AQRP Project 24-021

Improving WRF Representation of Coastal, Marine, and Residual Boundary Layers and Quantifying the Effects on Ozone Prediction

Final Report

August 31, 2025

Principal Investigators

Yuxuan Wang, University of Houston

James Flynn, University of Houston

Contributors

Shailaja Wasti, University of Houston

Travis Griggs, University of Houston

Geoffrey Roberts, University of Houston

Yong-Cheol Jeong, University of Houston

Semko Momeni, University of Houston

Acknowledgement

The preparation of this report was funded by a grant from the Texas Air Quality Research Program (AQRP) at The University of Texas at Austin through the Texas Emission Reduction Program (TERP) and the Texas Commission on Environmental Quality (TCEQ). The findings, opinions, and conclusions are the work of the author(s) and do not necessarily represent the findings, opinions, or conclusions of the AQRP or the TCEQ.

Executive Summary

The predictability of ozone and its precursors is strongly dependent on complex meteorological conditions that influence their development. Among the meteorological variables that affect air quality, the planetary boundary layer (PBL) plays a crucial role in atmospheric simulations, yet it remains challenging to predict. Leveraging new coastal and offshore observations of PBL height (PBLH) from the Tracking Aerosol Convection ExpeRiment-Air Quality (TRACER-AQ) campaigns during 2021-2023, this project was aimed at improving the Weather Research and Forecasting (WRF) model-based prediction of marine PBLH and the residual layer (RL) and assessing the associated effect on ozone prediction by the Comprehensive Air Quality Model with extensions (CAMx) in the Houston-Galveston-Brazoria (HGB) region.

First, WRF v4.6.0 was evaluated for its ability to simulate PBLH and key meteorological variables over Galveston Bay and the Gulf of America. The base configuration systematically underestimated PBLH, particularly over Galveston Bay in the afternoon, where observed PBLH frequently exceeded 1.5 km. In the Gulf, both observations and simulations showed weaker diurnal variability, with the model underestimating afternoon PBLH. Analysis along multiple boat tracks revealed location- and time-dependent biases, including overestimation along the northwest Houston Ship Channel (HSC) and underestimation along the southwest HSC. Temperature and pressure were well represented, while relative humidity and winds showed larger biases. This evaluation provided a baseline for tuning WRF to better capture marine PBL processes.

Second, the project identified sensitive physical parameters and constants in the local closure Mellor-Yamada-Nakanishi-Niino Level 2.5 (MYNN2.5) PBL scheme and explored different configurations in the WRF v4.6.0 base model that have the potential to improve the PBLH. Targeted perturbation experiments showed that marine PBLH predictions were strongly influenced by surface roughness, ocean coupling, and cumulus parameterizations. Among different sets of simulations, three configurations (PBLH23, PBLH24, and PBLH25) robustly outperformed the base simulation and increased the marine PBLH by up to 400 – 500 m.

Third, empirical and theoretical methods were developed to diagnose the RL, addressing the WRF base model's limitation of identifying only the shallow nighttime stable boundary layer. The empirical method relied on potential temperature gradients, while the theoretical method incorporated turbulence indicators from wind, temperature, and turbulent kinetic energy (TKE) profiles. Both captured RL structures, but the theoretical approach showed stronger correlation with observations. With the developed RL diagnostic, it was verified that the WRF configurations yielding better marine PBL prediction also offered better RL predictions. The diagnosed RL heights were later incorporated into CAMx perturbation simulations to represent layered boundary structures more realistically.

Fourth, the improved PBL simulations in WRF were passed to CAMx to assess their impact on ozone predictions and source attribution. Offshore ozone predictions improved substantially, with PBLH23 and PBLH25 reducing CAMx biases by 15-20 ppbv along ship tracks and increasing monthly mean ozone by ~5 ppbv. Changes at land-based monitors were smaller, but high-ozone days saw 20-30% higher locally produced ozone under improved PBL due to enhanced vertical transport. Vertical profile comparisons confirmed that PBLH23 and PBLH25 better matched ozonesonde observations than the Base case, while PBLH24 consistently

overpredicted ozone. Additionally, a diagnostic RL was implemented in CAMx by replacing WRF nighttime PBLH with diagnosed RL values, which increased surface ozone by ~0.5-1.5 ppbv across the years, reflecting enhanced redistribution of aloft ozone during stable nighttime conditions.

In summary, the project demonstrates that integrating offshore observations with targeted model development can substantially improve the representation of coastal PBLH and RL in the WRF model, which in turn improves photochemical model prediction of offshore ozone. The best-performing simulation, PBLH25, has the combined configuration of higher surface roughness over water, a 1-D ocean mixed layer model, and the New Simplified Arakawa-Schubert (NSAS) scheme. These configurations and RL implementation represent promising pathways to further advancing coupled WRF-CAMx modeling and improving coastal ozone predictions in the HGB region.

Table of Contents

1. Introduction

Meteorological processes, particularly those affecting the vertical structure of the lower atmosphere, play a crucial role in determining the distribution, transport, and accumulation of ozone and its precursors (Wang et al., 2021; J. Zhang & Rao, 1999). A key factor in this vertical structure is the planetary boundary layer (PBL), the lowest 1-2 km of the atmosphere, which regulates the exchange of heat, moisture, and momentum between the Earth's surface and the free troposphere (Stull, 1988). This layer is highly dynamic and responds rapidly to surface forcing such as heating, moisture, and land-water contrasts. The depth of the PBL, referred to as the planetary boundary layer height (PBLH), determines the extent of vertical mixing of near-surface emissions. It is challenging to accurately estimate PBLH in numerical weather prediction models, especially over coastal and marine areas due to the complexity stemming from various physical processes (Ghonima et al., 2017; Shaw et al., 2022). This uncertainty can directly affect photochemical modeling of ozone and its precursors driven by numerical weather prediction models.

The Houston-Galveston-Brazoria (HGB) region along the Texas Gulf Coast continues to experience persistent nonattainment of the National Ambient Air Quality Standards (NAAQS) for ozone as set by the U.S. Environmental Protection Agency (EPA), mainly due to the region's complex meteorology and emissions environment (TCEQ, 2022). To improve the understanding of marine meteorology and coastal boundary layer processes in the HGB region, multiple field campaigns were conducted in 2021-2023, including the Tracking Aerosol Convection ExpeRiment-Air Quality (TRACER-AQ) studies during July – October 2021 (TAQ1) (Jensen et al., 2021), April-October 2022 (TAQ2, TCEQ PGA# 582-22-32022-021), and the 2023 Mobile and Offshore Air Quality Monitoring Project during May-October 2023 (TAQ3, TCEQ PGA# 582-23-43296-028). These field campaigns provided, for the first time, observations of PBLH at multiple coastal and marine locations.

This Air Quality Research Program (AQRP) project aimed at validating and improving the Weather Research and Forecasting (WRF) model-based coastal and marine PBLH predictions over Galveston Bay and the Gulf by using the abundance of offshore observation data from TRACER-AQ campaigns during 2021-2023 and assessing the associated effect on ozone prediction. This aim was achieved through the successful execution of four related tasks.

First, the project evaluated the performance of the WRF model in simulating PBLH over coastal and marine environments. Offshore observations from field campaigns revealed that the baseline WRF v4.6.0 model configuration systematically underestimated PBLH over water, particularly in the afternoon, and failed to capture key aspects of the observed diurnal and spatial variability. This evaluation highlighted the limitations in the model's boundary layer parameterizations under complex coastal conditions.

The second task was to improve the prediction of marine PBLH by adjusting the physics parameters and exploring different configurations in the WRF v4.6.0 base model. The key physics parameters were identified and their ranges that most likely affect the evolution of PBLH in the Mellor-Yamada-Nakanishi-Niino Level 2.5 (MYNN2) scheme, along with some other additional parameters like surface roughness and albedo in the model. In addition to the physics

perturbations, numerous sensitivity simulations were also conducted by implementing different cumulus schemes and urban physics schemes, which are found to be effective in improving the PBLH from the base simulations. The perturbation simulations were evaluated by using selected PBLH observations from the TRACER-AQ field campaigns described above. The evaluation of these simulations helped to identify the most sensitive physics parameters with their corresponding values that were effective at producing the PBLH values closer to the observations. Three best-performing configurations of WRF (PBLH23, PBLH24, and PBLH25) were identified and used to drive the photochemical modeling conducted in the fourth task.

The third task focused on improving the representation of the residual layer (RL) in the WRF model. The WRF base configuration exhibits a notable gap in representing the RL height, as it only estimates the stable and convective boundary layers while omitting the elevated residual structure (Liu et al., 2023). In this task, first, both an empirical and a theoretical method were developed to diagnose RL height from the WRF base model. The empirical method uses the vertical gradient of potential temperature ($\partial\theta/\partial z$), while the theoretical method incorporates a combination of dynamic and thermodynamic parameters, including vertical profiles of wind speed, vertical velocity, potential temperature, turbulent kinetic energy (TKE), and their respective gradients. These diagnosed RL heights were then verified against available observational datasets. For the three best-performing perturbation simulations, predicted PBLH were also evaluated against observations. Both the empirical method and the theoretical method were used to calculate RL height across the entire simulation domain for both the WRF base model and the selected perturbation simulations to evaluate the changes in RL from the perturbation simulations.

The fourth task investigated the effects of improved PBL on ozone prediction and source attribution in the HGB region. To this end, Comprehensive Air Quality Model with Extensions (CAMx) v7.31 simulations driven by meteorological fields from three best-performing PBL perturbation simulations (PBLH23, PBLH24, and PBLH25) were conducted and quantified the resulting changes in CAMx-predicted surface ozone concentrations as compared to those CAMx outputs using the base WRF v4.6.0 model. Ozone predictions from the different simulations were also evaluated against the TRACER-AQ field campaign data to evaluate the extent to which improved PBL simulations affect the evaluation statistics of CAMx ozone prediction. The base CAMx simulation (i.e., CAMx driven by the base WRF v4.6.0 model) was also compared with our previous base CAMx simulations using an earlier version of WRF and CAMx to assess the differences caused by model version updates. Ozone source attribution was previously estimated based on CAMx emission perturbation experiments during the September 2021 period (AQRP 22-008). The same emission perturbation experiments were conducted in Task 4 to estimate the changes in ozone source attribution resulting from PBL perturbation simulations.

This report documents the analysis and findings from the project, and it is organized as follows. Section 2 describes the numerical models and the observational datasets used in the analysis. Section 3 evaluates the performance of the WRF v4.6.0 base model over the coastal and marine environment. In Section 4, perturbation simulations conducted by adjusting the physics ensembles in the WRF v4.6.0 base model are presented. Section 5 describes how the representation of the residual layer in the WRF model was improved. Section 6 assesses the effects of improved PBL representation on ozone prediction and source attribution in the HGB region.

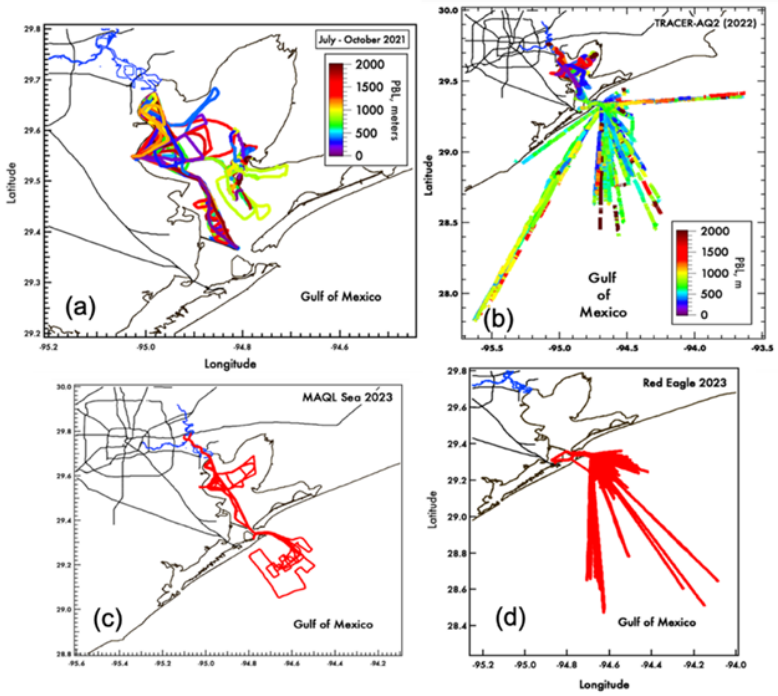


Figure 1. Offshore PBLH measurements taken by the 2021-2023 field campaigns. (a) PBL heights measured by UH Pontoon boat and a shrimp boat during TAQ1 (Jul – Oct 2021); (b) PBL heights measured by the UH Pontoon boat (Galveston Bay) and Red Eagle charter boat (the Gulf) during TAQ2 (Apr – Oct 2022); (c) PBL sampling locations by the UH Osprey in Sep 2023; (d) PBL sampling locations by Red Eagle during May – Oct 2023. Red lines in (c) and (d) show the offshore sampling tracks, not PBL heights.

2. Datasets and Models

2.1. Field campaigns and Observations

The project uses coastal and offshore PBLH observations from the TRACER-AQ field campaigns, which took place from July to October 2021 (TAQ1), April to October 2022 (TAQ2), and May to October 2023 (TAQ3). To measure the PBLH over the water, Vaisala CL-51 ceilometers were installed on three boats: the University of Houston pontoon boat (UHPB), the Red Eagle (RE), and the UH Osprey boat (UHOB). The RE primarily operates in the Gulf, while the Osprey and the pontoon boat mainly operate around Galveston Bay, Trinity Bay, and the Houston Ship Channel (HSC). The ceilometers onboard these boats are manufactured by Vaisala and provided valuable PBL data for both Galveston Bay and the Gulf.

The Vaisala software internally processes the measurements to identify three distinct boundary layers based on gradients of aerosol backscatter. In the 2022 field campaign (TAQ2), the consistent low PBLH values over the Gulf (~0.1 km) were noticed in the lowest retrieved PBL layer measured by the ceilometer onboard the RE (**Figure 2**) after the TAQ2 campaign concluded. To determine whether this was due to a malfunction in the instrument or if the PBLH over the Gulf was genuinely low, the ceilometer onboard the RE was replaced with a unit from

the UHPB during TAQ3. In 2023, the lowest retrieved PBL layer over the Gulf ranged from 0.5 to 1 km, indicating that the low PBLH values derived by the software in 2022 were due to an instrument malfunction. The problematic ceilometer was placed on UHOB in the 2023 campaign, the measurements from which are subject to the same issues.

Because the PBL retrieved by the ceilometer over the RE in 2022 and over UHOB in 2023 were found to be incorrect, the raw aerosol backscatter data from the ceilometer were reprocessed to more accurately characterize the PBL. The raw daily aerosol backscatter data were plotted on a logarithmic scale and analyzed to determine hourly PBL heights by identifying significant aerosol gradients. An example of the raw aerosol backscatter along with the reprocessed PBLH from the raw aerosol backscatter data for September 10 and 21, 2022 is shown in **Figure 2**. The re-processed PBLH shows more reasonable results and matches well with the aerosol backscatter gradients. **Figure 3a** compares the average diurnal patterns of the original ceilometer-identified PBLH with the reprocessed PBLH for September 2022 over the Gulf. The reprocessed data gives nighttime PBLH around 0.5 km, which increases to about 0.9 km during the daytime. For September 2023 over the Bay, the reprocessed PBLH shows a prominent diurnal cycle with a nighttime PBLH around 0.5 km, increasing to 1.2 km during the daytime (**Figure 3b**). All subsequent analysis used the re-processed PBLH values from the RE in 2022 and from the UHOB in 2023. The only drawback of the re-processed PBLH values is that its temporal resolution is hourly and the corresponding measurement locations are all condensed onto hourly time scales in the subsequent plots.

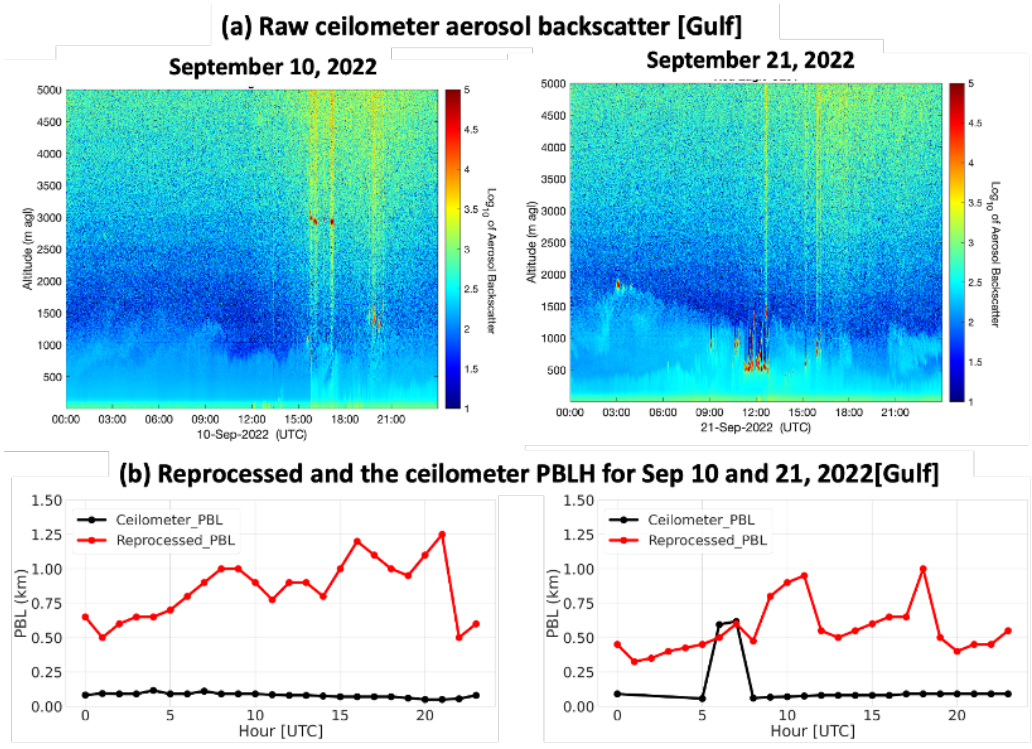


Figure 2. (a) Raw ceilometer aerosol backscatter plot and (b) the corresponding reprocessed PBLH, along with the ceilometer retrieved first layer of PBLH for September 10 and 21, 2022

measured by the RE. In the line plot, the red line represents the reprocessed PBLH, while the black line represents the ceilometer PBLH.

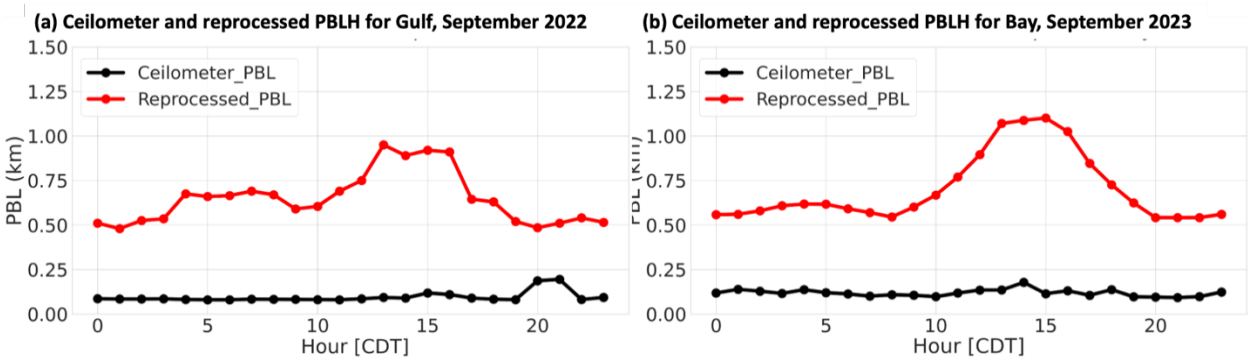


Figure 3. (a) Average diurnal boundary layer heights retrieved by the Vaisala CL51 ceilometers onboard the Red Eagle and the PBLH derived from the raw aerosol backscatter data for September 2022. (b) Similar to (a) but for UHOB in September 2023 over Galveston Bay. In the line plot, the red line represents the reprocessed PBLH, while the black line represents the ceilometer PBLH.

2.2. WRF model

The Weather Research and Forecasting (WRF) model version 4.6.0 with the Advanced Research WRF (ARW) solver was employed to simulate meteorological fields for three domains (as shown in **Figure 4)** over the contiguous United States (d01 - 12km × 12km), Southeast Texas (d02 - 4km × 4km), and the Houston-Galveston region (d03 - 1.33km × 1.33km), respectively. All three domains have identical vertical resolutions with 30 vertical levels from the surface to ~100 hPa. The WRF configurations included the local closure Mellor-Yamada-Nakanishi-Niino Level 2.5 (MYNN2) PBL scheme (Nakanishi & Niino, 2009), Morrison double moment (2M) microphysics scheme (Morrison et al., 2009), Rapid Radiative Transfer Model (RRTMG) longwave and shortwave radiation schemes (Iacono et al., 2008), Monin-Obukhov similarity surface layer scheme (Chen et al., 1997), Noah land-surface module (Chen & Dudhia, 2001), and the New Tiedtke cumulus parameterization (Tiedtke, 1989; C. Zhang et al., 2011). The initial and boundary conditions (IC/BC) for domain d01 were derived from the High-Resolution Rapid Refresh (HRRR) model, while the IC/BC for d02 were generated from d01, and those for d03 were generated from d02.

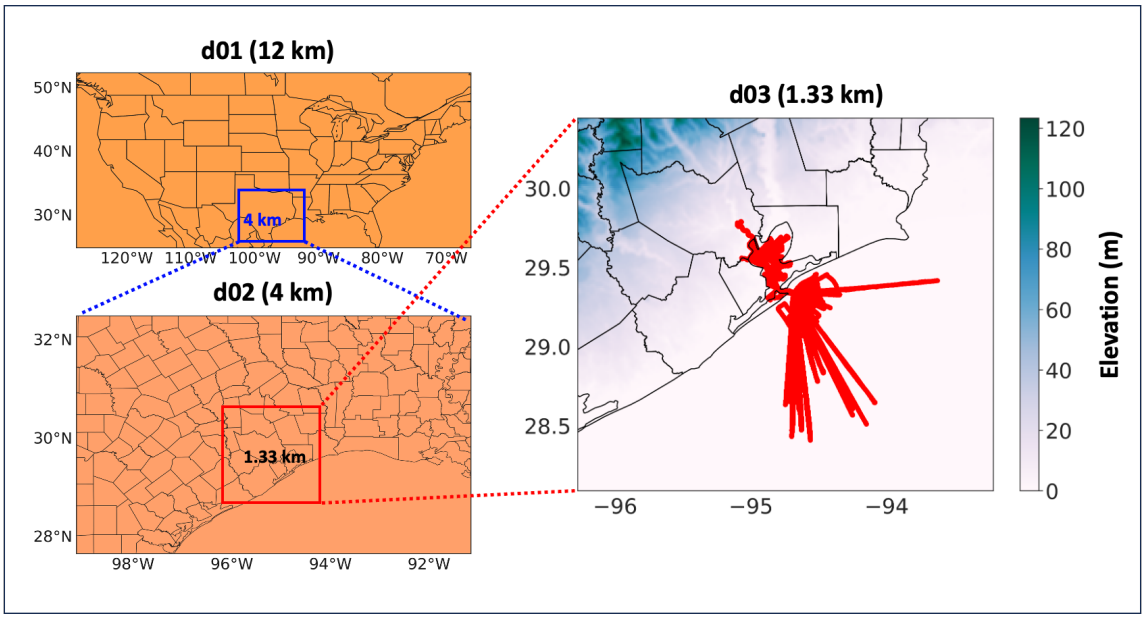


Figure 4. WRF domains showing d01 (contiguous United States), d02 (Southeast Texas), and d03 (Houston-Galveston) regions. The red lines in d03 show the ship tracks where PBLH measurements were made during the 2021-2023 period.

2.3. CAMx model

Meteorological fields from WRF were directly used as the input to the Comprehensive Air Quality Model with extensions (CAMx) version 7.31 over d01, d02, and d03 domains. The initial and boundary conditions inputs for the d01 domain were obtained from the GEOS-Chem (v14.1.1) global simulation, while IC/BC for d02 and d03 were taken from d01 and d02, respectively. The first-order eddy viscosity (K-theory) diffusion scheme was selected for vertical mixing within the PBL, and the Carbon Bond version 6 revision 5 (CB6r5) was used for gas-phase chemistry (Burkholder et al., 2019). Dry deposition was implemented according to the Wesely scheme (Wesely, 2007). Anthropogenic emissions data from the 2019 State Implementation Plan (SIP) modeling platform provided by TCEQ were converted and used in the model as discussed in (Li et al., 2023). Biogenic emissions are generated from the Biogenic Emission Inventory System (BEIS), and wildfire emissions are based on the Fire INventory from NCAR (FINNv2) from the SIP 2019 emission inventory.

For both WRF and CAMx models, the model performances were compared to observations using the metrics in **Table 1**. The impact of PBL perturbation simulations on each model's performance was evaluated using difference plots and the same metrics from **Table 1** to quantify the between-model differences.

Table 1. Performance metrics of the WRF and CAMx model.

Performance Metrics	**Formulas**
Mean Bias (MB)	$MB = 1/N\sum_{i=1}^{N}(M_i - O_i)$
Mean Absolute Error (MAE)	$MAE = 1/N\sum_{i=1}^{N}\lvert M_i - O_i\rvert$
Normalized Mean Bias (NMB)	$NMB = \frac{\sum_{i=1}^{N}(M_i - O_i)}{\sum_{i=1}^{N}O_i} \times 100\%$
Correlation Coefficient (Corr. R)	$Corr.R = \frac{\sum_{i=1}^{N}(M_i - \overline{M})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{N}(M_i - \overline{M})^2}\sqrt{\sum_{i=1}^{N}(O_i - \overline{O})^2}}$
Root Mean Square Error (RMSE)	$RMSE = \sqrt{1/N\sum_{i=1}^{N}(M_i - O_i)^2}$

Note: M is the model output, O is the observation, N is the number of samples, and $\overline{M} = 1/N\sum_{i=1}^{N}M_i, \overline{O} = 1/N\sum_{i=1}^{N}O_i$.

3. Marine PBL synopsis and model evaluation

3.1. Comparison of observed and modeled PBLH variation over the water

This section presents the evaluation of the WRF performance on PBLH estimation over Galveston Bay and the Gulf across the different years, with an emphasis on understanding the spatial and temporal patterns of PBLH over these open-water regions and the ability of the WRF model to capture those patterns. WRF v4.6.0 was run for the selected months of the 2021–2023 field campaign, where the availability of mobile observations over Galveston Bay and/or the Gulf spanned at least 10 observation days. These months are July–October 2021, August–September 2022, and May–October 2023, covering a total of 12 months over the three-year period. The WRF simulated outputs had a temporal resolution of an hour, while the observation data were recorded at 5-minute intervals for 2021 and 1-minute 2022 and 2023, except for the re-processed PBLH observations from the RE in September 2022 and the UHOB in September 2023, which were hourly as described above in Section 2.1. To obtain a visual representation of the modeled and observed PBLH, the native resolution of the observation data was retained

without averaging, considering the rapid changes in latitude and longitude of the mobile measurements.

The spatial variations of observed and simulated PBLH for selected months from 2021-2023 over Galveston Bay are shown in **Figure 5**. The figure shows daytime (7 AM to 7 PM) PBLH only. Periods when the boats were docked at night were excluded to minimize the impact of land-based influences. Over Galveston Bay, observations show that the PBLH is generally higher around the Houston Ship Channel (HSC) and lower toward the eastern part of the Bay. The model tends to capture these spatial variations, though with a slight underestimation. It is important to note that PBLH variations depicted in **Figure 5** could be influenced by differences in observation time and location. For example, in 2022, observations were mainly conducted in the afternoon over the northwest HSC, a narrow water body surrounded by land. During this period, the observed PBLH was approximately 1.5 km, and the model slightly overestimated the PBLH by about 0.3 km. This overestimation may be attributed to the model grid being closer to the land. In contrast, the model underestimates the afternoon PBLH over the southwest HSC, where the boat was moving farther from land toward open water. Earlier in the day, the boat mainly sampled the eastern part of the Bay, revealing lower PBLH values (**Figure 5b (i**)), and the model tends to slightly overestimate the PBLH in these morning times (**Figure 5b (ii**)). The summary statistics for each year's model evaluation are provided in the figures.

Figure 6 shows the spatial variation of observed and model PBLH over the Gulf for September 2022 and May-October 2023. The spatial map shows that the observed PBLH values are lower in the morning and higher in the afternoon over the Gulf, indicating a diurnal pattern of PBLH. The model captures this variation well for 2022, with a slight underestimation (-0.14 km) for 2023. The model prediction of PBLH is influenced by the time of day. In the morning, as observed from **Figure 6**b (iv), e.g., in 2023, the model slightly overestimated the PBLH, while in the afternoon, the model underestimated the PBLH. Compared to Galveston Bay, PBLH values over the Gulf are typically lower (~0.3 km) during the afternoon. The summary statistics for each year's model evaluation are provided in the figure. Detailed investigations on PBLH based on different boat routes and observation times will be presented in the next section.

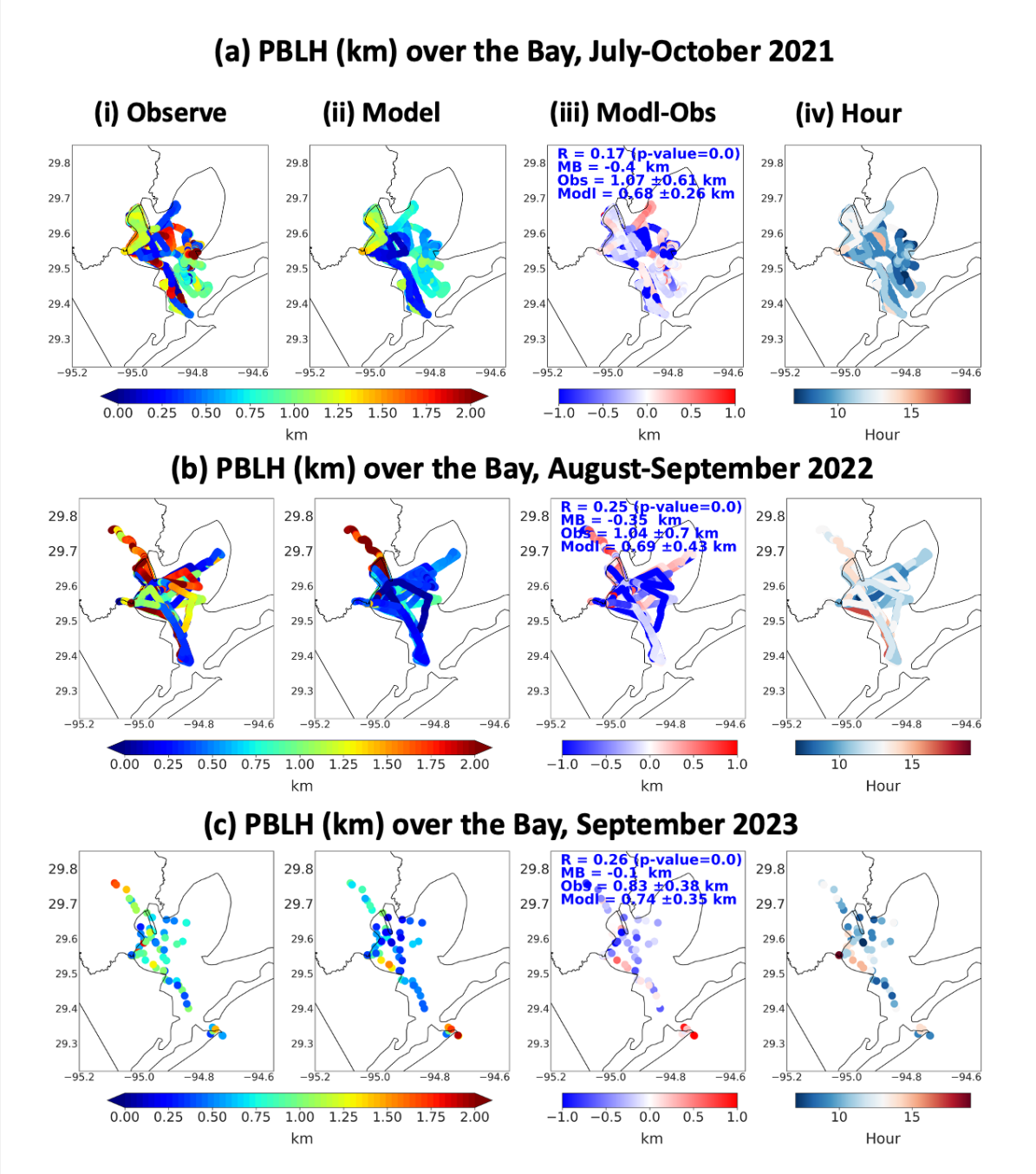


Figure 5. (a) Map of the observed, modeled, and modeled-observed PBLH (km) over Galveston Bay from July-October 2021, along with the time/hour when the boat was mobile. (b) Same as (a), but for 2022 from August-September. (c) Same as (a), but for September 2023. The plot shows daytime (7:00 CDT – 19:00 CDT) mobile data only.

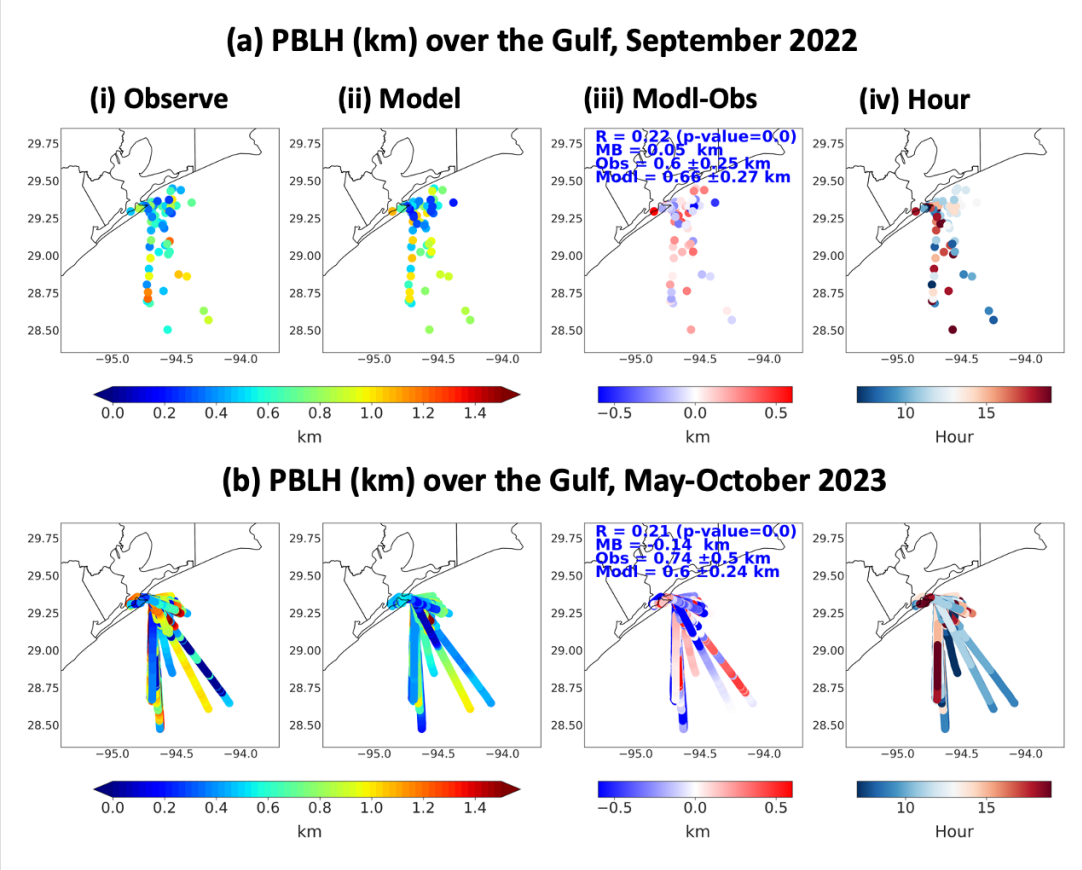


Figure 6. (a) Map of the observed, modeled, and modeled-observed PBLH (km) over the Gulf for September 2022, along with the time/hour when the boat was mobile. (b) Same as (a), but for 2023 from May-October. The plot shows daytime (7:00 CDT – 19:00 CDT) mobile data only.

