Final Report

A synthesis study of the role of mesoscale and synoptic-scale wind on the concentrations of ozone and its precursors in Houston

Project #18-010

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

By

Qi Ying and Jie Zhang
Zachry Department of Civil Engineering
Texas A&M University

And

John Nielsen-Gammon and David Coates
Department of Atmospheric Sciences
Texas A&M University

November 8, 2019

Executive Summary

The study placed the local-scale wind structures observed by long-term wind profilers into the context of large-scale winds and the processes that drive those structures. Clockwise wind rotation was found to be an overwhelming presence in the lower atmosphere over Texas. Disturbances tied to the time of day can operate in resonance with the inertial rotation to amplify the local-scale wind variations. In Texas, two such disturbances are the daily growth and collapse of the daytime convective boundary layer and the daytime heating of air over land relative to air over water. The latter also produces the sea breeze, and the coastal inertial oscillation can be viewed as a part of the sea breeze cycle. In the Houston area, local winds are strong when large-scale winds are light and large differences in the timing of wind variations at different altitudes are observed. This suggests that the coastal heating contrast as the dominant mechanism. Stagnation and recirculation take place when the local-scale winds are strong enough to equal or exceed the large-scale wind speeds. This happens over a range of wind speeds during the warm season (April-September), mostly speeds less than 2.5 m/s. The one main exception is that southwesterly winds can produce larger local-scale wind variations, so stagnation is likely with large-scale winds up to 5 m/s when the wind is from the southwest. Sample wind profiler back-trajectories illustrate a variety of wind patterns under light wind conditions. Near the transition zone between stagnation and non-stagnation, the coastal inertial cycle is sometimes strong enough to cause recirculation over a few hours. When large-scale winds are nearly calm, the wind rotation tends to dominate, and air can carry out large loops up to 100 km in diameter before returning to roughly the same location 24 hours later. At night, pollutant transport is strongly affected by vertical wind shear. Some of the same conditions that favor robust recirculation also favor rapid nighttime dispersal of the Houston pollution plume. The plume is most likely to remain intact when summertime winds are from the south-southeast and moderately strong.

In order to investigate model simulations of wind rotation and resultant pollutant evolution, four specific cases in 2000, 2013 and 2016 were chosen. The suite of cases was selected to include a wide variety of wind rotation behaviors, all of which were associated with ozone exceedances, during years of high interest by Texas Commission on Environmental Quality (TCEQ) in which radar wind profiler measurements were available. Because the recirculation of the sea breeze is strongly impacted by the background flow through turbulent mixing, the planetary boundary layer (PBL) scheme parameterization is a key component to how well the Weather Research and Forecasting (WRF) model resolves the recirculation, three different PBL schemes, namely the Yonsei University Scheme (YSU), the Mellor-Yamada-Janjic Scheme (MYJ), and the Quasi-normal Scale Elimination Scheme (QNSE), are tested for their capability of reproducing the back trajectories based on profiler data. The WRF models as configured struggle to show any recirculation in the boundary layer near Houston at all. It was often seen that rotation was weaker than observed or more uniform across the depth of the boundary layer than observed, and more often the winds were higher speed than observed, with trajectory segments and total lengths much longer than those of the observations.

The age-resolved Community Multiscale Air Quality (CMAQ) model developed in this study has been demonstrated through the case studies to be a useful tool to help elucidating the cause of the high O_3 concentrations in the study area. The coastal area in Texas is influenced by high O_3 over the Gulf of Mexico. The cases studies in this project show that the contributions of aged O_3 can be as high as 50% (or ~20 ppb) of the peak time non-background O_3 at Galveston. This level of aged non-background O_3 is almost as high as the fresh O_3 predicted in the vicinity of the urban Houston area on high O_3 days. While the WRF/CMAQ system applied in this study

gives an acceptable performance of O_3 at Galveston, it is obvious that correct assessment of the impact of aged O_3 to local O_3 events depends on the correct prediction of the wind patterns in the coastal area. As a sharp land/sea gradient of O_3 often exist, particularly at nighttime hours when O_3 over the urban areas are titrated by high NOx emissions, small errors in the wind (esp. the wind directions) can lead to significant differences in predicted O_3 concentrations at coastal areas. The current study demonstrates that the WRF model does not correctly reproduce the observed wind circulation patterns derived from the wind profiler data. Further studies are needed to better understand the cause of the errors in the WRF model in the Gulf of Mexico coastal area.

Regrading the budget adjustment: We requested last-minute budget adjustments on August 28, 2019, to change funds previously allocated for two postdoc researchers to two graduate students. Dr. John Nielsen-Gammon previously thought he would hire David Coates as a postdoc to work on the project but David was not able to graduate in time. Dr. Ying's postdoc in-mind was not able to come because she was in China and she was not able to obtain a proper J-1 visa due to her previous visit to the US as a visiting Ph.D. student, holding a J-1 visa. Since then, Dr. Ying opted to work with Jie Zhang, a new Ph.D. student, on this project. The two graduate students have been working on the project throughout the duration of the project. This statement is included in the final report, as requested by RoseAnna Goewey, the Program Manager.

Table of Contents

List of Figures

List of Tables

1. Background and Research Objectives

Around the time of the Texas Air Quality Study 2000 (TexAQS-2000) field program, the importance of mesoscale wind patterns in the Houston-Galveston-Brazoria (HGB) airshed had been recognized, even as understanding of those wind patterns has grown over time (Banta et al., 2005; Darby, 2005; Daum et al., 2003; Davis et al., 1998). Initially, the dominant mesoscale process was thought to be an ordinary sea breeze, modified by the complex coastline of HGB and conceptually separated into a (Galveston) Bay breeze and a Gulf (of Mexico) breeze. During the day under light synoptic-scale wind conditions, the Bay breeze is first to affect Houston, followed by the much stronger and larger Gulf breeze in the late afternoon. A land breeze develops at night (Day et al., 2010).

The next advance in meteorological understanding was the recognition that sea breeze behavior in the HGB area was unlike that observed at higher latitudes due to the inertial oscillation, which is in near-resonance with the daily surface heating cycle in HGB (Nielsen-Gammon et al., 2005; Parrish et al., 2009). The combined sea breeze-inertial oscillation dominates pollutant transport patterns during both day and night, leading to recirculation of pollutants and emission of "double doses" of pollutants into the already polluted air.

While the sea breeze-inertial oscillation seems key to understanding local pollution, low-level jet patterns have been found to play important roles in local and regional transport (Daum et al., 2004; Tucker et al., 2010). In a sense, the sea breeze itself features a low-level wind maximum and has been treated as a low-level jet by Tucker et al. (2010). Besides this, three types of low-level jets distinct from conventional coastal sea breeze-inertial oscillation appear to be relevant in southeast Texas. First, under southwest wind conditions, a coastal low-level jet is established that can interact with the sea breeze to produce extremely large diurnal wind oscillations, such as were observed during late August 2000. Second, an inland surge of the sea breeze has been found to produce a local low-level wind maximum in the interior of eastern Texas that may be important for regional transport. Third, the Great Plains low-level jet is the dominant diurnal wind feature of western, central, and southern Texas. Recent work by Nielsen-Gammon (2016) has shown that the Great Plains low-level jet is at times more influential than the sea breeze in inducing an inertial oscillation in HGB and even over the open Gulf of Mexico. Also important in producing high ozone in HGB are cold front passages (Langford et al., 2009; Ngan and Byun, 2011; Rappengluck et al., 2008), though it is not yet clear how mesoscale circulations evolve during such events and interact with the changing larger-scale weather patterns.

Figure 1 is a composite analysis of mean air parcel trajectories under light wind conditions at 500 m above ground level, based on TexAQS-II profiler observations. The classic daytime stagnation and recirculation under light wind conditions are apparent in the HGB profiler observations from La Porte. The composites from other profilers in the region show that this phenomenon is not unique to Houston, although it is perhaps more important in HGB than elsewhere because of the concentrated volatile organic compound (VOC) emissions from chemical processing facilities on or near the Houston Ship Channel. The inertial oscillation is simultaneously a mechanism for concentrating emissions and a strong modifier of regional transport patterns.

