first date to be processed in YYYYMMDD format
dateend = '20190824' # last date to be processed in YYYYMMDD format
sitestr = 'HGX' # Radar site to be processed
workdir = '/opt/src/' # home directory
wrf_indir = '/opt/data/wrf/' # Location of WRF output relative to the
Docker structure
outdir = '/opt/src/' # Location where output will be saved relative to
the Docker structure
```

5.5 Exisiting installation on TCEQ machine

Here is a list of the files present after a successful installation and execution of project software on a TCEQ machine. Here we show the contents of the docker/main directory after moving the contents of docker/step2, which is part of the training instructions.

```
[jhenderson@eira ~]$ pwd
/home/jhenderson/docker/main

[jhenderson@eira ~]$ ls -1
aggregate_plot_comparison.py
```

```
apply_qc_qvp.py
Build.log
Dockerfile
HGX-20190801-1200-wrfPBLH-map.png
HGX-20190801-1300-wrfPBLH-map.png
HGX-20190801-1400-wrfPBLH-map.png
HGX-20190801-1500-wrfPBLH-map.png
HGX-20190801-1600-wrfPBLH-map.png
HGX-20190801-1700-wrfPBLH-map.png
HGX-20190801-1800-wrfPBLH-map.png
HGX-20190801-1900-wrfPBLH-map.png
HGX-20190801-2000-wrfPBLH-map.png
HGX-20190801-20190802-mean-stdev-PBLH-timeseries.png
HGX-20190801-20190802-pblh_latlon-5.0976562deg_azimuths_270-360.nc
HGX-20190801-20190802-spaghetti-PBLH-timeseries.png
HGX-20190801-2100-wrfPBLH-map.png
HGX-20190801-2200-wrfPBLH-map.png
HGX-20190801-2300-wrfPBLH-map.png
HGX-20190801-mean-stdev-PBLH-timeseries.png
HGX-20190801-NEXRAD_WRF_pblh_estimates.csv
HGX-20190802-1200-wrfPBLH-map.png
HGX-20190802-1300-wrfPBLH-map.png
HGX-20190802-1400-wrfPBLH-map.png
HGX-20190802-1500-wrfPBLH-map.png
HGX-20190802-1600-wrfPBLH-map.png
HGX-20190802-1700-wrfPBLH-map.png
HGX-20190802-1800-wrfPBLH-map.png
HGX-20190802-1900-wrfPBLH-map.png
HGX-20190802-2000-wrfPBLH-map.png
HGX-20190802-2100-wrfPBLH-map.png
HGX-20190802-2200-wrfPBLH-map.png
HGX-20190802-2300-wrfPBLH-map.png
HGX-20190802-mean-stdev-PBLH-timeseries.png
HGX-20190802-NEXRAD_WRF_pblh_estimates.csv
out
pblh_validation.py
__pycache__
retrieve_qvp_aws.py
retrieve_ring_latlon.py
time-vs-height-radar-data.py
wrf_validation.py
```

6 Data Sources

6.1 Radar data

Time-vs-height-radar-data.py has been modified to automatically access NWS Level 2 radar data through Amazon's registry of open data (https://registry.opendata.aws/noaa-nexrad/). Data are freely available through this service and the full Level 2 dataset is available starting in 1991. Differential reflectivity, which is a necessary component of the PBLH radar retrieval code, became available at select radars as part of a radar hardware upgrade in 2010 and was fully available for all sites in 2013.

6.2 WRF data

The WRF data used during this study was supplied by TCEQ at 4 km resolution from March – October 2019. Use of other periods or sources of WRF simulations requires only that the variable "PBLH" is output from the model at runtime. Please see the above section for instructions on how to change the model data used by the software.

7 Analysis

We have conducted comparisons between NWS derived and WRF simulated PBLH in Houston, El Paso and Dallas from March – October 2019. These locations span a variety of climates and also exhibit local characteristics that require accommodation in the radar technique.