To understand temporal variations of PBLH over the water, the diurnal variations of PBLH over Galveston Bay and the Gulf for all three years are presented in **Figure 7** and **Figure 8**, respectively. The month of September was selected for the diurnal variation analysis because the observation data were available for all three years in this month for Galveston Bay and for two years for the Gulf. Over the Bay, the model underestimates the PBL with a mean bias (MB) of -0.11 km, -0.17 km, and -0.16 km for September 2021, 2022, and 2023, respectively. The time series plot includes both mobile and stationary data (when the boat was docked) to see the overall diurnal variation, as the boat is mobile only during the daytime. The observed PBL shows a prominent diurnal variation over the Bay, and the model tends to capture these variations well (**Figure 7**). The diurnal pattern of PBLH over the Bay is similar to land-based PBLH measurements observed at LaPorte in September 2021 (Liu et al., 2023). The model's diurnal variation of PBLH shows consistent interannual variability over Galveston Bay across the years. However, the diurnal variations from the observed data show a low diurnal variation in 2023. The difference for the observed diurnal variation in 2023 might be due to the differences in the method used to obtain PBLH. In 2021 and 2022, PBLH was the ceilometer-retrieved first layer of the boundary layer, while in 2023, it was the reprocessed PBLH based on the raw aerosol backscatter data.

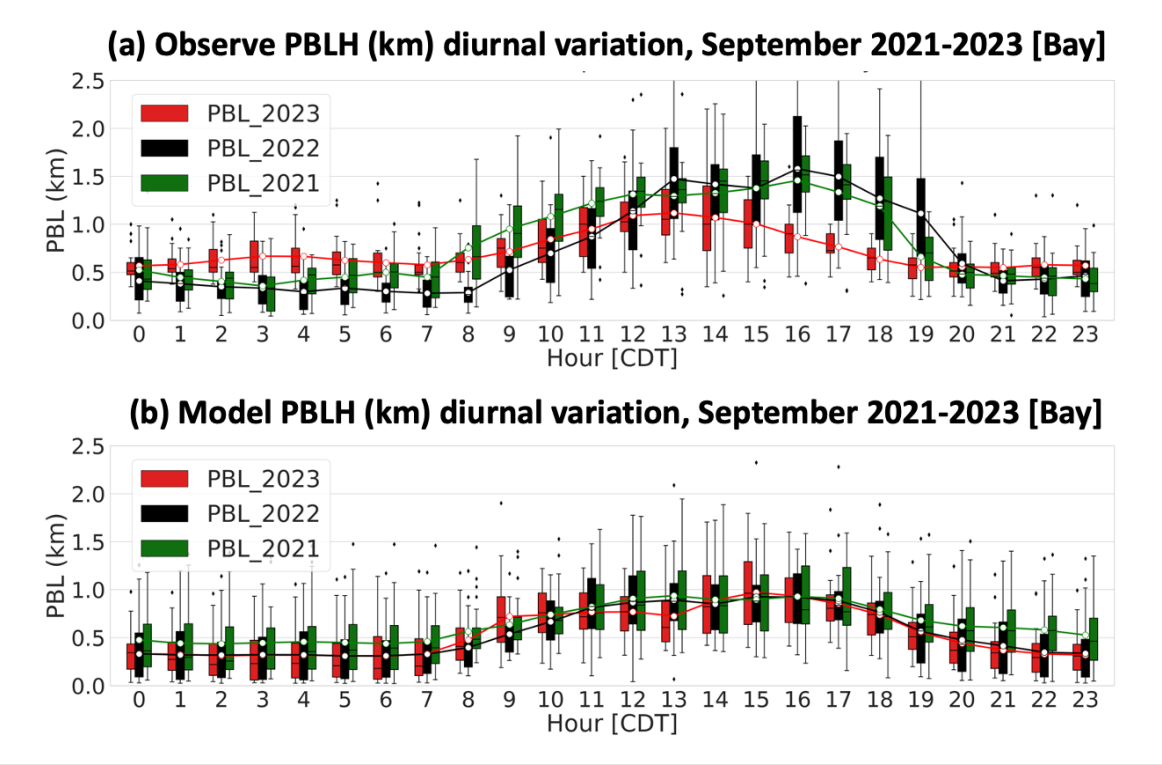


Figure 7. Boxplot of the diurnal variation of (a) Observed and (b) Model PBLH over Galveston Bay for September 2021 (green), 2022 (black), and 2023 (red)

Figure 8 shows the observed and modeled diurnal variations of PBLH for September 2022 and 2023 over the Gulf. The model slightly overestimated mean PBLH, with an MB of 0.092 km for September 2022 and 0.03 km for 2023, respectively. Observed PBLH over the Gulf exhibits a pronounced diurnal variation in 2023, while the variation is less noticeable in 2022. In contrast, the model's diurnal variation is less pronounced in both years.

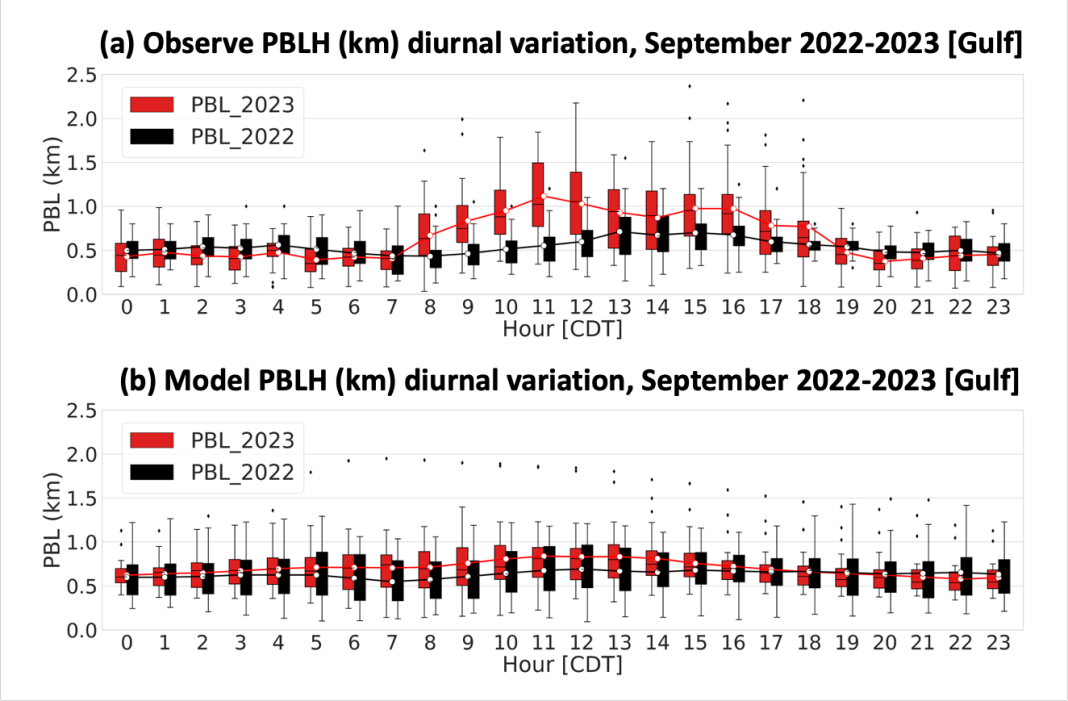


Figure 8. Boxplot of the diurnal variation of (a) Observe and (b) Model PBLH over the Gulf for September 2022 (black) and 2023 (red)

Month-by-month variations in PBLH were also analyzed over the Galveston Bay and the Gulf regions, as shown in **Figure 9** and **Figure 10**, respectively. Over Galveston Bay, the highest mean observed PBLH is found in July, with a gradual decrease to October when the PBLH was the lowest on average (**Figure 9a**). Model PBLH follows a similar seasonal trend as the observed data but with slightly less variability (**Figure 9b**). The observed PBLH values in July, August, and October show good agreement across the years between 2021 and 2022. However, the variability across years is large for September compared to other months in that year, potentially due to the significant number of mobile observation days (up to 13 days) compared to other months. The box-plot comparison between the observed and modeled PBLH reveals that the model is better at reproducing the lower quantile of the observed PBLH, including their month-to-month and year-to-year variability (e.g., September), but mostly underestimates the upper quantile and extremely high values of the observed PBLH in almost all months. The model also fails to reproduce the month-to-month and year-to-year variations of those higher ends. For example, observations show that PBLH was higher in September 2021 than in September 2022, but the model showed the opposite trend.

Figure 10 shows the observed and modeled month-to-month variations of PBLH over the Gulf for September 2022 and May-October 2023. The observed monthly variation shows that the mean PBLH in 2023 is highest in September and lowest in October, with similar mean values for all other months. In contrast, the model shows that September and October have slightly higher mean PBLH than the other months. The observed PBLH variability is higher for each month compared to the modeled PBLH. For example, in May 2023, the lower quantiles are around 0.3 km, and the upper quantiles are approximately 0.9 km, indicating diurnal variation over the Gulf. However, the model shows less variability, with lower quantiles around 0.4 km and upper

quantiles around 0.6 km over the same period. These results suggest that the model underestimates the diurnal variation compared to the observations over the Gulf.

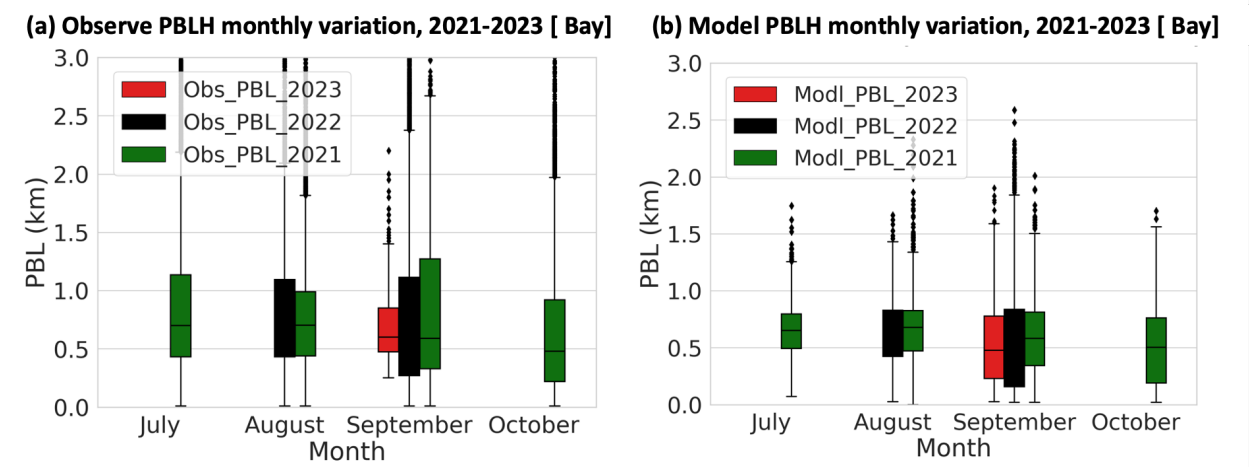


Figure 9. (a) Observe and (b) Model monthly variation of PBLH over the Galveston Bay in 2021 (green), 2022 (black), and 2023 (red)

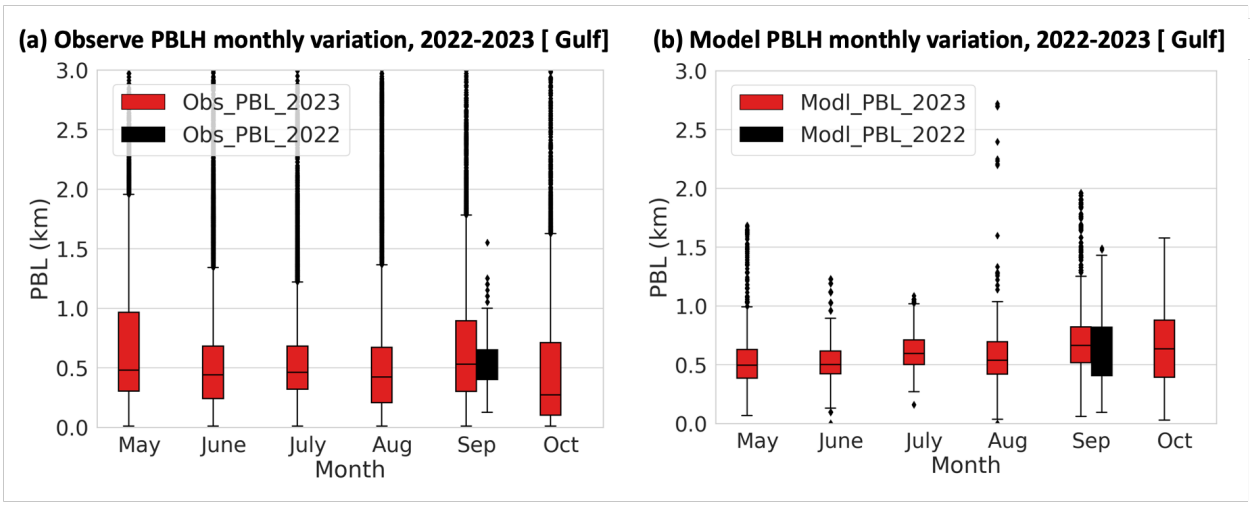


Figure 10. (a) Observe and (b) Model monthly variation of PBLH over Gulf in 2022 (black), and 2023 (red)

3.2. Spatial variability of PBLH over the water

In the previous section, the diurnal and monthly variations of PBLH over the entire Galveston Bay and the Gulf were presented. This section further examines the spatial variations of PBLH over Galveston Bay and the Gulf. Given that the observed PBLH at each observation point was sampled at different time and location (**Figure 5**) and PBL has distinct diurnal pattern, the different observation time in each observation route needs to be considered. For example, there were two similar observation routes in Galveston Bay on 09/08/2021 and 09/25/2021 (**Figure 11**). On 09/08/2021, the boat moved toward the south in the morning and moved north in the afternoon, while the boat moved in the opposite direction on 09/25/2021. Simple averages of the PBL measured by locations would misrepresent the spatial variability by overlooking the large

diurnal cycle of PBL (c.f. **Figure 7**). To properly represent the spatiotemporal variability of PBL from measurements, seven representative routes were picked where the boats frequently visited (**Figure 12**): (1) the southern part of Houston Ship Channel (SHSC); (2) the northern part of Houston Ship Channel (NHSC); (3) the eastern part of Galveston Bay; (4) near-bay southern Gulf (NBSGFM); (5) near-bay eastern Gulf (NBEGFM); (6) the southern Gulf (SGFM); and (7) the eastern Gulf (EGFM). Then, the research team divided the observed PBLH in each route into two temporal sets: Morning (07:00 CDT – 11:59 CDT) and Afternoon (12:00 CDT – 18:59 CDT).

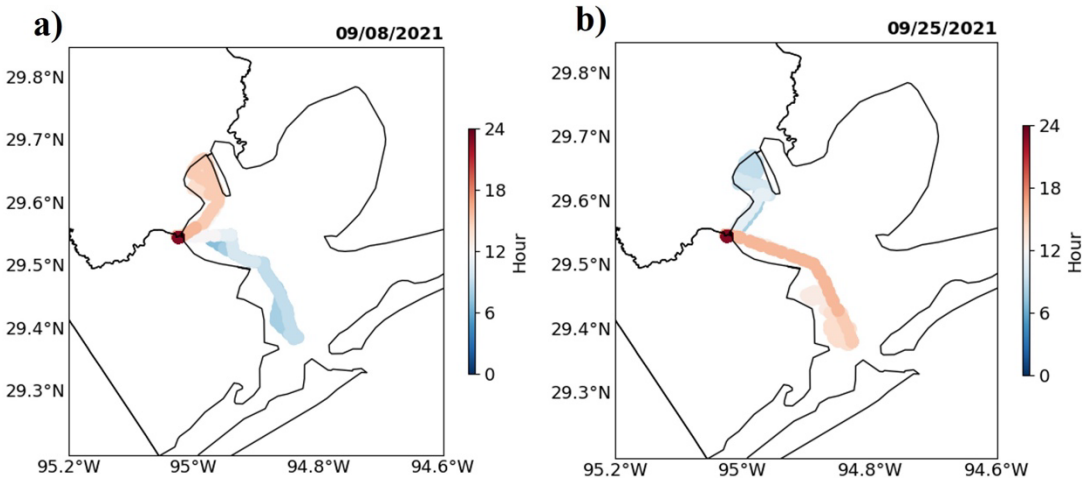


Figure 11. (a) Boat observation points over the Houston Ship Channel on 09/08/2021 along with the time/hour when the boat was mobile. (b) Same as (a) but for observation points on 09/25/2021.

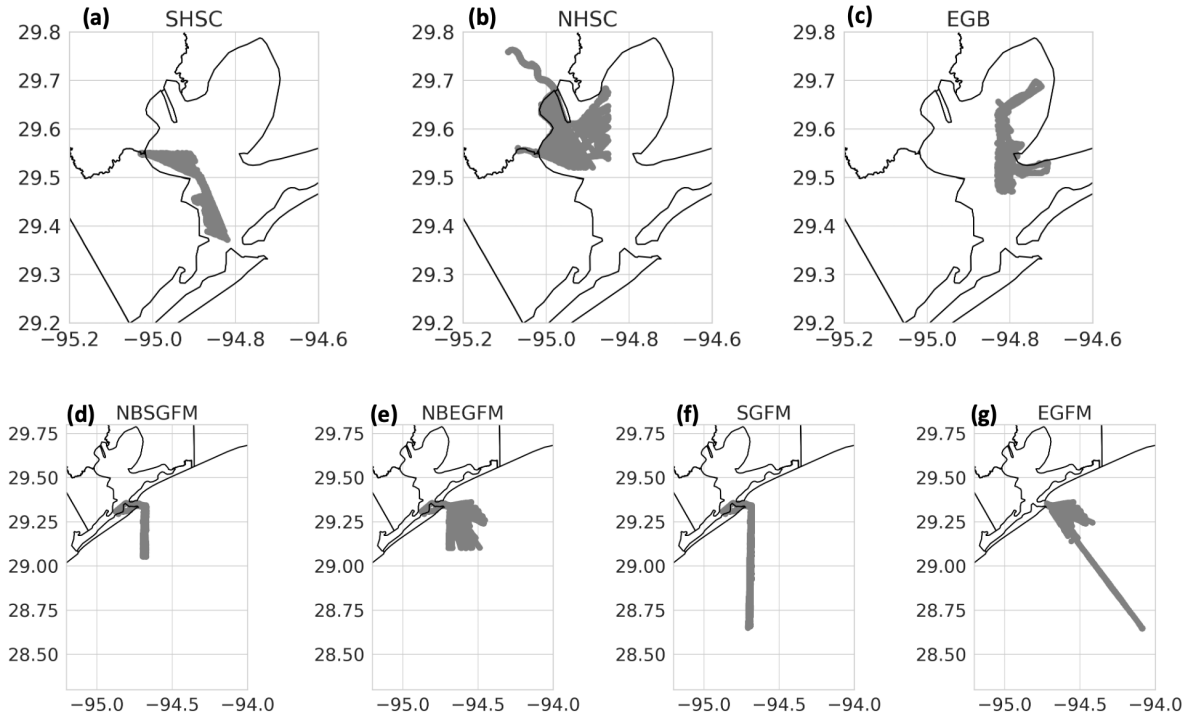
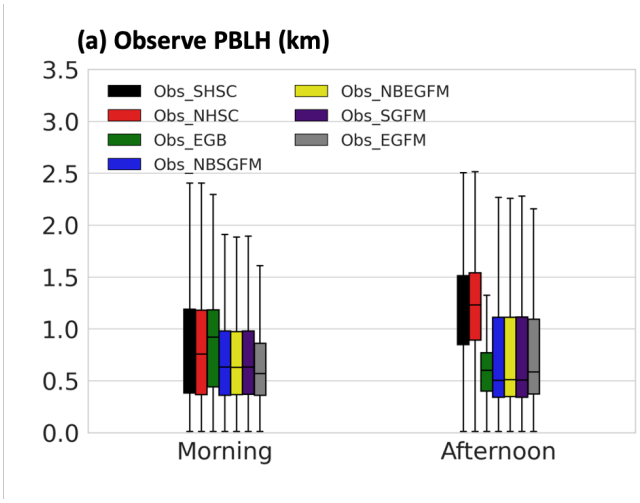


Figure 12. All available boat observation points (gray dots) for (a) the southern part of Houston Ship Channel (SHSC), (b) the northern part of Houston Ship Channel (NHSC), (c) the eastern part of Galveston Bay, (d) near-bay southern Gulf (NBSGFM), (e) near-bay eastern Gulf (NBEGFM), (f) the southern Gulf (SGFM), and (g) the eastern Gulf (EGFM).

Figure 13 and **Figure 14** summarize the observed and modeled PBLH in the morning and the afternoon in each route. In the morning, the highest PBLH is found in the EGB (observed:0.91 ± 0.55 km, model: 0.70 ± 0.18 km) route. Over Galveston Bay, the morning PBLH is higher over its northern and eastern parts (NHSC and EGB routes). Three different routes (NBSGFM, NBEGFM, and SGFM) over the Gulf have almost identical mean PBLHs in the morning. This result indicates the Gulf has less spatial variation of morning PBLHs than Galveston Bay, which is consistent with the results in the previous section (**Figure 5**). The lower observed PBLH over the Gulf is found in EGFM route which is reproduced by the model. Smaller standard deviations are also found in the routes over the Gulf than in the three routes over Galveston Bay (**Figure 13**), indicating fewer temporal variations of morning PBLH over the Gulf. This feature is also reproduced by the model (**Figure 14**).

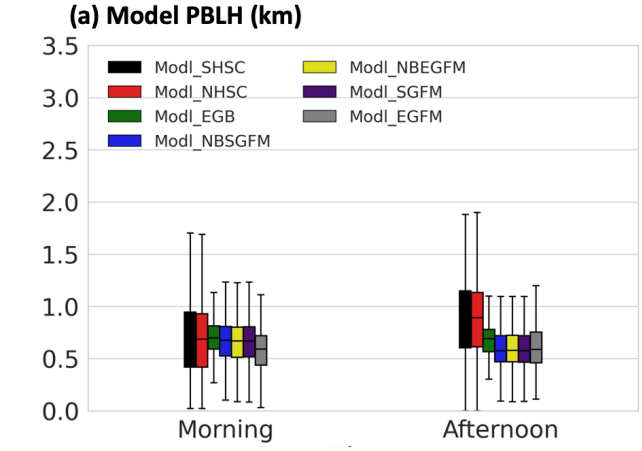
In the afternoon, the highest PBLH is found in NHSC route (1.20 ± 0.57 km) while the lowest PBL is found in EGB route (0.62 ± 0.39 km). As described in the previous section, PBLHs in the two routes over Galveston Bay exhibit significant increases (0.35 km in SHSC and 0.39 km in NHSC) in the afternoon compared to those in the morning (**Figure 13**). The model reproduces this increase but to a less extent (0.2 km in SHCS and NHSC) (**Figure 13**). Compared to Galveston Bay, PBLHs in four routes over the Gulf have smaller increases in the afternoon compared to the morning (0.05 km in SGFM, 0.04 km in NBSGFM, 0.06 km in NBEGFM, and 0.12 km in EGFM), which is also consistent with the previous result showing weaker diurnal variation over the Gulf **(Figure 7**b). While the model PBLH has smaller differences than

observed ones over the Gulf, the model simulates no change in PBLH between the morning and afternoon in all four routes over the Gulf (**Figure 14**). The difference in diurnal variation between Galveston Bay and the Gulf reflects a larger horizontal PBL gradient from Galveston Bay to the Gulf in the afternoon than in the morning. This feature is reproduced by the model.



Area	Observe PBLH (km)	
	Morning (7-12)	Afternoon (12-18)
SHSC	0.82 ± 0.52	1.17 ± 0.56
NHSC	0.81 ± 0.53	1.20 ± 0.57
EGB	0.91 ± 0.55	0.62 ± 0.39
NBSGFM	0.71 ± 0.49	0.75 ± 0.57
NBEGFM	0.70 ± 0.49	0.76 ± 0.57
SGFM	0.71 ± 0.49	0.76 ± 0.57
EGFM	0.64 ± 0.47	0.76 ± 0.50

Figure 13. The boxplot for observed PBLH for morning time (07:00 CDT – 11:59 CDT) and for afternoon time (12:00 CDT – 18:59 CDT) in each boat observation route. The black box represents SHSC route, the red box represents NHSC route, the green box represents EGB route, the blue box represents NBSGFM route, yellow box represents NBEGFM route, indigo box represents SGFM route, and gray box represents EGFM route. Mean PBLH and standard deviation in the morning and the afternoon for each route are also presented in the table.



Area	Model PBLH (km)	
	Morning (7-12)	Afternoon (12-18)
SHSC	0.69 ± 0.35	0.90 ± 0.38
NHSC	0.68 ± 0.34	0.89 ± 0.38
EGB	0.70 ± 0.18	0.67 ± 0.18
NBSGFM	0.67 ± 0.26	0.60 ± 0.24
NBEGFM	0.66 ± 0.25	0.61 ± 0.25
SGFM	0.67 ± 0.26	0.60 ± 0.24
EGFM	0.58 ± 0.20	0.66 ± 0.31

Figure 14. The boxplot for modeled PBLH for morning time (07:00 CDT – 11:59 CDT) and for afternoon time (12:00 CDT – 18:59 CDT) in each boat observation route. The black box represents SHSC route, the red box represents NHSC route, the green box represents EGB route, the blue box represents NBSGFM route, yellow box represents NBEGFM route, indigo box represents SGFM route, and gray box represents EGFM route. Mean PBLH and standard deviation in the morning and the afternoon for each route are also presented in the table.

3.3. Comparison of observed and modeled meteorological variables

In addition to PBLH, WRF simulated meteorological variables were evaluated across three years (2021-2023) over Galveston Bay and the Gulf. **Figure 15** compares the hourly temporal variation of the model and observed meteorological variables for September 2021 and 2022 over Galveston Bay. The model temperature closely follows the observed values with a slight underestimation of ~1°C and ~0.5°C in September 2021 and 2022, respectively (**Figure 15**a). Underestimation of temperature is higher in the daytime (~1°C) compared to nighttime (~0.1°C) for both years. During the night, the boat is mainly at the dock, so most of the data points are stationary and closer to the land, and the model captures the variation better compared to daytime when the boat was mobile over Galveston Bay. The model relative humidity is 5-10% higher than the observations (**Figure 15**b). The model captures the diurnal trend of relative humidity with a lower value in the afternoon and higher at night. The model overestimates the wind speed by ~2 m/s (**Figure 15**c) and shows less diurnal variability than the observation. The overestimate of surface winds is a common problem for WRF and was found in Liu et al. (2023) when using an older version of WRF. The model underestimates the wind direction by ~3 degrees at night. However, during the daytime, it overestimates the wind direction by ~80 degrees (**Figure 15**e). Model shows a prominent diurnal variation in the wind direction, but observation shows a minimal diurnal variation. The model slightly underestimated the pressure with an average of ~1.5 mb than the observation (**Figure 15**d). The model captures the diurnal variability of the pressure well.

Figure 16 shows the month-to-month variation of meteorological variables from July to October 2021 and August to September 2022 over Galveston Bay. These months were selected based on the availability of mobile observations, with the criterion of at least 10 days of mobile data per month. Boxplot of the month-to-month variations shows that the model consistently underestimated temperature in 2021 and 2022 (**Figure 16**a). On average, the model's temperature is 0.77°C lower than observed in 2021 and 0.45°C lower in 2022. The boxplots show that the model better captures the lower quantiles of temperature than the upper quantiles which suggests that the model is better at reproducing the nighttime temperatures than daytime when the temperature is high. This observation is consistent with the September hourly time series (**Figure 16**a). The model's underestimation of daytime temperature might be a reason for its underestimation of afternoon PBLH over the water. Model relative humidity is higher by ~12% compared to observation in 2022 and ~8% in 2022. The model exhibits an average overestimation of wind speed ~2.87 m/s in 2021 and ~2.72 m/s in 2022 compared to the observation. The model shows an average overestimation of wind direction ~12.41 deg in 2021 and ~7.65 in 2022 compared to the observation. The model has low biases in estimating the pressure, with a slight underestimation of ~1.35 mb in 2021 and ~1.03 mb in 2022. The summary statistics for each year's model evaluation are provided in the figures.

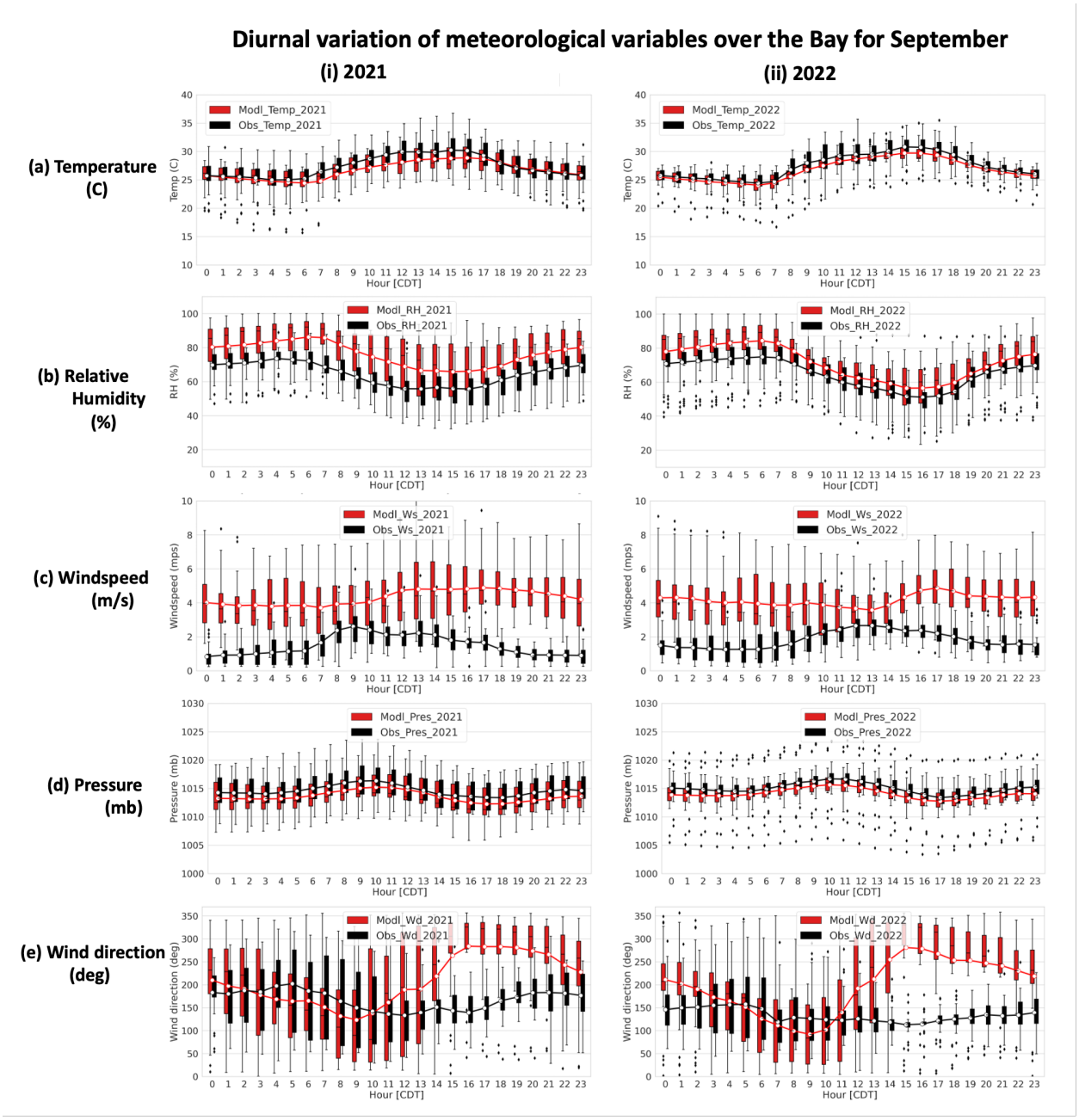


Figure 15. Boxplot of the diurnal variation of observed and model (a) Temperature, (b) Relative Humidity, (c) Wind speed, (d) Pressure, and (e) Wind direction for September (i) 2021 and (ii) 2022 over Galveston Bay. Black represents the observed values, and red represents the modeled values.

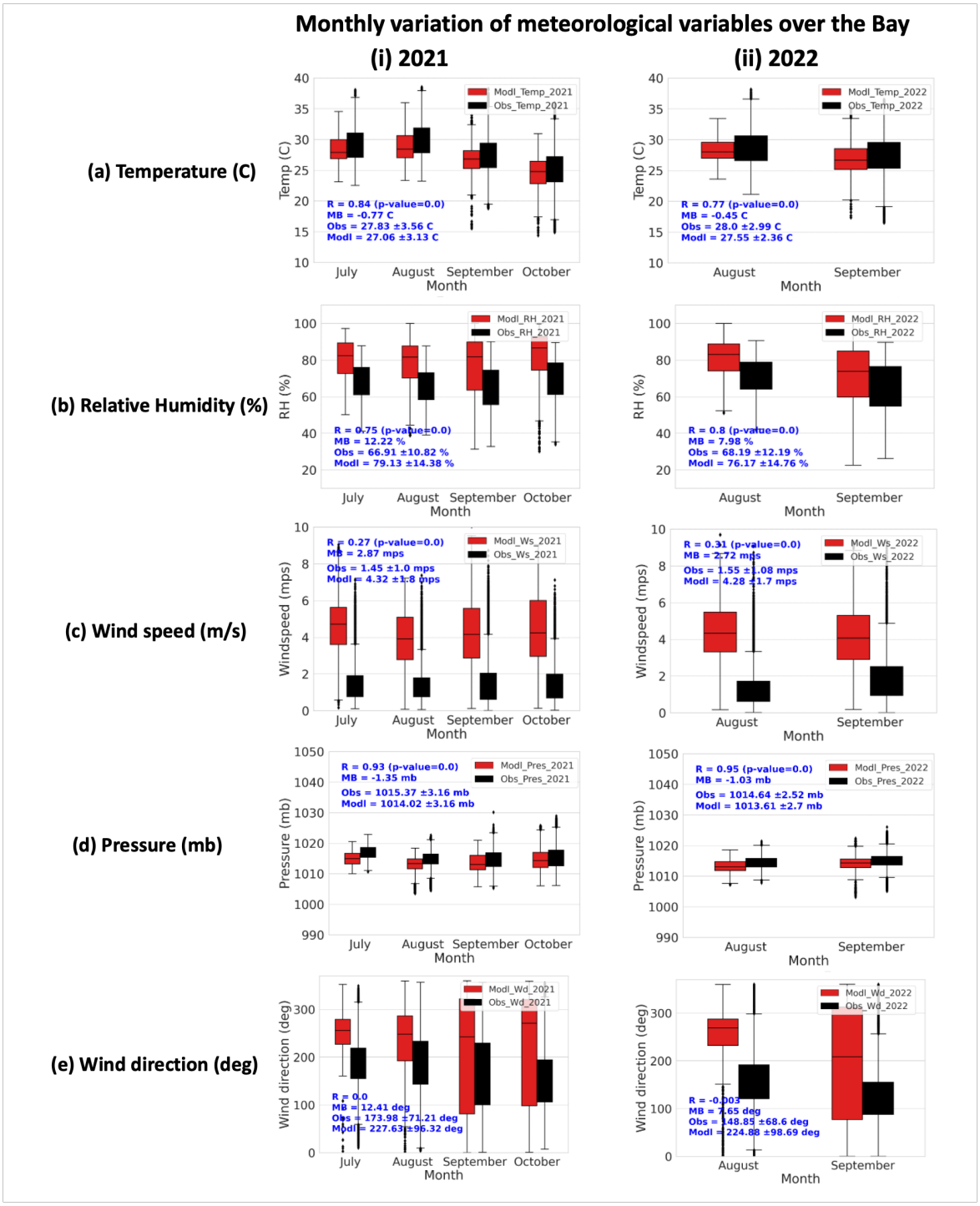


Figure 16. Observed and Model monthly variation of (a) Temperature, (b) Relative Humidity, (c) Wind speed, (d) Pressure, and (e) Wind direction over Galveston Bay in (i) 2021 and (ii) 2022

Figure 17 presents the hourly temporal variation of the model and observed meteorological variables for September 2022 and 2023 over the Gulf. During the day, the model slightly underestimated the temperature by ~1°C, while at night it overestimated the temperature by ~1°C in both years. The model shows less diurnal temperature variation compared to the observed

values, as shown in **Figure 17**a. For relative humidity (**Figure 17**b), the model overestimated the observed values, with a higher overestimation during the day (~10%) compared to night (~5%). The model overestimated wind speed by ~1.5 m/s (**Figure 17**c) and showed less diurnal variability for both years. The model overestimates the wind direction by ~70 degree (**Figure 17**e) and shows a prominent diurnal variability than the observation for September. The model captures the diurnal variability in the pressure well, although it slightly underestimated pressure in September 2022 compared to 2023 (**Figure 17**d).