Modern ozone source apportionment models such as the Ozone Source Apportionment Technique (OSAT) in CAMx (ENVIRON, 2015) and Integrated Source Apportionment Method (ISAM) in the Community Multiscale Air Quality (CMAQ) model (Kwok et al., 2015) can be used to differentiate ozone from long-range transport, local emissions, and adjacent regions. In these models, many non-reactive tracers are used to keep track of the amount of NO_x and VOCs emitted from different sources and/or source-regions. The in-situ ozone formed at each model

time step is then attributed to different sources and/or source regions based on the ozone formation sensitivity regime and the NO_x and VOC source contributions. For example, Kemball-Cook et al. used CAMx-OSAT and determined that both local and regional ozone are important in Houston and Dallas. In the Houston area, however, when recirculation happens aged emissions from the Houston area can reenter the area and contribute to high ozone events (Pierce et al., 2009). Ensemble Lagrangian trajectories were used to identify potential source regions of transported background ozone. However, the computation of Lagrangian trajectories have large uncertainties and cannot fully account for entrainment. To fully understand the timescale of this recirculation that affects the Houston area and the amount of aged pollutants and O_3 re-entered through recirculation, O_3 and its precursors with different atmospheric age (i.e. time spent in the air since release) should be quantified but the models tracking ozone precursors based on their emission locations and source sectors cannot differentiate the influence of freshly-emitted and aged local emissions on ozone.

The overall objective of this research is to synthesize existing data, previous analyses, and photochemical model experiments to provide a comprehensive and reconciled description of how mesoscale and synoptic-scale winds affect dispersion and accumulation of air pollutants emitted in the Houston area and from other regions, and how they contribute to high ozone events. The relationship between surface winds and boundary-layer mesoscale transport features will be clarified, and a novel source- and age-resolved regional air quality model with ozone source apportionment capability will be applied to investigate selected high ozone events with and without mesoscale circulations.

Chapter 2 focuses on using the long-term radar wind profiler observations from the Houston and Dallas areas to determine how the local-scale winds are affected by or even driven by large-scale winds, to identify the large-scale winds which, when combined with local-scale winds, lead to stagnation or recirculation, to document the vertical structure of these wind patterns, and to assess the implications of these winds for nighttime pollutant transport. Chapter 3 describes the development of a modified CMAQ with capability of determining the atmospheric age of ozone and its precursors and the testing of the age resolved model for a high ozone episode in the Houston area from August 20 to September 6, 2000, which is part of the Texas Air Quality Study 2000. Chapter 4 evaluates the capability of the Weather Research and Forecast (WRF) model in reproducing the recirculation of the sea breeze in the Houston area for three selected high ozone episodes. Applications of the source and age-resolved CMAQ model in analyzing the impact of synoptic wind and sea breeze on ozone concentrations for these three cases are also reported in this chapter.

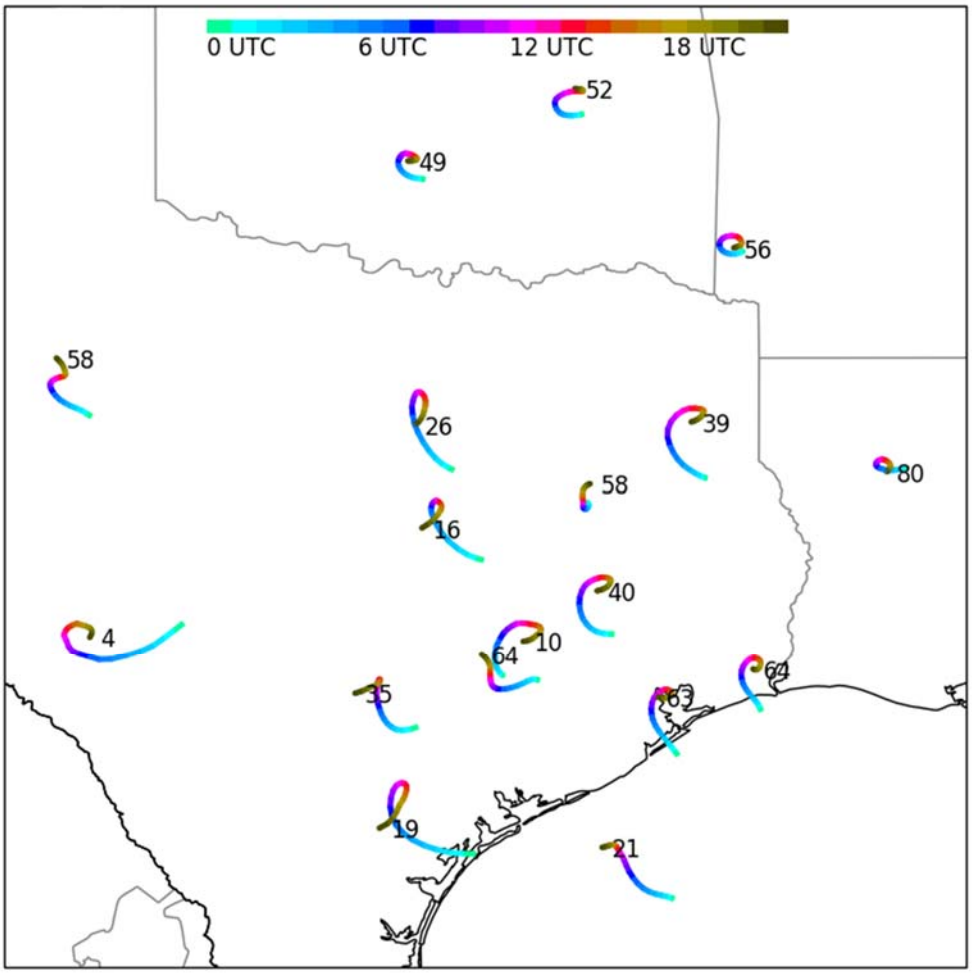


Figure 1.1 Composite analysis of mean air parcel trajectories under warm-season light wind conditions at 500 m above ground level, based on TexAQS-II profiler observations. Colors (bar at the top) correspond to the time of day (subtract six hours for LST), while numbers indicate the number of observed days meeting the low wind criterion.

2. Synthesis of mesoscale wind structures in synoptic-scale context

Many polluted cities lie within basins surrounded on many sides by mountains. The mountains inhibit transport, causing natural stagnation and pollutant buildup. Except for El Paso, the situation in Texas is different. The relatively low-relief topography cannot trap pollutants, so high pollution buildup requires wind stagnation to be brought about via other means. One such potential mechanism is large-scale stagnation. If a city is sitting directly under a high-pressure center, winds will be light and accumulation rates will be rapid.

Previous studies of Houston air quality have identified another mechanism: local wind structures that can bring about stagnation or double-dosing (injection of pollutants into the same air at different times) even when the large-scale winds are not stagnant. This greatly expands the number of days in which meteorology favors high ozone concentrations over and above what large-scale weather systems alone can provide. These local-scale winds predominantly follow a diurnal (daily) cycle. Proper prediction and simulation of ozone episodes require accurate understanding and modeling of such local-scale winds. Furthermore, the proper prediction and simulation of transport of pollutants from one airshed to another requires an understanding of the combination of wind conditions that lead to intact or sheared overnight plumes.

The purpose of the portion of the project described here is to use long-term radar wind profiler observations from the Houston and Dallas areas to determine how the local-scale winds are affected by or even driven by large-scale winds, to identify the large-scale winds which, when combined with local-scale winds, lead to stagnation or recirculation, to document the vertical structure of these wind patterns, and to assess the implications of these winds for nighttime pollutant transport.

This section is organized as follows. Section 2.1 discusses the observational data set and its processing. Section 2.2 lays out the basic physics of the sorts of diurnal wind variations that can commonly occur in Texas. Section 2.3 documents the observed diurnal wind variability, including its dependence on large-scale wind and season, and diagnoses the specific causes of this variability. Section 2.4 specifically considers the large-scale and seasonal settings of the combinations of large-scale and local-scale winds that produce stagnation and recirculation. Section 2.4 also addresses differences in those observations at different profilers. Section 2.5 considers the nighttime transport issue, discussing relationships between nighttime wind shear on large-scale winds, local winds, and whatever winds were present the previous afternoon.