It has been shown that PBLH varies diurnally, seasonally, and geograpahically. Zhang et al. (2020) found that the diurnal variability of PBLH in the United States was greatest in spring and summer, especially under cloud free conditions. Since our method works best during sunny periods and excludes times and locations with precipitation, our results could be biased towards the high end of the expected range of PBLH. Xi et al. (2022) used hourly data collected as airplanes take-off and land at 54 airports in the United States. They showed that the diurnal cycle of PBLH also depends on surface conditions. The highest diurnal variability was found in humid regions and weakest in coastal regions. Arid or mountainous areas had moderate variability. Thus, we anticipate that our results will vary based on location since Houston, El Paso, and Dallas have a variety of factors influencing their climates.

Houston is a coastal city whose weather is strongly influenced by the nearby Gulf of Mexico and Galveston Bay. Figure 4a shows a map of WRF simulated PBLH. Note that we specifically limited our NWS analysis to only the northwest quandrant of the radar domain to avoid water where the PBLH is considerably lower. (Additional studies of the performance of WRF for only marine locations would benefit, however, from use of azimuths reflective of the southeast quadrant.) Based on monthly summaries of conditions at Houston Interncontinental Airport, 71 days during this period had measurable rain. March and July were dry, but May was wet. September had one day that received over 9 inches of rain.

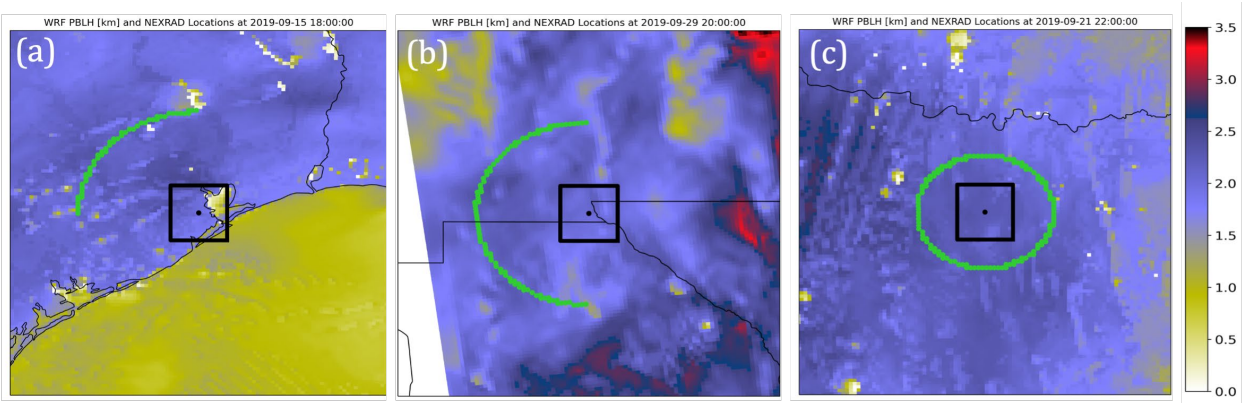

Figure 4 - Map of WRF PBLH field (km, shaded) near (a) Houston at 1800 UTC 25 September 2019, (b) El Paso at 2000 UTC 29 September 2019 and (c) Dallas at 2200 UTC 21 September 2019. Green dots indicate the locations of the radar PBLH estimates used in the WRF Point estimates. Black dot indicates the location of the radar in the center of the 30x30-km black square used in the Regional WRF estimates.

El Paso is a desert city located at 1100 m elevation. Additionally, the north-to-south- oriented Franklin Mountains are located to the north of the city and reach a maximum elevation of 2100 m. The Sierra de Juarez Mountians are located farther to the west in Mexico. The NWS radar is located northwest of El Paso and the Franklin Mountains generally block the eastern hemisphere of the radar, even at the 5 degree elevation angle scans. Figure 4b shows that we only used the western hemisphere of the El Paso radar domain during our NWS radar analysis. Monthly summaries from the El Paso International Airport indicate that the months of July – September 2019 were abnormally warm. The average temperature in each of these months was at least 3 °F above normal. July also had anormally dry conditions. The presence of rain on a total of 64 days for even this dry climate reflects the need for quality control at all locations to avoid low quality days.

Dallas is located distant from any substantial bodies of water and the local area is devoid of mountains. This uniform landscape allowed us to use the full radar domain in our NWS analysis, as is shown in Figure 4c. Monthly summaries from the Dallas Fort – Worth International Airport also indicate that Dallas was very warm and dry in September 2019. The average temperature was more than 7 °F above normal with total precipitation 2.7 inches below normal. However, March, June, and October were cool and had average temperatures greater than 2 °F below normal. April and May received more than 3 inches above the normal rainfall. 75 days had measureable rain during this 8 – month period in 2019.