Figure 18 shows the month-to-month variation of meteorological variables from May to October 2023 and September 2022 over the Gulf. The model overestimated the overall monthly mean values by 0.25°C and 0.27°C in 2022 and 2023, respectively. The boxplot for each month reveals that observed temperature variability was higher than that of the model (**Figure 18**a). Model relative humidity is higher by ~8% compared to observation in 2022 and ~10% in 2023 (**Figure 18**b). The model exhibited an overestimation of wind speed ~1.43 m/s in 2022 and ~2.66 m/s in 2023 compared to the observation (**Figure 18**c). The model shows an average overestimation of wind direction ~24.63 degrees in 2022 and ~25.0 degrees in 2023 compared to the observation (**Figure 18**e). The model has low biases in estimating pressure, with a slight underestimation of ~0.82 mb in 2022 and ~0.3 mb in 2023 (**Figure 18**d). The summary statistics for each year's model evaluation are provided in the figures.

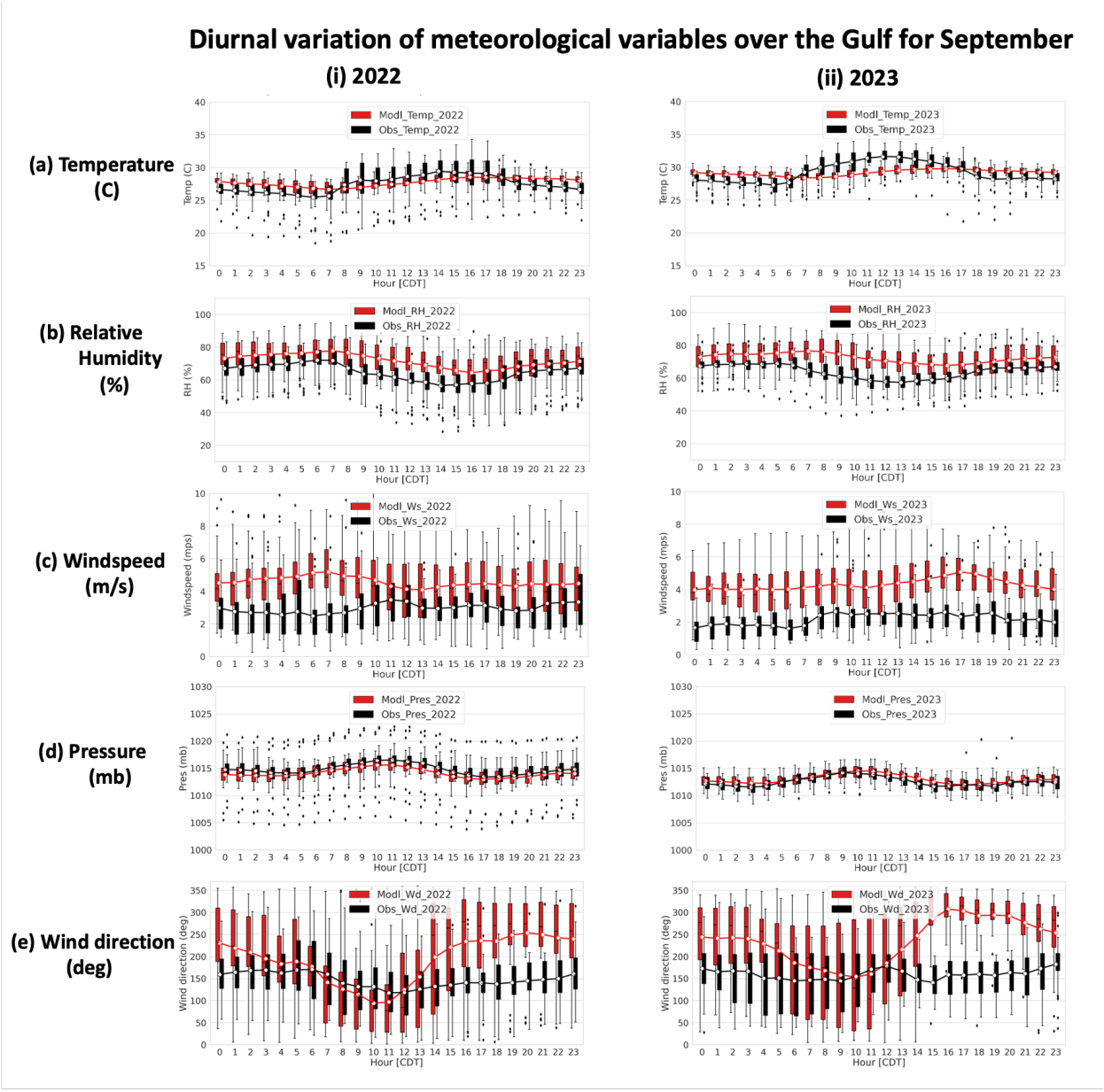


Figure 17. Boxplot of the diurnal variation of observed and model (a) Temperature, (b) Relative Humidity, (c) Wind speed, (d) Pressure, and (e) Wind direction for September (i) 2022 and (ii) 2023 over Gulf. Black color in the plot represents the observed values and red for model.

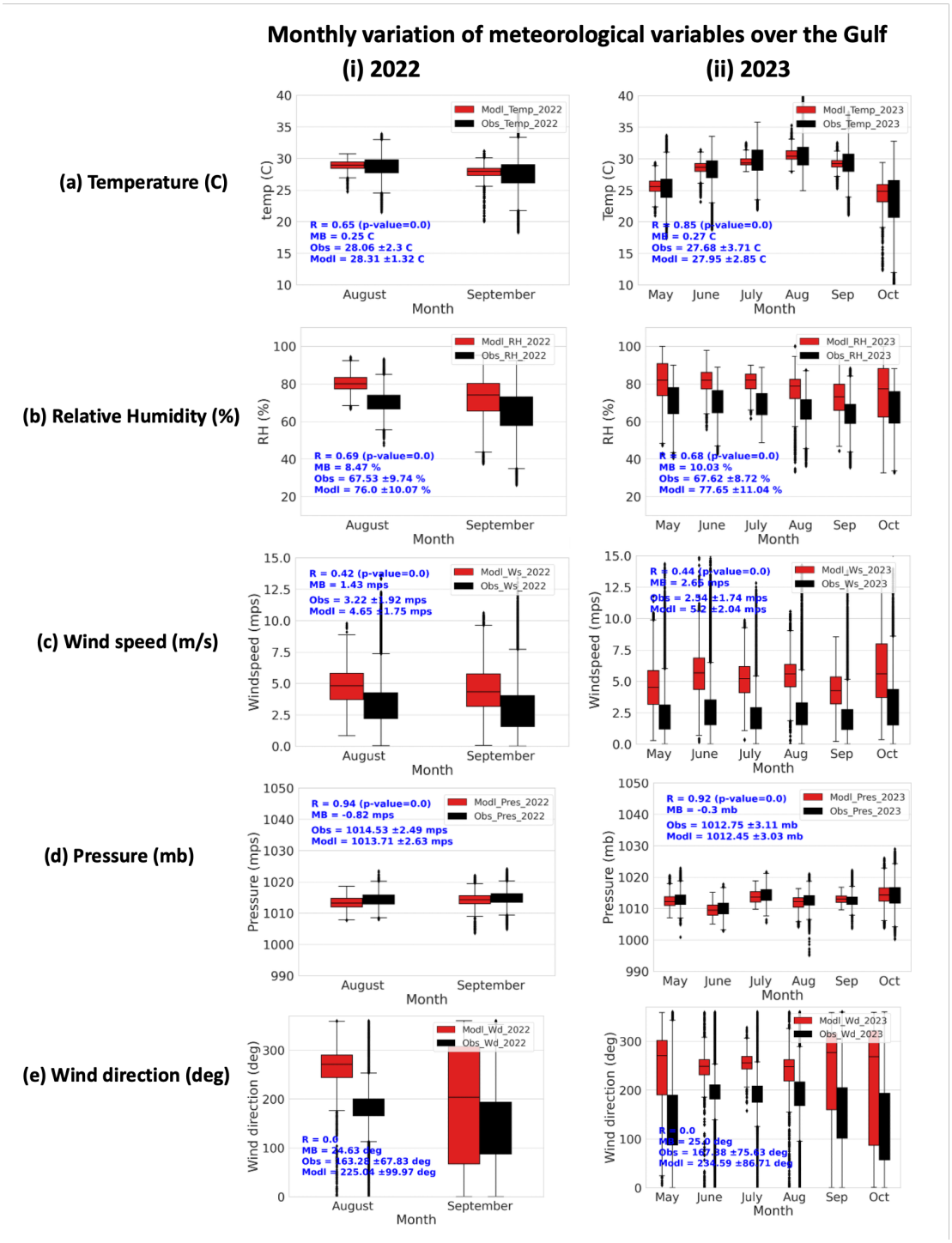


Figure 18. Observed and Model monthly variation of (a) Temperature, (b) Relative Humidity, (c) Wind speed, (d) Pressure, and (e) Wind direction over the Gulf in (i) 2022 and (ii) 2023

3.4. Summary

Though several previous studies have investigated the WRF model's estimation of PBLH and meteorological variables in the landmass region, such investigations have been limited over the

water. This study evaluated the WRFv4.6.0 model's estimation of PBLH and meteorological variables over Galveston Bay and the Gulf, two large offshore water bodies in southeast Texas. The observations from multiple field campaigns (2021-2023) were used for the evaluation. Different dynamics of PBLH and meteorological variables – characterized by differences in their diurnal variations, monthly variations, variation across different years, and different regions within the Galveston Bay and the Gulf - were obtained in the two studied regions. In analyzing the temporal variations of the PBLH, the observed PBLH showed a clear diurnal cycle over Galveston Bay, which the model reproduced, although with an underestimation of PBLH in the afternoon. Over the Gulf, both observed and simulated PBLH exhibited a less distinct diurnal variation, with the model underestimating the afternoon PBLH. The detailed analysis of different boat routes over the two regions further demonstrates that the afternoon increase in PBLH was more pronounced over Galveston Bay than in the Gulf, suggesting a larger PBL gradient between Galveston Bay and the Gulf in the afternoon than in the morning. The WRF model's performance was found to be location and time dependent. In the afternoon, the model overestimated PBLH over the northwest HSC, a narrow water body surrounded by land, while it underestimated PBLH over the southwest HSC, where the boat moved further from land toward open water. Diurnal and monthly variations in the meteorological variables are captured well in the model, with temperature and pressure showing less biases, and the relative humidity, wind speed, and wind direction with a slightly higher bias over Galveston Bay and the Gulf. Overall, this evaluation enhances our understanding of WRF model performance over water bodies and provides a baseline for subsequent tune-up of model parameters to improve PBL predictions over the water.

4. Improvements to marine PBL in WRF

Section 3 discussed the limitations of the base model's performance in simulating the marine PBL and its diurnal variations. In addition to those limitations, the base simulation of WRF was found to present the 'cloud-like' patterns in the PBLH distribution over the waters under certain conditions (**Figure 19**). The cloud-like features are caused by a sharp change in PBLH within neighboring grids, which is deemed unrealistic. Upon further investigation, this artifact was attributed to model limitations in simulating wind speed and direction under small-scale cyclonic conditions, which cause sudden updrafts and/or downdrafts of air parcels, leading to the observed cloud-like patterns in PBLH. This finding indicates that offshore PBLH prediction in WRF is affected by the choice of cumulus schemes. For example, after replacing the New Tiedtke Cumulus scheme used in baseline configurations with the New Simplified Arakawa-Schubert (NSAS) scheme, the unrealistic patterns in the PBLH distribution over the Gulf are not present (**Figure 19b**). Drawing from this exercise, different cumulus schemes were added as part of the sensitivity simulations carried out in the project, and the related results are described below.

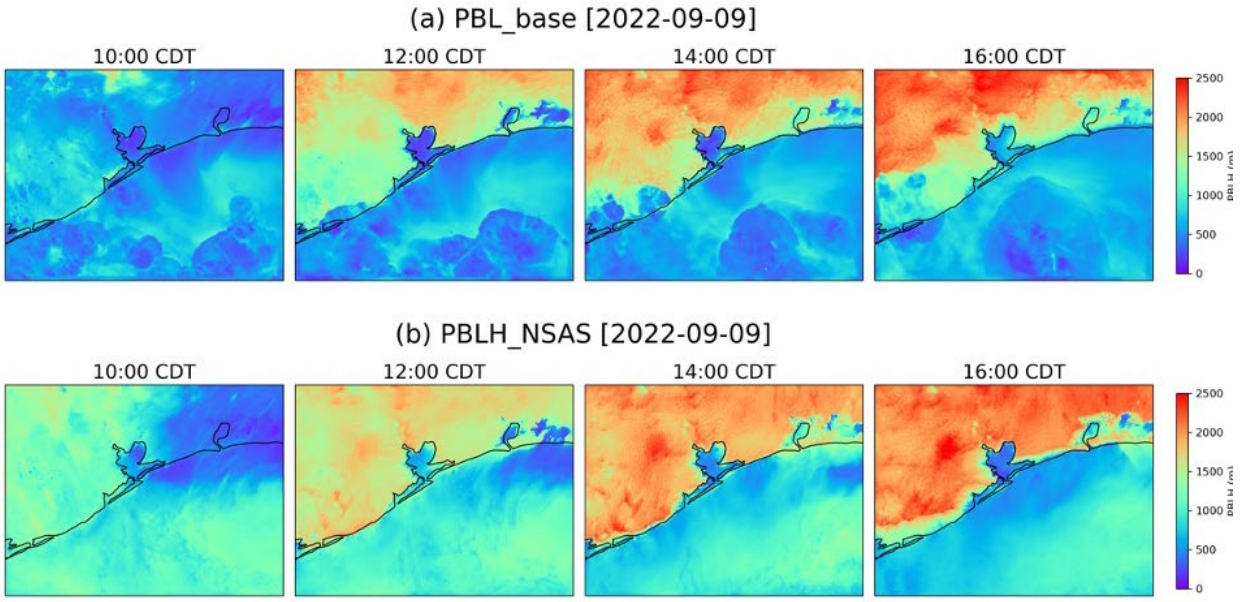


Figure 19. Spatial distribution of hourly PBLH (m) variation from (a) the base simulation and (b) base simulation with New Simplified Arakawa-Schubert (NSAS) cumulus scheme.

Another issue with PBLH was identified in the baseline configuration over the urban region. The model shows persistently higher PBLH over urban Houston during nighttime (**Figure 20a**). The PBLH hotspots over urban Houston exhibit a fixed boundary, which is unrealistic and appears due to the simplified treatment of urban land types in the model. One consequence of this PBLH artifact is that it causes the model prediction of NOx to be unrealistically low over urban Houston during nighttime hours, and thus it is of importance for the overall model performance. The issue could be resolved by implementing a more advanced scheme for treating urban areas, such as the Urban Canopy Model (UCM) and Building Effect Parametrization (BEP). For example, after adopting the UCM scheme in the simulations (**Figure 20b**), the high PBLH hotspots over urban Houston are mitigated. It was noted that the intriguing effect that UCM leads to changes in PBLH over the water, despite being implemented only in the urban region. It is probably because it affects the land-sea thermal gradient by changing the heat and moisture budgets in urban Houston. Because of this effect, the UCM was included as part of the sensitivity simulations, and the related results are described below.

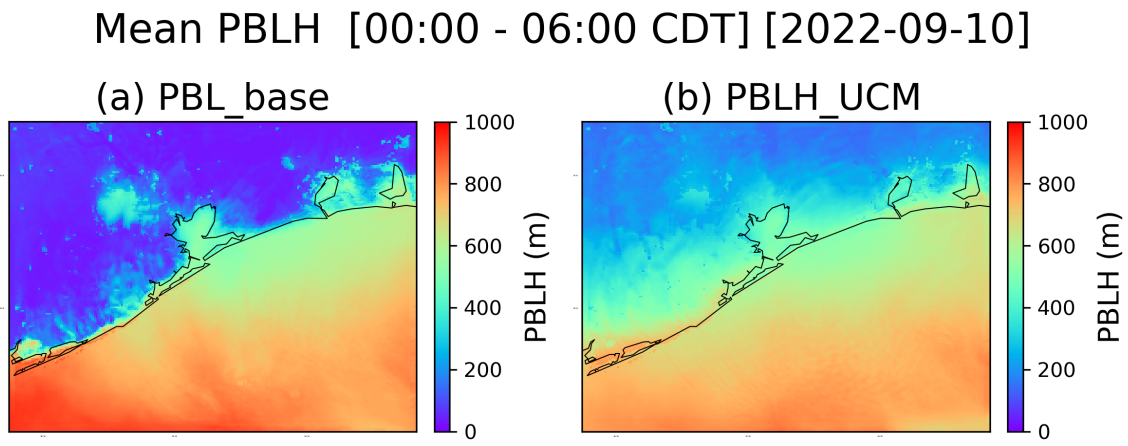


Figure 20. Spatial distribution of Midnight - 06:00 CDT mean PBLH from (a) the base simulation (b) the base simulation with Urban Canopy Model (UCM).

4.1. Selection of physics parameters and schemes

4.1.1. Physics Parameters

WRF incorporates various schemes to simulate the evolution and development of the PBL. The sensitivity of these schemes depends on the topographical and environmental conditions, making it essential to identify the most effective schemes for the region. Our previous study (Liu et al., 2023) evaluated different PBL schemes over the diverse onshore and offshore environments of the Houston-Galveston-Brazoria (HGB) region, including Galveston Bay and the Gulf of America, as sampled by the TAQ1 campaign. The MYNN with 2.5 closure (MYNN2) scheme was identified as the best choice for the region among all available PBL schemes in WRF with the least bias compared to PBLH observations from the TAQ1 campaign. The evaluation of WRF PBL with different field campaign observations, conducted in Task 3 (Section 3) of the present project, indicates that the model underestimates the afternoon PBLH over the water and fails to capture the diurnal cycle accurately. To address this gap, key parameters were identified and constants adopted in the MYNN PBL scheme that are likely to impact the PBLH, as shown in **Table 2**. The Table includes the values of the parameters used in the base model and their possible ranges reported from the literature that can be used to fine-tune the model predictions.

Table 2. Physics parameters identified for perturbation in the MYNN PBL scheme

Parameters /Constants	**Description**	**Default Values**	**Perturbation Range**	**References**
B1	Constant for Dissipation Rates	24	[12,36]	Nakanishi & Niino, (2009)
B2	Constant for Dissipation Rates	15	[7.5, 22.5]	Nakanishi & Niino, (2009)

α1	Constant for Turbulent length scale (L_T)	0.23	[0.2-0.3]	(Yang et al. 2017)
Pr	Turbulent Prandtl Number	0.74	[0.7 – 2.0]	(Pithan et al., 2015)
Z0	Surface roughness	0.0185	[0.01 – 0.03]	(Yang et al. 2017)
C3	Closure Constant	0.33	[0.33 – 0.5]	(Huang & Peng, 2017)
ALBD	Albedo	0.08	[0.06-0.09]	(Liu et al., 2022)

The MYNN PBL scheme is a turbulence closure model with several constants and unknown parameters that represent physical factors such as the dissipation rates of total kinetic energy (TKE), pressure co-variances, and turbulence fluxes, which are determined empirically. TKE dissipation rates measure turbulence strength and are proportional to the energy transfer from larger to smaller eddies. These rates are parameterized as a function of the mixing length (L) and the closure constants B1 and B2. Higher values of B1 and B2 result in lower TKE dissipation rates, which is expected to increase the PBLH (Yang et al., 2017). The default values of B1 and B2, along with their ranges, are detailed in **Table 2**.

The mixing length (L) typically determines the distance an air parcel with energy can travel before dispersing, and it is expressed as the harmonic sum of L_s, L_T, and L_B. L_s represents the length scale determining mixing near the surface, while L_B is effective for the stable boundary layer (Nakanishi & Niino, 2009). L_T dominates the mixing length in the middle to upper convective boundary layer. Our primary goal is to improve the model PBLH mostly during unstable cases, so L_T was chosen to perturb, which depends on the value of α1. Thus, α1 was perturbed for a range of values to test its sensitivity to PBL.

Turbulent Prandtl number (Pr) is the ratio of the diffusivity of momentum to that of heat. The values for Pr range from 0.74 to 2. The choice of Pr determines the near-surface gradients of temperature and humidity. Lower values of turbulent Prandtl number led to warmer convective boundary layers, especially in tropical and sub-tropical regions (Pithan et al., 2015). The choice of Pr was expected to be sensitive in the coastal regions where the convective boundary layer is usually seen during the daytime.

The closure constant (C3) represents the effects of shear and buoyancy in the MYNN2 scheme whose values are theoretically determined to be 0.33 for isotropic turbulence and 0.5 for the convective boundary layer (Nakanishi & Niino, 2009). The default value for the base case is 0.34, which is subject to subsequent perturbations of 0.42 and 0.5.

Roughness length (Z0) is a surface-dependent parameter that typically refers to the maximum height at which winds diminish to zero due to surface friction. This term is related to the transfer of heat, momentum, and moisture and is parameterized in surface schemes. Due to unresolved sea wave effects, roughness length may have higher uncertainties in coastal areas. An improved

representation of roughness length can affect wind speed near the surface, leading to increased turbulence and, consequently, a rise in PBL.

4.1.2. WRF configuration test

In addition to perturbing physical parameters described above, sensitivity simulations were conducted by selecting different cloud, urban physics schemes, and 1-D ocean physics aimed to mitigate the unrealistic 'cloud-like' features of PBL (**Figure 19**) and PBL hotspots in urban Houston (**Figure 20**) from the baseline simulations. To resolve the former, a single-layer urban canopy model (UCM) was implemented in WRF, which considers the momentum drag effects of buildings and the energy exchange between the urban structure and the atmosphere. UCMs are expected to better predict the PBLH over urban areas. To resolve the cloud-like features in the spatial distribution of PBL, the New Tiedtke cumulus scheme was replaced with the NSAS scheme. Furthermore, the MYNN3 scheme was tested, which is more advanced as it employs a prognostic approach to diagnose sub-grid scale vertical motions, rather than a diagnostic one in MYNN2. Also, further simulation was performed by turning on the 1-D ocean mixed layer model in WRF. There is limited research on the impact of coupling 1-D ocean mixed layer model on PBLH, but coupling is usually preferred to understand air-sea interaction during strong wind conditions. **Table 3** lists the additional sensitivity simulations conducted so far.

Table 3. Additional schemes adopted for the sensitivity runs

Schemes	**Description**	**Default Settings**	**Customized Settings**
MYNN	PBL scheme	2.5	3.0
UCM	Urban Canopy Model	off	on
Cumulus physics	Switched to different schemes	New Tiedtke	New Simplified Arakawa-Schubert (NSAS)
1-D ocean mixed layer model	Ocean physics	off	on

4.1.3. Selected days for model perturbation

The selected days for the model perturbation are based on the model evaluation in Section 3, as well as our previous modeling analysis of TAQ1 and TAQ2 campaign datasets. A total of 9 days were chosen for model perturbation/sensitivity simulations: July 27-28, September 9, and September 24-26, 2021; September 8-10, 2022; and May 19 and September 9-10, 2023. The selection criteria for these dates consider multiple factors, including various locations in Galveston Bay and the Gulf, both near and far from land, including the Houston Ship Channel (HSC). Other criteria include different times of day, varying pollutant levels, and the availability of both mobile and stationary observational data. These diverse factors help ensure a comprehensive evaluation of the model. **Figure 21** shows the locations for each selected day with the time-of-day for each observation point.

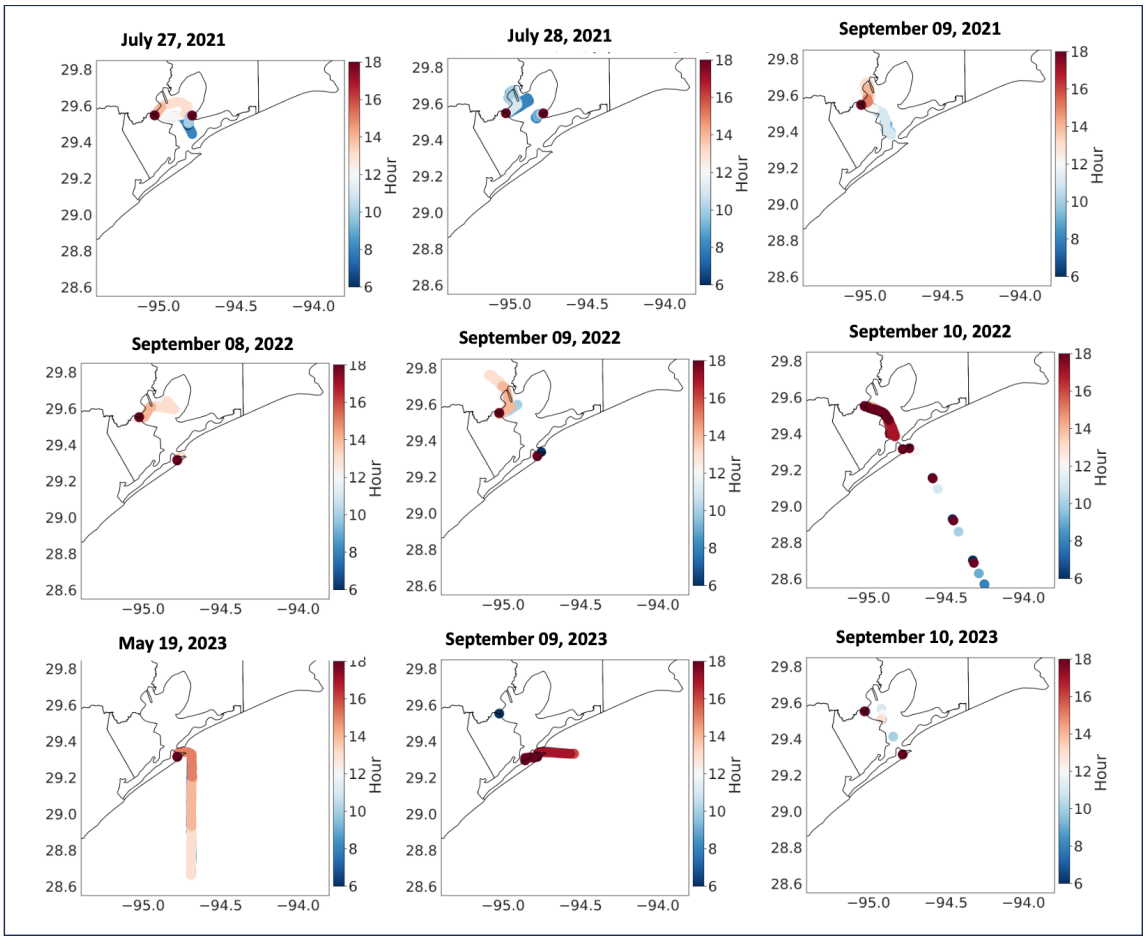


Figure 21. Location of the selected days for model perturbation. The color in the spatial map represents the time of observation.

4.2. Perturbation simulations and results

4.2.1. Experimental Setup for model perturbation

Table 4 lists 25 model perturbations conducted in the project. Each simulation was conducted for all the selected days listed in Section 4.1.3, including the relevant spin-up period, resulting in more than 350 simulation-days. The first group of simulations (PBLH1-PBLH10 and PBLH25) are designed to test the sensitivity of identified physics parameters across a range of values. The aim is to identify optimal parameters that improve the model's ability to simulate marine PBL and its diurnal variation. The second group of simulations (PBLH11-25) tested various cumulus schemes and urban physics schemes, along with their potential combinations, to resolve issues seen in the base simulations described above.

Table 4. Different sets of simulations designed by perturbing the physics parameters and changing the schemes.

Simulation ID	PBL Scheme	Cumulus Physics Scheme	Urban Physics Scheme	Parameter Perturbation
PBL_base	MYNN2.5	New Tiedtke	off	None
PBLH1	MYNN2.5	New Tiedtke	off	B1 = 36, B2 = 22.5
PBLH2	MYNN2.5	New Tiedtke	off	B1= 30, B2 = 18.5
PBLH3	MYNN2.5	New Tiedtke	off	$\alpha 1 = 0.27$
PBLH4	MYNN2.5	New Tiedtke	off	$\alpha 1 = 0.50$
PBLH5	MYNN2.5	New Tiedtke	off	C3 = 0.50
PBLH6	MYNN2.5	New Tiedtke	off	C3 = 0.42
PBLH7	MYNN2.5	New Tiedtke	off	Pr = 2
PBLH8	MYNN2.5	New Tiedtke	off	Pr = 1.37
PBLH9	MYNN2.5	New Tiedtke	off	Z0 = 0.02 (for water)
PBLH10	MYNN2.5	New Tiedtke	off	albedo = 0.09
PBLH11	MYNN3.0	New Tiedtke	off	MYNN 2.5 to MYNN3.0
PBLH12	MYNN2.5	NSAS	off	New Tiedtke to NSAS
PBLH13	MYNN2.5	New Tiedtke	On	UCM
PBLH14	MYNN2.5	NSAS	On	NSAS+UCM
PBLH15	MYNN2.5	NSAS	on	NSAS+UCM+PBLH5
PBLH16	YSU	New Tiedtke	on	UCM
PBLH17	BouLac	New Tiedtke	on	UCM
PBLH18	YSU	New Tiedtke	on	BEP scheme
PBLH19	BouLac	New Tiedtke	on	BEP scheme

PBLH20	YSU	New Tiedtke	on	BEM scheme
PBLH21	BouLac	New Tiedtke	on	BEM scheme
PBLH22	MYNN2.5	New Tiedtke	off	1-D ocean mixed layer model
PBLH23	MYNN2.5	NSAS	off	PBLH12+1-D ocean mixed layer model
PBLH24	MYNN2.5	NSAS	on	PBLH14+1-D ocean mixed layer model
PBLH25	MYNN2.5	NSAS	off	1-D ocean mixed layer model, Z0 = 0.02 (for water) (i.e., PBLH9+ PBLH23)

4.2.2. PBLH changes from model perturbations compared to observations

Outputs from each of the 25 perturbation simulations were compared to the base simulations and with onshore and offshore observations from TAQ1, TAQ2, and TAQ3. The impact of the perturbations was found to be variable throughout the domain and the selected days. Some of them have a significant impact on PBLH over the land with small changes over the water, while others degraded PBL over land despite improving it over the water. After conducting a thorough analysis, a few of the simulations were selected that increased PBLH over the water with less impact on land and performing better on most of the selected days. The six best simulations identified are PBLH5, PBLH12, PBLH14, PBLH23, PBLH24, and PBLH25. PBLH5 refers to the perturbed value of closure constant C3. PBLH12 refers to the simulations that implement newer cumulus schemes called NSAS. PBLH14 includes a combination of the UCM with the NSAS cumulus scheme. PBLH23 refers to the simulations that combined PBLH12 with the 1-D ocean physics. PBLH24 refers to the simulations that combined PBLH14 with the 1-D ocean physics. PBLH25 refers to the perturbed value of the surface roughness length over the water combined with the PBLH23. The details of their simulation setups are listed in **Table 4**.

Figure 22 compares the frequency distribution of daytime (08:00 – 18:00 CDT) PBLH between the observations, the base simulation, and the chosen perturbations (PBLH5, PBLH12, PBLH14, PBLH23, PBLH24, and PBLH25) for all the selected days combined. The comparison is shown separately for Galveston Bay (**Figure 22a**) and the Gulf (**Figure 22b**). In Galveston Bay, observed PBLH shows two distinct peaks, the first one corresponding to the increased PBLH shortly after sunrise and the second one mostly during the afternoon. In all the model experiments, the first peak is skewed towards higher values than in the observations, meaning the model persistently overestimates the morning rise of PBLH in the bay. The second peak is of more interest because it predominantly occurs in the afternoon when photochemistry is most

active. The base simulation misses this peak by underestimating afternoon PBLH. PBLH23, PBLH24, and PBLH25 closely replicate the second peak, indicating they improve upon the base simulation by predicting higher afternoon PBLH in the bay that matches better with the observations. In the case of higher values, PBLH5 and PBLH12 have a longer tail toward the higher values than other simulations.

In the Gulf, the PBLH simulated by all the perturbation runs is primarily concentrated around the lower values. This is the same way as the base simulation, indicating a fundamental difficulty of the model in reproducing the higher PBL occasions in the Gulf. PBLH23 and PBLH25 have a longer tail toward the higher values than other simulations, making them the closest to the observations. Based on the frequency distribution evaluation, PBLH5, PBLH23, PBLH24, and PBLH25 were selected for further analysis in the subsequent sections, and additional diagnostic evaluations were conducted to understand the impact of the perturbations from those simulations.

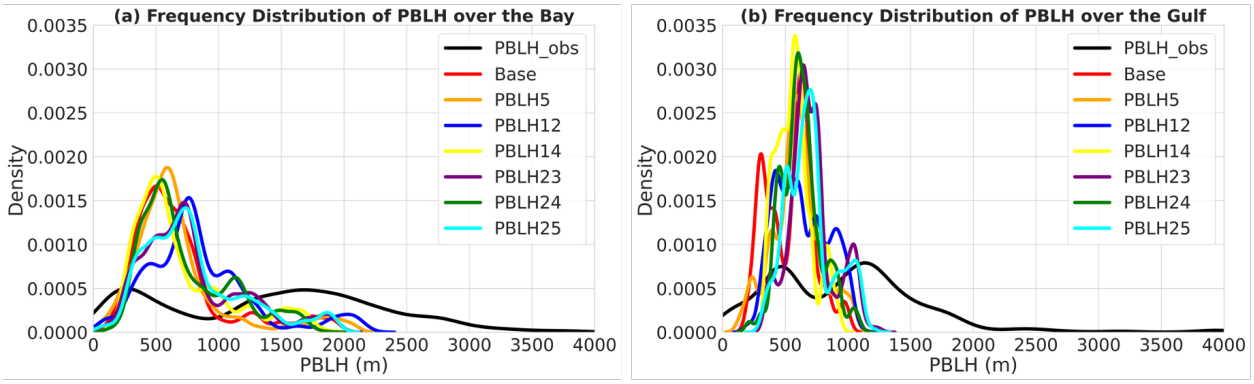


Figure 22. Frequency distribution of daytime (08:00 - 18:00 CDT) PBLH for all the selected days over the (a) Galveston Bay and (b) Gulf.

Figure 23 shows the spatial variability of the ceilometer-observed PBLH compared to the base simulation, along with the differences between the perturbed and base simulations for the three selected best simulations (PBLH5, PBLH23, PBLH24, and PBLH25) over Galveston Bay and the Gulf. The spatial variation shown in **Figure 22** covers only the selected days (July 27-28 and September 9, 2021; September 8-10, 2022; and September 9-10, 2023) for the model perturbations. Across all the selected days through the three years, the base model consistently underestimates the PBLH for both the Bay and Gulf.

Over most of Galveston Bay, all four perturbation simulations show increased PBLH compared to the base simulation. However, variations in observed PBLH, the base simulation, and the perturbed simulations differ across years and locations, likely due to differences in observation time and location. For example, in 2021, in the northwest Bay where the boat was mobile mostly in the afternoon, PBLH increased by ~300 m in PBLH5 relative to the base, and by ~200 m in both PBLH23 and PBLH25, while PBLH24 showed a slight decrease. In contrast, in the northeast Bay, sampled mostly in the morning, the base simulation slightly overestimated PBLH. In this region, all the perturbation simulations performed better than the base, showing values closer to observations. PBLH24 performed better during the morning hours, while PBLH5, PBLH23, and PBLH25 performed better in the afternoon.

Over the three years, the ceilometer-observed PBLH in 2022 was much higher than in 2021 and 2023. During 2022, the base simulation significantly underestimated PBLH over the Bay. All perturbation simulations showed increased PBLH compared to the base, with PBLH23, PBLH24, and PBLH25 demonstrating large improvements (200–400 m), especially during the afternoon. PBLH5 also showed an increase in PBLH in the northeastern Bay, but the improvement was smaller than that from PBLH23, PBLH24, and PBLH25.

In 2023, the boat sampling focused mostly on the Houston Ship Channel (HSC) and southern Bay areas. Observed PBLH was moderate relative to 2022, and the base simulation again slightly underestimated PBLH in and around the southern HSC. All perturbation simulations showed increased PBLH compared to the base, with the largest increases from PBLH23 and PBLH25, particularly during afternoon hours when PBLH is typically at its peak.

Over the Gulf, observed PBLH is generally lower than over Galveston Bay. The base model performed better in capturing PBLH variations for the Gulf compared to the Bay. Over the Gulf, in 2022, the observations were mostly from the eastern part of the Gulf, except when the boat was docked. The ceilometer-observed PBLH in these regions ranged around ~1000 m, and the base simulation slightly underestimated these PBLH values. All four perturbation simulations show increased PBLH compared to the base, with PBLH23 and PBLH25 showing the best agreement with observations.