2.1. Data and Methods

2.1.1. Data

The data for this study consists of radar wind profiler data from three sites: La Porte (LPT), University of Houston Coastal Research Center (HSN), and Cleburne (CLE). LPT and HSN are located about 30 km east-southeast and 45 km southeast of downtown Houston, respectively, while CLE is about 80 km southwest of downtown Dallas. All three of these profilers are boundary-layer profilers. Unlike the National Profiler Network, the three profilers have enough vertical resolution (and sufficient proximity to Houston and Dallas) to examine local circulations in detail. Radar wind profiler data from Beaumont was also collected but was not used in this study because of its distance from Houston.

The data were downloaded from the Meteorological Assimilation Data Ingest System (MADIS), hosted by the National Centers for Environmental Prediction (NCEP). The data period requested was January 1, 2005, through August 31, 2018. None of the profilers had complete data during that period. LPT had the equivalent of about four years of data, HSN five

years, and CLE nine years. The data from LPT and HSN partially overlap, so that there are about seven years of observations in total from the Houston area.

2.1.2. Methods

Observations are available at hourly or half-hourly intervals and can be interpreted as average conditions for the period between the previous observation time and the current observation time. Each observation consists of header information followed by nominal pressure, altitude, temperature, humidity, wind speed, and wind direction. Temperature data may be provided by a radio-acoustic sounding system (RASS) but was not used in this study.

The profiler winds were interpolated to a regular vertical grid, with levels at increments of 50 m below 800 m above ground level, then increments of 100 m to 1600 m, and then increments of 200 m up to 3000 m. Ignoring data levels with temperature (since RASS does not measure winds, wind data is always missing at RASS levels), wind components were linearly interpolated between adjacent data levels. If one or both adjacent data levels had missing winds, the interpolated winds were set to missing at this stage. Also, if no winds were reported either above or below the grid level, the grid level winds were set to missing.

For some periods, interpolation yielded winds at some grid levels that were systematically different (weaker or stronger) than winds at adjacent levels. Closer inspection revealed that the wind data were reported on two separate sets of observation grids: a lower-resolution set and a higher-resolution set, and one set had systematically stronger or weaker winds than the other. The two sets were interleaved according to altitude in each observation report. Since it was not known which set was more reliable for local wind purposes, both sets were used for interpolation according to the nearest observation levels to the grid level. The resulting artifacts in the profiler composites serve as a reminder of the magnitude of errors associated with radar wind profiler measurements.

For creating composites or averages of profiler observations, it is essential that data be present at all levels and times. Gaps in data can lead to additional artifacts when, for example, data is present on a high-wind day on only some of the levels. Since the basic unit of analysis of diurnal wind variations was a day, each day (defined as 0000 Coordinated Universal Time (UTC) of one day through 0000 UTC of the next day, with 0000 UTC being 7 PM Central Daylight Saving Time (CDT) or 6 PM Central Standard Time (CST)), days were deleted if there were not 18 hours with at least one wind observation within the layer 1000 m – 1500 m, inclusive. These criteria were chosen to ensure that data would be present up to the top of the typical daytime boundary layer and nighttime residual layer in most cases.

The next step was to determine which levels to exclude. A grid-level was excluded (i.e., all observations set to missing) if there were less than 25% non-missing observations over the available data. Remaining grid levels extended from 150 m to 2000 m for LPT and CLE and 250 m to 2000 m for HSN.

The final step was to perform interpolation to fill in missing data on the remaining included levels. Linear interpolation in time will tend to suppress diurnal wind variations if the gap is more than a few hours long. To avoid this, interpolation was performed using principal component analysis (PCA), also known as empirical orthogonal functions (EOFs). PCA interpolation is common in oceanography to fill spatiotemporal gaps in surface temperatures while being true to the structures of common modes of variability such as El Niño. In the profiler case, it was desirable to retain the structures of diurnal and higher-order modes of variability.

The process proceeded as follows. A first guess of missing values was supplied through linear interpolation. Then PCA was performed to identify the leading EOF. Then, on each day, the EOF loading (amplitude) was calculated using only points with data, and the resulting EOF structure and amplitude were used to obtain new guesses for the missing values. The procedure was repeated nineteen times or until the updates to the missing data were below a specified size threshold. Next, PCA was performed and two EOFs were retained, missing values were updated, and the process repeated. The same succession of steps was performed until analysis using ten EOFs had been completed. In order to avoid nonrobust PCA at the starting and ending times of the day, the same analysis was performed simultaneously using data shifted by 12 hours and the missing data were estimated using a weighted average of the non-shifted and shifted data that favors data from near the middle of a 24-hour period.

Screening of days was done using winds at 400 m above ground level (AGL). A weak wind day had an average 400 m wind speed of less than 4 m/s, while a strong wind day had an average wind speed of more than 10 m/s. A steady wind day had a vector difference of less than 4 m/s between the 400 m wind at the beginning of the day and the 400 m wind at the beginning of the next day. (We call this a trend of less than 4 m/s.) The perturbation (local) wind is defined as the wind with its mean and trend removed. The u and v wind components are direct toward the east and north, respectively, while onshore is defined as toward 330° and alongshore as toward 60°.

All coding described in this session was performed using Python, generally embedded in Jupyter notebooks.

2.2. Physics of Diurnal Variability

2.2.1. The Inertial Oscillation

Diurnal variability can arise through processes that affect the atmosphere on a daily basis as well as physical situations that support an intrinsic diurnal frequency. In other words, the atmosphere can either be nudged once a day or it can swing back and forth once a day on its own. If the atmosphere is capable of swinging back and forth on its own, daily nudging can produce a resonant, high-amplitude response. This is the case in Texas, particularly near the latitude of Houston.

The cause is something called the inertial oscillation. If air is moving but there are no horizontal pressure gradients, the air will tend to rotate in a complete circle due to the rotation of the Earth. The length of time it takes the air to complete that rotation is dependent on latitude. At 30°N or 30°S, the period of rotation is exactly one day.

In the real world, the horizontal pressure gradient is almost never exactly zero for long. Yet, the air will always tend to execute a day-long loop while generally moving in the appropriate direction, much like a spot on the side of a rolling wheel keeps going around in circles as it moves along.

This sort of motion is always clockwise in the Northern Hemisphere. Other important processes in Texas do not lead to daily local wind rotation preferring one direction over the other. A plot of the sign and coherency of the observed wind rotation at LPT (Fig. 2.1) illustrates the overwhelming tendency for clockwise wind rotation caused by the inertial oscillation.

2.2.2. Causes of Diurnal Forcing

The tendency for inertial oscillation means that just about any disturbance or change to the air pressure distribution will trigger an oscillation. However, if atmospheric disturbances are occurring with daily regularity, they will resonate with the inertial oscillation and amplify it.

One such trigger is the daytime growth and evening collapse of the boundary layer. Each day, the growth of the convective boundary layer over land means that air from the ground is constantly and rapidly being carried aloft and being replaced by other air within the boundary layer. This causes the wind, water vapor mixing ratio, and virtual potential temperature to tend toward uniform conditions within the boundary layer. This mixing imposes an effective horizontal force on the winds due in part to surface drag. When the mixing ceases in the late afternoon or early evening, the forces are suddenly out of balance and an inertial oscillation ensues.

Another trigger is the daytime heating contrast across the coast. When the air over land warms up while the air over water maintains its temperature, a time-varying pressure gradient is produced. Air responds to this pressure gradient by, in part, undergoing an inertial oscillation, though other structures such as sea breeze fronts can develop as well.

2.2.3. Conditions Favoring Diurnal Variability and Wind Stagnation/Recirculation

Diurnal variability is largest when the wind disturbance is largest. In the case of diurnal variability triggered by boundary layer growth and collapse, the maximum oscillation occurs when the strongest pressure gradient is at the bottom of the atmosphere. When the daytime boundary layer develops, this air is slowed both by surface drag and by mixing of lower-momentum air from above. When the mixing stops in the evening, the imbalance at low levels is largest, and the combination of strong pressure gradient and large inertial oscillation can cause the wind speed to be largest near the ground: a low-level jet.