It is important to note that the radar technique works best in clear, sunny conditions without cloud layers in order to promote steady development of a convectively-driven boundary layer. The retrieval should not be used overnight, during preciptation or when the convective boundary layer has been disrupted by recent precipitation. We apply the technique only between 1200 and 0000 UTC, omitting locations with precipitation in the WRF grid boxes. This approach is straightforward but represents only a preliminary level of mitigation. A variety of other factors impact the quality of the retrieved PBLH, such as the presence of residual layers in the morning and also, for example, the sensitivity of the radar as determined by the choice of operation mode (clear air or precipitation) by the radar operator. In the future, additional screening methods that take advantage of the plethora of WRF fields and also independent datasets, such as radiosondes, could be developed to more accurately define the environment and

aid in interpretation of the radar signal both to support PBLH estimation and identify periods of uncertainty.

Figures 5a, 6a, and 7a show the 8 – month average PBLH time series as retrieved from NWS radars and simulated in WRF for Houston, El Paso, and Dallas, respectively. The other panels in Figures 5, 6, and 7 show the average time series for each individual month from March – October 2019. As stated above, the NWS retrieval and point WRF PBLH were only calculated for the northwest quadrant of the Houston radar to minimize the influence of the Gulf. For El Paso, the NWS retrieval and point WRF PBLH was only conducted for the western half of the radar domain that is not contaminated by the nearby mountains.

The following conclusions can be drawn:

- For each site, the diurnal cycle of the PBLH in the WRF points and WRF region is greater than the cycle observed by the NWS radar.
- The WRF points and WRF region estimates of PBLH during the morning (1200-1500 UTC) is lower than observed by the NWS radar at all locations for the majority of months. The radar retrieved PBLH in the morning often appears too high and too flat over time. The WRF diurnal cycle is representative of the growing boundary layer with a maximum in the late afternoon and appears reasonable.
- The estimates provided by WRF points and WRF region are most different from each other in Houston. This difference is attributed to the WRF region methodology covering portions of the Gulf of Mexico and Galveston Bay (Figure 4a), which were excluded by design in the radar-based approach.
- The NWS retrieved and WRF simulated PBLH are most similar in Houston.
- A distinct seasonal trend is seen in El Paso and Dallas. In each location, the NWS and WRF PBLH are most similar in spring and fall. The differences are largest in summer when WRF simulated PBLH are considerably higher than the NWS retrieved PBLH.
- The WRF PBLH estimates exhibit a reasonable geographical/climatological sensitivity to the maximum PBLH height, with El Paso having the highest PBLH of all the regions. Geographic variability in the NWS PBLH is comparably small, likely influenced by frequent application during suboptimal time periods.
- The WRF simulations conducted by TCEQ use the Yonsei University (YSU) planetary boundary layer parameterization. This scheme is commonly used within the scientific community and is known to have stronger mixing than other schemes. This can result in a simulated PBL that grows more quickly and reaches greater depths [Xie et al. 2012; Milovac et al. 2015]. This may explain some of the differences between the lower radar-diagnosed PBLHs and those from WRF.
- In summary, the differences between the WRF and radar PBLH estimates are attributable to a combination of inherent forecast errors associated with all numerical weather prediction models at this spatial scale; the choice of PBL scheme and PBLH estimation approach used by WRF; limitations of the radar technique itself, and, importantly, application of the radar technique on days with a poorly developed or contaminated PBL structure.