In 2023, observations over the northeast Gulf (closer to the Bay) show slightly higher PBLH than over the southern Gulf (farther offshore). The base model underestimates PBLH in both regions. PBLH5, PBLH23, and PBLH24 slightly decrease PBLH over the northeast Gulf relative to the base, while PBLH25 shows a subtle increase. In the southern Gulf, all perturbation simulations increase PBLH, with PBLH25 showing the largest improvement and the best agreement with observations.

In summary, when observed PBLH is high over water, all four selected perturbation simulations produce better PBLH estimates than the base model. However, when observed PBLH is low, as is often the case over the Gulf, the perturbation simulations show only minor differences from the base. Overall, PBLH24 performed better during the morning hours, while PBLH23 and PBLH25 performed better during the afternoon. PBLH23 includes a combination of the NSAS cumulus scheme and 1-D ocean physics. PBLH24 includes both the urban canopy model (UCM) and NSAS cumulus scheme, along with 1-D ocean physics. PBLH25 represents a perturbation to the surface roughness length over water and includes the NSAS cumulus scheme and 1-D ocean physics.

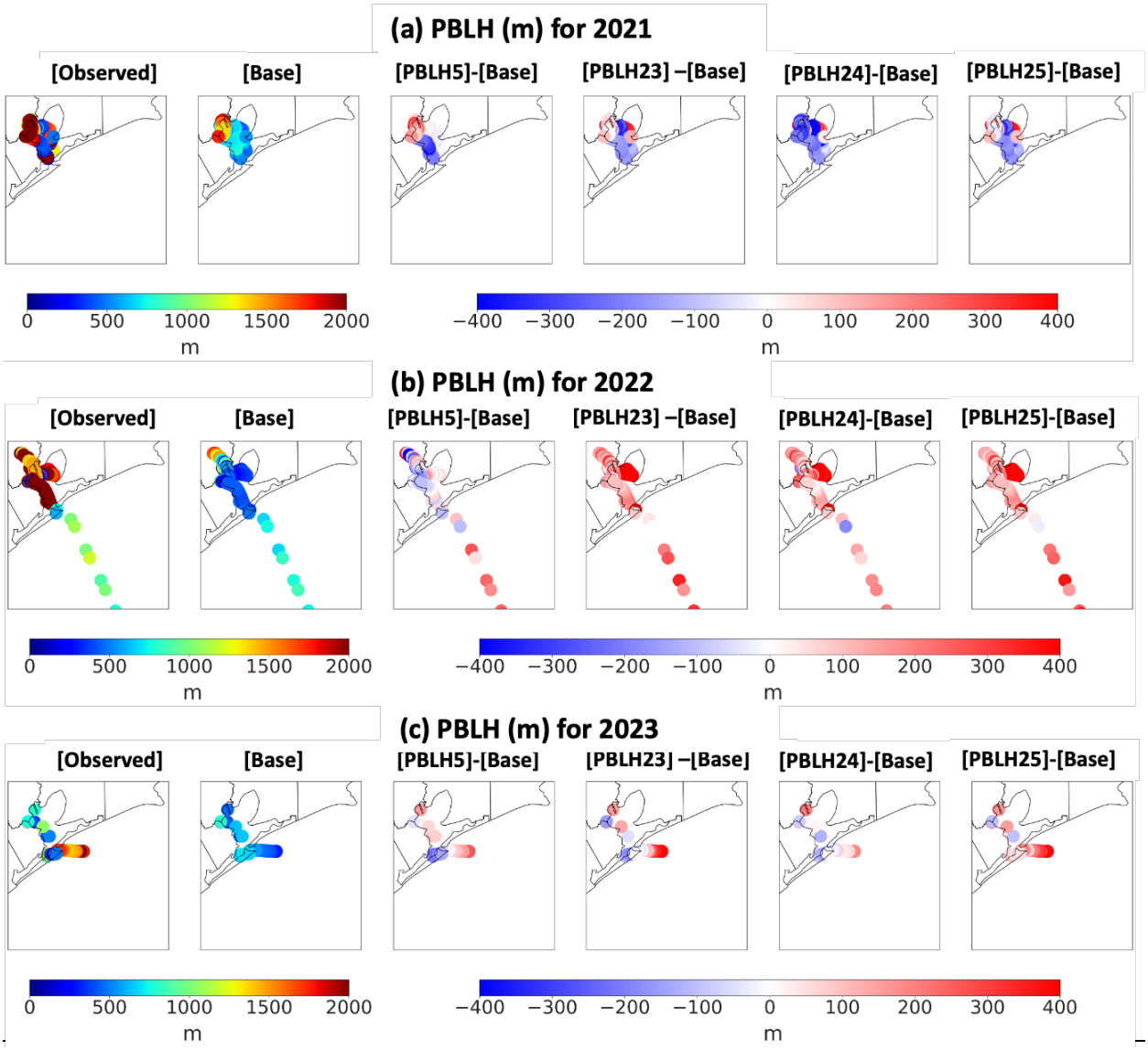


Figure 23. Spatial variation of observed, WRF [Base], [PBLH5] minus [Base], [PBLH23] minus [Base], [PBLH24] minus [Base], and [PBLH25] minus [Base] PBLH (m) for the selected days (a) July 27-28 and September 09, 2021, (b) September 8-10, 2022, and (c) September 9-10, 2023, for the Galveston Bay.

4.2.3. Temporal and spatial changes in PBLH from model perturbations

After establishing that the selected perturbation models better match the observations over the offshore regions for the selected days, the modeled spatial and temporal variations of PBLH were further analyzed over land, Galveston Bay, and the Gulf. For this analysis, September of 2021, 2022, and 2023 were selected due to abundant observations in September compared to other months, and all available observations from each September were used to compare with the selected perturbation experiments.

Given that PBL has distinct diurnal cycles, the changes in the hourly spatial variation of PBLH from the selected perturbation simulations were first examined and compared with the base

simulations and observation data. There is one common day, September 9, from the selected days in each of the three years. This day was used as an example to reveal the spatial and temporal effects of perturbation experiments. **Figure 24** shows the hourly (9:00, 12:00, and 15:00 CDT) spatial distribution of PBLH from the base simulation and the four selected perturbation experiments (PBLH5, PBLH23, PBLH24, and PBLH25) by taking their differences from the base. The ceilometer observed PBLH is overlaid as filled circles in the base plot. The circles include the stationary observations from the LaPorte and Battleground sites (median values for the selected hour) and mobile observations from the different boats and MAQL1 which are overlaid at the sampling time without averaging.

The spatial variation plot shows that the effect of the perturbation is not homogeneous across the region or over the years. These year-to-year differences might be due to year-to-year variations in meteorology. For example, on September 9 (**Figure 24**a) at 9:00 CDT, the perturbation simulation PBLH23 and PBLH25 shows an increase in PBLH of ~100 m over the water, and PBLH5 and PBLH24 don't show much change compared to the base simulation. At 12:00 CDT, the base simulation underestimates PBLH over both the Bay and LaPorte. All perturbation simulations show increases in PBLH (~200 m) over both water and land, improving agreement with observations. In the afternoon, at 15:00 CDT, the base model shows again an underestimation of PBLH over both Bay and LaPorte. PBLH5, PBLH23, and PBLH25 show substantial increases over both the Bay and LaPorte, while PBLH24 shows a decrease over land but maintains better agreement with observations over the Bay. This suggests that the perturbed simulation improved the PBLH estimation over both the water and the surrounding land close to the water for September 2021, with a better improvement from PBLH23 and PBLH25.

For 2022 and 2023, the hourly spatial variation of PBLH in the Base simulation shows unrealistic 'cloud-like' patterns, with a sharp change in PBLH within neighboring grids. As discussed in Section 4, these patterns were influenced by the choice of cumulus schemes used in the model. When the New Tiedtke Cumulus scheme used in the baseline configurations is replaced with the NSAS scheme in PBLH23, PBLH24, and PBLH25, the unrealistic patterns in the PBLH distribution are resolved. When comparing the PBLH from the base simulation with the perturbed simulations for both 2022 and 2023, an increase in PBLH is observed in the perturbed simulations (**Figure 24**b&c), mostly in regions where a sharp decrease in PBLH occurs in the base simulation.

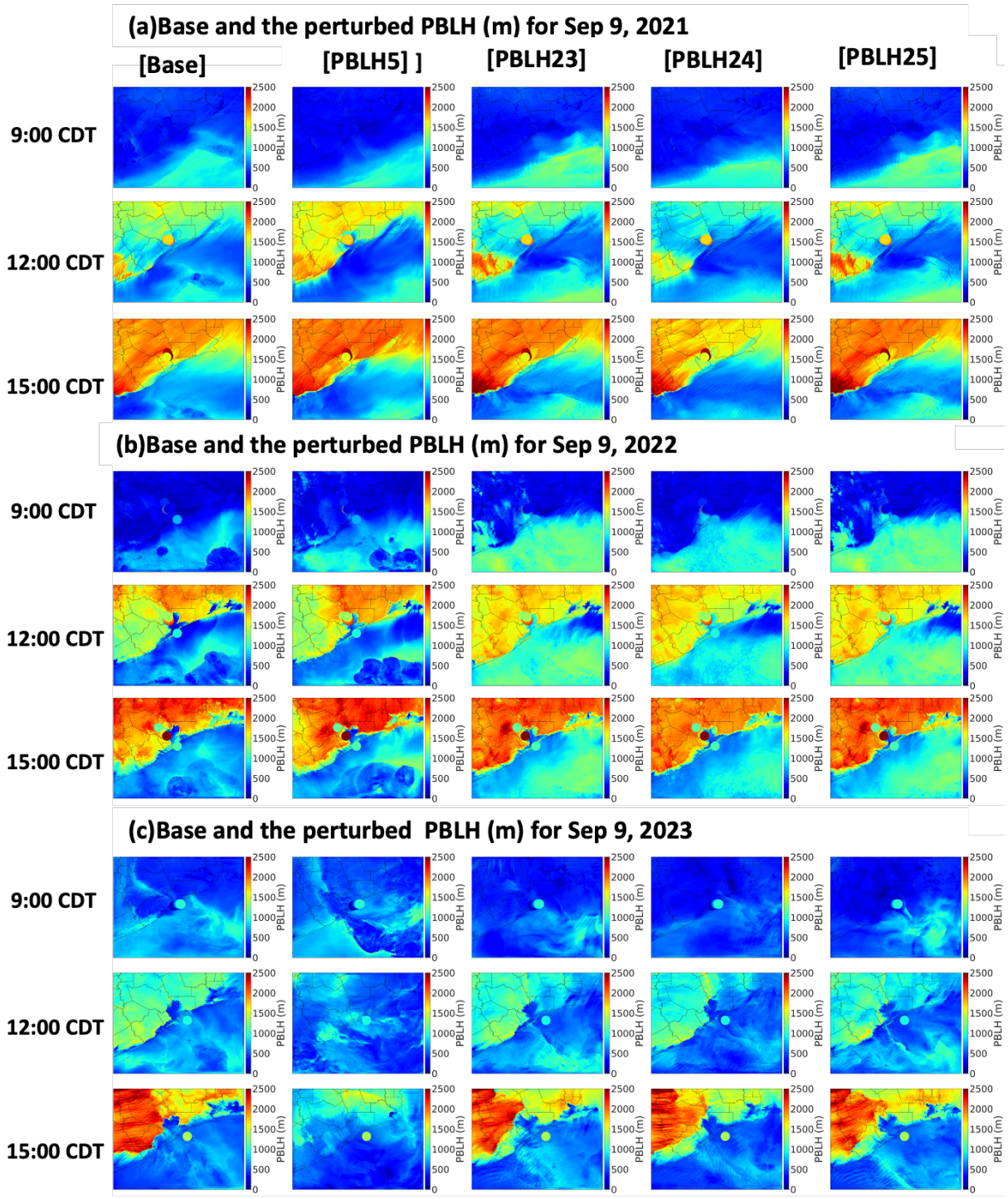


Figure 24. Spatial variation in PBLH (m) from [Base], [PBLH5], [PBLH23], [PBLH24], and [PBLH25] over the land, Galveston Bay, and Gulf for September 9, (a) 2021, (b) 2022, and (c) 2023. Dot overlaid on the map is the ceilometer observed PBL from the LaPorte, University of Houston Pontoon Boat (UHPB), Red Eagle boat (RE), MAQL1, MAQL2, and Osprey Boat (OB).

In September 2022, the base simulation shows very low PBLH during noon and the afternoon over the water, which contrasts with the ceilometer-observed PBLH. The perturbation simulations increase PBLH over the water, yielding values more consistent with observations. At 12:00 CDT, the base simulation significantly underestimates PBLH over the water, whereas all three NSAS-based perturbation simulations show increased PBLH (~400–600 m), better aligning with ceilometer observations. The increase in PBLH over the water is more pronounced in PBLH23 and PBLH25, where the NSAS cumulus scheme is used, compared to PBLH5, where the cumulus scheme is the same as in the base, but the closure constant value is perturbed. By 15:00 CDT, similar trends continue with PBLH23 and PBLH25 showing the strongest and most spatially consistent improvements over both the Bay and coastal land.

In September 2023, the base simulation shows very low PBLH even over land throughout the day. At 9:00 CDT, all simulations, including the base and all perturbation cases, underestimate PBLH compared to observations. By 12:00 CDT, the base simulation continues to underestimate the PBLH regionally, while PBLH25 shows modest increases over land, bringing it closer to observed values. The other perturbation simulations show limited changes during this hour.

In summary, PBLH23 and PBLH25 show consistent daytime improvements in PBLH predictions across years, especially over Galveston Bay. In 2022, all four perturbation simulations improved PBLH compared to the base, particularly during the morning and midday hours. However, PBLH23 and PBLH25 demonstrate stronger improvements during afternoon peaks. The cloud-like artifacts seen in the base simulation are resolved in PBLH23, PBLH24, and PBLH25, highlighting the importance of using the NSAS cumulus scheme. The patterns in 2023 are similar to those in 2022, although the magnitude of the improvements is smaller.

4.2.4. Effect of model perturbations on meteorology

This section presents the effects of the perturbation simulations (PBLH5, PBLH23, PBLH24, and PBLH25) on the meteorological variables for the selected days described in Section 4.1.3, namely: July 27-28 and September 9, 2021; September 8-10, 2022; and September 9-10, 2023, over Galveston Bay and the Gulf.

Figure 25 shows the spatial variability of temperature compared to the base simulation, along with the differences between the perturbed and base simulations for the four selected perturbation simulations. Over Galveston Bay, the base simulation temperature closely follows the observed values, with a slight overestimation of ~0.5-1°C in and around the HSC. Compared to the base simulation, all four perturbation simulations show a temperature decrease (~0.5°C) around the HSC, suggesting that the perturbation simulations slightly improve temperature prediction over the HSC. For other regions of the Bay, the base simulation slightly overestimated the temperature by ~0.5°C in 2021, and the PBLH5 simulation better captures the temperature compared to the others.

In 2022, all perturbation simulations outperform the base in reproducing observed temperatures. Over inland areas near the HSC, the base simulation overestimates temperatures by ~0.5–1°C, while PBLH5, PBLH23, and PBLH24 reduce this bias, with PBLH5 showing the closest agreement with observed temperatures. In the southern HSC region, the base simulation underestimates temperature by ~0.5°C. All perturbation simulations improve upon this, with PBLH5 and PBLH23 correcting the underestimation by about 0.5°C, and PBLH25 showing the largest improvement, increasing temperatures by ~0.5–1°C and aligning well with observed

values. Over the Gulf, the base simulation slightly underestimates temperature by ~0.5°C. All perturbation simulations provide improvements, with PBLH25 showing the best improvement, ~0.7°C, bringing simulated temperatures closer to observations. PBLH5 and PBLH23 also yield improvements of ~0.3–0.5°C, still outperforming the base simulation.

In 2023, the base simulation generally underestimated temperature by ~0.5-1.0°C over Galveston Bay and the Gulf. Among the perturbation simulations, PBLH5 performed the best in the Bay, reducing the underestimation and providing a closer match to observed temperatures compared to the other perturbations. PBLH23 and PBLH24 also showed some improvements. Over the Gulf of America, all four perturbation simulations improved upon the base, with PBLH25 showing the most substantial improvement. PBLH25 increased temperatures by ~0.7–1°C, bringing them closer to the observed values. PBLH5 and PBLH23 provided smaller corrections of ~0.3–0.5°C but still performed better than the base simulation. Overall, PBLH25 was the most effective in correcting the temperature bias over the Gulf.

In summary, PBLH5 performed best in reducing temperature bias over Galveston Bay, especially in 2021 and 2023. PBLH25 was the most effective in correcting underestimation over the Gulf across all years.

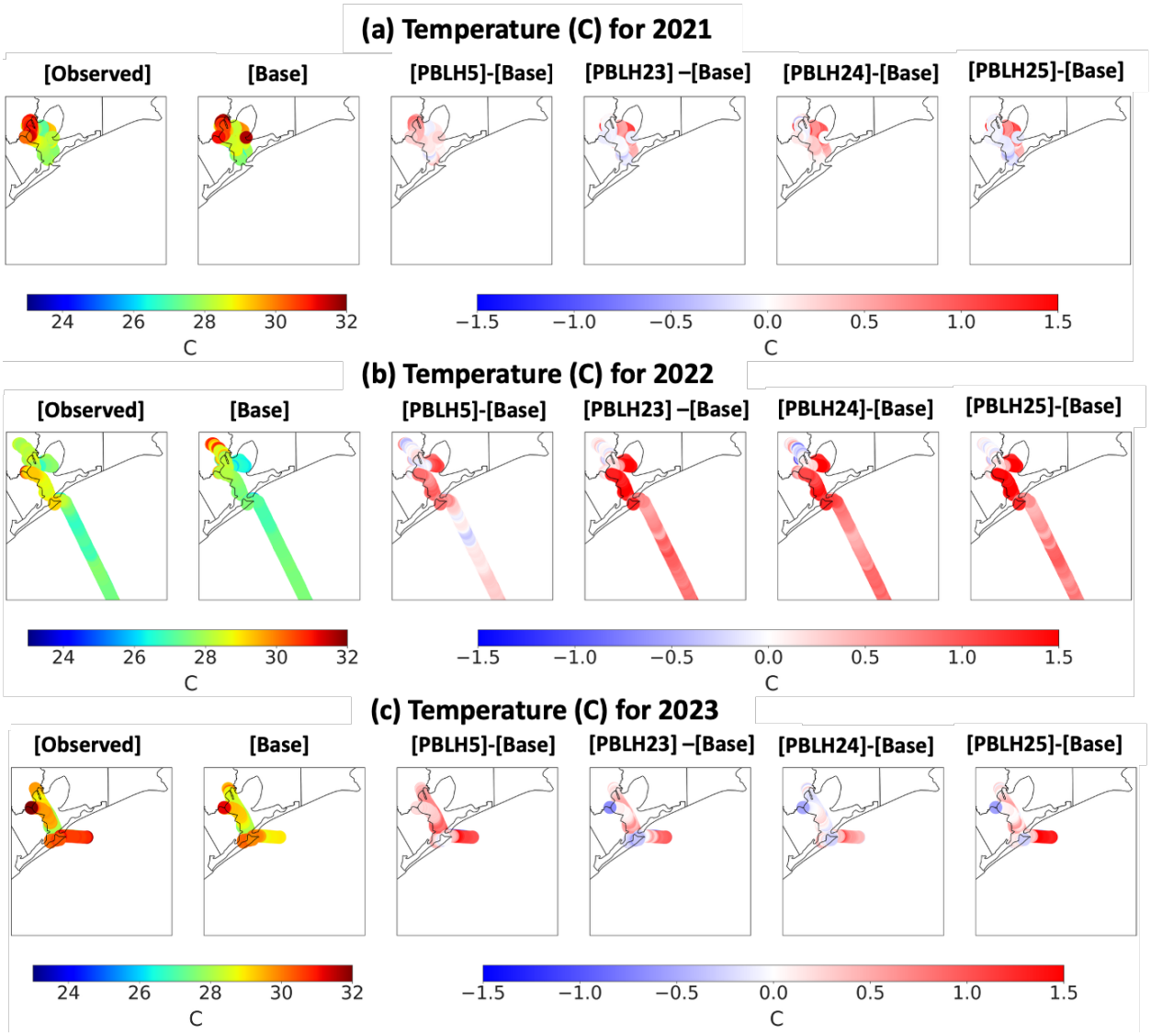


Figure 25. Spatial variation of observed, WRF [Base], [PBLH5] minus [Base], [PBLH23] minus [Base], [PBLH24] minus [Base], and [PBLH25] minus [Base] Temperature (C) for the selected days (a) July 27-28 and September 09, 2021, (b) September 8-10, 2022, and (c) September 9-10, 2023, for the Galveston Bay.

Figure 26 shows the spatial variability of wind speed compared to the base simulation, along with the differences between the base and the four selected perturbation simulations (PBLH5, PBLH23, PBLH24, and PBLH25). For the selected simulation days, the base simulation generally overestimated wind speed compared to the observations. In 2021, the base simulation underestimated wind speed by ~1–1.5 m/s. Among the perturbation simulations, PBLH24 performed the best, providing a more accurate representation of wind speed. In particular, over the eastern part of Galveston Bay, PBLH24 corrected the underestimation seen in the base simulation and aligned more closely with observed values.

In 2022, the base simulation underestimated wind speed in the northern and central Bay regions. In these areas, PBLH5 performed better, reducing the bias relative to the base. In the southern

Bay, the base simulation overestimated wind speed by ~1–1.5 m/s. Here, PBLH5, PBLH23, and PBLH25 all reduced wind speed, improving agreement with observations. Over the Gulf, the base simulation also underestimated wind speed, and PBLH24 provided the best match to observed values, particularly in the southern Gulf.

In 2023, the base simulation overestimated wind speed over the northern Bay by ~1–1.5 m/s. All four perturbation simulations reduced the wind speed in this region, improving agreement with observations. In the southern Bay, where the base model underestimated wind speed, PBLH5, PBLH23, and PBLH24 increased wind speeds and brought the simulation closer to observed values. Over the Gulf, the base simulation overestimated wind speed, especially in the southern Gulf, by ~2–3 m/s. All perturbation simulations improved performance, with PBLH24 showing the greatest reduction in wind speed bias and the closest alignment with observed data. PBLH5, PBLH23, and PBLH25 also reduced the overestimation but were slightly less effective than PBLH24 in this region.

In summary, PBLH24 performed best overall in improving wind speed predictions, especially in regions with overestimated wind, such as the southern Gulf. PBLH25 also showed improvements, particularly in the southern Bay and Gulf, but its performance was slightly less consistent than PBLH24 across all regions and years.

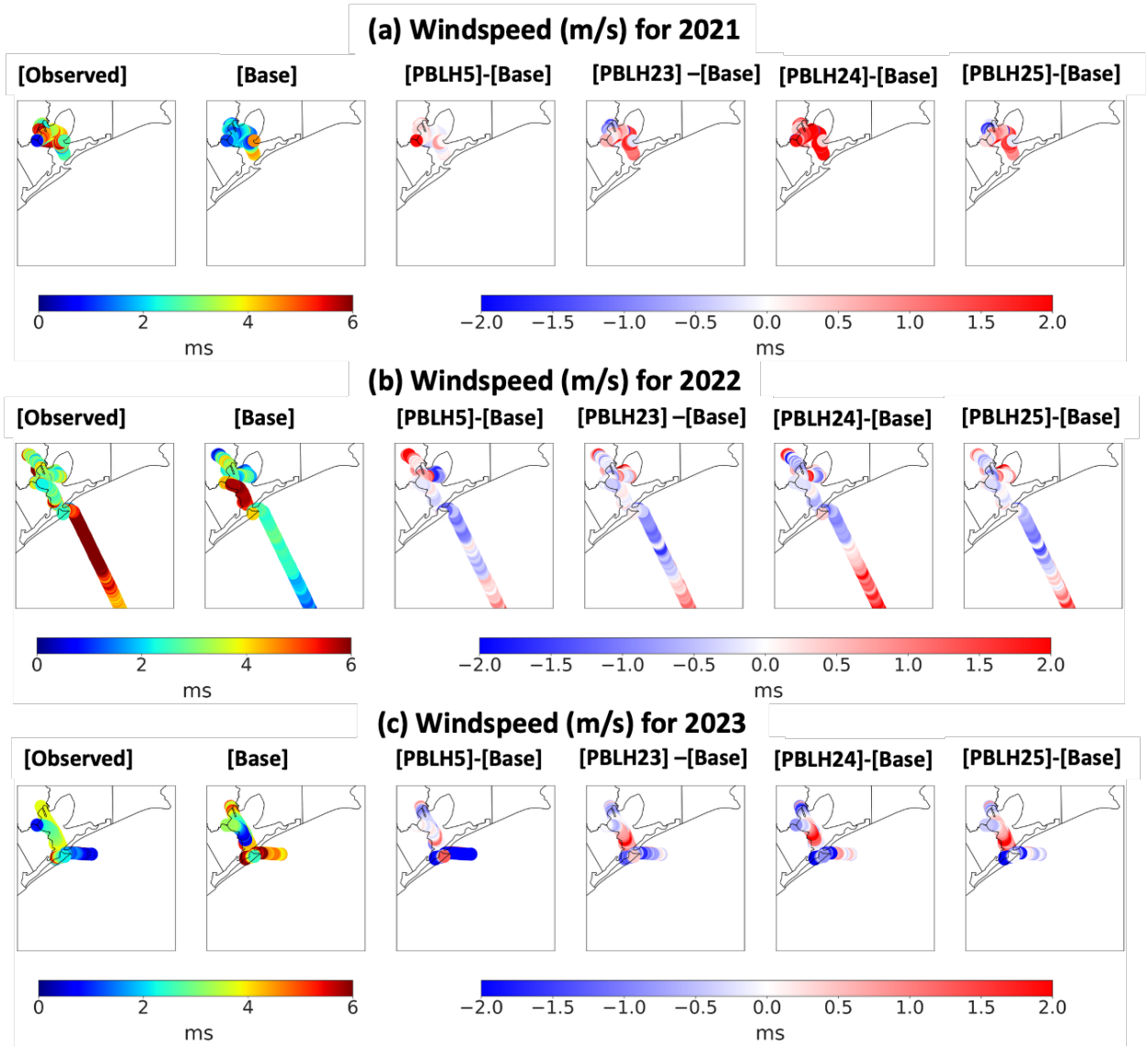


Figure 26. Spatial variation of observed, WRF [Base], [PBLH5] minus [Base], PBLH23 minus [Base], PBLH24 minus [Base], and PBLH25 minus [Base] Windspeed (m/s) for the selected days (a) July 27-28 and September 09, 2021, (b) September 8-10, 2022, and (c) September 9-10, 2023, for the Galveston Bay.

4.3. Evaluation by SCOAPE Campaign

Given the similarity between PBLH23, PBLH24, and PBLH25 in configuration and performance over the water, PBLH24 was chosen to be further evaluated by independent PBL measurements over Southern Louisiana and the Gulf from the Satellite Coastal and Oceanic Atmospheric Pollution Experiment (SCOAPE) 2019 (SCOAPE-19) campaign (Thompson et al., 2023). The SCOAPE-19 campaign was run in May 2019 and was composed of onshore and offshore observations. This report used only the offshore observations from the University of Southern Mississippi's Research Vessel (R/V) Point Sur, which cruised in the Gulf from the 10th to the 18th

of May 2019, as this was the only observation that gave directly calculated PBLH. The PBL calculation was done automatically from the Lufft Cloud Height Meter (CHM) 8k ceilometer on board the R/V Point Sur and gave the heights of three PBLs per measurement. Only the first/lowest layer was used and was run through a filter to remove any possible artifacting and to temporally average the data per hour, as shown in **Figure 27b**. High noise between 05-14 18:00 to 05-16 06:00 was present in the unfiltered observed data (**Figure 27a**). Due to the PBLH values being calculated from the integrated controller and the widespread presence of noise around that time, the issue was attributed either to an instrument error or an issue with the integrated controller. This period was removed from the analysis.

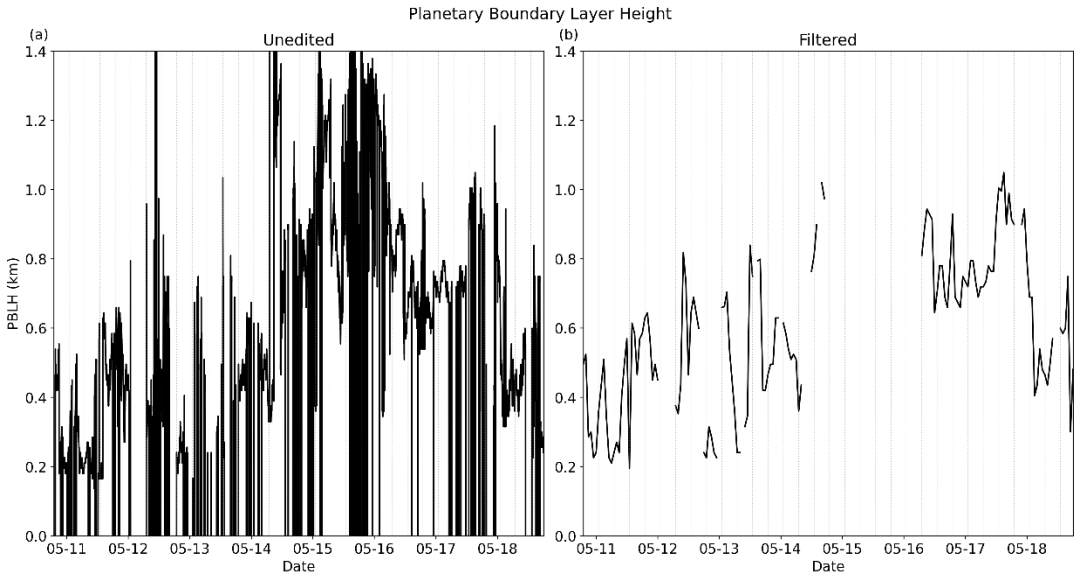


Figure 27. (a) Unprocessed layer 1 PBLH (km) data and (b) Filtered PBLH (km) data from the Lufft CHM 8k ceilometer on board the R/V Point Sur in 2019.

The WRF v4.6.0 model, as described above, was set up for the SCOAPE campaign area as shown in **Figure 28**. The three domains are over the contiguous United States (d01 – 12 km x 12km), Southeastern Texas/Louisiana (d02 – 4 km x 4 km), and Southern Louisiana (d03 – 1.33 km x 1.33 km), respectively. Two different model runs were done, the base model and a run using the settings from PBLH24 (PBLH24). The initial and boundary conditions (IC/BC) were generated in the same way as in Section 2.2, such that IC/BC for domain d01 were generated from the HRRR meteorology model, and IC/BC conditions for domains d02 and d03 were generated from d01 and d02, respectively. Both WRF runs were run from May 10th, 2019, to May 18th, 2019, with May 10th used as a spin-up day.

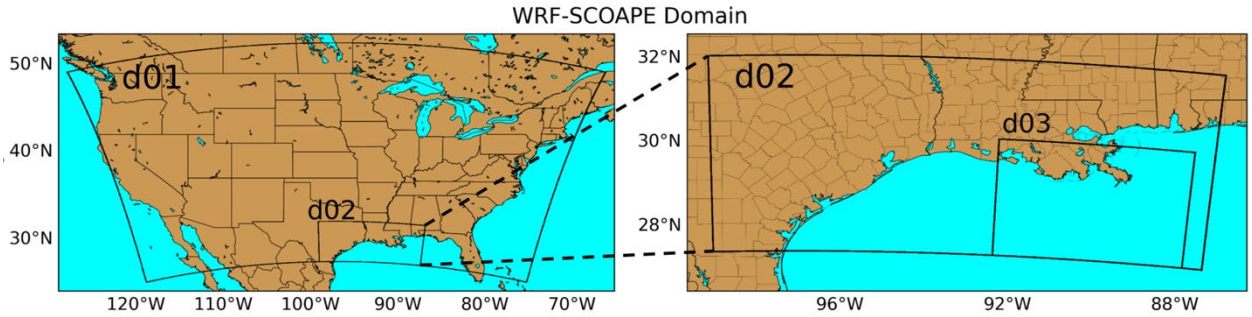


Figure 28. WRF domains for SCOAPE showing d01 (contiguous United States), d02 (Southeastern Texas/Louisiana), and d03 (Southern Louisiana).

The performance of PBLH for both the base run and PBLH24 were evaluated by comparing to the filtered observation data from R/V Point Sur over the campaign period. Both the filtered observation data and WRF simulations have a temporal resolution of an hour. The model PBL values were taken based on the average latitude and longitude for each hour of the R/V Point Sur tracks. Observed and simulated PBLH values are **Figure 29** for the diurnal and in **Figure 30** over the track of R/V Point Sur. As observed in **Figure 29**, the diurnal profile of PBLH over the track of R/V Point Sur is relatively consistent with mean values ranging from ~0.5 km to ~0.8 km. Both models tend to overestimate early morning hours (3:00 – 10:00 CDT) while keeping close to the observed for the remainder of the hours. The base model's overestimation also correlates with its higher day-to-day variation, as denoted by the large color bars in **Figure 29**. PBLH24 has a lower bias and smaller day-to-day variation than the base model in the morning hours, making it better match the observations. The two models are similar in performance in the afternoon hours, while PBLH24 has a slower rate of PBL decreasing from late afternoon to the evening hours (18:00 – 20:00), like the observations.

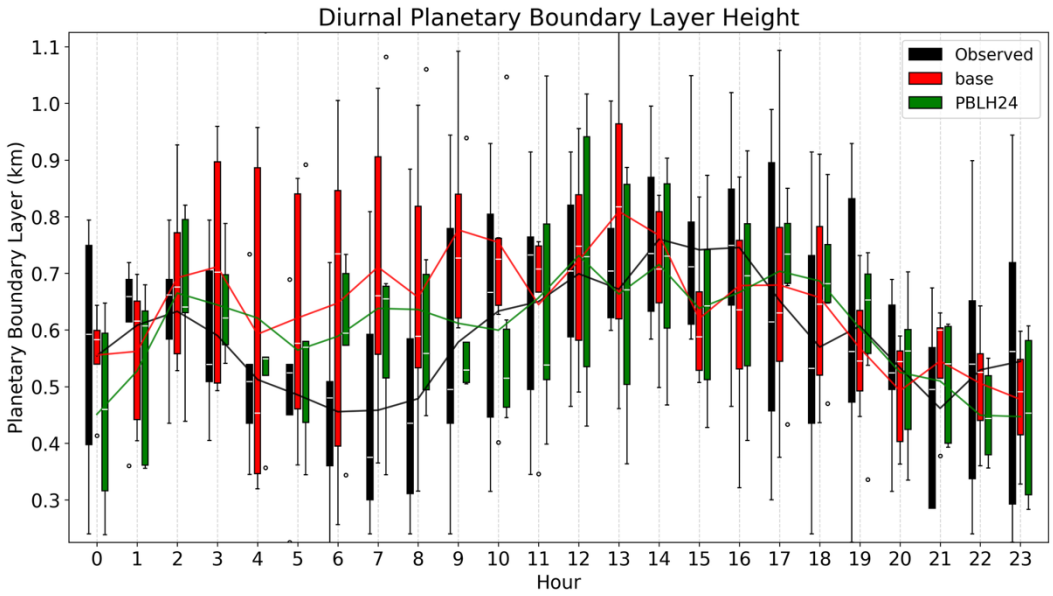


Figure 29. Boxplot of the diurnal variation of PBLH over the ship track of R/V Point Sur for observed data (black), base run (red), and PBLH24 (green). Data is from 05-10 19:00 to 05-18 19:00, with 05-14 18:00 to 05-16 06:00 removed.

Overall performance of the models is quite similar with the mean PBLH values of both WRF runs being within 50 m of each other (0.648 km and 0.611 km for base and PBLH24, respectively) as seen in **Figure 30**. When compared to the mean observed values, both runs are within 100 m, and both tend to slightly overestimate PBLH, with PBLH24 outperforming the base model by lowering the bias by 0.037 km. When looking at the correlation coefficient, PBLH24 also shows a higher value (0.52) than the base run (0.43), showing better performance both spatially and temporally.