Such a situation develops when the air is warmer to the left of the wind vector than to its right. Along the Texas coast, this can occur in the summertime when the wind is blowing parallel to the coast toward the northeast, and in the wintertime when the wind is blowing parallel to the coast toward the southwest. More broadly, the topography and rainfall patterns of the Great Plains and the Rocky Mountains facilitate just such a situation when the winds are from the south, as they commonly are in the summertime. This produces a well-known phenomenon called the Great Plains low-level jet.

This type of diurnal oscillation can be maximized through any of the three mechanisms. First, a larger surface wind speed produces a larger oscillation because the time-varying effect of surface friction is greater. Second, a strong horizontal temperature gradient, oriented properly with respect to the wind direction, produces a larger oscillation because the daytime vertical mixing of momentum is larger. Third, dry surface conditions lead to a deeper daytime boundary layer, which allows for more mixing of low-level momentum with air aloft.

Indirectly, the time of year also contributes, because the greater solar heating in the summertime also leads to a deeper daytime boundary layer.

With the sea breeze mechanism, the maximum oscillation occurs when the heating is largest, so the sea breeze is also favored in the summertime. Dry conditions on land would lead to faster, larger heating and thus a larger sea breeze response. It is expected that light winds favor a strong sea breeze by virtue of the resonance effect, as air is hit by the coastal pressure gradient day after day.

Nielsen-Gammon (2016) showed that the inertial oscillation responses do not just happen locally. Indeed, he found evidence that much of the inertial oscillation observed in the atmosphere over the Gulf of Mexico was caused by the low-level jet overland rather than by the sea breeze.

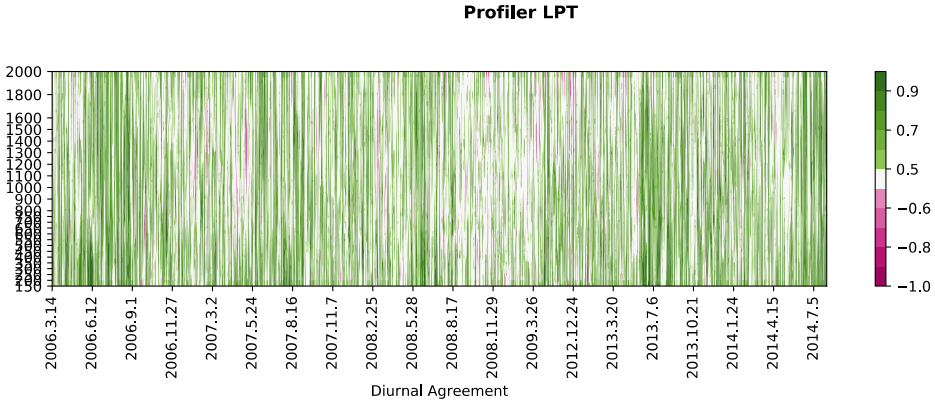
It is difficult to get wind stagnation locally via the low-level jet mechanism. One would need to have so much variation of wind with height that the wind direction at low-levels at night would be opposite the direction of the average daytime wind. However, nonlocal stagnation is

possible. Because of remote effects of the Great Plains low-level jet, it is possible that the resulting wind oscillation can be felt (and affect ozone) as far away as Houston, where wind speeds are generally weaker.

Stagnation and recirculation associated with the sea breeze are much more likely than with the low-level jet. Because the coastal wind oscillation is favored when the large-scale winds are light, the local winds are stronger when they need to be to force recirculation.

2.3. Observed Diurnal Variability

To identify strong diurnal variability in the observation record, two metrics are used here. The first, introduced in Figure 2.1, is the correlation between the smoothed, time-shifted northward wind and the eastward wind, with positive values indicating the expected clockwise rotation. The second is the amplitude of the projection of the time-shifted northward wind onto the eastward wind.



Figu
shift
-1 in
tend
stror
to be
ference between winds at the start of one period and the start of the next) are shown.

Figure 2.1 showed a tendency for stronger correlations during the summertime at LPT. Figures 2.2 and 2.3 show similar information for HSN and CLE. The second panel in each figure gives a measure of the amplitude of the diurnal variation. All three have correlations and amplitudes (not shown for LPT) that peak in the summertime. The most noticeable difference between CLE and the other two profilers is that the diurnal variation at CLE is predominantly below 1200 m, whereas it extends up at least to 2000 m at LPT and HSN.

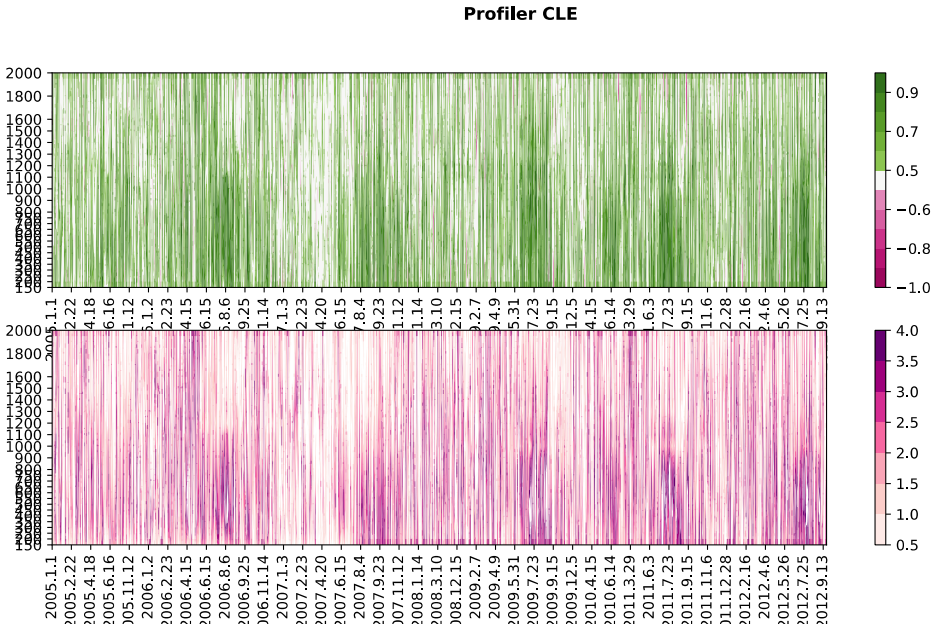


Figure 2.2 Time-height section of correlation, as in Figure 2.1, and of the amplitude of the local diurnal wind variation (m/s, bottom panel), for CLE. The correlations are predominantly positive, and the strongest correlations tend to be below 1200 m AGL in the summertime. Diurnal wind amplitude tends to be largest where and when the correlations are largest, with some amplitudes exceeding 4 m/s.

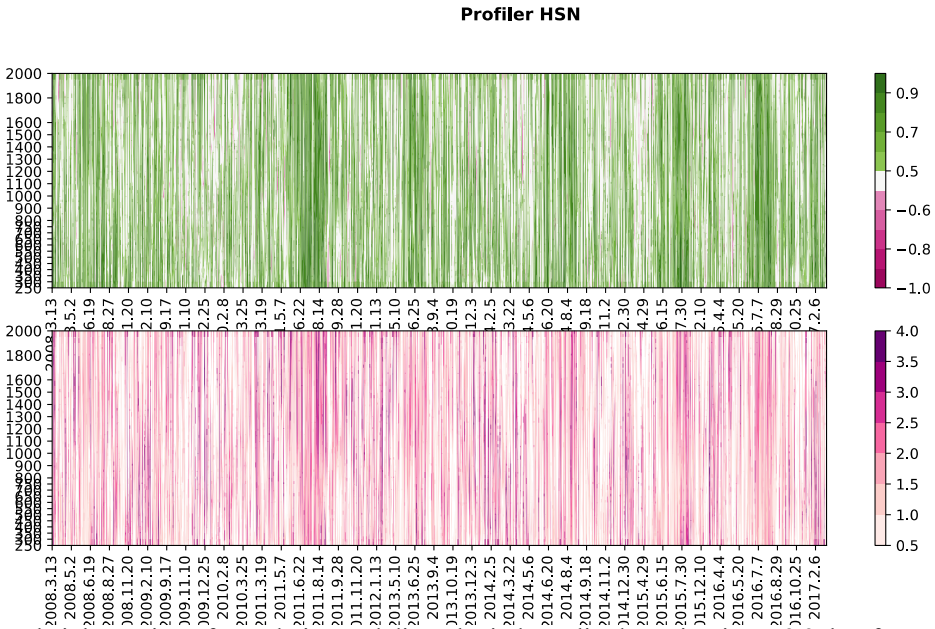


Figure 2.3 Time-height section of correlation and diurnal wind amplitude, as in Figure 2.3, but for HSN. The vertical structure and seasonality of high correlations are similar to LPT (Figure 2.1), and amplitudes tend to be largest where the correlations are largest. Amplitudes are generally not as large as at CLE.