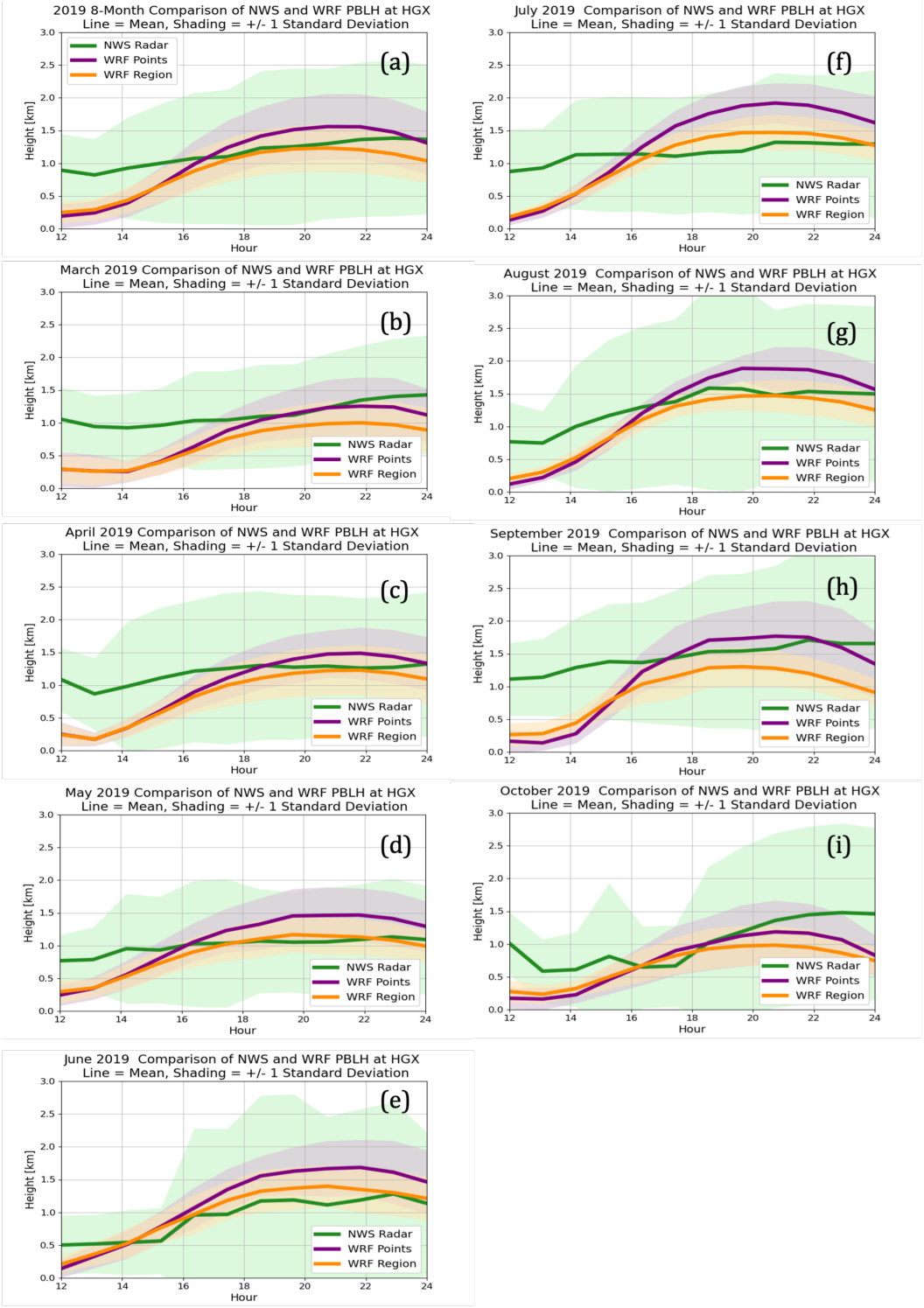

Figure 5 - Time series of the mean (lines) and +/- one standard deviation (shading) of the NWS radar retrieved and WRF simulated PBLH for (a) March – October, (b) March, (c) April, (d) May, (e) June, (f) July, (g) August, (h) September, and (i) October 2019 in Houston, TX. NWS radar retrieved PBLH in green, point – based WRF PBLH estimates in purple, and region– based WRF estimates in orange.