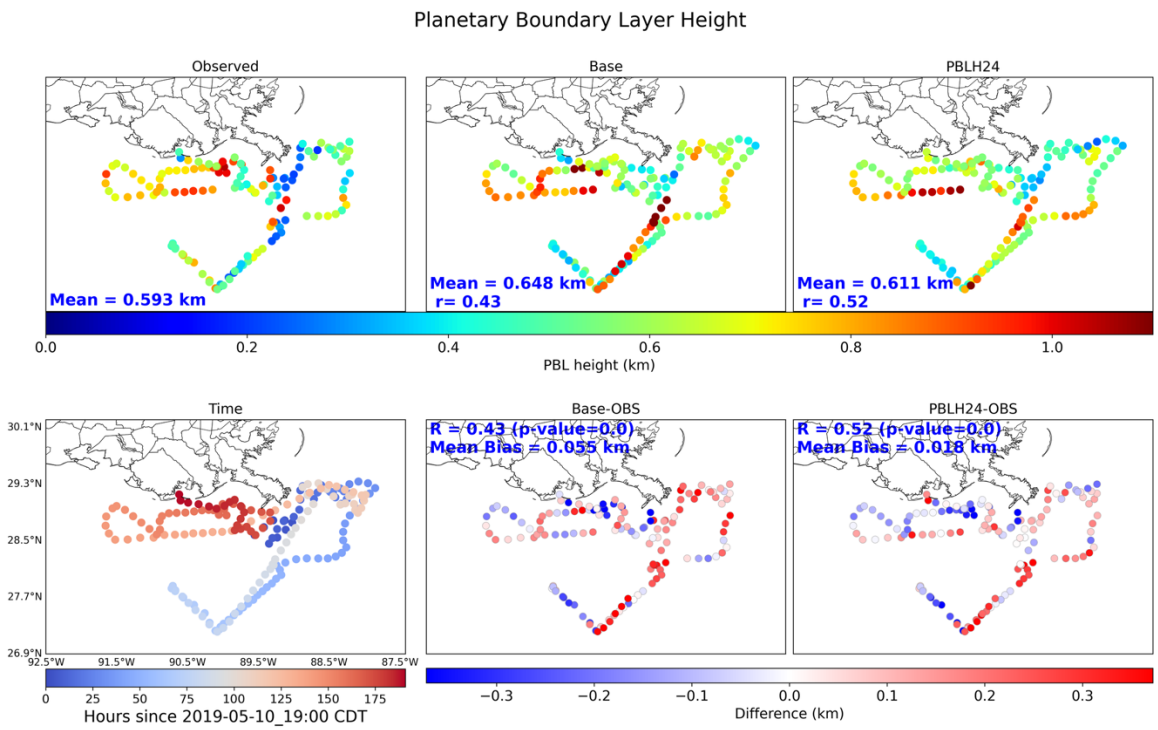


Figure 30. PBLH over the R/V Point Sur ship track for (a) Observed, (b) Base run, (c) PBLH24 with time distribution shown in (d), and modeled-observed differences shown for the Base run (e) Base modeled-observed and (f) PBLH24-observed.

4.4. Summary

PBL-related parameterizations in WRF need improvements and tuning to better match observations over water bodies around Houston/Galveston. This task identified key parameters and constants in the MYNN PBL scheme that influence the evolution and development of PBLH over water. Based on the sensitivity of these parameters to PBLH, multiple perturbation simulations were performed to improve the model's performance by fine-tuning the parameters. In addition to perturbing physical parameters in the model, sensitivity simulations were also conducted by selecting different cloud and urban physics schemes because the base WRF model was found to exhibit unrealistic PBLH features that were traced back to those configurations. A total of 9 days were chosen for model perturbation/sensitivity simulations, with each simulation run for all 9 days, resulting in more than 200 simulation days.

PBLH from different WRF4.6.0 sensitivity simulations was evaluated by comparing with the onshore and offshore observation data from TAQ1, TAQ2, and TAQ3. From various perturbation simulations, four sets of simulations were selected: PBLH5, PBLH23, PBLH24, and PBLH25, which improved the PBLH prediction over the water without considerably affecting performance over the land. PBLH5 refers to the perturbed value of closure constant C3. PBLH23 incorporates the NSAS cumulus scheme with 1-D ocean physics; PBLH24 includes both the UCM and NSAS cumulus scheme with ocean coupling; and PBLH25 adds a surface roughness length perturbation over water in addition to NSAS and ocean physics. Each of these perturbations improved WRF's ability to simulate marine PBLH across different years and conditions.

By comparing the perturbed PBLH experiments with the base simulation and observations, all four simulations (PBLH5, PBLH23, PBLH24, and PBLH25) were consistently more successful in improving WRF's PBLH representation over Galveston Bay and the Gulf without significantly degrading performance over land than the base model (by increasing PBLH up to 400 m). Frequency distribution analysis showed that PBLH23, PBLH24, and PBLH25 captured the afternoon peak in observed PBLH over the Bay more accurately than the base model. Spatial and temporal comparisons further showed that PBLH23 and PBLH25 performed best during afternoon hours, especially in 2022 and 2023, when the base simulation showed the largest underestimation. These two configurations also resolved unrealistic "cloud-like" spatial artifacts seen in the base, underscoring the benefit of replacing the New Tiedtke cumulus scheme with the NSAS scheme and including ocean coupling. As a result, PBLH23 and PBLH25 were identified as the most robust options for improving offshore PBLH and are prioritized for further evaluation in subsequent tasks.

The perturbation simulations also showed varying impacts on meteorological variables such as temperature and wind speed. PBLH5 consistently reduced temperature bias over Galveston Bay, particularly in 2021 and 2023, while PBLH25 was the most effective in correcting temperature underestimation over the Gulf. For wind speed, PBLH24 outperformed the other perturbations across all years, especially in reducing overestimation in the southern Gulf. PBLH25 also improved wind predictions in the southern Bay and Gulf but was slightly less consistent than PBLH24. These results highlight the importance of carefully selected physics configurations not only for improving PBLH but also for enhancing broader meteorological accuracy in coastal and offshore regions.

The analysis performed in this section presents that the WRF model performance in offshore PBL can be improved by fine-tuning the physical parameters of the PBL scheme within the published ranges and by different configurations of the cloud and urban physics schemes. The four selected perturbation models all improve upon the base model in predicting offshore PBL and meteorology, although the specific best perturbations are dependent on the time of day and region. To build on these findings, PBLH23, PBLH24, and PBLH25 are selected as the three most robust configurations to carry forward into subsequent analyses. These simulations consistently improved PBLH representation, particularly during the afternoon over Galveston Bay and the Gulf and also reduced temperature and wind speed biases compared to the base. Each incorporates key physical modifications, including the NSAS cumulus scheme, 1-D ocean coupling, the urban canopy model, and surface roughness length adjustments. Among them, PBLH25 is recommended as the lead configuration based on its balanced and consistent performance across all evaluated meteorological variables. However, all three configurations will be further assessed for their ability to represent the residual layer and the impact on ozone prediction. The final selection of the optimal configuration will be based on comprehensive performance across both meteorological and air quality outcomes.

To independently evaluate these configurations beyond the Houston region, PBLH24 was tested against offshore PBL observations from the SCOAPE-19 field campaign over the Gulf. Results showed that PBLH24 reduced bias and improved temporal correlation relative to the base model, particularly during early morning hours when overestimation was most pronounced. This independent validation suggests that the perturbation simulations are effective not only in the Houston-Galveston-Brazoria region but also in other coastal and offshore environments. Thus, the improvements demonstrated by these perturbation configurations may have broader regional

applicability for enhancing marine PBL and air quality modeling across the Gulf Coast and similar settings.

5. Improvements to WRF representation of Residual Layer

The PBL, typically extending up to ~2 km, is the lowest part of the troposphere directly influenced by surface forcings. During the day, surface heating generates turbulence, forming a convective boundary layer (CBL) characterized by vertical mixing. After sunset, turbulence decays and the CBL transitions into a residual layer (RL) aloft (~1–2 km) while the surface layer cools and forms a shallow, stable boundary layer (SBL) below (~0–0.5 km). From our previous work (Liu et al., 2023), which analyzed PBL variations in the Houston region during September 2021, it was observed that the diurnal structure of daytime CBL followed by nighttime SBL is typically dominant under less polluted or low-ozone conditions. However, during high-ozone episodes in the Houston area, a RL was found to persist overnight. This persistence of RL was often capped by a strong capping inversion (CI). The elevated RL can trap pollutants aloft and potentially contribute to next-day ozone buildup.

In the WRF base model, only a single-layer PBL height is diagnosed. The model identifies only the shallow SBL at night and fails to capture the overlying residual layer (Liu et al., 2023). To address this limitation, both an empirical method and a theoretical method were developed, described below, to diagnose the residual layer height. The empirical method uses the vertical gradient of potential temperature ($\partial\theta/\partial z$) to calculate the RL, which is capable of identifying strong capping inversions and layered thermal structures within the model. The theoretical method uses a combination of dynamic and thermodynamic parameters, including vertical profiles of wind speed, vertical velocity, potential temperature, turbulent kinetic energy (TKE), and their gradients, to assess the balance between mechanical and buoyant turbulence.

5.1. Thermodynamic calculation of residual layer

To diagnose the residual layer, vertical profiles of the potential temperature gradient ($\partial\theta/\partial z$) from the WRF base model at LaPorte for September 2021 were first analyzed. The LaPorte site was selected due to the availability of continuous, stationary ceilometer-observed PBLH observations throughout the month, which allowed for abundant datapoints to develop the empirical method based on comparing the modeled thermodynamic structure with observed boundary layer features. Our previous work (Liu et al., 2023) showed that the LaPorte ceilometer observations clearly identified the presence of a stable boundary layer (SBL) on clean-air days and a well-defined residual layer (RL) on high-ozone days. By comparing the modeled $\partial\theta/\partial z$ profiles with these observed structures, it was found that the model reproduced sharp vertical gradients in potential temperature that corresponded well with the observed RL heights. The observed agreement provided confidence in developing an empirical-based approach to diagnose the RL height by analyzing the thermodynamic structure in the model.

Comparison of the observed PBL with the modeled $\partial\theta/\partial z$ showed that strong temperature inversions associated with the top of the RL typically occurred where $\partial\theta/\partial z$ exceeded 6.6 °C/km at LaPorte. This threshold was applied in our empirical algorithm to identify temperature inversion layers within the model as follows. At each model grid column in time and space, the algorithm searched for the strongest thermal inversion defined as a continuous sequence of at least three vertical levels where $\partial\theta/\partial z$ exceeded 6.6 °C/km; the search was limited to heights

below 3 km above the surface. The height at which this sharp inversion occurred (the first point in the three consequent thermal inversion layers) was then recorded as the residual layer height. If no strong inversion was found, the algorithm then searched for weaker but still meaningful gradients using lower thresholds of 3.0, 2.0, and 1.0 °C/km, each requiring at least four continuous vertical levels exceeding these thresholds. The search first happens with a threshold of 3.0 °C/km and moves to a lower threshold if no residual layer assessment is made for the given threshold. These additional thresholds allowed the method to capture more diffuse or less well-defined boundary layer structures, especially under transitional or weakly stable conditions.

There could still be some limited days, e.g., clean days, where none of the gradient thresholds will be met. To still have a PBL assessment for these days, which will associate with boundary layer height, the algorithm identifies a boundary layer by checking for any noticeable increase in potential temperature with height. Specifically, it searched for the first level where $\partial\theta/\partial z$ exceeded 3.0, 2.0, or 1.0 °C/km. Once such a value was found, the corresponding height was recorded as the estimated boundary layer height. This method is useful for identifying shallow surface-based stable boundary layer (SBL) associated with nocturnal radiative cooling.

The combined approach with sharp inversion identification or the identification of noticeable increase in $\partial\theta/\partial z$ allowed us to diagnose different boundary layer structures, including the stable boundary layer (SBL), the convective boundary layer (CBL), and the residual layer (RL). On clean-air days, the diurnal cycle is typically characterized by a shallow SBL at night and a well-developed CBL during the day. In contrast, under polluted conditions, a persistent residual layer capped by a strong inversion often forms overnight.

Figure 31 shows example plots comparing the thermodynamically-diagnosed residual layer (RL) height (green markers) based on the vertical gradient of potential temperature ($\partial\theta/\partial z$), observed PBLH from the ceilometer (red markers), base model-predicted PBLH (black line) for September 2 and September 24, 2021, at LaPorte. The background contour represents the modeled $\partial\theta/\partial z$ profile, which highlights key atmospheric structures, such as capping inversions. On September 2, no strong capping inversion was observed, and both the observations and model indicate a relatively shallow boundary layer throughout the day. In contrast, on September 24, a distinct capping inversion is present, and the ceilometer data clearly shows the existence of a nocturnal residual layer. While the WRF base model fails to capture this structure by simulating only a shallow stable boundary layer at night, the diagnosed RL height from the potential temperature gradient closely matches the observed RL. This demonstrates the utility of the diagnostic approach for identifying residual layers that are otherwise missing in the model output. In the afternoon, the diagnosed layer (green markers) is comparable to model-derived PBLH, which is an indirect verification of our method of thermodynamically-diagnosed RL.

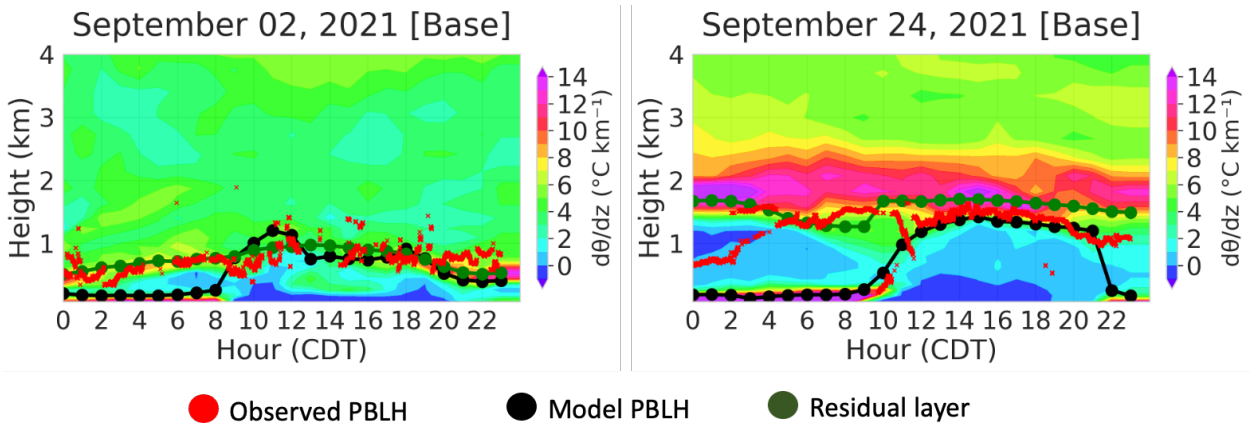


Figure 31. Thermodynamically identified residual layer (RL), along with the observed and base modeled planetary boundary layer heights (PBLH) for September 02 and 24, 2021, at La Porte. The background shading represents the modeled vertical gradient of potential temperature ($\partial\theta/\partial z$). Red dots indicate ceilometer-observed PBLH, black lines show model-predicted PBLH, and green markers represent the diagnosed residual layer top based on potential temperature gradient analysis.

5.1.1. Evaluation of residual layer

To quantitatively validate our approach for diagnosing the residual layer, the thermodynamically calculated residual layer heights were compared with observed PBLH data at LaPorte for September 2021, where continuous ceilometer-based observations were available throughout the month. **Figure 32** presents a statistical comparison of how well the WRF base model and our thermodynamic method capture residual layer structures. During nighttime hours (10 PM to 6 AM), the base model significantly underestimates the PBLH and fails to capture the elevated residual layer, as shown in **Figure 32**ii. In contrast, the thermodynamically-diagnosed residual layer heights based on potential temperature gradients show substantially better agreement with observations (**Figure 32**i), with higher correlation and reduced bias.

To further assess the applicability of our diagnostic approach, the performance during daytime hours (8 AM to 8 PM) was also evaluated. **Figure 32**iii and 4iv compare the thermodynamically calculated PBLH and the model-predicted PBLH, respectively, against observations. Both approaches show better agreement with the observations during the day compared to nighttime, with stronger correlation and lower bias. These results confirm that our method based on vertical gradients of potential temperature effectively captures the residual layer during nighttime conditions, when the WRF base model tends to underestimate boundary layer height, while also performing well for daytime PBLH estimation.

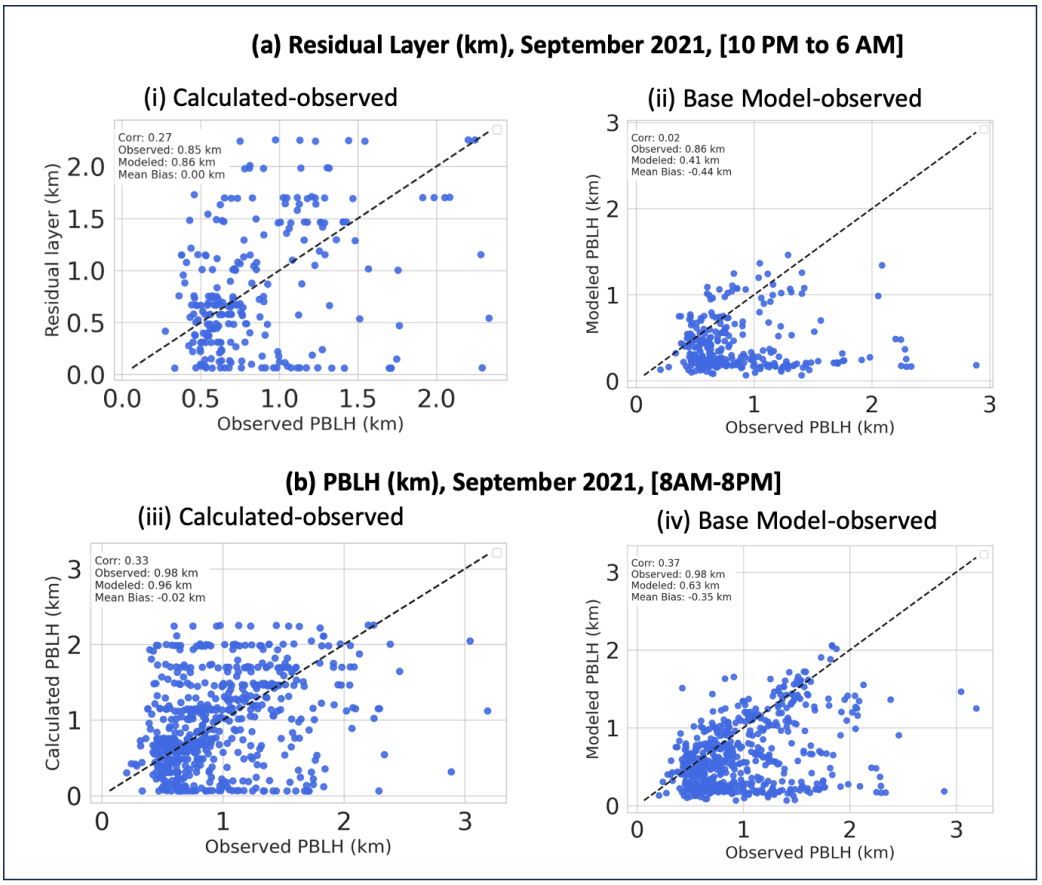


Figure 32. Scatterplots of (i) observed and thermodynamically-diagnosed residual layer, (ii) observed and modeled residual layer (km), (iii) observed and thermodynamically calculated PBLH (km), along with (iv) observed-modeled PBLH. The dashed line represents the 1:1 reference line. Statistical metrics such as correlation, mean bias are included in each plot.

5.1.2. Spatial variation of the residual layer

The residual layer height was diagnosed thermodynamically, as described in Section 5.1, using the vertical gradient of potential temperature ($\partial\theta/\partial z$) for selected nighttime periods across 2021, 2022, and 2023. The diagnostic method was applied to the WRF base model (Base) and three selected best perturbation simulations: PBLH23, PBLH24, and PBLH25 to assess how well each configuration represents the spatial structure of the residual layer over land, Galveston Bay, and the Gulf.

Figure 33 shows the spatial distribution of thermodynamically-diagnosed RL height, calculated from the WRF model potential temperature gradients from midnight to 6 AM for (a) July 27–28, 2021, (b) September 8–9, 2021, (c) September 8–9, 2022, and (d) September 8–9, 2023. RL heights are shown for the WRF base model and three perturbation simulations: PBLH23, PBLH24, and PBLH25. The ceilometer observed mean PBLH for the selected time is overlaid as filled circles in all the plots. The circles include the stationary observation from LaPorte and mobile observations from UHPB, RE, OB, and MAQL1. Across most days, the diagnosed RL heights from the WRF base and all perturbation simulations show reasonable agreement with observations, with the perturbed simulations showing higher RL over water. However, for

September 8–9, 2022, PBLH23, PBLH24, and PBLH25 overestimate the RL height by ~1000 m compared to the observed PBLH from UHPB, RE, and MAQL1, while the base model matched better with the observations.

The thermodynamic approach of calculating RL is empirically developed based on the potential temperature gradient profile from the base run and fine-tuned with the ceilometer observations at the LaPorte site for the year 2021. This approach captured the RL height well for the base run, but the results are not consistent between the perturbation runs PBLH23, PBLH24, and PBLH25. This is probably because the base uses the same cumulus scheme from the base configuration, but other perturbation runs use the NSAS cumulus scheme. This suggests that our empirically developed RL diagnosis method needs to be tuned for the NSAS cumulus scheme, which will be conducted in subsequent work of the project.

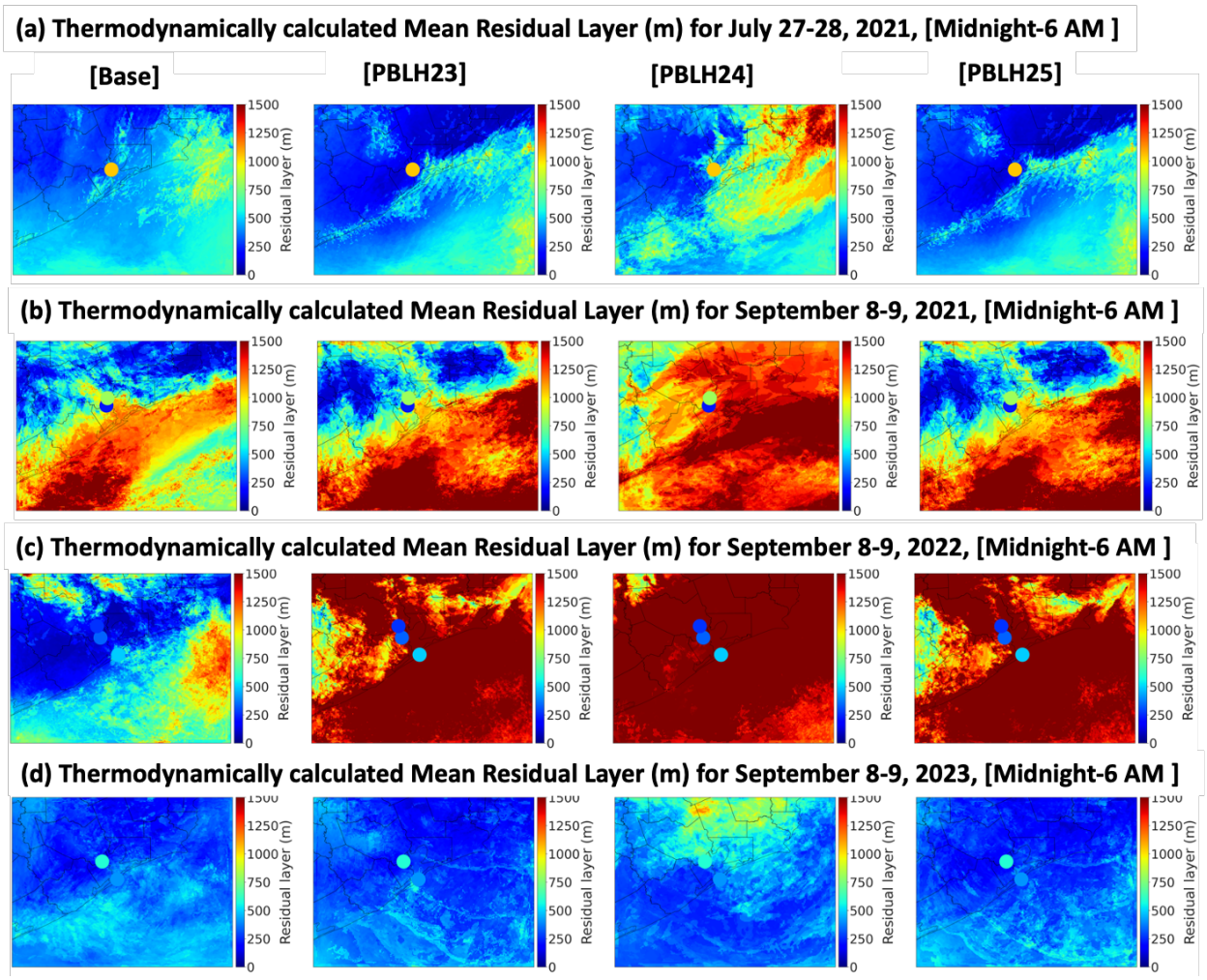


Figure 33. Spatial variation in thermodynamically calculated residual layer from WRF [Base], [PBLH23], [PBLH24], and [PBLH25] over the land, Galveston Bay, and Gulf for (a) July 27-28, 2021, (b) September 8-9, 2021, (c) September 8-9, 2022, and (d) September 8-9, 2023, during the midnight to 6 AM. Dot overlaid on the map is the ceilometer observed mean PBLH from the LaPorte, University of Houston Pontoon Boat (UHPB), Red Eagle boat (RE), MAQL1, and Osprey Boat (OB).

5.2. Development of a theoretical model for the residual layer model

Considering the shortcomings of the empirical model described above, an attempt was made to develop an observation-guided theoretical model to diagnose the RL. In the atmospheric boundary layer, both thermal (buoyancy-driven) and mechanical (shear-driven) turbulence are key determinants of atmospheric stability. These processes directly influence the PBLH and the potential formation of a RL under varying meteorological conditions. One important concept derived from this interaction is the flux Richardson number (Rif), which quantifies the ratio of buoyant production (or suppression) of turbulent kinetic energy (TKE) to its shear production. In essence, Ri reflects the balance between thermal suppression and mechanical generation of turbulence.

Building upon this theoretical foundation, a model was developed that integrates multiple meteorological parameters indicative of both thermal and mechanical turbulence. Specifically, the algorithm utilizes vertical profiles of wind speed, vertical wind speed, potential temperature, TKE, and their vertical gradients to diagnose RL height during nighttime hours (from 9 PM to 5 AM). The final algorithm is based on composite metrics that capture the combined influence of dynamic and thermodynamic processes. A series of logical conditions and thresholds are then applied to estimate the RL height, aiming to robustly identify the residual layer structure under various atmospheric conditions. The theoretical model is defined as:

$$F = |\mathbf{U}| \times \left| \frac{d\theta}{dz} - \min\left(\frac{d\theta}{dz}\right) \right| \times |\Delta\theta| \times \left| \frac{dw}{dz} - \min\left(\frac{dw}{dz}\right) \right| \times |\mathrm{TKE} - \max(\mathrm{TKE})| \times \left| \frac{d\mathrm{TKE}}{dz} - \max\left(\frac{d\mathrm{TKE}}{dz}\right) \right|$$

where |U| is wind speed, dθ/dz is vertical gradient of potential temperature, |Δθ| is deviation of potential temperature from a reference profile, dw/dz is vertical gradient of vertical wind speed, TKE is Turbulence Kinetic Energy (TKE), and d(TKE)/dz is vertical gradient of TKE.

The workflow begins by calculating *F* parameter across all vertical levels within the specified range. An initial estimate of the residual layer height denoted is assigned to the level where the *F* parameter attains its maximum value. To assess whether elevated RL values are physically justified, the algorithm incorporates a wind speed threshold of 6 m/s. This threshold is grounded in the Pasquill-Gifford stability classification, which identifies wind speeds exceeding 6 m/s as indicative of neutral atmospheric conditions, irrespective of diurnal variation. By applying this criterion, the algorithm avoids misclassifying strong mechanical turbulence as stable classification, thereby preventing the underestimation of the RL height under dynamically active conditions.

When RL height exceeds 1700 m, the algorithm evaluates whether wind speeds below 1000 m surpass the 6 m/s threshold. If this condition is not met, it re-examines *F* parameter within two focused altitude bands: 300–1200 m and 1200–1500 m. If the cumulative param values in these sub-ranges remain low, wind-based RL height is reassigned based on alternative indicators, such as the mean of reference local minima and maxima in the potential temperature profile. The algorithm further checks that if the maximum potential temperature gradient occurs below 500 m, this level is prioritized as the most probable RL height.

When wind speeds remain below 6 m/s throughout the lowest 500 m, the algorithm defaults to turbulence-based criteria. In such cases, it identifies the altitude corresponding to the minimum gradient of turbulent kinetic energy (TKE) as the primary fallback estimate. If this criterion is insufficient, the level of the maximum potential temperature gradient is used as an alternative. Upon completing all adjustments and cross-validations, the reliable estimate is finalized as of the RL height. This approach offers several advantages.

The cumulative parameterization ensures that only layers exhibiting strong and coherent signals are selected as candidate boundary layers. The incorporation of multiple independent diagnostics, such as wind speed, potential temperature gradients, and turbulence metrics, enhances the robustness of the estimation. Additionally, the use of dynamic thresholds mitigates the risk of both overestimating and underestimating the RL height under variable atmospheric conditions. Over the water, the algorithm can capture the residual layer height, which is below 150 meters.

Despite the strengths of the developed RL model, the method has limitations. The detection of elevated RL heights, particularly those exceeding 1700 m, can be problematic when the profiles lack clear evidence of strong turbulence or wind shear to support such depths. Another limitation arises during periods when the vertical structure of key variables (e.g., wind speed, potential temperature, TKE) appears similar across consecutive time steps (e.g., hours 2, 3, and 4), yet the observed RL heights vary significantly. In these cases, the algorithm cannot resolve the discrepancies, as the input signals do not contain sufficient information to account for the observed variability, an issue largely attributable to uncertainties in the simulated model variables.

5.2.1. Evaluation of the developed RL model

To perform a quantitative assessment of our method for diagnosing the RL and to assess the performance of the theoretical model relative to the empirical model, the RL heights estimated by the theoretical model were compared with observed planetary boundary layer heights at the LaPorte station. The scatter plot in **Figure 34** illustrates the performance of the theoretical model, showing a moderate positive correlation (R = 0.38) between estimated residual layer heights and observed PBLH over LaPorte during August and September 2021.

The theoretical model, applied over both August and September 2021, showed a stronger correlation with observations (R = 0.38) compared to the empirical model (R = 0.27), which was only evaluated for September 2021. The theoretical model exhibited a regression slope of 0.47 and an intercept of 0.39, indicating a more dynamic response to variability in observed PBLH.

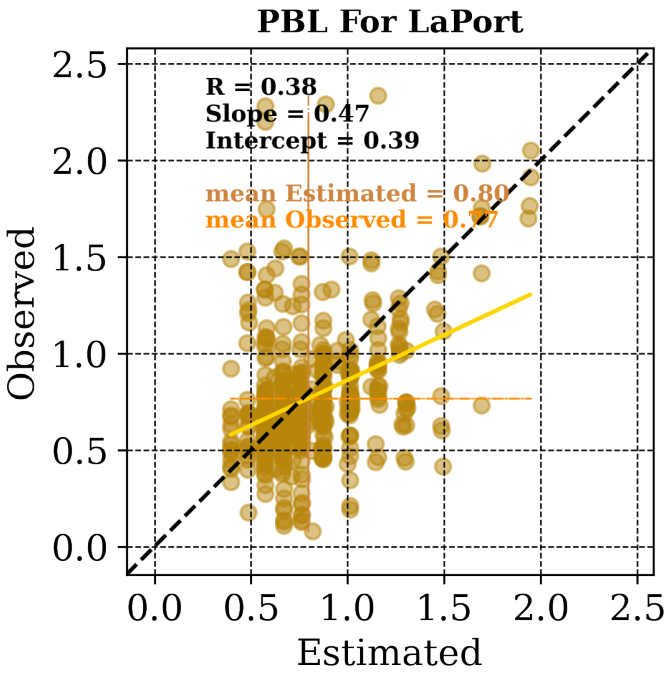


Figure 34. Scatter plot comparing theoretical model–estimated residual layer heights with observed PBL heights at the LaPorte station for August and September 2021. The black dashed line represents the 1:1 line, while the solid gold line shows the linear regression fit (R = 0.38, Slope = 0.47, Intercept = 0.39).

The mean estimated RL from the theoretical model was 0.80 km, compared to a mean observed PBLH of 0.77 km. In contrast, the empirical model yielded nearly identical mean values for modeled and observed heights (0.86 km and 0.85 km, respectively) with zero mean bias, but its predictions were more clustered and less sensitive to higher PBLH values. Overall, the theoretical model better captures the structure and variability of the residual layer, particularly during nocturnal conditions in late summer, suggesting improved capability over the empirical approach. The theoretical model, applied over both August and September 2021, showed a stronger correlation with observations (R = 0.38) compared to the empirical model (R = 0.27), which was only evaluated for September 2021. The theoretical model exhibited a regression slope of 0.47 and an intercept of 0.39, indicating a more dynamic response to variability in observed PBLH. The mean estimated RLH from the theoretical model was 0.80 km, compared to a mean observed PBLH of 0.77 km. In contrast, the empirical model yielded nearly identical mean values for modeled and observed heights (0.86 km and 0.85 km, respectively) with zero mean bias, but its predictions were more clustered and less sensitive to higher PBLH values. Overall, the theoretical model better captures the structure and variability of the residual layer, particularly during nocturnal conditions in late summer, suggesting improved capability over the empirical approach.

Figure 35 compares theoretical model estimates (red line) with observed RL heights (blue bars) at LaPorte from August 31-September 21-28, 2021. The model generally captures the observed trends and magnitudes, showing reasonable agreement in both periods. Some underestimations

occur during high RL events, especially around September 22-24, likely due to local meteorological effects not captured by the model. The additional plot for September 7-14, 2021, also shows that the model follows the general trend but underestimates several peaks in observed RL heights. Overall, the theoretical model provides a consistent and reasonable estimation of RL height.

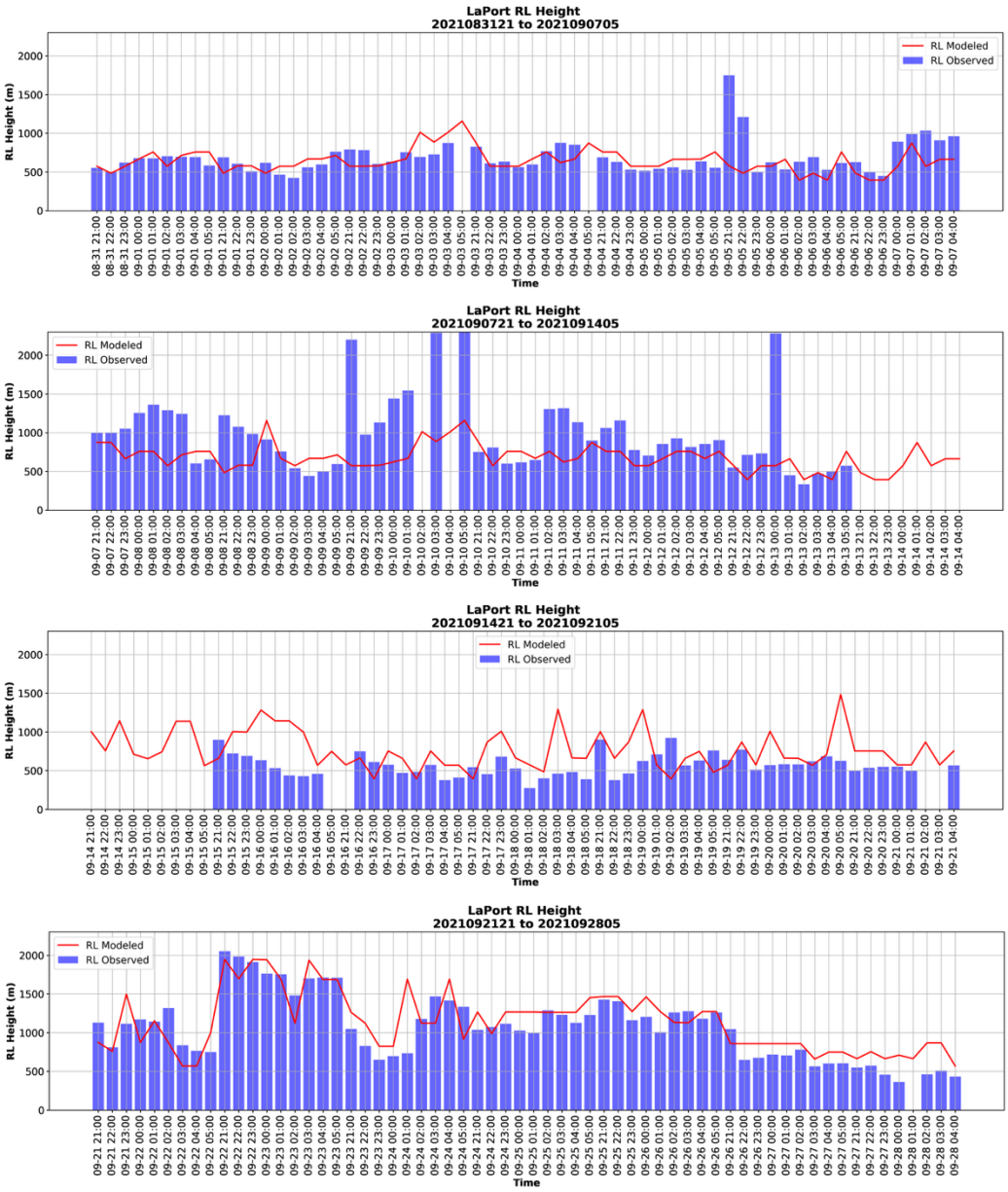


Figure 35. Comparison of Theoretical Modeled and Observed Residual Layer Heights at LaPorte, from August 31, 2021 to September 28, 2021.