The seasonality in amplitude and structure is more clearly seen in composite analyses of the nearly-steady cases. Figures 2.4 and 2.5 show the warm season (Apr.-Sep.) and cool season (Oct.-Mar.) composite for CLE. The magnitudes of the perturbation winds are about twice as

large in the warm season as in the cool season. The magnitudes peak at around 500 m and are considerably weaker above 1000 m. The summer composite features a low-level wind maximum. Changes of phase with height are weak or nonexistent.

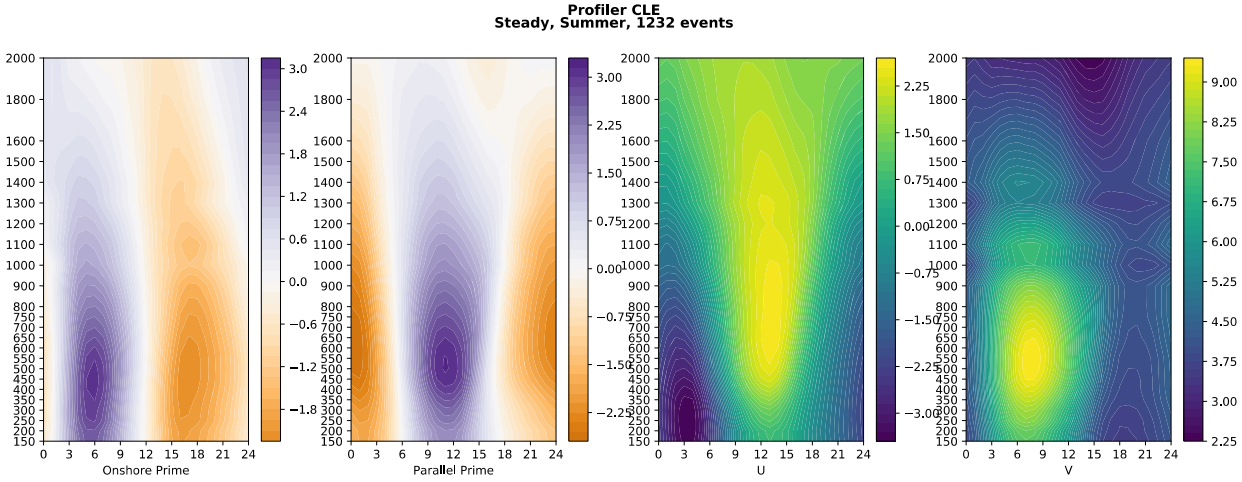


Figure 2.4 Composite of warm-season steady days at CLE. The left panels show the perturbation (local) onshore and shore-parallel winds, while the right panels show the total u and v winds. The perturbation amplitudes peak around 500 m, with an approximately 6-hour offset between the onshore peak and shore-parallel peak. The total wind is on average strongest between 0600 and 0900 UTC (0000-0300 LST) at about 600 m, close to the perturbation amplitude peak and indicating the composite presence of a low-level jet. The wind perturbations have a similar phase through the lowest 1400 m.

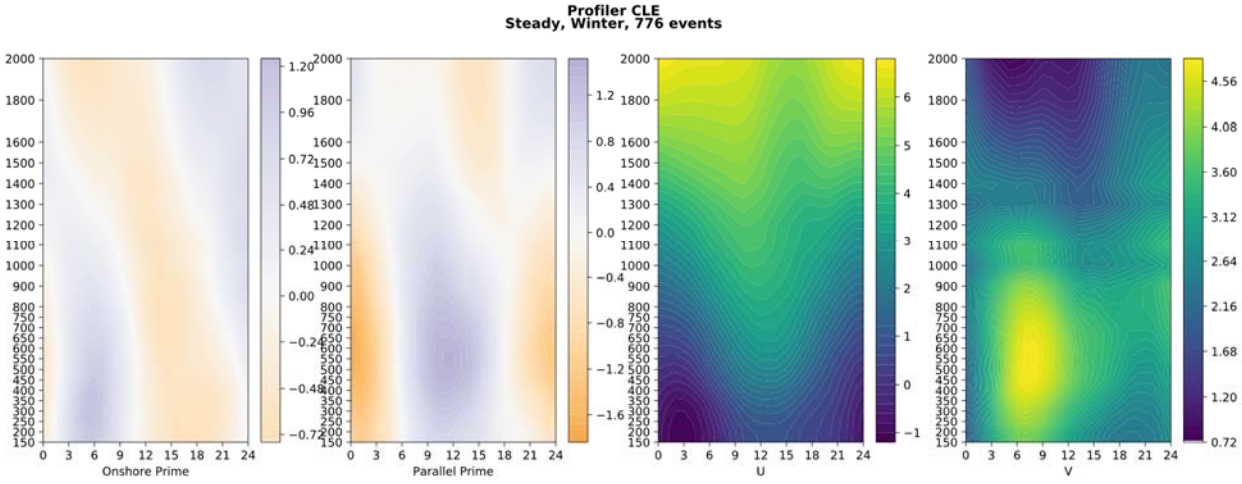


Figure 2.5 Composite of cool-season steady days at CLE. The phasing of the diurnal cycle is similar to that of the summer composite, but the amplitude is about half as large.

The winds at LPT are similar with respect to seasonality but dissimilar with respect to amplitude and vertical structure (Figures 2.6 and 2.7). The LPT local winds are weaker, but more importantly, they are of similar magnitude at all levels, with a peak around 1400 m. Also, the local winds change phase with height, so that peak northward perturbation flow occurs about seven hours later at 1400 m in the composite than at the surface.

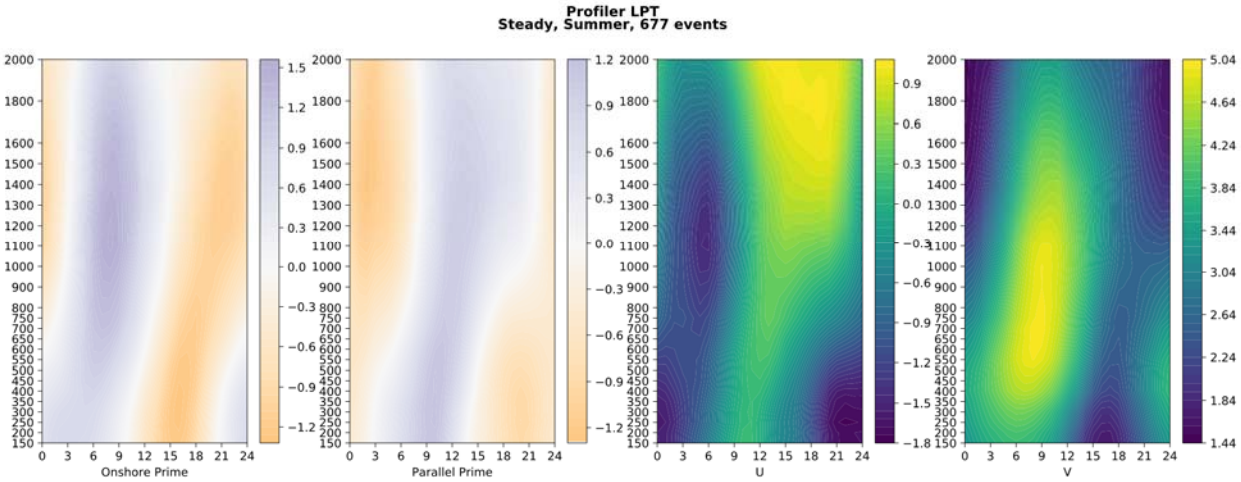


Figure 2.6 Composite of warm season steady days at LPT. The diurnal amplitude is similar to the cool season diurnal amplitude at CLE, but the largest amplitude is around 1400 m AGL, well above the CLE peak. Also, the LPT local wind phase tilts strongly with height, so that the low-level (150 m) phase leads the 1400 m phase by about six hours.