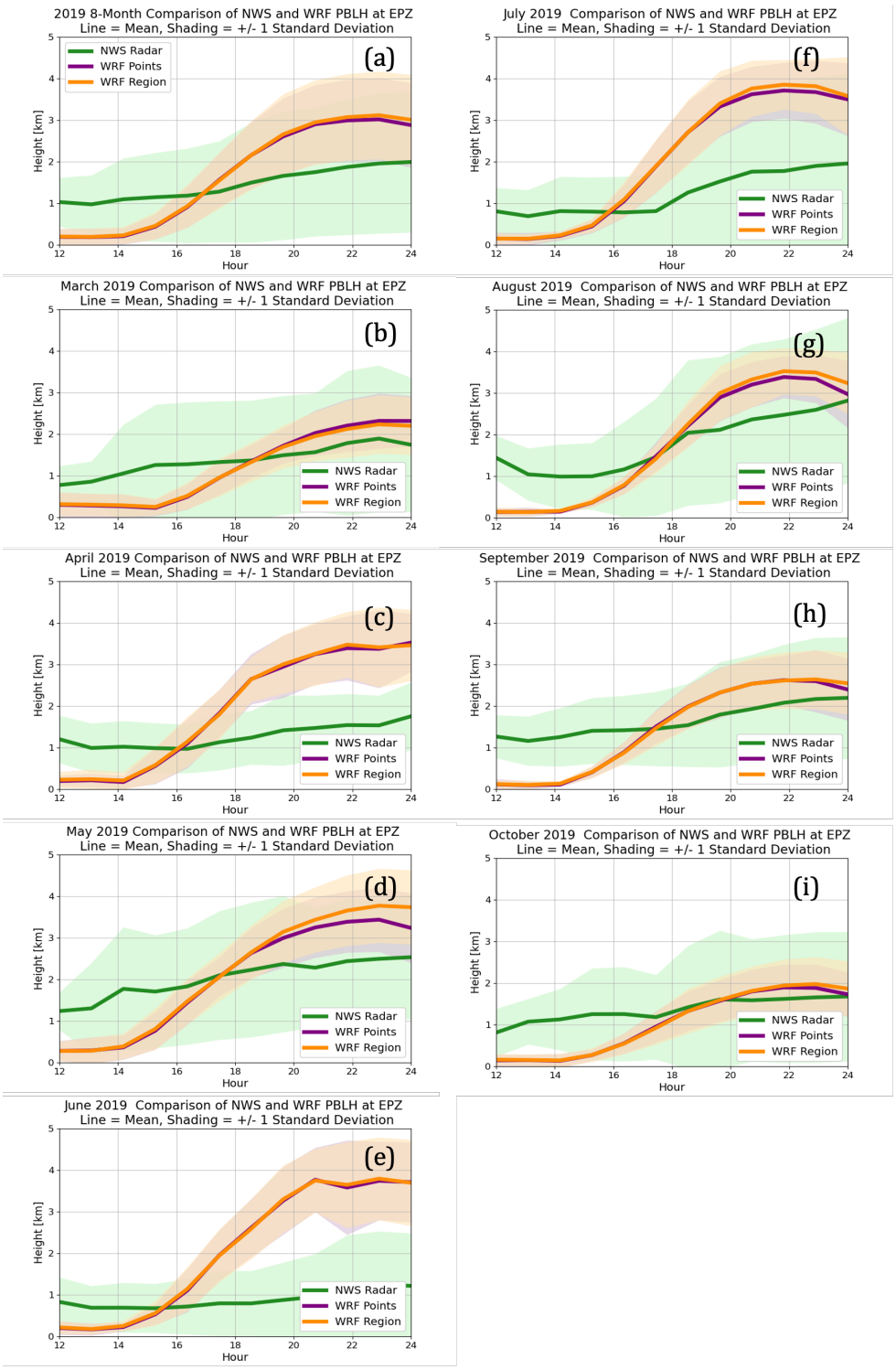

Figure 6 - Time series of the mean (lines) and +/- one standard deviation (shading) of the NWS radar retrieved and WRF simulated PBLH for (a) March – October, (b) March, (c) April, (d) May, (e) June, (f) July, (g) August, (h) September, and (i) October 2019 in El Paso, TX. NWS radar retrieved PBLH in green, point – based WRF PBLH estimates in purple, and region– based WRF estimates in orange.