5.2.2. Spatial variation of the residual layer

Figure 36 shows the spatial distribution of RL height calculated by the theoretical model from WRF [Base], [PBLH23], [PBLH24], and [PBLH25] from 9:00 PM to 5:00 AM for September 8–

9, 2021 (left), September 8–9, 2022 (middle), and September 8–9, 2023 (right). Filled circles represent the ceilometer-observed mean PBLH during this period at LaPorte, Bay, and Gulf stations.

The results indicate that the theoretical model effectively distinguishes between land and open water, consistently producing lower RL heights over water. The model also closely reproduces the observed PBLH values for [Base], [PBLH23], [PBLH24], and [PBLH25] at the selected sites. However, capturing the observed PBLH in [PBLH23] and [PBLH25] for the year 2022 remains challenging. Compared to the empirical model, the theoretical model shows better agreement with observations and more accurately captures spatial differences, particularly the contrast between RL heights over land and water. Overall, the theoretical model demonstrates stronger performance in representing residual layer structure across the domain.

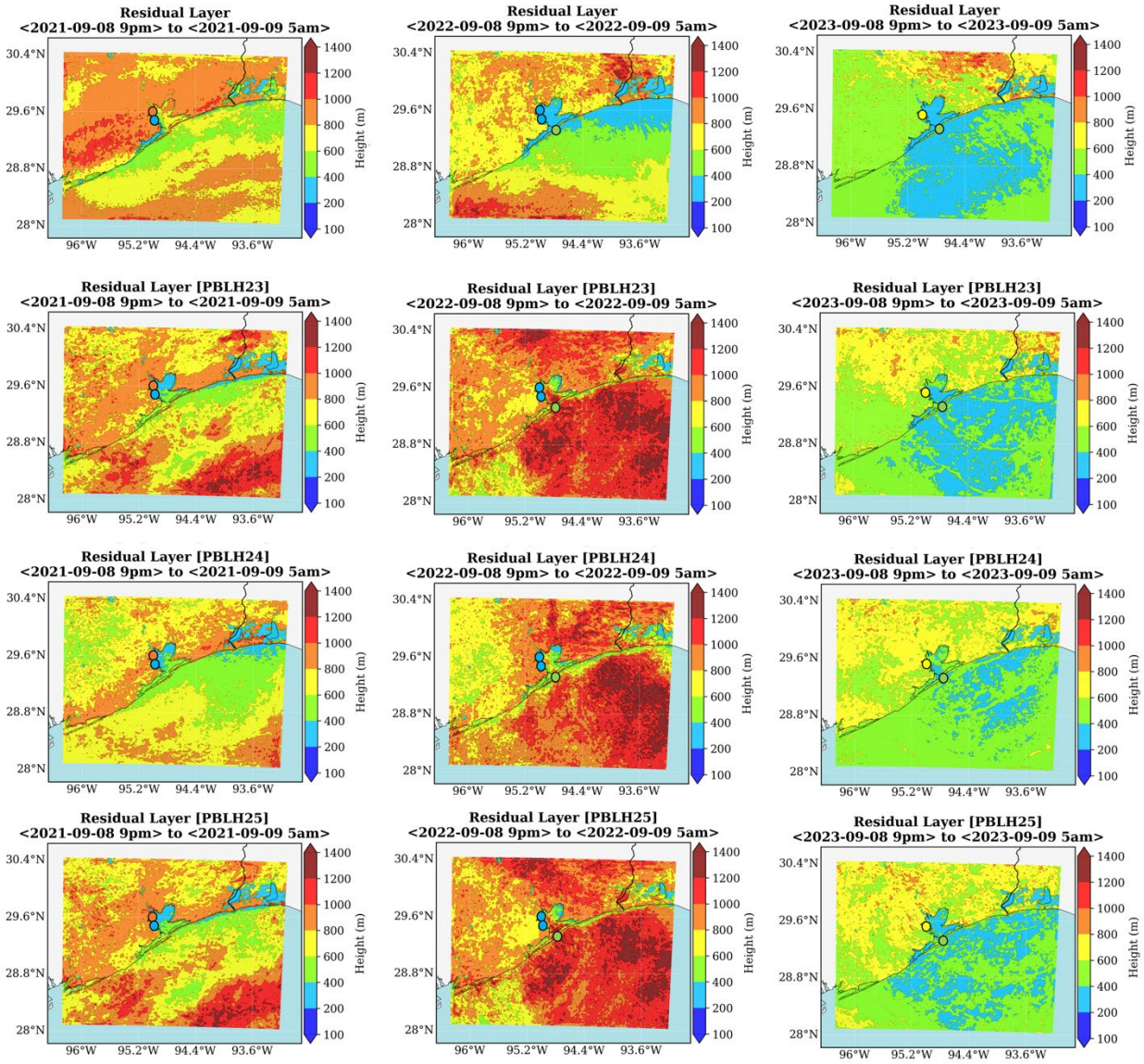


Figure 36. Spatial distribution of residual layer height calculated the theoretical model from WRF [Base] (first row), [PBLH23] (second row), [PBLH24] (third row), and [PBLH25] (fourth

row) over the land, Galveston Bay, and Gulf from 9:00 PM to 5:00 AM (local time) on September 8-9 2021 for the left column, September 8-9 2022 for the middle column, and September 8-9 2023 for the right column. Dot overlaid on the map is the ceilometer observed mean PBLH from the LaPorte, Bay, and Gulf measurements.

5.3. Summary

Estimating the PBLH is important in understanding atmospheric processes, air quality, weather forecasting, and climate modeling. The WRF base model diagnoses only a single PBLH, which typically represents the shallow stable boundary layer (SBL) only at night and fails to capture the overlying residual layer (RL). To address this limitation, an empirical method was developed to calculate RL height using vertical gradients of potential temperature ($\partial\theta/\partial z$). The diagnosed RL heights showed a good agreement with the observed PBLH at LaPorte in September 2021, especially at night where the base model fails to capture the layered boundary structures, demonstrating the method's utility. The RL height estimation approach also performed well for daytime PBLH, confirming its broader applicability in capturing layered boundary structures. The RL estimation method was applied across the full model domain for both the WRF base model and the three best-performing perturbation simulations. Comparison with field campaign observations from TAQ1, TAQ2, and TAQ3 shows that the method successfully captured the residual layer structure in the base simulation across all eight selected days. PBLH23, PBLH24, and PBLH25 also performed well on most days, with the exception of September 8–9, 2022, where the calculated RL was overestimated. Overall, our analysis shows that potential temperature gradient-based RL estimation is capable of identifying RL at night. However, the method is found sensitive to the different cumulus cloud schemes in the model.

To address the limitations of the empirical model, a theoretical model was developed to improve the estimation of RL height using physical principles and vertical meteorological profiles. The developed RL model estimates nighttime RL height (9 PM–5 AM) by integrating thermal and mechanical turbulence indicators, using vertical profiles of wind speed, vertical wind speed, potential temperature, TKE, and their gradients. The theoretical model employs a composite parameter to identify the level of maximum turbulent activity and applies a 6 m/s wind threshold to avoid underestimating RL under neutral or mechanically driven conditions. Additional checks refine the estimate, including gradient-based logic, fallback criteria using TKE and temperature structure, and tailored rules for high RL cases (>1700 m). The approach captures RL heights over both land and water (even <150 m), offering a robust, multi-variable method that improves detection accuracy. However, its performance can be limited when vertical profiles lack strong gradients or show little variability over time despite changes in observed RL height, often due to uncertainties in model-simulated inputs. The theoretical model showed a mean estimated RL height of 0.80 km, closely matching the observed mean of 0.77 km and demonstrating better sensitivity to variability, especially during nocturnal conditions in late summer. Compared to the empirical model, which had minimal bias but produced more clustered predictions, the theoretical model showed a stronger correlation with observations (R = 0.38 vs. 0.27) and captured the structure and dynamics of the residual layer more effectively.

6. Effect of improved PBL on CAMx ozone prediction and source attribution

This section evaluates the influence of improved PBL representation on ozone prediction and source attribution over the Houston-Galveston-Brazoria (HGB) region using the CAMx model.

The assessment begins with evaluating the performance of the CAMx base simulation, which is driven by meteorological inputs from the WRF v4.6.0 base configuration. This provides a benchmark for comparison against ozone observations and prior model versions. Next, the impact of the three selected WRF perturbation simulations, PBLH23, PBLH24, and PBLH25, on CAMx ozone predictions is examined to assess how changes in boundary layer dynamics affect photochemistry. Finally, background simulations are conducted using these perturbation cases to isolate the contributions of anthropogenic and background ozone sources, providing insight into the relative influence of local emissions versus transported pollution.

6.1. CAMx Base Model Evaluation

The project team previously used CAMx version 7.10 in conjunction with meteorological inputs from WRF version 3.9.1.1 to simulate ozone photochemistry for the TAQ1 and TAQ2 campaigns (Li et al., 2023; Soleimanian et al., 2023). The present work used CAMx version 7.31, driven by meteorological fields from the more recent WRF version 4.6.0. The assessment first focused on how version updates in both the meteorological and photochemical models affect ozone predictions over the HGB region.

Figure 37 compares the time series of ozone concentrations simulated by CAMx v7.10 and v7.31 against boat observations for September 8–10, 2021, over Galveston Bay. CAMx v7.10 (green line), using WRF v3.9.1.1, produces higher peak ozone levels in the afternoon but overestimates morning and evening ozone, with a mean bias of +1.34 ppbv and a correlation coefficient (R) of 0.60 compared to the observations. In contrast, CAMx v7.31 (red line), using WRF v4.6.0, better captures the timing of ozone variability and the morning and evening ozone concentrations, with a slightly stronger correlation (R=0.64), but underestimates peak ozone levels and has a negative mean bias of -4.18 ppbv. These results indicate that the updated WRF-CAMx system improves temporal representation but underpredicts peak ozone under certain conditions.

Even when using the same physical configuration, the updates between WRF v3.9.1.1/CAMx v7.10 and WRF v4.6.0/CAMx v7.31 influenced the model performance. In WRF, key improvements included bug fixes to the MYNN PBL scheme, resolving issues with wind speed, TKE, and stability, yielding more reliable meteorological inputs and improved offshore boundary layer dynamics (Skamarock et al., 2021). Similarly, CAMx v7.31 incorporated substantial code enhancements, such as improved memory handling, streamlined MPI communication, and updates to the CB6r5 chemistry mechanism, that better support modern parallel computing architectures. These updates provided a more stable meteorological driver from WRF and more robust data processing within CAMx, which contributed to more consistent simulations and notably improved ozone prediction accuracy.

The subsequent analysis and results in the report are based on CAMx v7.31, driven by meteorology from WRF v4.6.0, and the model version numbers will not be specified for simplicity. CAMx simulations driven by the WRF base model are referred to as the CAMx base model, and those driven by the selected PBL perturbation simulations as CAMx-PBLH23 and CAMx-PBLH24, respectively.

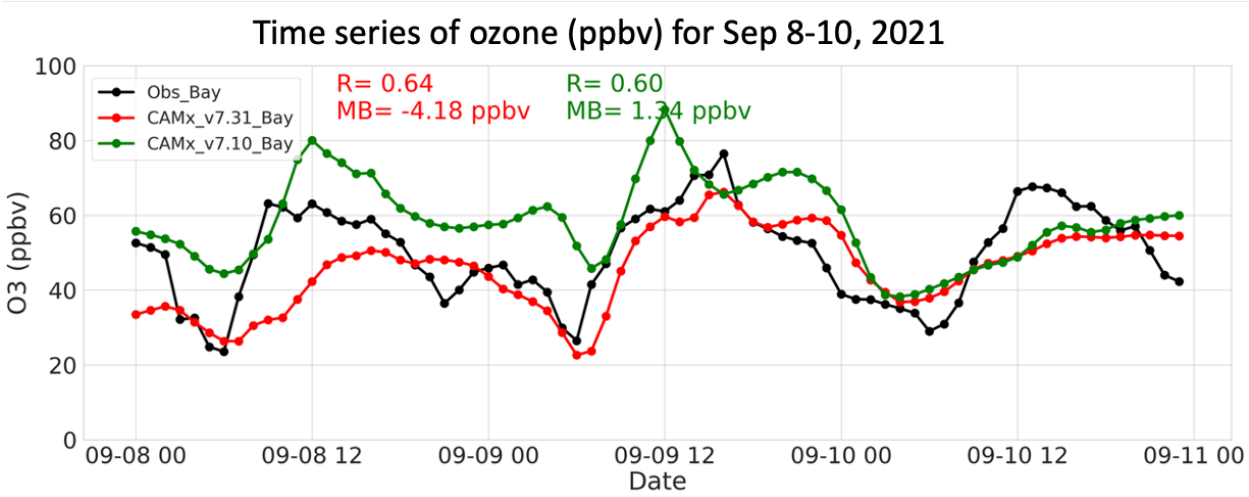


Figure 37. Time series of ozone (ppbv) for September 8-10, 2021, over Galveston Bay. Black lines represent observed ozone, red lines show CAMx v7.31 simulations, and green lines show CAMx v7.10 simulations. Statistical metrics (correlation, mean bias) are shown for each model version.

The evaluation of CAMx base model performance on ozone predictions focused on September 8-10 for 2021, 2022, and 2023. These dates represent both high ozone and clean periods, allowing us to assess model performance under varying air quality conditions. These dates also aligned with periods of available mobile boat observations during the TRACER-AQ campaigns. In 2021, the boat sampling was active primarily within Galveston Bay, so the evaluation for that year is limited to the Bay region. In 2022 and 2023, the observational coverage expanded to include both Galveston Bay and the Gulf, providing a broader spatial domain for comparison. This contrast in measurement coverage between the three years allows for a more complete assessment of CAMx performance over the coastal and marine environment.

Figure 38 presents the spatial distribution of observed and CAMx base-simulated ozone concentrations, along with the model-observation difference, for September 8-10, 2021, 2022, and 2022, over Galveston Bay. The plots include only daytime (07:00 to 19:00 CDT) mobile measurements to emphasize photochemically active periods and exclude nighttime docked observations, which are more influenced by land-based processes.

Over Galveston Bay in 2021, CAMx captured the general spatial gradient of ozone but tended to underestimate concentrations across most of the Bay, with a mean bias of -12.73 ppbv. The underestimation appears more pronounced during morning hours in the southern Bay, while in the afternoon, the model slightly overestimated ozone in the northern Bay near the Houston Ship Channel (HSC). In 2022, CAMx showed regionally varying performance in the spatial variability of ozone across Galveston Bay, with a mean bias of -9.80 ppbv. Most of the sampling in the southern Bay occurred during afternoon hours, where the model tended to underestimate ozone in this more marine-influenced region. In northwestern Bay and around the HSC, the model often overestimated ozone in the afternoon, similar to the 2021 pattern. In 2023, CAMx generally overestimated ozone throughout the region, although slight underestimations were observed in the morning over the southern HSC, particularly during morning sampling hours near the Gulf. These patterns highlight spatial and temporal biases in the base model that motivate the use of improved meteorological inputs from PBL perturbation simulations.

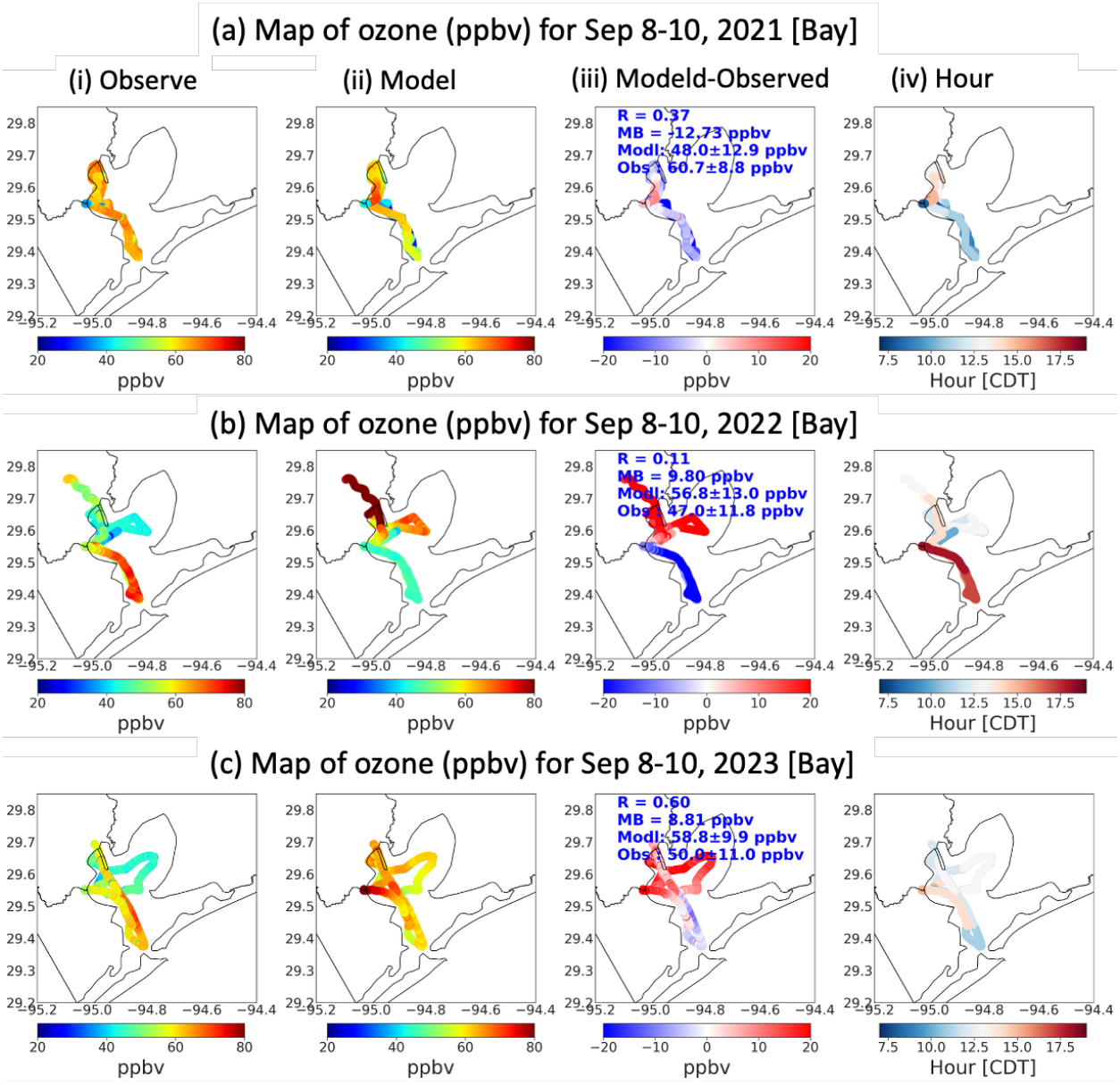


Figure 38. (a) Map of the observed, CAMx base modeled, and modeled-observed ozone (ppbv) over Galveston Bay for September 8-10, 2021, along with the time/hour when the boat was mobile. (b) Same as (a), but for September 8-10, 2022. (c) Same as (b) but for September 8-10, 2023. The plot shows daytime (7:00 CDT – 19:00 CDT) mobile data only.

Figure 39 shows the spatial distribution of observed and CAMx base-simulated ozone concentrations, along with the model-observation difference, for September 8-10, 2021, and 2022, over the Gulf. In 2022, CAMx slightly underestimated ozone in the northern Gulf and near the Bay, where the boat was mobile mostly around noon. In the morning, when the boat sampled in the southern Gulf (farther offshore), the model slightly overestimated ozone concentrations. While the magnitude of the bias was smaller than over Galveston Bay, the model struggled to capture the observed variability, especially during afternoon sampling periods. This indicates

potential limitations in representing ozone dynamics over open water under the current base model configuration.

In 2023, the boat was primarily mobile in the northeastern part of the Gulf, mostly near the coast. During the morning hours, when the boat conducted measurements before noon, the CAMx base simulation generally overestimated ozone concentrations. However, in the late afternoon, particularly after 3-4 PM, the model underestimated ozone levels relative to observations. These time-dependent biases suggest that the base model struggles to represent the diurnal evolution of ozone over the Gulf, especially under changing wind and mixing conditions near the land-water interface.

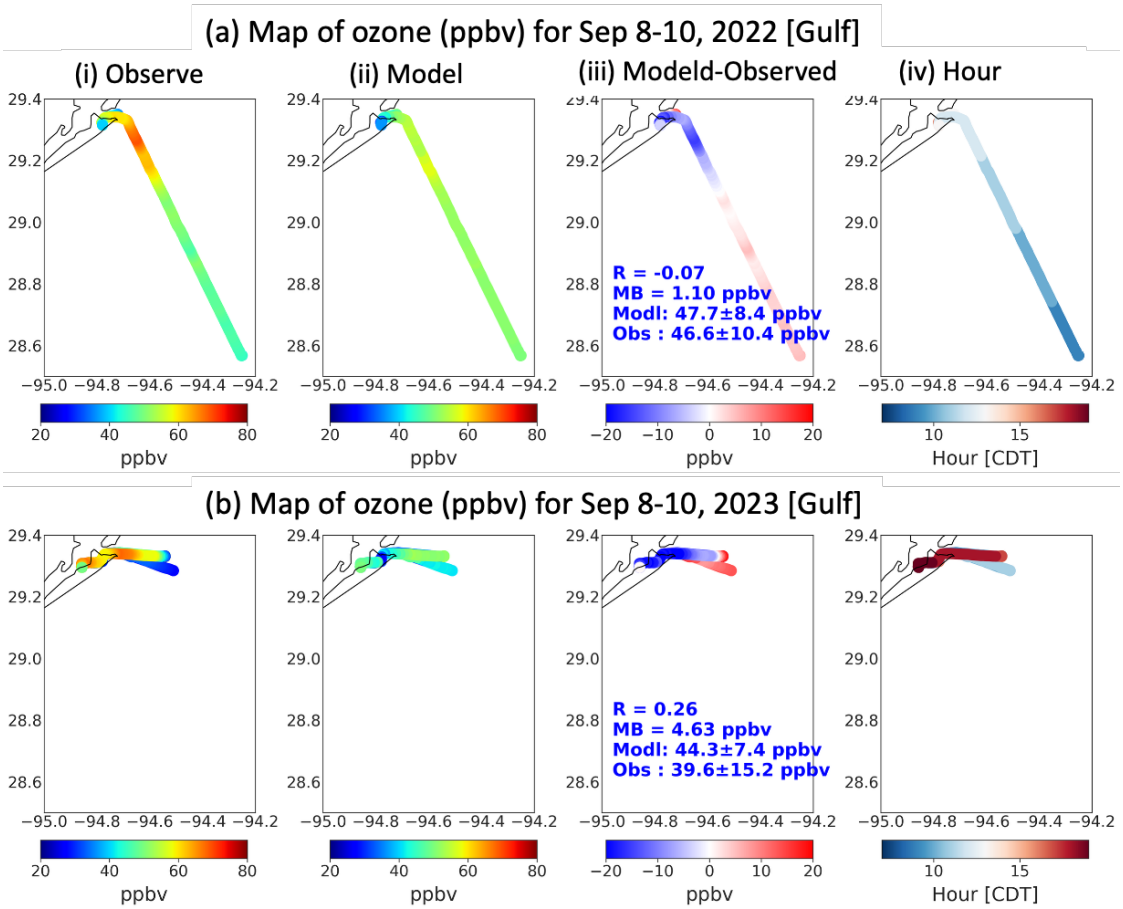


Figure 39. (a) Map of the observed, CAMx base modeled, and modeled-observed ozone (ppbv) over the Gulf for September 8-10, 2022, along with the time/hour when the boat was mobile. (b) Same as (a), but for September 8-10, 2023. The plot shows daytime (7:00 CDT – 19:00 CDT) mobile data only.

6.2. Effects of Improved PBL on CAMx Ozone Prediction

To evaluate the effect of improved PBL representation on ozone prediction, CAMx simulations were conducted using meteorology from the WRF base model as well as three best-selected PBL perturbation configurations: PBLH23, PBLH24, and PBLH25. These simulations were performed for September 8-10 in both 2021, 2022, and 2023 and were compared to the CAMx

base model as well as observed ozone concentrations from the TRACER-AQ mobile platform data over Galveston Bay and the Gulf.

Figure 40 shows the average surface ozone concentrations simulated by the Base, PBLH23, PBLH24, and PBLH25 configurations for the selected periods. In 2021 and 2022, CAMx simulations driven by PBLH23, PBLH24, and PBLH25 produced higher average ozone surface concentrations compared to the base simulation, particularly over Galveston Bay and coastal regions. This increase in modeled ozone is associated with the deeper and more realistic PBL heights produced by PBLH23, PBLH24, and PBLH25. Enhanced vertical mixing over the water would both bring down ozone precursors from aloft (i.e., the sea breeze return flow), leading to increased photochemical production of ozone, and increase the entrainment of free-tropospheric ozone into the boundary layer. All three perturbation cases used the NSAS cumulus scheme and the 1-D ocean mixed layer model. PBLH24 includes the UCM, which contributes to ozone increases over land, while PBLH25 is distinguished by an additional perturbation to the surface roughness length over water, enhancing vertical exchange specifically over offshore areas. PBLH24 resulted in the most widespread ozone enhancement, with average increases up to ~10 ppbv across the domain. Both PBLH23 and PBLH25, which exclude the UCM, showed ozone increases primarily in offshore regions, with PBLH25 producing slightly higher enhancements than PBLH23, suggesting that the roughness length modification in PBLH25 further amplifies ozone production over water compared to PBLH23.

In 2023, the CAMx simulations driven by the perturbation configurations showed only minor differences in ozone concentrations compared to the base. This limited change is consistent with the relatively low WRF v4.6.0 simulated PBL heights during this period, as discussed in Section 4. Unlike in 2021 and 2022, when the perturbation simulations substantially increased the PBLH and enhanced vertical mixing, the effect of the perturbations on PBLH in 2023 was small. As a result, the associated changes in vertical transport of ozone and precursors were minimal, and CAMx ozone predictions remained largely unchanged. This suggests that the effectiveness of PBLH perturbations in influencing ozone depends strongly on background meteorological conditions, particularly the extent of vertical development in the boundary layer, and that limited PBL growth can constrain the impact of even physically enhanced configurations.

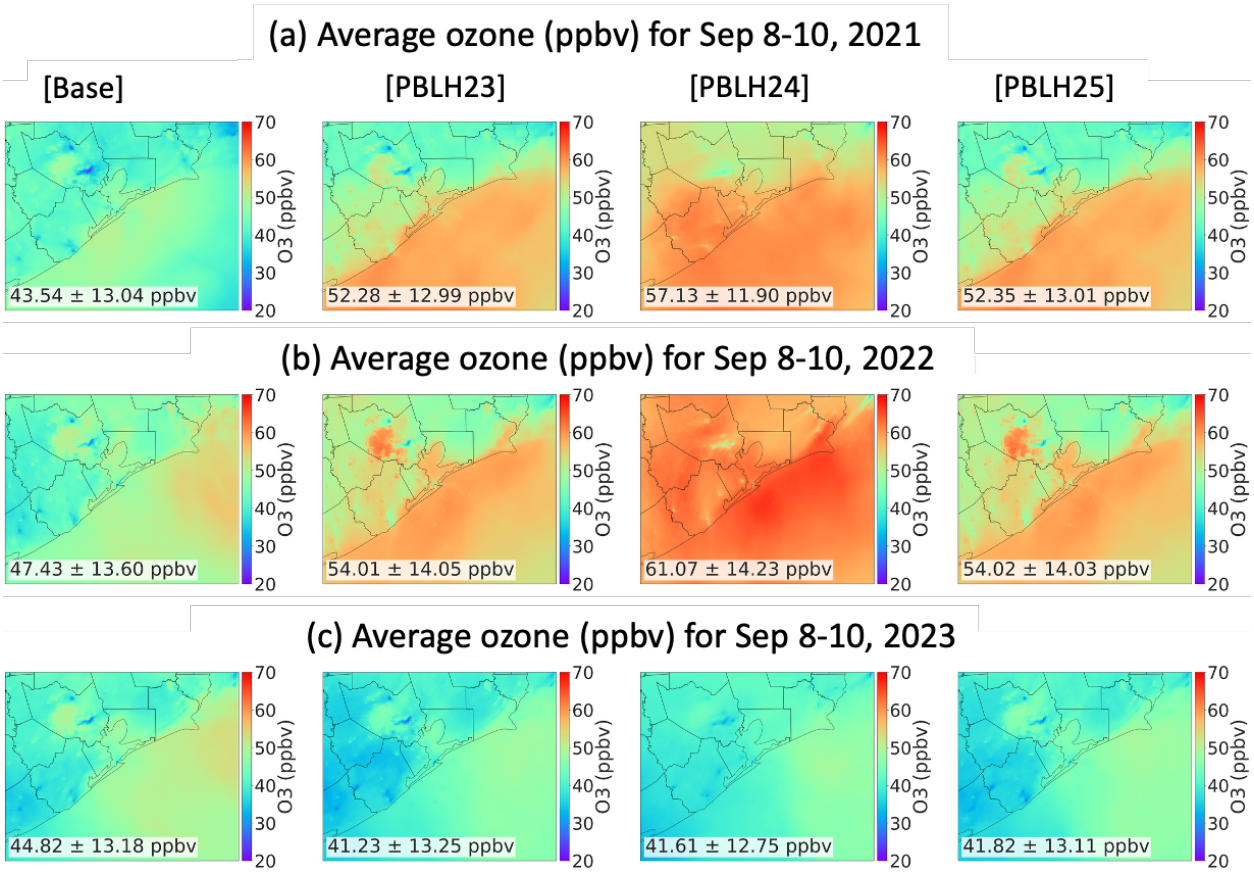


Figure 40. Average ozone (ppbv) surface concentration for CAMx Base, PBLH23, PBLH24, and PBLH25, for (a) September 8-10, 2021, (b) September 8-10, 2022, and (c) September 8-10, 2023.

Figure 41 compares the time series of observed and simulated ozone concentrations at the offshore sampling locations during September 8-10, 2021, 2022, and 2023 for the base, PBLH23, PBLH24, and PBLH25 simulations. In 2021, consistent with the spatial variation shown in **Figure 40**, PBLH23, PBLH24, and PBLH25 simulated higher ozone concentrations at offshore locations compared to the base, effectively eliminating the base model's negative bias over Galveston Bay (**Figure 41a**). In 2022 (**Figure 41b**), while the base model already exhibited a positive bias of ~7 ppbv, PBLH23, PBLH24, and PBLH25 further increased ozone concentrations to ~14 ppbv, ~23 ppbv, and ~14 ppbv, respectively. Despite the amplified daytime bias, PBLH23 and PBLH25 captured the early morning ozone minimum more accurately, predicting 10–15 ppbv lower ozone than the base at ~6:00 AM on September 8 and 9, aligning more closely with observed values. (**Figure 41b)**. Across the Bay, PBLH23 consistently struck the best balance between reducing early morning overestimation and improving afternoon peaks, while PBLH24 tended to overestimate throughout the day. PBLH25 generally performed between PBLH23 and PBLH24 in terms of bias, but importantly, it showed the highest correlation with observations among all simulations in 2022, indicating that it best captured the spatiotemporal variability in ozone. In 2023 (**Figure 41**c), differences between the base and perturbation simulations were minor. All three perturbation cases - PBLH23, PBLH24, and PBLH25 - produced ozone levels similar to the base, with only subtle improvements. This muted

response reflects the low PBLH and limited vertical mixing observed in 2023 over the Bay, which constrained the impact of boundary layer enhancements on surface ozone.

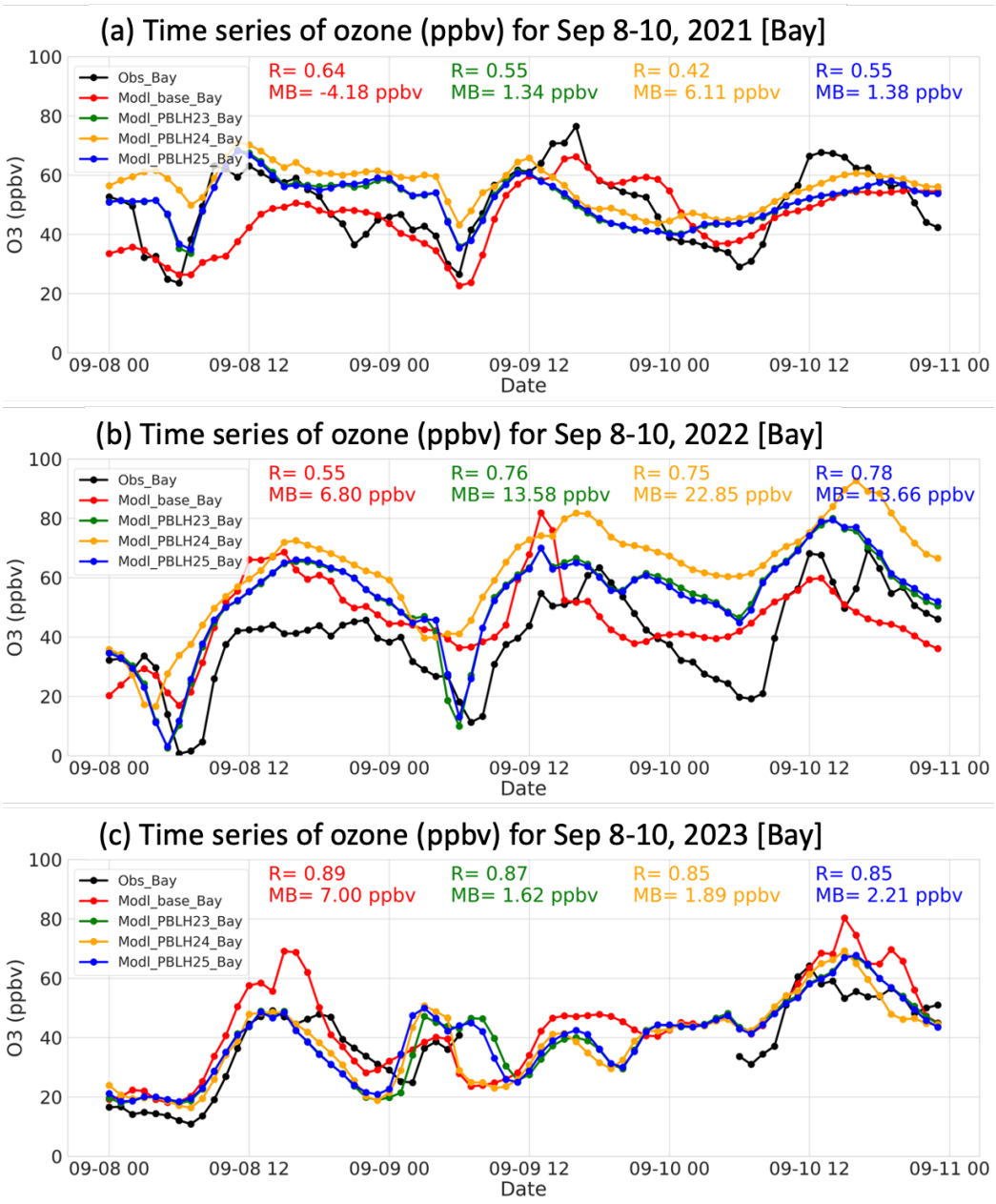


Figure 41. (a) Time series of ozone (ppbv) for September 8-10, 2021, over Galveston Bay at the ship tracks where measurements were made, (b) Same as (a) but for September 8-10, 2022, and (c) Same as (a) but for September 8-10, 2023. Black lines represent observed ozone, red lines show base simulation, green lines show PBLH23, orange lines show PBLH24, and blue lines show PBLH25. Statistical metrics (correlation, mean bias) are shown for each plot.