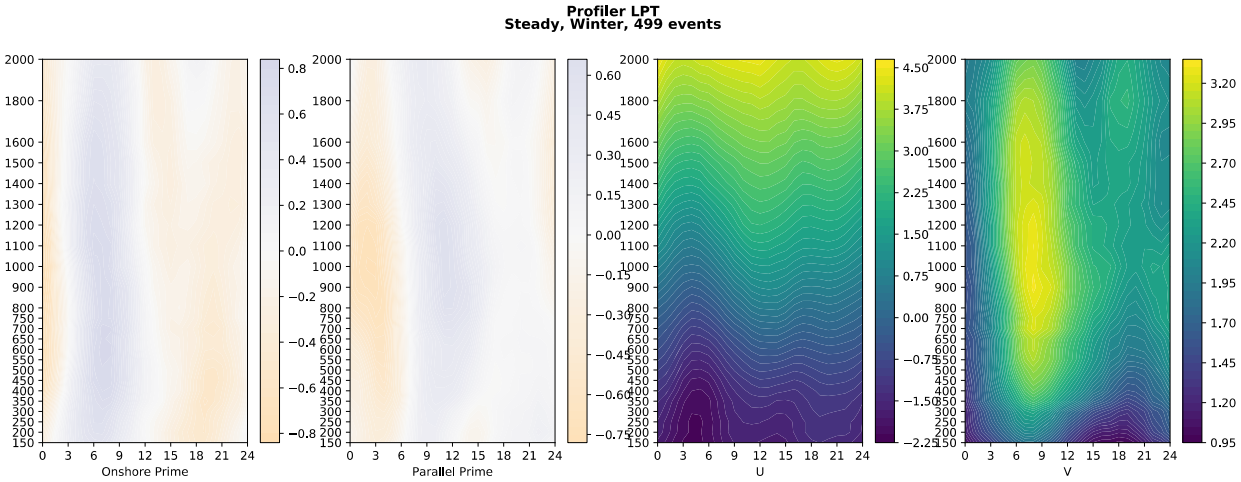


Figure 2.7 Composite of cool season steady days at LPT. Amplitudes are weaker than with cool season days at CLE, but structures are otherwise similar.

For background wind dependence, this report considers warm season winds only because the diurnal signal is strongest there. In the light wind cases (Figure 2.8), weak flow from the south is associated with twice as much diurnal variation as weak flow from the north. There is a strong diurnal wind signal in summer with an amplitude exceeding 1.5 m/s.

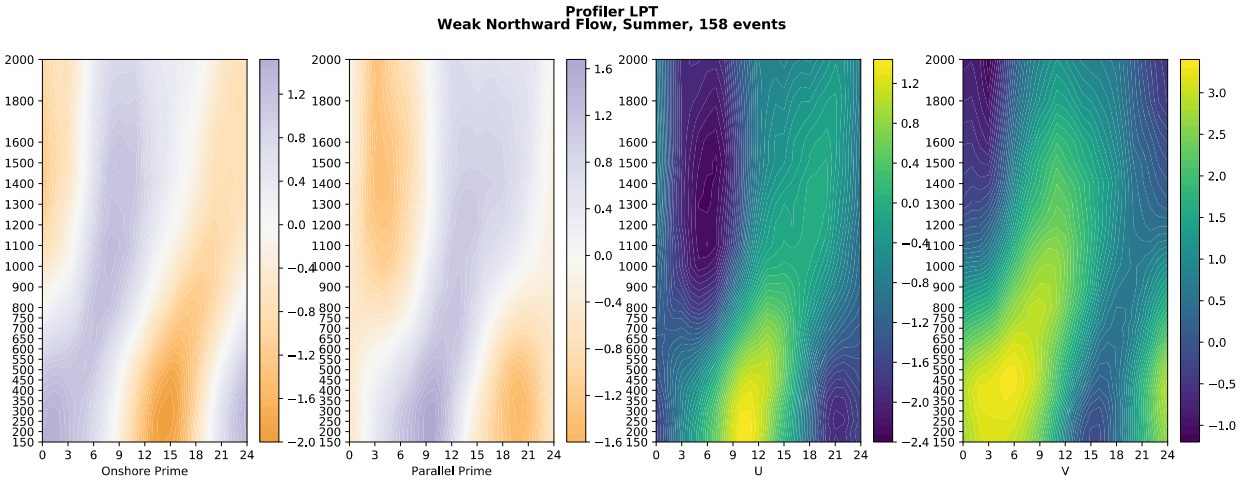


Figure 2.8 Composite of weak wind days (steady, mean 400 m wind less than 3 m/s) at LPT. Top row: weak southward flow. Bottom row, weak northward flow. In the weak wind flow composites, the strongest perturbation winds show up around 250 m AGL and there is a strong and pronounced change in wind structure with height. The strongest offshore flow occurs around 9:00 AM local time near the ground and several hours above 1000 m.

With moderate wind cases (Figures 2.9 and 2.10), background winds with a southward component produce a disorganized, low-amplitude composite. Northward component cases are roughly similar to the weak wind northward case, except with stronger perturbation wind amplitude and less vertical shear.

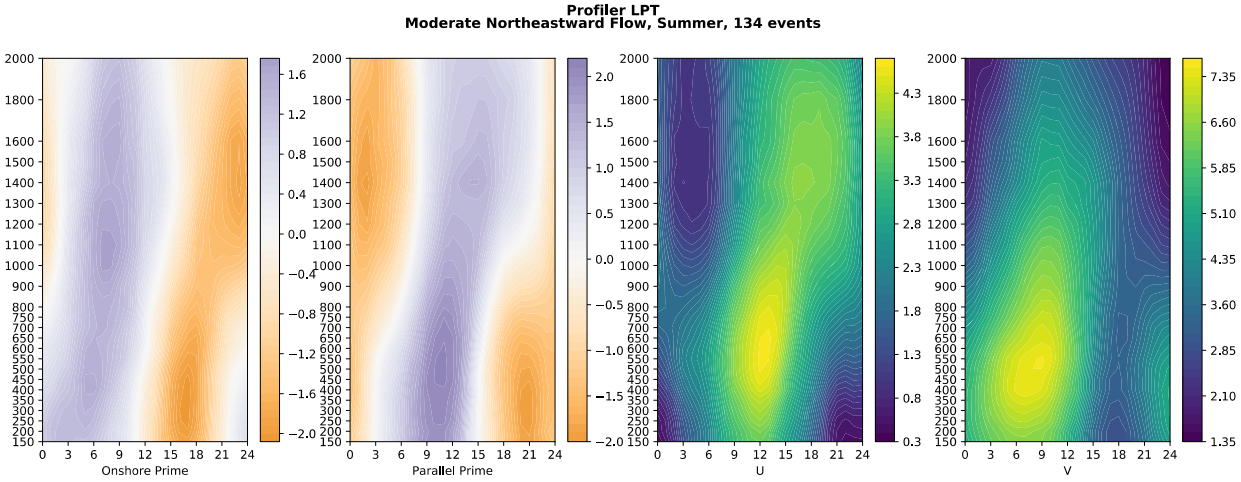


Figure 2.9 Composite of moderate winds toward the northwest (top) and northeast (bottom) at LPT. The patterns are similar to those in Figures 2.8 bottom, except that there is less vertical tilt and the diurnal amplitude has strengthened in particular at around 1400 m.