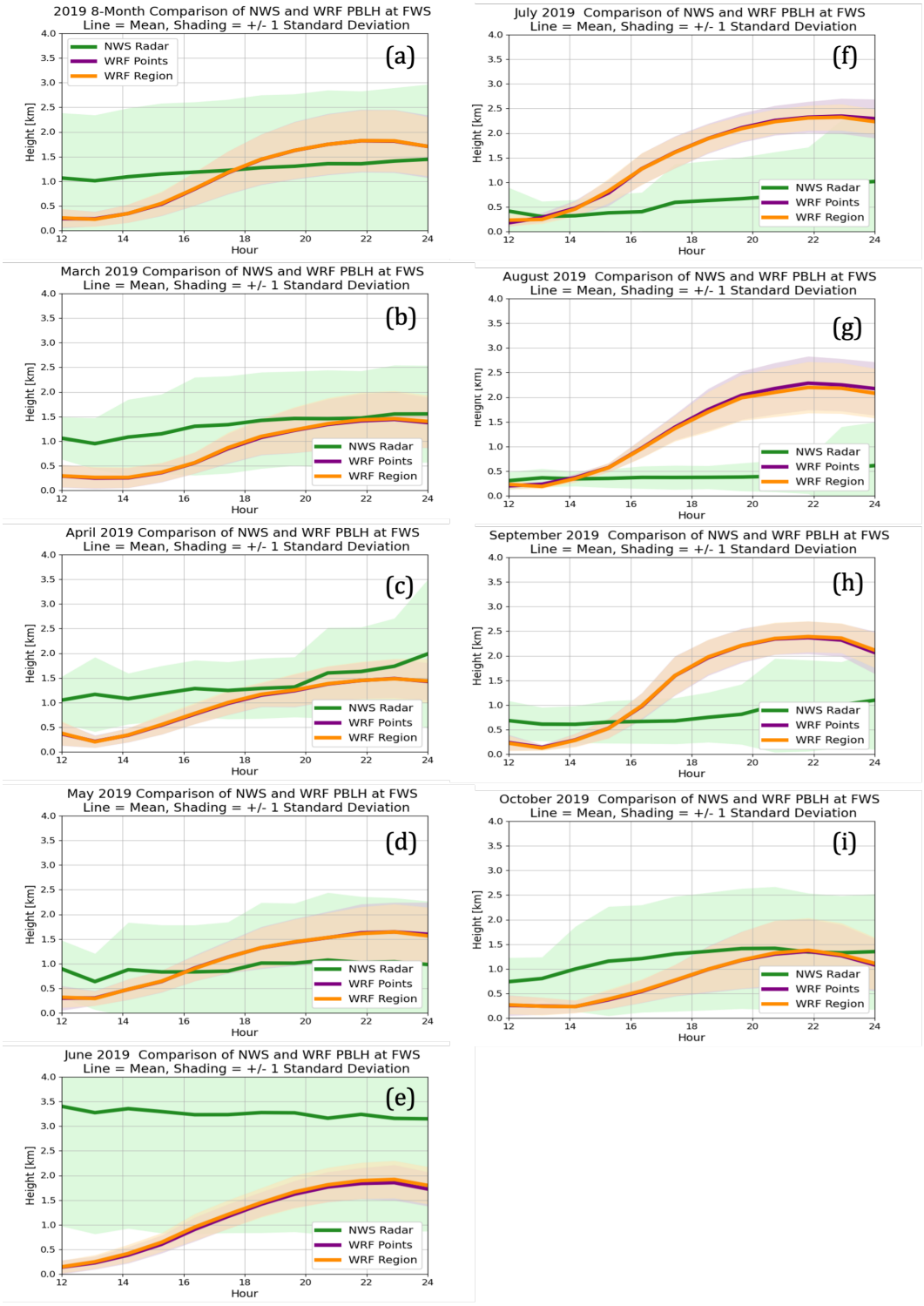

Figure 7 - Time series of the mean (lines) and +/- one standard deviation (shading) of the NWS radar retrieved and WRF simulated PBLH for (a) March – October, (b) March, (c) April, (d) May, (e) June, (f) July, (g) August, (h) September, and (i) October 2019 in Dallas, TX. NWS radar retrieved PBLH in green, point – based WRF PBLH estimates in purple, and region– based WRF estimates in orange.

8 Analysis of Project Results

8.1 Major Activities and Key Findings

This project refined a radar-based approach to estimating PBLH that was delivered to TCEQ by AER in 2022. Additional capabilities were added, the most important of which was an observation operator that can be applied to WRF PBLH fields in order to compute a single PBLH height estimate at the same location as the radar data. This facilitated direct comparison between WRF and radar PBLH estimates for an 8-month period in 2019 for the Houston, El Paso and Dallas radars. Software was delivered on the TCEQ machine 'eira' and training was given to TCEQ staff.