Figure 42 presents the time series of observed and simulated ozone concentrations over the Gulf during September 8-10 for 2022 (**Figure 42**a) and 2023 (**Figure 42**b), comparing the base simulation with PBLH23, PBLH24, and PBLH25. In 2022, the base simulation significantly

underestimated ozone throughout the period, with a mean bias of –9.00 ppbv and low correlation (R = 0.28), failing to capture the observed variability. All three perturbation simulations, PBLH23, PBLH24, and PBLH25, substantially improved ozone predictions. PBLH24 produced the largest increase in ozone (~14 ppbv mean bias), often overestimating observed values, while PBLH23 and PBLH25 provided more balanced improvements. Among them, PBLH25 showed the best overall agreement with observations, achieving the highest correlation (R = 0.66) and a modest mean bias (+3.41 ppbv). In 2023 (**Figure 42**b), the differences among the simulations were smaller. All configurations, including the base, showed similar performance, with lower biases and higher correlations. The perturbation simulations still slightly improved correlation, with PBLH25 again performing best (R = 0.66, MB = 0.59 ppbv). This limited impact reflects the relatively shallow PBLH observed in 2023, reducing the influence of boundary layer processes on ozone transport and production. As a result, the benefit of the perturbations is less pronounced compared to 2022.

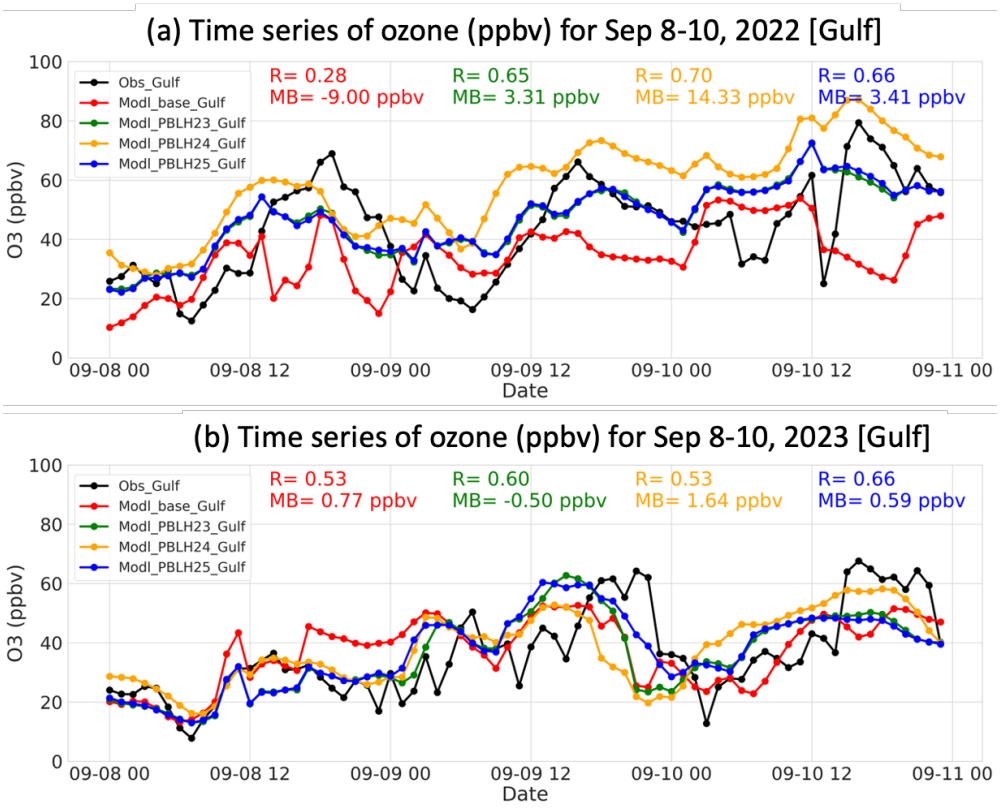


Figure 42. (a) Time series of ozone (ppbv) for September 8-10, 2022, over the Gulf at the ship tracks where measurements were made, (b) Same as (a) but for September 8-10, 2023. Black lines represent observed ozone, red lines show base simulation, green lines show PBLH23, orange lines show PBLH24, and blue lines show PBLH25. Statistical metrics (correlation, mean bias) are shown for each plot.

Ozone simulated by the CAMx Base, PBLH23, PBLH24, and PBLH25 cases was further evaluated against the TCEQ continuous ambient monitoring station (CAMS) sites to assess the effects of boundary layer perturbations over land. **Figure 43** compares the time series of observed and simulated ozone concentrations during September 8–10, 2021, 2022, and 2023.

In 2021 (**Figure 43a**), the base simulation slightly overestimated ozone by ~1 ppbv on average, with a correlation of 0.82. The perturbation cases PBLH23 and PBLH25 also showed positive biases (~6 ppbv), but with slightly higher correlations (R = 0.83). Since PBLH25 applies the same PBLH adjustment as PBLH23 but with an added surface roughness perturbation over the water, the two cases performed almost identically over land, indicating that the roughness perturbation over the water had less impact over the land. By contrast, PBLH24 produced a much larger positive bias (~12 ppbv) and a weaker correlation (R = 0.63), reflecting a systematic overestimation of ozone throughout the diurnal cycle. In 2022 (**Figure 43b**), compared to 2021, all four simulations exhibited a stronger positive bias. The Base case showed the lowest mean bias (~13 ppbv), whereas the perturbation simulations further amplified the overprediction, with mean biases of ~21 ppbv for PBLH23 and PBLH25 and ~27 ppbv for PBLH24. However, all perturbation cases showed improved correlations (~0.88) compared to the Base (0.83), suggesting that while they overpredicted the magnitude, they captured the temporal variability more accurately. In 2023 (**Figure 43c**), compared to 2021 and 2022, mean biases were small across all simulations (ranging from -1 to 3 ppbv). The perturbation cases also showed improved correlations relative to the Base. Among them, PBLH25 achieved the lowest bias (0.49 ppbv) and the highest correlation (R = 0.89), indicating the best overall performance during this year.

In summary, considering bias, correlation, temporal, and spatial variability over Galveston Bay, the Gulf, and the Land, PBLH25 demonstrated the most consistent overall performance in capturing the ozone variation across years.

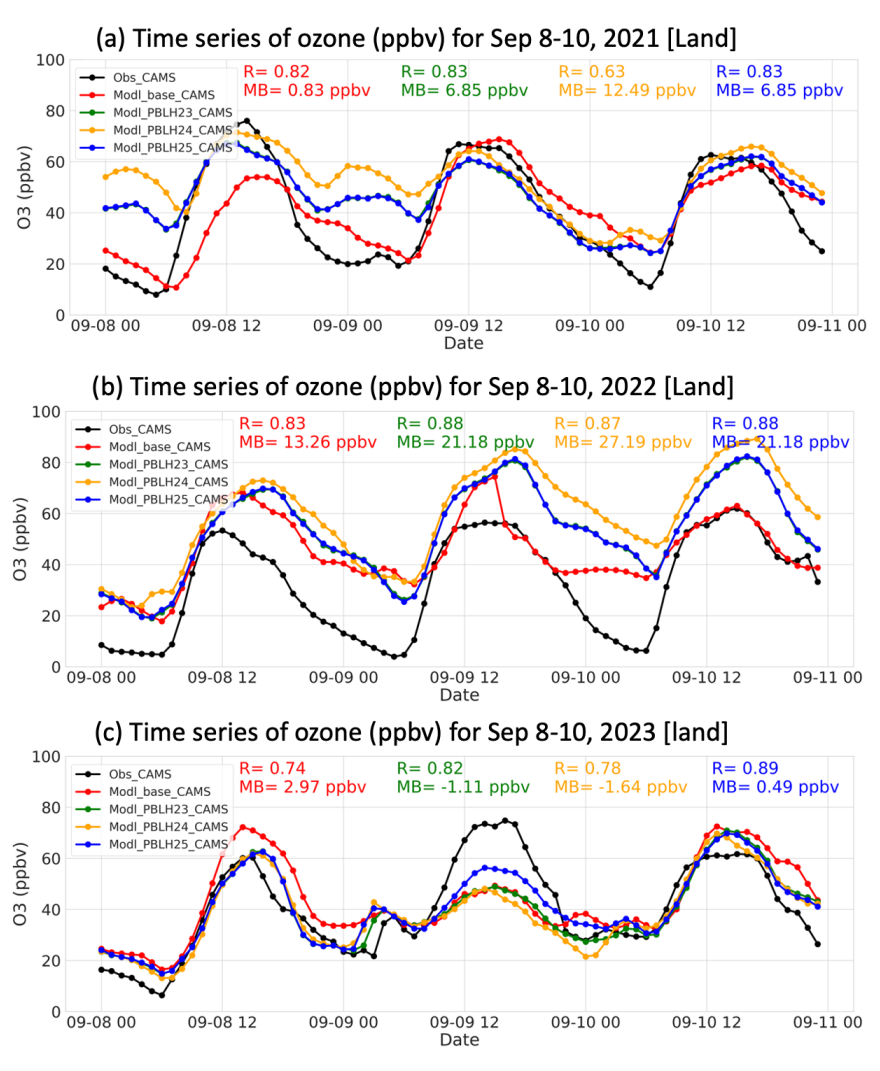


Figure 43. (a) Time series of ozone (ppbv) for September 8-10, 2021, over the land averaged across all the locations of available CAMS sites within the d03 domain; (b) Same as (a) but for September 8-10, 2022; (c) Same as (a) but for September 8-10, 2023. Black lines represent observed ozone, red lines show base simulation, green lines show PBLH23, orange lines show PBLH24, and blue lines show PBLH25. Statistical metrics (correlation, mean bias) are shown for each plot.

Further, CAMx simulations were conducted for the Base, PBLH23, and PBLH25 for the entire month of September 2022 to evaluate model performance between episodes and clean days against the CAMS site observations. The purpose was to ensure the perturbed model simulations, which were constrained by offshore observations, did not degrade the model performance over the land. The episode days were Sep 13-15 and Sep 20-26, 2022, where at least one monitors recorded hourly ozone higher than 70 ppbv. The other days were defined as clean days. **Figure 44** shows the mean surface ozone concentrations (ppbv) across the HGB region, with observation CAMS sites overlaid, separated into episode days (left), clean days (middle), and their difference (episode minus clean, right). All three models correctly predicted surface ozone concentrations to be consistently higher during episode days than clean days. The models generally overestimated

observed concentrations under both conditions. During the episode days, the Base produced a mean ozone of 54.43 ± 11.8 ppbv, compared to the observed 47.86 ± 21.1 ppbv, while clean days were lower (42.13 ± 9.5 ppbv vs. observed 37.78 ± 14.5 ppbv). PBLH23 further increased ozone during the episode days (56.09 ± 12.9 ppbv), while PBLH25 yielded comparable magnitudes as the base (55.4 ± 13.1 ppbv) but showed the strongest improvement in correlation with observations (r = 0.84 vs. 0.82 for PBLH23 and 0.80 for Base). During clean days, the models remained biased high, though correlations weakened (r = 0.63–0.66) relative to episode days. The episode-minus-clean maps show urban Houston has the largest ozone differences, where episode concentrations exceeded clean days by 15–20 ppbv. The magnitude and distribution of such differences were best predicted by PBLH25. Overall, this analysis confirms that PBLH23 and PBLH25 have similar, if not better, performance over the land as the CAMx base model. While all three CAMx simulations overpredicts ozone levels, the PBLH25 configuration best represents the observed variability and improves the episode-to-clean day contrast.

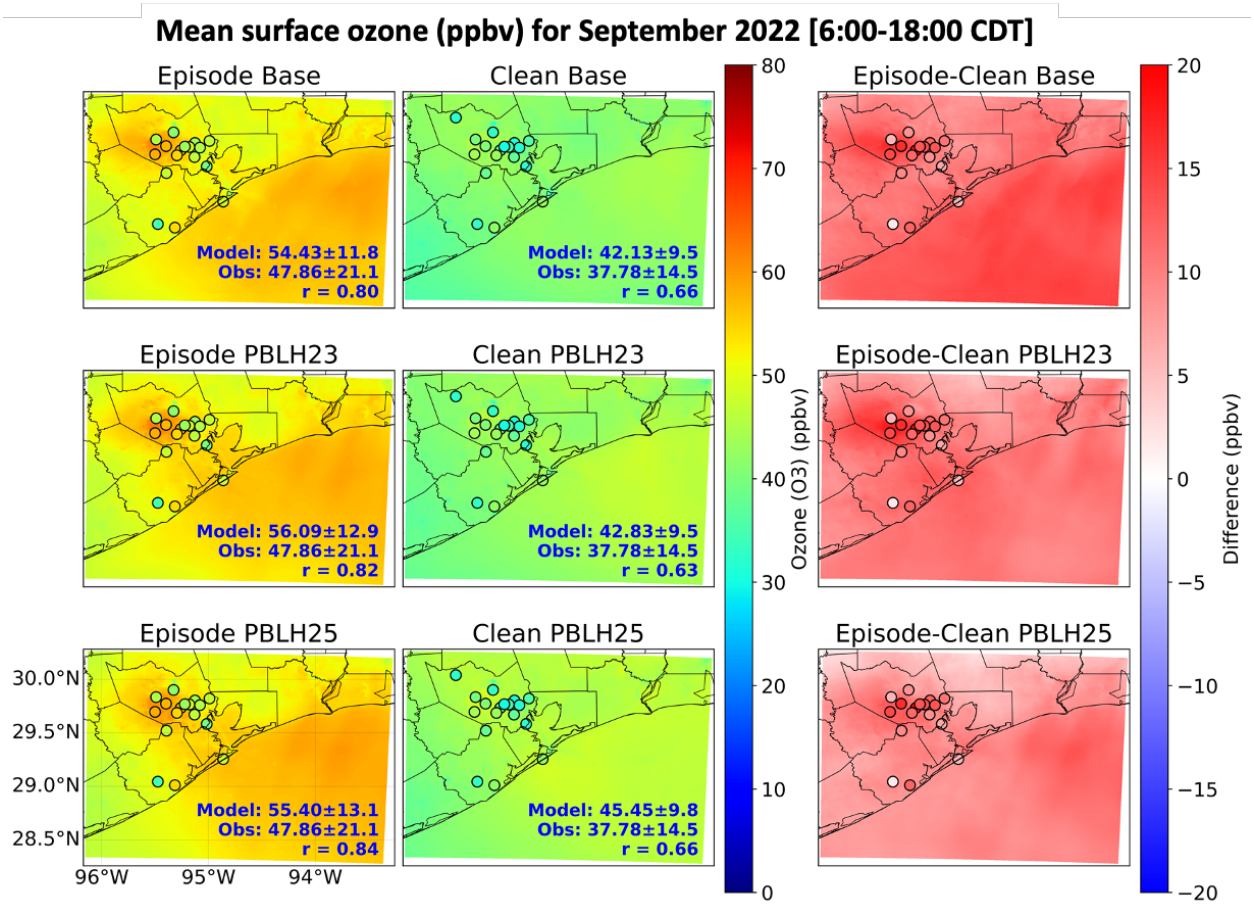


Figure 44. Maps mean ozone (ppbv) surface concentration for CAMx Base, PBLH23, and PBLH25 for September 2022 during ozone episodes (left), clean days (middle), and episode minus clean days (right). Dots overlaid on the map are the observed mean values from the CAMS sites.

6.3. Residual Layer Effects on Ozone

The diagnosed RL heights from Section 5 were implemented in CAMx to evaluate their effects on ozone over the HGB region. Between the two approaches, empirical (Section 5.1) and theoretical (Section 5.2), the theoretical method showed better agreement with observations. Therefore, RL heights calculated from the theoretical method were used in CAMx.

In this implementation, the original PBL heights from the WRF output were replaced with the diagnosed RL heights during nighttime hours (21:00–05:00 CDT). These modified PBL heights were then used in CAMx to calculate vertical diffusion based on k-theory, thereby controlling the extent of vertical mixing within the boundary layer. By incorporating RL information into the PBL representation, the model was expected to more realistically account for residual layer ozone aloft and its influence on nighttime surface ozone concentrations.

Figure 45 compares the nighttime ozone concentrations (21:00–05:00 CDT) for September 8–10, 2021, 2022, and 2023 across the Base, PBLH23, PBLH24, and PBLH25 simulations, with and without the residual layer (RL) implementation. The third row in each panel shows the spatial differences between the RL and non-RL cases. In 2021 (**Figure 45a**), the Base simulation produced a mean nighttime surface ozone of ~38 ppbv, while the perturbation cases increased ozone substantially, with PBLH23 and PBLH25 around 48 ppbv and PBLH24 exceeding 54 ppbv. With the RL implementation, nighttime surface ozone increased by ~0.3–1.1 ppbv across the domain, especially over coastal and inland regions. The largest increases occurred in the Base and PBLH23/PBLH25 cases, suggesting that RL mixing enhanced surface ozone by maintaining residual ozone aloft and redistributing it downward during stable nighttime conditions. PBLH25 showed little additional change because ozone levels were already high prior to the RL adjustment. In 2022 (**Figure 45b**), nighttime ozone levels were higher than in 2021. The Base averaged ~41 ppbv, while the perturbation cases produced 48–57 ppbv. Adding the RL increased surface ozone by ~0.2–1.1 ppbv, with the strongest enhancement in the Base, similar to 2021. In 2023 (**Figure 45c**), surface ozone concentrations were lower than in 2021 and 2022, averaging ~36–41 ppbv across cases. Perturbation runs produced only modest increases relative to the Base. With RL, nighttime ozone increased by ~1.1–1.5 ppbv, but the effect was more spatially uniform than in the earlier years.

In summary, across all years and meteorological configurations, the implementation of RL consistently increased nighttime surface ozone by ~0.5–1.5 ppbv. These results demonstrate that incorporating the RL into CAMx modifies nighttime vertical mixing, sustaining higher surface ozone concentrations and better representing the role of residual ozone aloft. The impact is most pronounced under conditions with stronger vertical mixing (2021–2022) and is less influential during suppressed boundary layer development (2023).

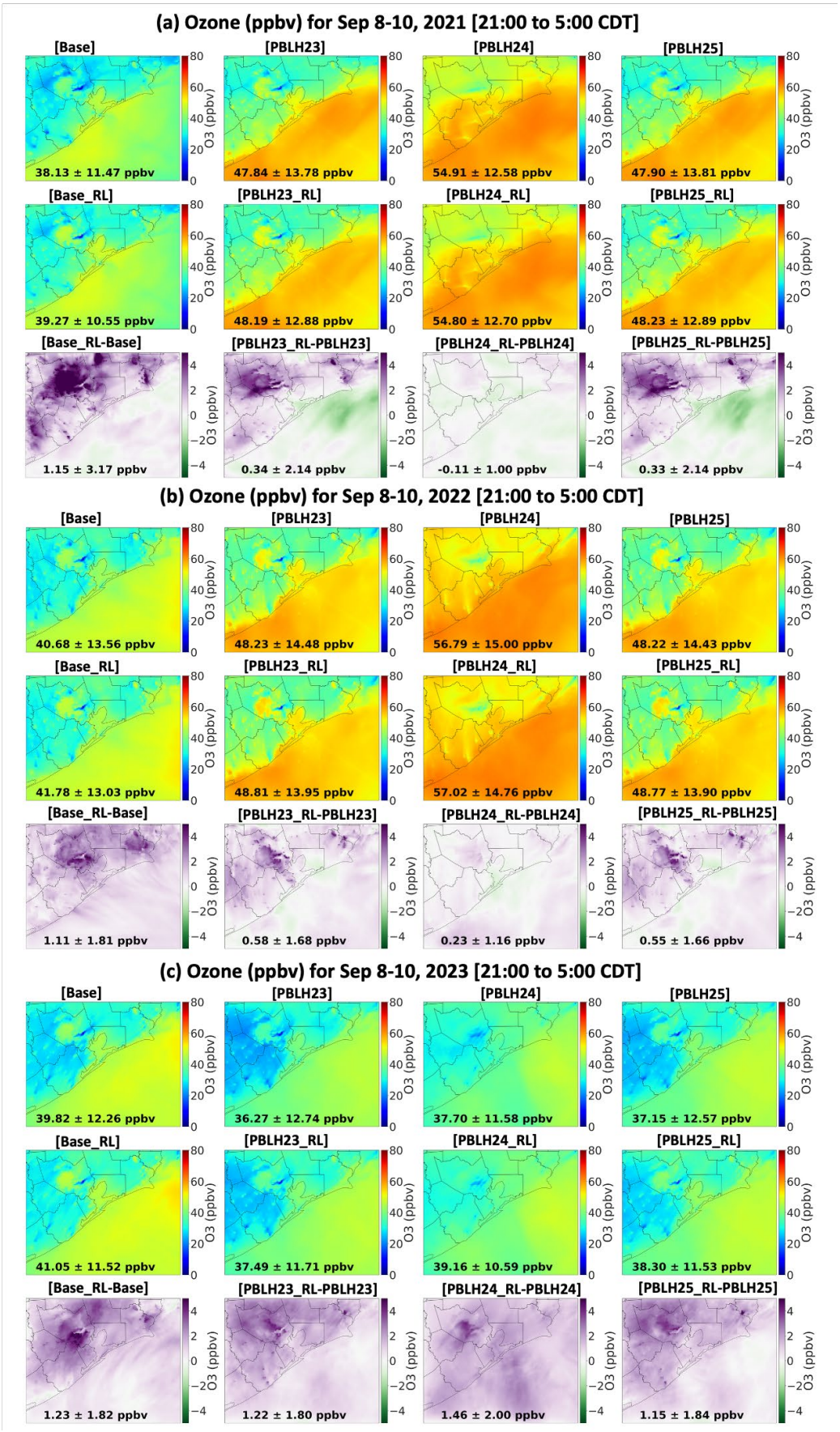

(a) Ozone (ppbv) for Sep 8-10, 2021 [21:00 to 5:00 CDT]
[Base]
38.13 ± 11.47 ppbv
[PBLH23]
47.84 ± 13.78 ppbv
[PBLH24]
54.91 ± 12.58 ppbv
[PBLH25]
47.90 ± 13.81 ppbv
[Base_RL]
39.27 ± 10.55 ppbv
[PBLH23_RL]
48.19 ± 12.88 ppbv
[PBLH24_RL]
54.80 ± 12.70 ppbv
[PBLH25_RL]
48.23 ± 12.89 ppbv
[Base_RL-Base]
1.15 ± 3.17 ppbv
[PBLH23_RL-PBLH23]
0.34 ± 2.14 ppbv
[PBLH24_RL-PBLH24]
-0.11 ± 1.00 ppbv
[PBLH25_RL-PBLH25]
0.33 ± 2.14 ppbv
(b) Ozone (ppbv) for Sep 8-10, 2022 [21:00 to 5:00 CDT]
[Base]
40.68 ± 13.56 ppbv
[PBLH23]
48.23 ± 14.48 ppbv
[PBLH24]
56.79 ± 15.00 ppbv
[PBLH25]
48.22 ± 14.43 ppbv
[Base_RL]
41.78 ± 13.03 ppbv
[PBLH23_RL]
48.81 ± 13.95 ppbv
[PBLH24_RL]
57.02 ± 14.76 ppbv
[PBLH25_RL]
48.77 ± 13.90 ppbv
[Base_RL-Base]
1.11 ± 1.81 ppbv
[PBLH23_RL-PBLH23]
0.58 ± 1.68 ppbv
[PBLH24_RL-PBLH24]
0.23 ± 1.16 ppbv
[PBLH25_RL-PBLH25]
0.55 ± 1.66 ppbv
(c) Ozone (ppbv) for Sep 8-10, 2023 [21:00 to 5:00 CDT]
[Base]
39.82 ± 12.26 ppbv
[PBLH23]
36.27 ± 12.74 ppbv
[PBLH24]
37.70 ± 11.58 ppbv
[PBLH25]
37.15 ± 12.57 ppbv
[Base_RL]
41.05 ± 11.52 ppbv
[PBLH23_RL]
37.49 ± 11.71 ppbv
[PBLH24_RL]
39.16 ± 10.59 ppbv
[PBLH25_RL]
38.30 ± 11.53 ppbv
[Base_RL-Base]
1.23 ± 1.82 ppbv
[PBLH23_RL-PBLH23]
1.22 ± 1.80 ppbv
[PBLH24_RL-PBLH24]
1.46 ± 2.00 ppbv
[PBLH25_RL-PBLH25]
1.15 ± 1.84 ppbv
O3 (ppbv)

Figure 45. Nighttime average (21:00–5:00 CDT) surface ozone concentrations (ppbv) from CAMx simulations for (a) September 8-10, 2021, (b) September 8-10, 2022, and (c) September 8-10, 2023. Each panel includes four columns corresponding to the Base, PBLH23, PBLH24, and PBLH25 meteorological configurations. The first row shows the Base, PBLH23, PBLH24, and PBLH25, the second row shows Base, PBLH23, PBLH24, and PBLH25 with residual layer implementation, and the third row shows the spatial difference between with RL and without RL implementation.

6.4. Effects on Vertical Ozone Profiles

PBL can bring air aloft toward the surface, and vice versa, leading to an uneven vertical distribution of ozone that directly influences surface ozone concentrations. The PBL perturbations, as well as the implementation of RL in the model, have the potential to affect this vertical mixing. **Figure 46** compares vertical ozone profiles from CAMx Base, PBLH23, PBLH24, and PBLH25 simulations (with and without RL) against ozonesonde observations over Galveston Bay and the Gulf during September 8–10, 2021–2023. Since RL was applied only at night and the ozonesonde launches occurred mostly during the daytime, the RL and non-RL results appear nearly identical. In 2021, both over the Bay and Gulf, the Base simulation underestimated ozone at both the surface and aloft, while PBLH23 and PBLH25 reduced these biases and more closely matched the observed vertical gradients. PBLH24, however, consistently overpredicted ozone throughout the column, producing the largest positive bias. The effect of RL implementation in CAMx can be seen in the Gulf profile at 09:00 CDT on September 9, 2021, where Base_RL showed modest improvement compared to the Base simulation. In 2022, the Base overestimated ozone in the morning and underestimated it in the afternoon, while the perturbation cases generally overpredicted observations. During the morning, PBLH23 and PBLH25 provided the best agreement with observations both at the surface and aloft, whereas in the afternoon they tended to overestimate ozone below ~2.5 km but matched observations more closely above that level. In 2023, similar to 2022, the Base model underestimated ozone at the surface and aloft up to ~2.5 km, while overestimating in the afternoon; PBLH23 and PBLH25 again showed better overall agreement with observations compared to the Base and PBLH24.

Overall, across all three years, PBLH23 and especially PBLH25 provided the most robust performance, demonstrating improved capability in capturing both surface and vertical ozone distributions relative to observations.

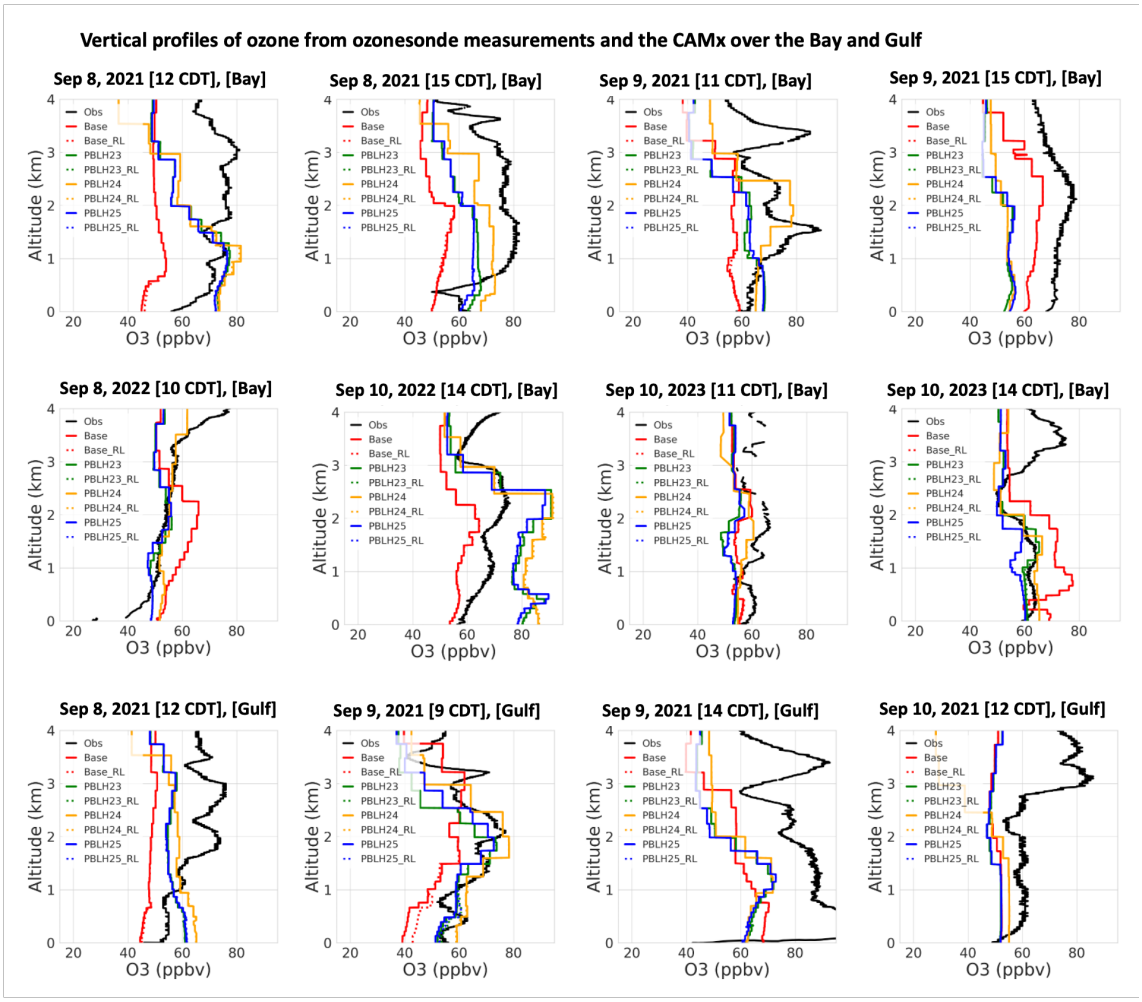


Figure 46. Vertical profiles of ozone from ozonesonde measurements and the CAMx over the Bay and Gulf during September 8-10, 2021-2023 for Base, PBLH23, PBLH24, and PBLH25. Black lines represent observed ozone, solid and dotted red lines show the Base and Base with RL simulation, solid and dotted green lines show the PBLH23 and PBLH23 with RL simulation, solid and dotted orange lines show the PBLH24 and PBLH24 with RL simulation, and solid and dotted blue lines show the PBLH25 and PBLH25 with RL simulation.

6.5. Effects of Improved PBL on Regional vs. Local Contributions to Ozone

Background ozone refers to the surface ozone concentration in a region influenced primarily by long-range transport, biogenic emissions, and free-tropospheric entrainment, in the absence of local anthropogenic emissions. To estimate regional background ozone in the HGB area, CAMx zero-emission simulations (referred to as the background [BGD] cases) were performed, in which all anthropogenic emissions were turned off within the innermost domain (d03). These simulations retained biogenic emissions and used the same meteorological and chemical boundary conditions as their corresponding full-emission control (CTR) runs. Background simulations were conducted for each meteorological configuration of CAMx: Base, PBLH23, PBLH24, and PBLH25, for September 8-10 in 2021, 2022, and 2023. Comparing the BGD and CTR results allows us to quantify the contribution of anthropogenic emissions to surface ozone

within the HGB region and assess how changes in meteorology influence background ozone levels.

Figure 47 presents the daytime average (08:00–20:00 CDT) surface ozone for the CTR and BGD simulations across all four meteorological cases and three years. In 2021, for the Base simulation, the average surface ozone in the HGB region was 49.22 ppbv, while the background ozone was 40.21 ppbv, indicating that ~81% of the total ozone was attributed to regional background. These results are consistent with findings from previous work (AQRP Project 22-008), which also reported a dominant role of background processes over the HGB region. In PBLH23, the total ozone increased to 56.68 ppbv, and background ozone rose to 45.35 ppbv, resulting in a slightly lower relative background contribution of ~80%. PBLH25 exhibited similar behavior to PBLH23. In PBLH24, total ozone further increased to 59.62 ppbv, with background ozone at 48.24 ppbv, yielding a background contribution of approximately ~81%. This suggests that while PBLH24 substantially enhanced background ozone through stronger mixing and entrainment, it produced a similar influence of local emissions as the PBLH23 and PBLH25.

In 2022, the contribution of local anthropogenic emissions to total ozone varied across meteorological configurations. For the Base simulation, total ozone averaged 53.79 ppbv, while background ozone was 45.71 ppbv, indicating that ~ 84% of the total ozone was attributed to regional background. In PBLH23, total ozone increased to 59.99 ppbv, and background ozone to 48.31 ppbv, resulting in a background contribution of ~80%. PBLH25 exhibited similar behavior to PBLH23. In PBLH24, total ozone reached 66.05 ppbv, and background ozone was 54.58 ppbv, yielding a background contribution of ~82%.

In 2023, total and background ozone levels were slightly lower across all cases. The base simulation showed total ozone of 49.87 ppbv, while the background ozone was 40.61 ppbv, indicating ~81% of the total ozone was attributed to the regional background. In PBLH23-PBLH25, total ozone was ~45–46 ppbv, while background ozone remained around ~37 ppbv, keeping the background contribution close to ~80-81%.

The largest differences between CAMx base and perturbation simulations (PBLH23-25) are revealed in the locally produced ozone (third row in **Figure 47a-c**). Here, locally produced ozone was estimated by taking the differences between the total ozone and background ozone. While simulations with enhanced marine PBL representation (PBLH23, PBLH24, and PBLH25) increased the background ozone due to stronger vertical mixing and entrainment, they also amplified the surface influence of local anthropogenic emissions through the same processes. The perturbation experiments (PBLH23-25) predicted local emissions contributed to an average of 11-12 ppbv in the 2021 and 2022 days, compared to 8-9 ppbv from the base simulation. The spatial distribution of the locally produced ozone also changes significantly between the perturbation experiments and the base in 2021 and 2022, with the former predicting more local contributions over urban Houston, Galveston Bay, and the Gulf. In Sep 8-10, 2023, PBLH23-25 predicted a slightly smaller contribution from local emissions compared to the base, likely because of less active photochemical conditions for those days as demonstrated by lower total ozone. Overall, these findings demonstrate that meteorological conditions and PBL structure significantly affect the estimation of regional background and local emissions on ozone in the HGB region, with the latter being more sensitive to the PBL perturbations.