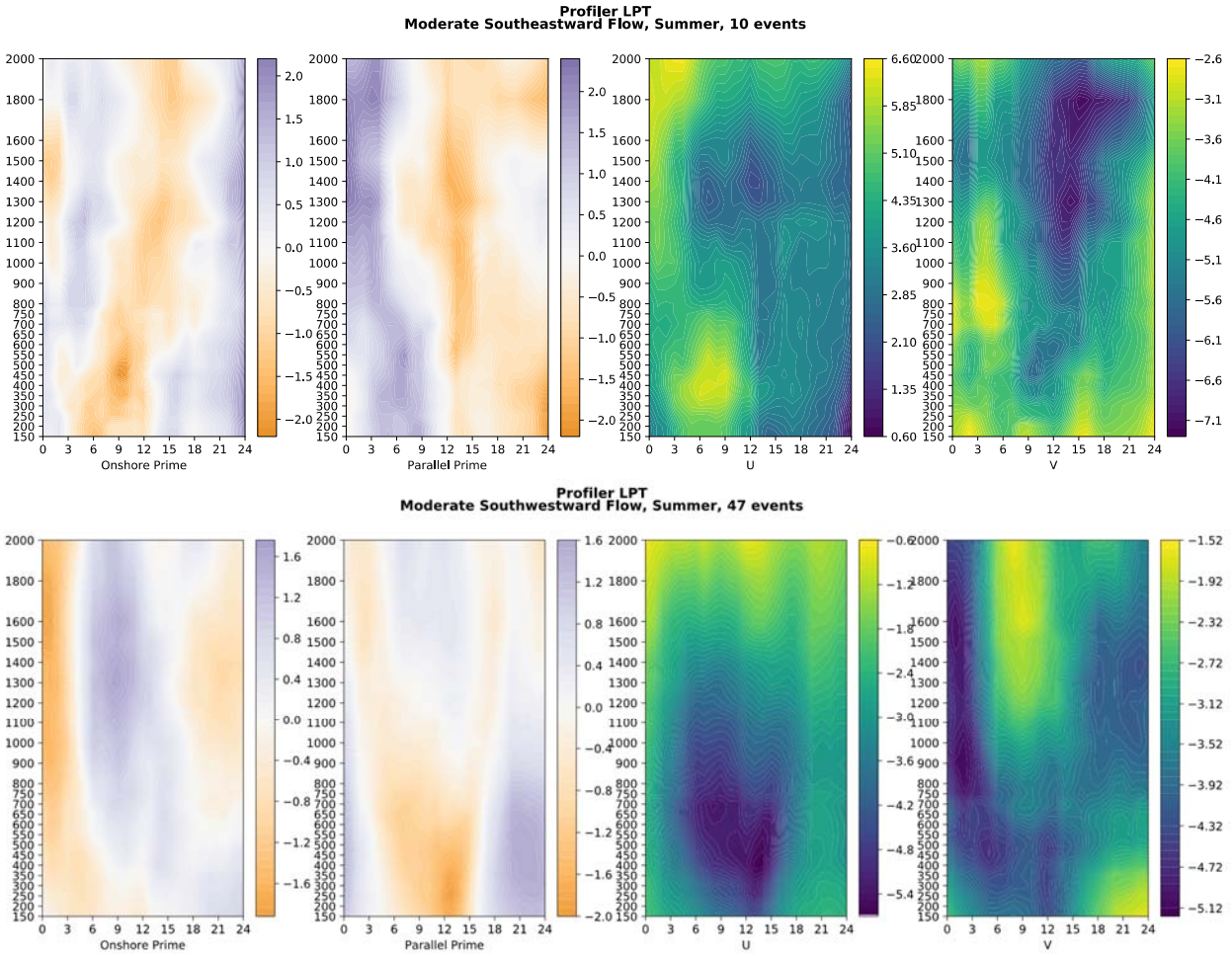


Figure 2.10 Composite of moderate winds toward the southeast (top) and southwest (bottom) at LPT. Most correlations are weak and noisy.

There are too few strong wind cases during summertime from the north to construct a meaningful composite. With strong winds from the south at LPT there is moderate tilt and again moderate amplitude (Figure 2.11).

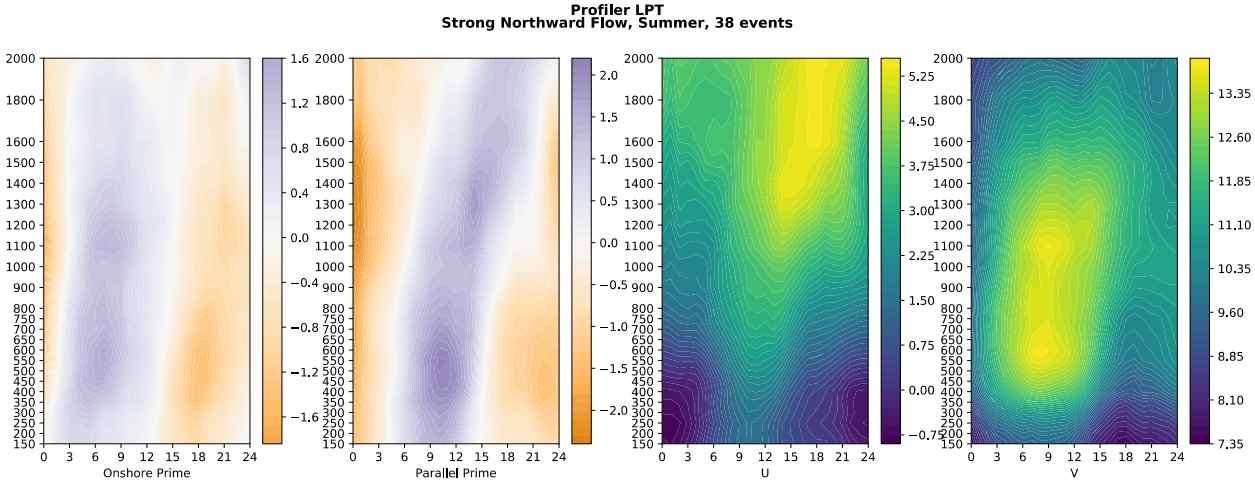


Figure 2.11 Composite of strong winds toward the north at LPT. Correlation pattern weaker than in the moderate northward flow cases, but otherwise a close match.

The two weak wind composites for CLE (Figure 2.12) have a stronger amplitude than the corresponding LPT composites. In the composite with winds from the south, there is essentially no vertical tilt in phase. The strong CLE composites (Figure 2.13) show the classic low-level jet structure with a strong flow from the south. The southward flow case still has substantial amplitude, but there are few events so the phase tilt here (opposite to that at LPT) may not be robust.