The following key findings highlight our impression of the comparison between WRF and radar PBLH estimates:

A – The overarching result is that the WRF simulated PBLH values are physically reasonable and are consistent with the meteorological experience of AER researchers. They exhibit appropriate diurnal behavior and pronounced sensitivity, as expected, to the marine boundary layer over the Gulf of Mexico. The PBLH estimates are higher in the arid El Paso area as expected by climatology.

B - The differences between the NWS retrieved and WRF simulated PBLH are increased by application on days and time periods when the PBLH is poorly defined or is otherwise contaminated by innumerable meterological factors. These can include recent precipitation or thick cloud cover. Developing automated routines to limit the application to appropriate time periods is challenging.

C – The automated radar algorithm performs well on high quality days, but requires additional adjustments through further validation to perform optimally. Recent adjustments to the processing details during the early morning time period has enhanced the ability to accommodate a variety of environmental profiles but more work is needed.

9 Problems Encountered and Mitigation Strategies

AER encountered some minor problems with rebuilding the Dockerfile on 'eira' but these were resolved relatively quickly with no downstream consequences. AER was only provided with WRF simulations for El Paso late in the period of performance, however, AER delivered the project results on time.

10 Recommendations

We believe the potential of the radar technique to aid in evaluation of WRF simulations has great potential that has not yet been fully realized. Implementation of these recommendations will greatly enhance its value to TCEQ.

The first recommendation is to apply further improvements to the radar algorithm. This would, for example, minimize identification of elevated minima of the differential reflectivity field in the morning and afternoon shoulder periods. We recommend use of additional validation datasets, such as radiosonde PBLH estimates and observations through TRACER-AQ, to first identify how well the radar approach performs under a variety of conditions. This information will guide refinements to the radar algorithm. The code could also be further optimized so that large datasets associated could be processed more efficiently.

Second, we recommend use of additional WRF fields as another layer of quality control to discriminate when to apply the comparison between radar and WRF fields. A plethora of fields could be used to evaluate the structure of the PBL and aid in determining whether the PBLH in the WRF dataset is of sufficiently high quality to use in comparisons with the radar data.

11 References

Banghoff, J. R., Stensrud, D. J., & Kumjian, M. R. (2018). Convective boundary layer depth estimation from S-band dual-polarization radar. Journal of Atmospheric and Oceanic Technology, 35(8), 1723-1733.

Eure, K. C., P. D. Mykolajtchuk, Y. Zhang, D. J. Stensrud, F. Zhang, S. J. Greybush, and M. R. Kumjian, 2023: Simultaneous Assimilation of Planetary Boundary Layer Observations from Radar and All-Sky Satellite Observations to Improve Forecasts of Convection Initiation. Mon. Wea. Rev., 151, 795–813, https://doi.org/10.1175/MWR-D-22-0188.1.

Milovac, J., K. Warrach-Sagi, A. Behrendt, F. Späth, J. Ingwersen, and V. Wulfmeyer, 2016: Investigation of PBL schemescombining the WRF model simulationswith scanning water vapor differentialabsorption lidar measurements,J. Geophys. Res. Atmos.,121, 624–649,doi:10.1002/2015JD023927

Xi, X., Zhang, Y., Gao, Z., Yang, Y., Zhou, S., Duan, Z., & Yin, J. (2022). Diurnal climatology of correlations between the planetary boundary layer height and surface meteorological factors over the contiguous United States. International Journal of Climatology, 42(10), 5092– 5110. https://doi.org/10.1002/joc.7521

Xie, B., J. C. H. Fung, A. Chan and A. Lau, 2012: Evaluation of nonlocal and local planetary boundary layer schemes in the WRF model, *J. Geophys. Res.*, 117, D12103, doi:10.1029/2011JD017080.

Zhang, Y., Sun, K., Gao, Z., Pan, Z., Shook, M. A., & Li, D. (2020). Diurnal climatology of planetary boundary layer height over the contiguous United States derived from AMDAR and reanalysis data. Journal of Geophysical Research: Atmospheres, 125, e2020JD032803. https://doi.org/10.1029/2020JD032803