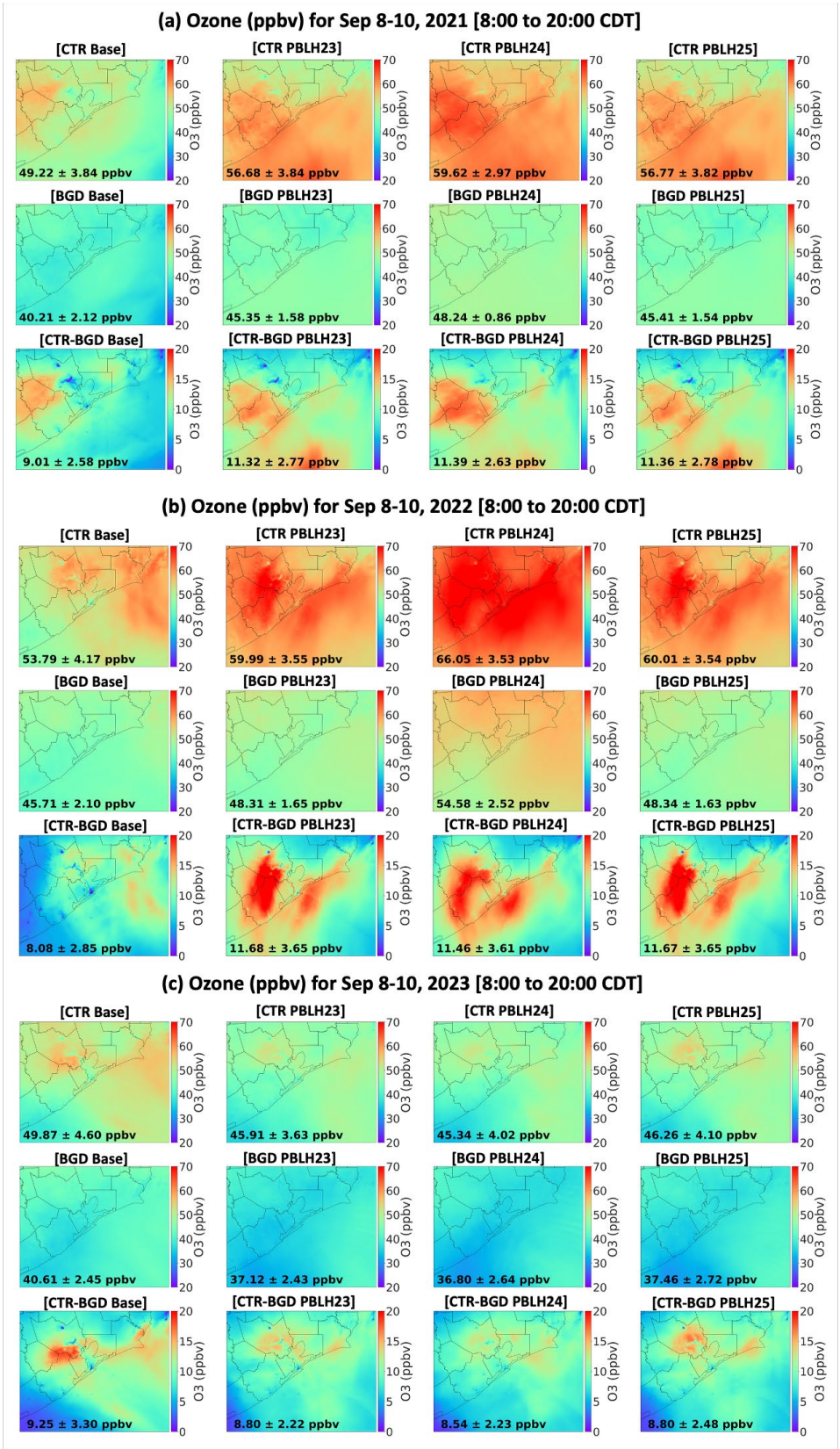

(a) Ozone (ppbv) for Sep 8-10, 2021 [8:00 to 20:00 CDT]
[CTR Base]
49.22 ± 3.84 ppbv
[CTR PBLH23]
56.68 ± 3.84 ppbv
[CTR PBLH24]
59.62 ± 2.97 ppbv
[CTR PBLH25]
56.77 ± 3.82 ppbv
[BGD Base]
40.21 ± 2.12 ppbv
[BGD PBLH23]
45.35 ± 1.58 ppbv
[BGD PBLH24]
48.24 ± 0.86 ppbv
[BGD PBLH25]
45.41 ± 1.54 ppbv
[CTR-BGD Base]
9.01 ± 2.58 ppbv
[CTR-BGD PBLH23]
11.32 ± 2.77 ppbv
[CTR-BGD PBLH24]
11.39 ± 2.63 ppbv
[CTR-BGD PBLH25]
11.36 ± 2.78 ppbv
O3 (ppbv)
(b) Ozone (ppbv) for Sep 8-10, 2022 [8:00 to 20:00 CDT]
[CTR Base]
53.79 ± 4.17 ppbv
[CTR PBLH23]
59.99 ± 3.55 ppbv
[CTR PBLH24]
66.05 ± 3.53 ppbv
[CTR PBLH25]
60.01 ± 3.54 ppbv
[BGD Base]
45.71 ± 2.10 ppbv
[BGD PBLH23]
48.31 ± 1.65 ppbv
[BGD PBLH24]
54.58 ± 2.52 ppbv
[BGD PBLH25]
48.34 ± 1.63 ppbv
[CTR-BGD Base]
8.08 ± 2.85 ppbv
[CTR-BGD PBLH23]
11.68 ± 3.65 ppbv
[CTR-BGD PBLH24]
11.46 ± 3.61 ppbv
[CTR-BGD PBLH25]
11.67 ± 3.65 ppbv
O3 (ppbv)
(c) Ozone (ppbv) for Sep 8-10, 2023 [8:00 to 20:00 CDT]
[CTR Base]
49.87 ± 4.60 ppbv
[CTR PBLH23]
45.91 ± 3.63 ppbv
[CTR PBLH24]
45.34 ± 4.02 ppbv
[CTR PBLH25]
46.26 ± 4.10 ppbv
[BGD Base]
40.61 ± 2.45 ppbv
[BGD PBLH23]
37.12 ± 2.43 ppbv
[BGD PBLH24]
36.80 ± 2.64 ppbv
[BGD PBLH25]
37.46 ± 2.72 ppbv
[CTR-BGD Base]
9.25 ± 3.30 ppbv
[CTR-BGD PBLH23]
8.80 ± 2.22 ppbv
[CTR-BGD PBLH24]
8.54 ± 2.23 ppbv
[CTR-BGD PBLH25]
8.80 ± 2.48 ppbv
O3 (ppbv)

Figure 47. Daytime average (08:00–20:00 CDT) surface ozone concentrations (ppbv) from CAMx simulations for (a) September 8-10, 2021, (b) September 8-10, 2022, and (c) September 8-10, 2023. Each panel includes four columns corresponding to the Base, PBLH23, PBLH24, and PBLH25 meteorological configurations. The first row shows control (CTR) simulations with full anthropogenic and biogenic emissions, the second row shows background (BGD) simulations with anthropogenic emissions turned off, and the third row shows the difference between control and background (CTR − BGD), representing the modeled contribution of local anthropogenic emissions to surface ozone.

6.6. Summary

This study evaluates the impact of improved planetary boundary layer height representation on CAMx ozone predictions and source attribution for the HGB region, focusing on the period of September 8-10 for 2021, 2022, and 2023. The work builds on Sections 4 and 5 and uses meteorological inputs from the WRF v4.6.0 model under four configurations: a Base, and three improved configurations, PBLH23, PBLH24, and PBLH25, which were selected from the WRF perturbation/sensitivity experiments for their better performance in simulating PBLH over the water. Meteorological inputs from the WRF base model, as well as these improved PBLH cases, were used to drive CAMx v7.31 simulations to assess the influence of boundary layer dynamics on surface ozone distribution and source contributions.

Evaluation of the CAMx base simulation revealed a tendency to underestimate surface ozone over Galveston Bay and the Gulf, particularly during morning and early afternoon hours. In contrast, simulations using the PBLH23, PBLH24, and PBLH25 configurations produced higher ozone levels, improved spatial variation, and significantly reduced model bias. PBLH24 shows the largest ozone increases over land, attributed to the inclusion of the Urban Canopy Model (UCM). PBLH23 and PBLH25 showed better balance by reproducing low ozone in the early morning and minimizing afternoon overestimations. PBLH25, in particular, provided the most consistent improvements across all years, especially offshore, due to the additional perturbation to surface roughness length over water. In 2023, when PBLH development was weaker, the effect of all perturbations on ozone was more limited; however, PBLH25 still delivered the best correlation with observed values. Overall, PBLH25 emerged as the most robust configuration for simulating ozone across both coastal and marine environments.

Building on these results, the diagnosed RL heights from Section 5 were implemented in CAMx by replacing nighttime WRF PBL heights (21:00–05:00 CDT). Across 2021–2023, RL incorporation consistently increased nighttime surface ozone by ~0.5–1.5 ppbv, particularly over coastal and inland regions, by redistributing residual ozone aloft downward. While the effect was limited during daytime ozonesonde launches, PBLH23 and PBLH25 combined with RL showed the best agreement with observed vertical ozone distributions, sustaining higher nighttime surface ozone and better capturing vertical gradients. Overall, RL implementation enhances CAMx representation of vertical mixing and surface ozone concentrations, complementing the improvements achieved by the optimized PBLH configurations.

To estimate the contributions of regional background and local anthropogenic emissions, zero-emission simulations were performed. Background ozone increased with improved PBL representation due to enhanced vertical mixing and entrainment of ozone-rich air. However, the largest differences between CAMx base and perturbation simulations (PBLH23-25) are revealed

in the locally produced ozone. In 2021 and 2022, local contributions were increased by up to 30% in PBLH23-25 compared to the base, reflecting that enhanced vertical mixing also resulted in a greater sensitivity of surface ozone to local emissions. Among the tested configurations, PBLH25 provided the best trade-off between bias reduction and realistic attribution of local versus regional contributions.

In summary, this work highlights the importance of meteorology, particularly the structure of the PBLH, in driving ozone variability and source attribution in coastal environments. Improved PBLH configurations not only enhanced model agreement with observations but also shifted the balance between regional background and locally produced ozone, with the latter being more enhanced than the former under polluted conditions (e.g., Sep 8-10, 2021, and 2022). These results suggest that the expected outcome of local emission control strategies to mitigate ozone air pollution is dependent on the representation of PBL dynamics in photochemical models that are used to estimate such outcomes.

7. Quality Assurance/Quality Control

Quality assurance and quality control (QA/QC) were ensured for both the secondary data and modeling data used and generated in the project. The secondary data includes onshore and offshore observations from the TRACER-AQ field campaigns during 2021 – 2023. Monitoring data from these projects includes ship-based ozone concentrations, PBL and other meteorological parameters, and ozonesondes. Additional offshore data came from the SCOAPE-19 campaign, which was used to conduct an independent model evaluation of offshore PBL. Other secondary data used in this project include trace gas concentrations and meteorological parameters measured at the continuous ambient monitoring stations (CAMS) inside the d03 modeling domain. All secondary data were obtained from the collections or versions of final, publishable quality, which have gone through QA/QC processes of the individual investigators who collected the datasets. For example, Dr. James Flynn and Dr. Travis Griggs were the primary investigators who collected offshore observations during the TRACER-AQ campaigns. They have reviewed those observations used in the project and verified the data quality. The ceilometer data re-processing process described in Section 2.1 was an outcome of the QA/QC procedure.

Quality of modeling data was ensured by in-depth evaluation of the models' performance in simulating meteorology, ozone, and precursors against both on- and offshore observations, including those from the TRACER-AQ field campaigns and routine meteorology/air quality monitoring sites. At least 10% of the model data were audited for QA/QC, with the specific focus on the model data of the innermost domain (1.33 km x 1.33 km). The model performances compared to observations were documented throughout the project using the metrics in **Table 1**. The impact of PBL improvements on the model performance of meteorology and ozone was also thoroughly evaluated and documented using difference plots and the same metrics from **Table 1** to quantify the differences. For example, **Figure 48-51** shows the QA/QC plots of meteorological parameters produced by the WRF base model and two perturbation simulations (PBLH23 and PBLH25) for September 2022. The model parameters evaluated include temperature, pressure, wind speed, and dew point. Observations were taken from the CAMS sites. As shown by the evaluation statistics in the figure, the models passed the recommended benchmarks from literature for model performance (Emery et al., 2017). **Figure 44** is another example of QA/QC conducted for the CAMx model.

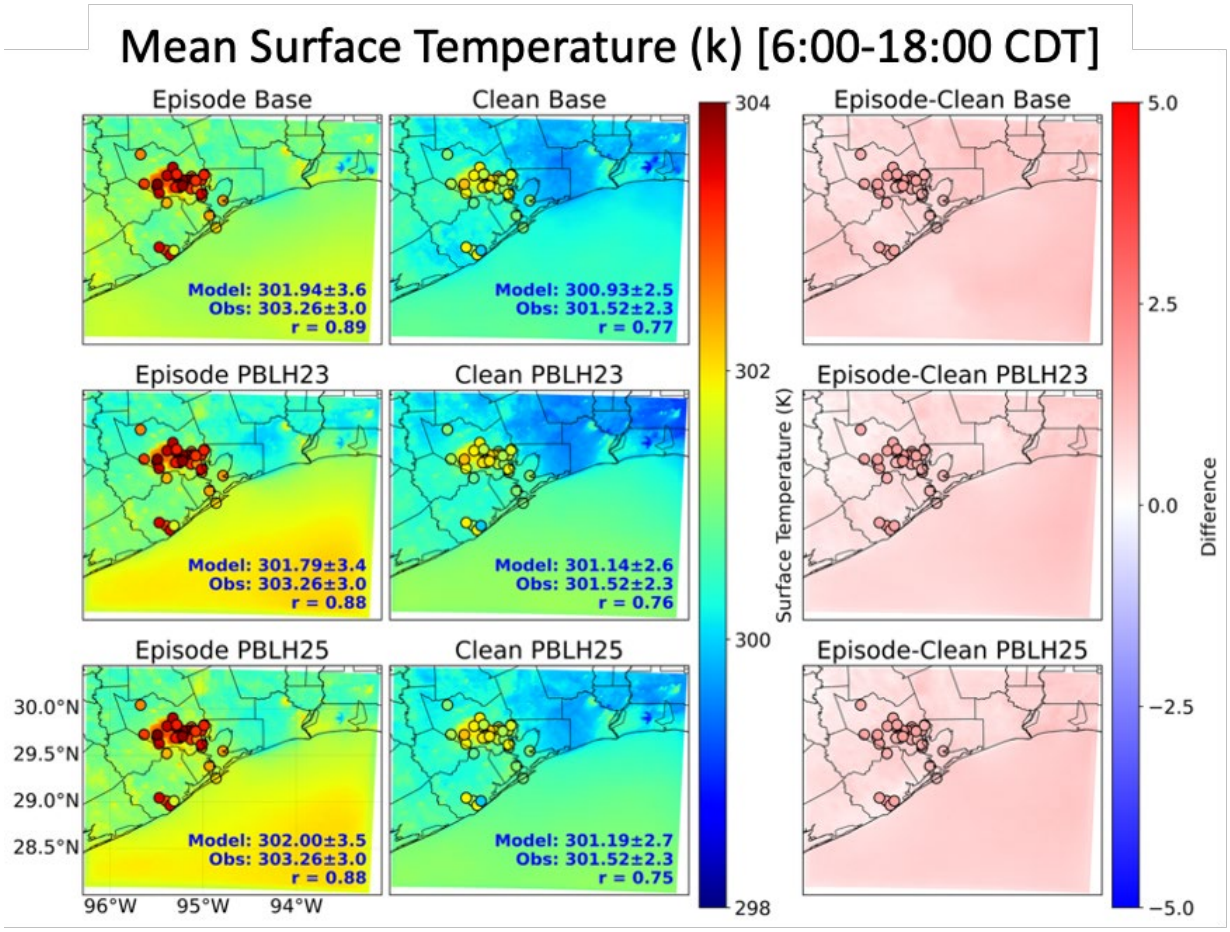


Figure 48. Maps of mean surface temperature (k) for CAMx Base, PBLH23, and PBLH25 for September 2022 during ozone episodes (left), clean days (middle), and episode minus clean days (right). Dots overlaid on the map are the observed mean values from the CAMS sites.

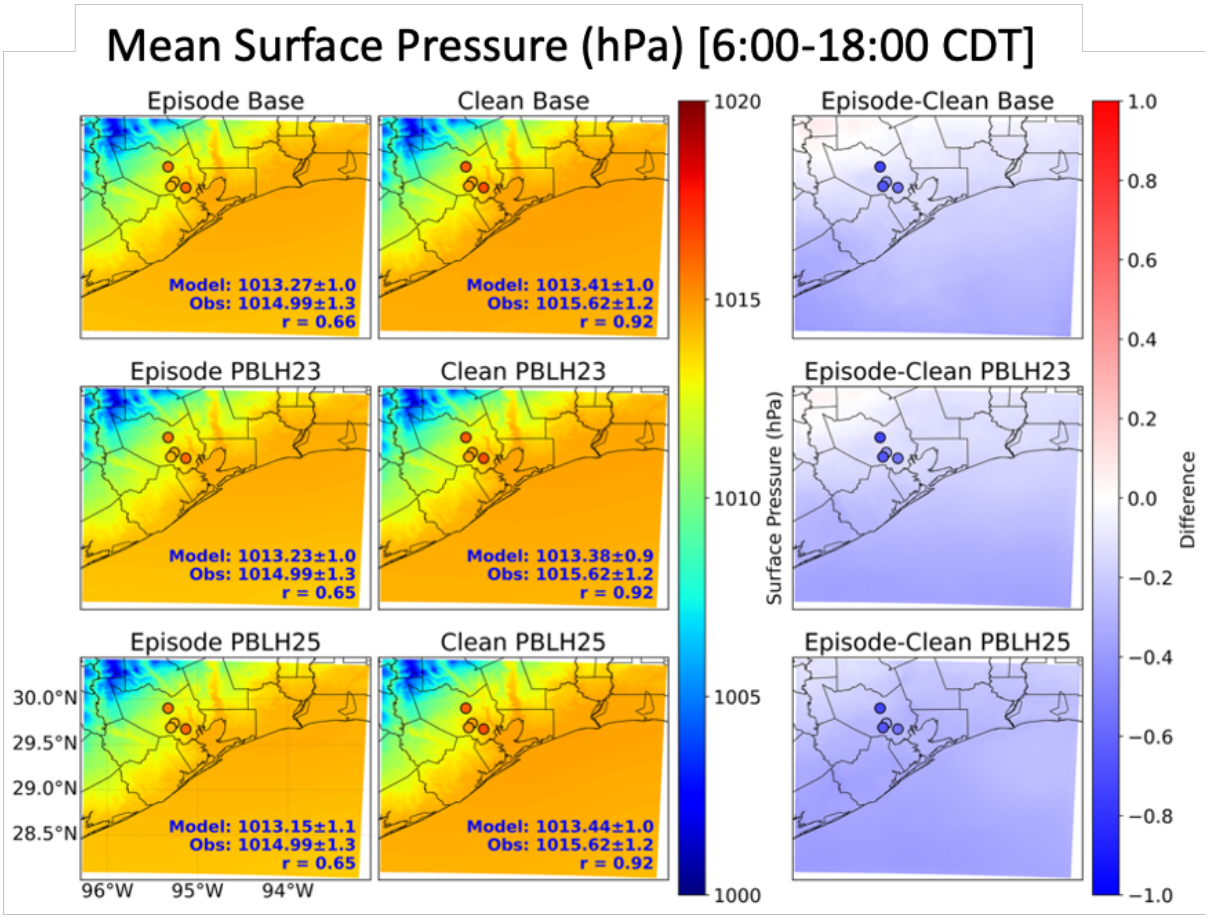


Figure 49. Maps of mean surface pressure (hPa), for CAMx Base, PBLH23, and PBLH25 for September 2022 during ozone episodes (left), clean days (middle), and episode minus clean days (right). Dots overlaid on the map are the observed mean values from the CAMS sites.

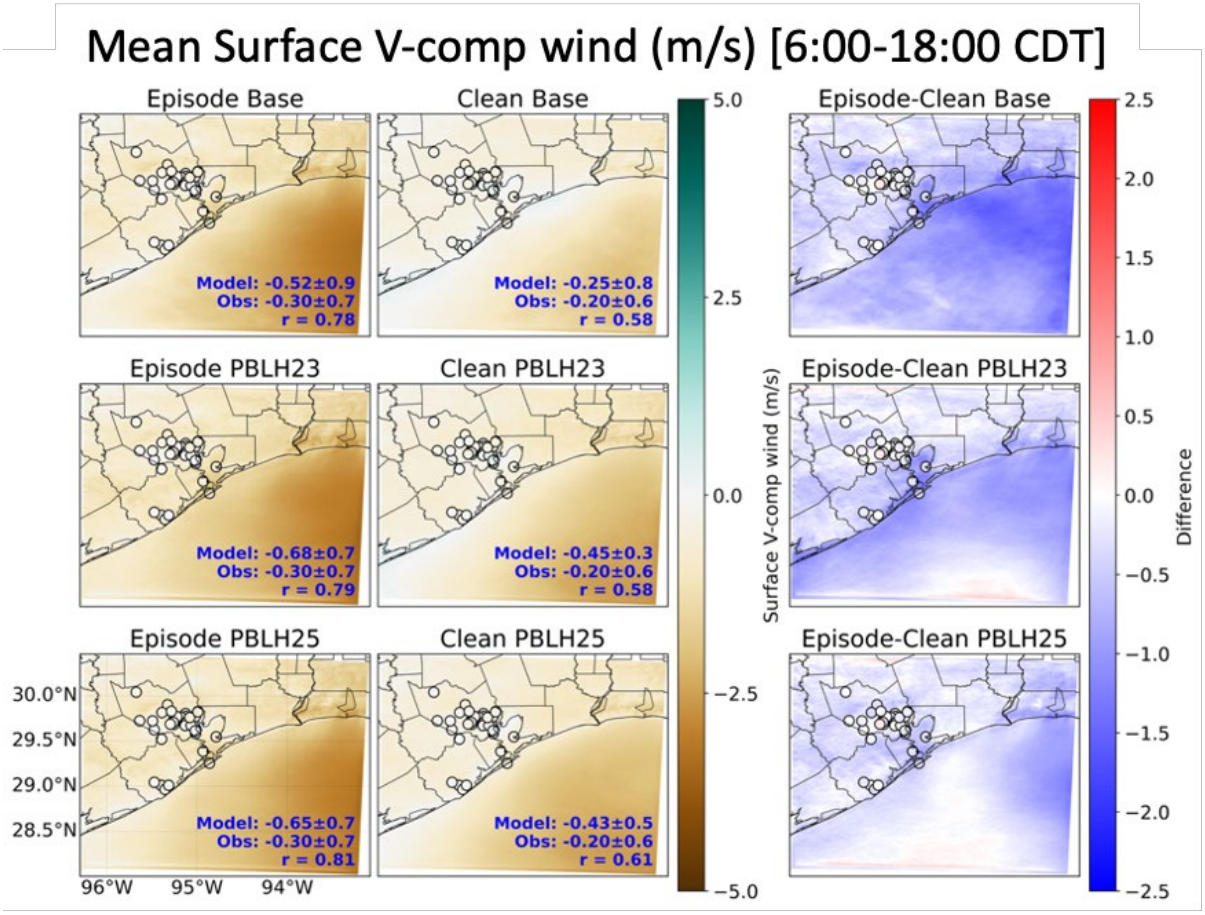


Figure 50. Maps of mean surface V-component (m/s), for CAMx Base, PBLH23, and PBLH25 for September 2022 during ozone episodes (left), clean days (middle), and episode minus clean days (right). Dots overlaid on the map are the observed mean values from the CAMS sites.

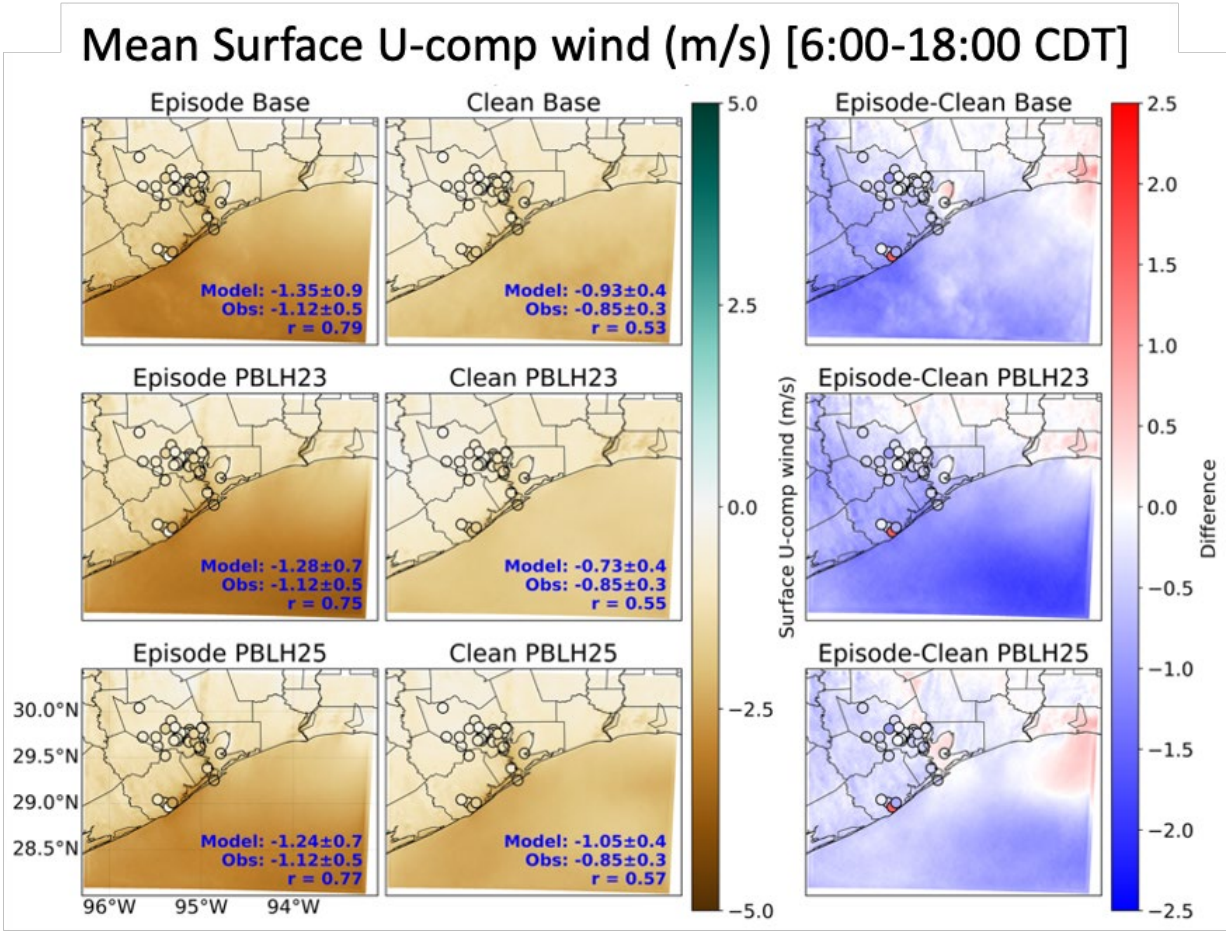


Figure 51. Maps of mean surface U-component (m/s), for CAMx Base, PBLH23, and PBLH25 for September 2022 during ozone episodes (left), clean days (middle), and episode minus clean days (right). Dots overlaid on the map are the observed mean values from the CAMS sites.

8. Summary and Future Directions

To improve the understanding of marine meteorology and coastal boundary layer processes in the HGB region, multiple field campaigns were conducted from 2021 to 2023. These campaigns provided, for the first time, observations of PBLH at multiple coastal and marine locations, enabling the evaluation of air quality model performance over Galveston Bay and the Gulf of America. Utilizing these datasets, this AQRP project validated and improved WRF model–based marine PBLH predictions over Galveston Bay and the Gulf of America. In particular, the study used extensive offshore observations from TRACER-AQ to assess how improved representation of marine and coastal PBL processes influenced ozone formation and transport in the CAMx. The project activities have produced robust analyses that answer the following primary science questions:

Q1: How well does mesoscale meteorological grid modeling replicate coastal/maritime boundary layers observations from the 2021-2023 offshore observations?

Answer: The WRF v4.6.0 base model captured the general diurnal and monthly variations but consistently underestimated marine PBL height, particularly over Galveston Bay during

afternoon hours, where observed PBLs frequently exceeded 1.5 km. Model performance was found to be location- and time-dependent, with larger biases over open water compared to nearshore regions. These results provide a baseline for improving WRF parameterizations for PBLH predictions over offshore water bodies.

Q2: How sensitive is WRF's prediction of coastal/maritime boundary layers to model parameters? To what extent do the 2021-2023 offshore observations constrain those parameters?

Answer: Sensitivity/perturbation simulations show the strong dependence of marine PBLH on physical parameterizations and the cumulus scheme selected. Higher surface roughness over the water improved vertical mixing near the coast, coupling of a 1-D ocean model better represented marine PBLH, and the NSAS cumulus scheme enhanced cloud-PBL interactions. When combined (PBLH25), these settings produced the most robust improvements to marine PBL. The offshore observations proved essential in constraining these sensitivities, confirming that parameter tuning without observational benchmarks would have been insufficient to resolve the biases.

Q3: How will the simulation of residual layer ozone be improved by explicitly parameterizing the entrainment of free tropospheric ozone into the residual layer?

Answer: Implementation of a diagnosed RL into CAMx improved representation of nighttime vertical mixing by redistributing aloft ozone downward, increasing surface ozone by ~0.5–1.5 ppbv. Although ozonesondes were launched during the daytime (when RL impacts were minimal), comparisons showed localized improvements (e.g., over the Gulf in 2021). These results indicate that explicit RL treatment enhances model realism but remains limited by CAMx treatment of entrainment between the surface layer, RL, and free troposphere. Further development is needed to allow RL products to dynamically affect vertical mixing processes inside CAMx.

Q4: What are the effects of improved PBL and RL simulation on offshore ozone prediction and source attribution in CAMx?

Answer: Improved PBL representations show the positive effects on the CAMx ozone predictions at offshore locations, reducing model biases by up to 15–20 ppbv along ship tracks and raising monthly mean ozone by ~5 ppbv. Locally produced ozone increased by up to 30% under improved PBL configurations due to enhanced vertical transport. RL implementation further elevated nighttime surface ozone, sustaining higher background levels. Together, these results highlight the importance of boundary layer dynamics in shaping offshore ozone levels and demonstrate that improving PBLH and RL representation can refine source attribution and photochemical assessments in CAMx. The study reveals that improved PBL representation in CAMx enhances locally produced ozone more than regional background ozone under polluted conditions (e.g., Sep 8-10, 2021, and 2022). Thus, the expected outcome of local emission control strategies to mitigate ozone air pollution is dependent on the representation of PBL dynamics in photochemical models that are used to estimate such outcomes.

The project results suggest five key directions for future research to improve meteorological and photochemical modeling over marine environments. First, further refinement of PBLH representation in WRF is needed, including improvements to surface roughness, ocean coupling, and cumulus parameterizations under diverse meteorological conditions. Second, studies should

evaluate how enhanced PBLH simulations influence top-down emissions estimates and vertical pollutant transport in air quality models, as these processes are highly dependent on vertical mixing. Third, a detailed investigation of differences between Galveston Bay and the Gulf is recommended to better understand the physical and chemical drivers of their distinct boundary layer structures and how they are represented in models. Fourth, further work is required to integrate diagnosed residual layer information into CAMx to improve vertical mixing and surface ozone predictions. Fifth, while the urban canopy model improves PBLH predictions over water, it can overestimate ozone by up to 20 ppbv in some cases, and its impacts on air quality require additional investigation.

References:

Burkholder, J. B., Sander, S. P., Abbatt, J., Barker, J. R., Cappa, C., Crounse, J. D., Dibble, T. S., Huie, R. E., Kolb, C. E., Kurylo, M. J., Orkin, V. L., Percival, C. J., Wilmouth, D. M., & Wine, P. H. (2019). Chemical kinetics and photochemical data for use in atmospheric studies, Number 19. In *Jet Propulsion Laboratory, Pasadena: Vol. JPL Public* (Issue II).

Chen, F., & Dudhia, J. (2001). Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4). https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Chen, F., Janjić, Z., & Mitchell, K. (1997). Impact of atmospheric surface-layer parameterizations in the new land-surface scheme of the NCEP mesoscale Eta model. *Boundary-Layer Meteorology*, *85*(3), 391–421.

Emery, C., Liu, Z., Russell, A. G., Odman, M. T., Yarwood, G., & Kumar, N. (2017). Recommendations on statistics and benchmarks to assess photochemical model performance. *Journal of the Air and Waste Management Association*, *67*(5). https://doi.org/10.1080/10962247.2016.1265027

Ghonima, M. S., Yang, H., Kim, C. K., Heus, T., & Kleissl, J. (2017). Evaluation of WRF SCM Simulations of Stratocumulus-Topped Marine and Coastal Boundary Layers and Improvements to Turbulence and Entrainment Parameterizations. *Journal of Advances in Modeling Earth Systems*, *9*(7). https://doi.org/10.1002/2017MS001092

Huang, Y., & Peng, X. (2017). Improvement of the Mellor–Yamada–Nakanishi–Niino Planetary Boundary-Layer Scheme Based on Observational Data in China. *Boundary-Layer Meteorology*, *162*(1). https://doi.org/10.1007/s10546-016-0187-0

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: Atmospheres*, *113*(D13).

Jensen, M. P., Flynn, J. H., Judd, L. M., Kollias, P., Kuang, C., McFarquhar, G., Nadkarni, R., Powers, H., & Sullivan, J. (2021). A Succession of Cloud, Precipitation, Aerosol, and Air Quality Field Experiments in the Coastal Urban Environment. *Bulletin of the American Meteorological Society*, *102*(2). https://doi.org/10.1175/BAMS-D-21-0104.1

Li, W., Wang, Y., Liu, X., Soleimanian, E., Griggs, T., Flynn, J., & Walter, P. (2023). Understanding offshore high-ozone events during TRACER-AQ 2021 in Houston: Insights from WRF-CAMx photochemical modeling. *Atmospheric Chemistry and Physics*, *23*(21). https://doi.org/10.5194/acp-23-13685-2023

Liu, L., Menenti, M., & Ma, Y. (2022). Evaluation of Albedo Schemes in WRF Coupled with Noah-MP on the Parlung No. 4 Glacier. *Remote Sensing*, *14*(16). https://doi.org/10.3390/rs14163934

Liu, X., Wang, Y., Wasti, S., Li, W., Soleimanian, E., Flynn, J., Griggs, T., Alvarez, S., Sullivan, J. T., Roots, M., Twigg, L., Gronoff, G., Berkoff, T., Walter, P., Estes, M., Hair, J. W., Shingler, T., Scarino, A. J., Fenn, M., & Judd, L. (2023). Evaluating WRF-GC v2.0 predictions of boundary layer height and vertical ozone profile during the 2021 TRACER-AQ campaign in Houston, Texas. *Geoscientific Model Development*, *16*(18). https://doi.org/10.5194/gmd-16-5493-2023

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes. *Monthly Weather Review*, *137*(3), 991–1007.

Nakanishi, M., & Niino, H. (2009). Development of an improved turbulence closure model for the atmospheric boundary layer. *Journal of the Meteorological Society of Japan*, *87*(5). https://doi.org/10.2151/jmsj.87.895

Pithan, F., Angevine, W., & Mauritsen, T. (2015). Improving a global model from the boundary layer: Total turbulent energy and the neutral limit Prandtl number. *Journal of Advances in Modeling Earth Systems*, *7*(2). https://doi.org/10.1002/2014MS000382

Shaw, W. J., Berg, L. K., Debnath, M., Deskos, G., Draxl, C., Ghate, V. P., Hasager, C. B., Kotamarthi, R., Mirocha, J. D., Muradyan, P., Pringle, W. J., Turner, D. D., & Wilczak, J. M. (2022). Scientific challenges to characterizing the wind resource in the marine atmospheric boundary layer. In *Wind Energy Science* (Vol. 7, Issue 6). https://doi.org/10.5194/wes-7-2307-2022

Skamarock, W. C., Klemp, J. B., Dudhia, J. B., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., & Powers, J. G. (2021). A Description of the Advanced Research WRF Model Version 4.3. *NCAR Technical Note*, *TN–556+STR*(July).

Soleimanian, E., Wang, Y., Li, W., Liu, X., Griggs, T., Flynn, J., Walter, P. J., & Estes, M. J. (2023). Understanding ozone episodes during the TRACER-AQ campaign in Houston, Texas: The role of transport and ozone production sensitivity to precursors. *Science of the Total Environment*, *900*. https://doi.org/10.1016/j.scitotenv.2023.165881

Stull, R. B. (1988). An introduction to boundary layer meteorology. *An Introduction to Boundary Layer Meteorology*. https://doi.org/10.1007/978-94-009-3027-8

TCEQ. (2022). *Texas Commission on Environmental Quality (2022)*. https://www.tceq.texas.gov/airquality/sip/hgb/hgb-status

Thompson, A. M., Kollonige, D. E., Stauffer, R. M., Kotsakis, A. E., Abuhassan, N., Lamsal, L. N., Swap, R. J., Blake, D. R., Townsend-Small, A., & Wecht, H. D. (2023). Two Air Quality

Regimes in Total Column NO2 Over the Gulf of Mexico in May 2019: Shipboard and Satellite Views. *Earth and Space Science*, *10*(3). https://doi.org/10.1029/2022EA002473

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, *117*(8), 1779–1800.

Wang, X., Xiang, Y., Liu, W., Yun, L., Lv, L., Zhang, M., Xu, J., Fan, G., Ou, J., Wang, Q., Duan, D., & Zhang, T. (2021). Characteristics and Source Apportionment of the Vertical Distribution of Ozone at a Site of the Pearl River Delta Region of China. *Earth and Space Science*, *8*(6). https://doi.org/10.1029/2020EA001578

Wesely, M. L. (2007). Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmospheric Environment*, *41*, 52–63.

Yang, B., Qian, Y., Berg, L. K., Ma, P. L., Wharton, S., Bulaevskaya, V., Yan, H., Hou, Z., & Shaw, W. J. (2017). Sensitivity of Turbine-Height Wind Speeds to Parameters in Planetary Boundary-Layer and Surface-Layer Schemes in the Weather Research and Forecasting Model. *Boundary-Layer Meteorology*, *162*(1). https://doi.org/10.1007/s10546-016-0185-2

Zhang, C., Wang, Y., & Hamilton, K. (2011). Improved representation of boundary layer clouds over the southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme. *Monthly Weather Review*, *139*(11), 3489–3513.

Zhang, J., & Rao, S. T. (1999). The role of vertical mixing in the temporal evolution of ground-level ozone concentrations. *Journal of Applied Meteorology*, *38*(12). https://doi.org/10.1175/1520-0450(1999)038<1674:TROVMI>2.0.CO;2