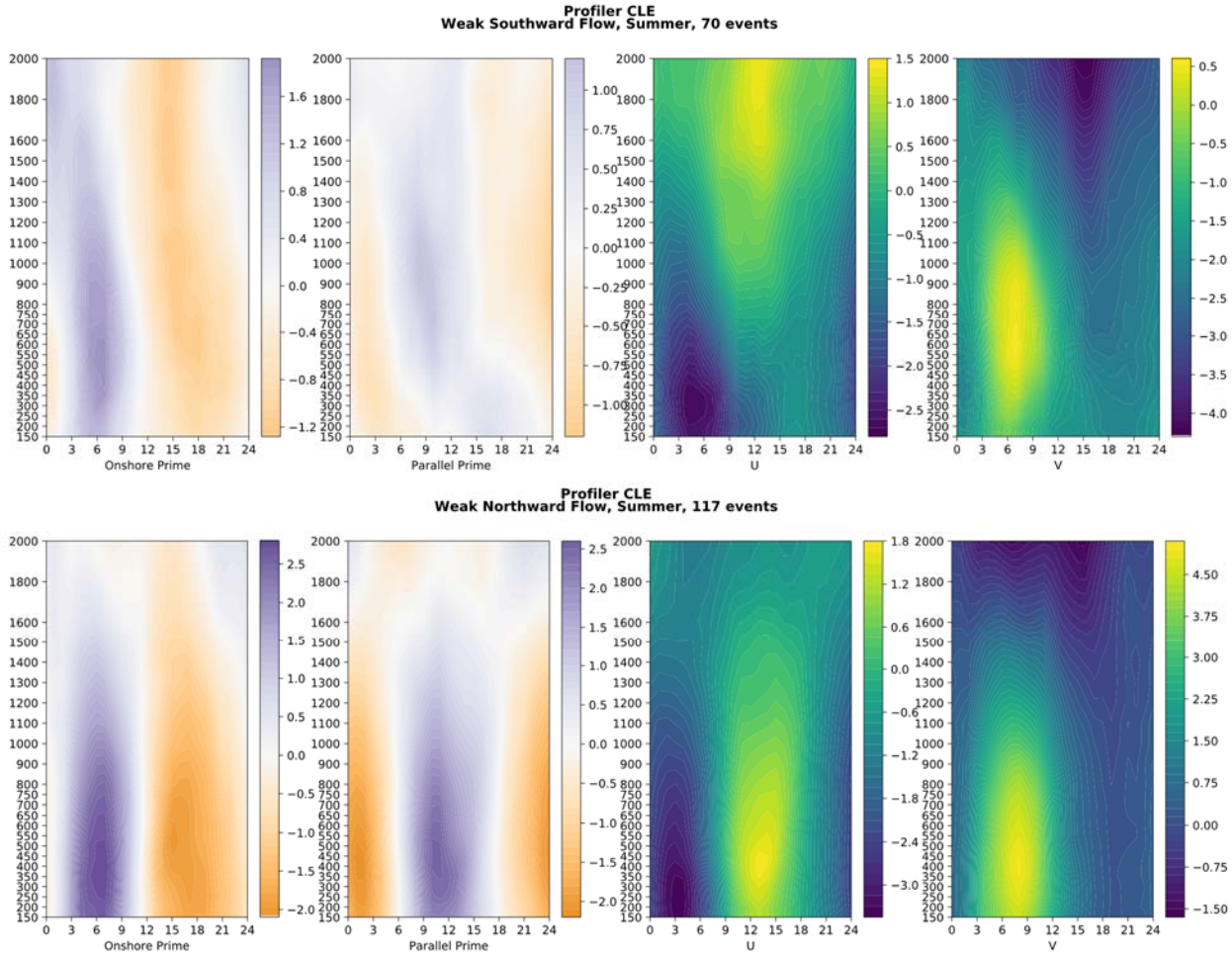


Figure 2.12 Weak wind composites for CLE. There is almost no vertical wind shear. Both perturbation wind patterns have a slightly larger amplitude than their corresponding patterns at LPT (Fig. 2.8).

At CLE, the lack of vertical tilt, the tendency for larger perturbations with the flow from the south, and the increase in perturbation amplitude with an increase in wind speed all point to the Great Plains low-level jet as the driver of the diurnal wind variations.

In Houston, the situation is more nuanced. The strongest diurnal cycle occurs under light to moderate wind conditions, pointing toward coastal heating (the sea breeze) as the driving mechanism. However, the presence of diurnal variations at high wind speeds (with low tilt) points to a role for boundary layer growth and collapse, either locally-driven or (as found by Nielsen-Gammon 2016) connected to the larger-scale Great Plains low-level jet. Meanwhile, the preference for a strong diurnal circulation with winds blowing along the coast toward the northeast suggests a hybrid driving mechanism. With such a wind direction in the summertime, the pressure gradient should tend to be largest at low levels, so the boundary layer growth and collapse mechanism should drive a diurnal circulation in tandem with the sea breeze heating contrast.

Even though the diurnal wind cycle in Houston under light wind conditions is primarily a manifestation of the heating contrast that drives the sea breeze, the term "sea breeze" is commonly used to refer to an onshore component of wind that develops in the afternoon, often with a sea breeze front at its leading edge. Henceforth, this report will reserve the terms "bay breeze" and "gulf breeze" to these particular local circulations. The diurnal wind rotation that includes a tendency for onshore flow during the afternoon and evening and offshore flow during the morning will be referred to specifically as the coastal inertial cycle and generically as diurnal wind rotation.

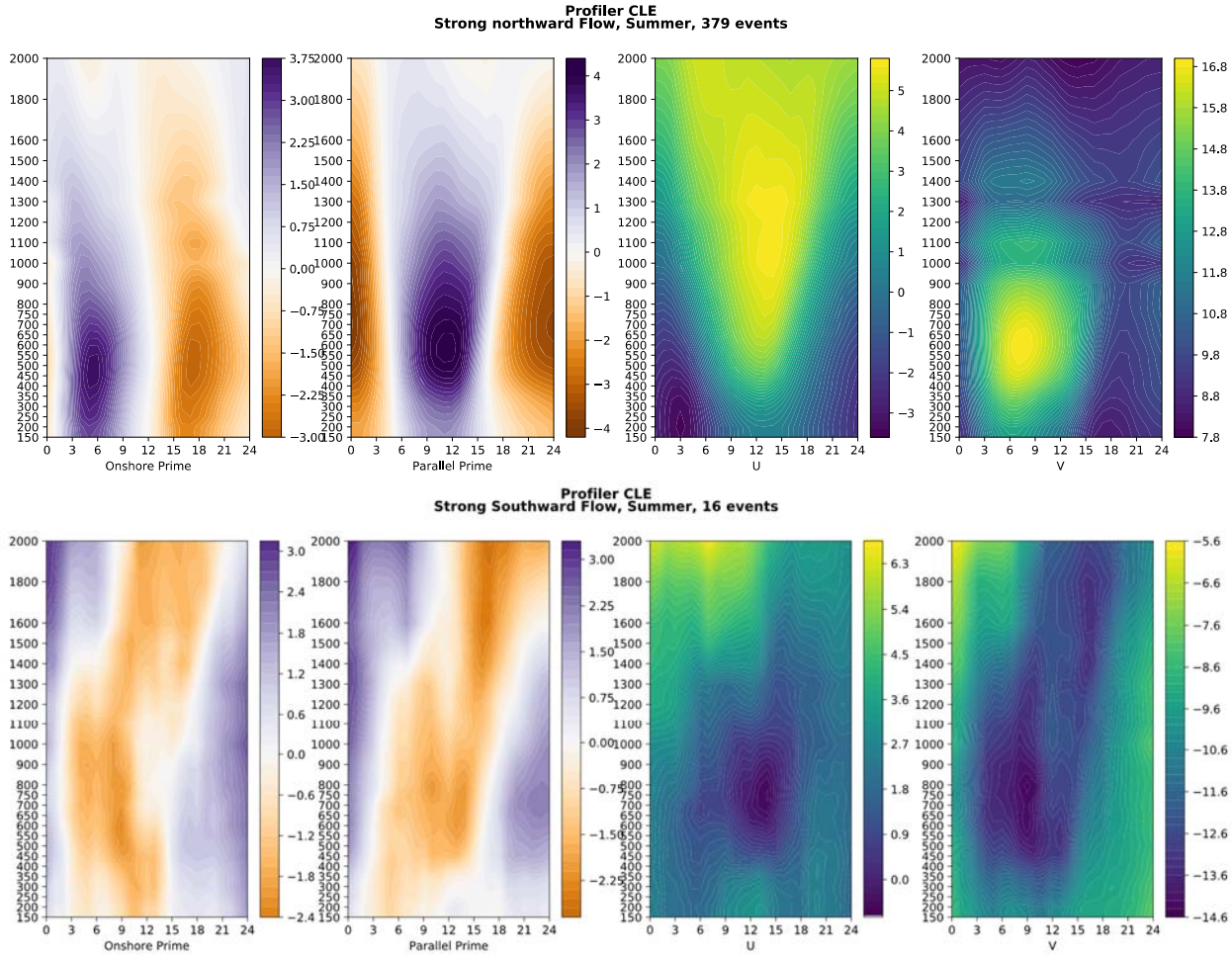


Figure 2.13 Strong wind composites for CLE. The composite for northward flow has a classic southerly low-level jet structure, with maximum diurnal variations at 500-600 m. The southward flow composite is less well organized and has weaker perturbation winds.

2.4. Recirculation

2.4.1. Diagnostic Technique

The previous diagnostic analysis showed how the variation of winds within a day are influenced by the inertial cycle, the sea breeze heating contrast, and the growth and collapse of the boundary layer. These variations are especially relevant if they produce stagnation or recirculation. The enhanced buildup of pollutants caused by stagnation and recirculation is known to have played an important role in many ozone exceedance episodes in the past.

To diagnose stagnation and recirculation, this report utilizes a metric that is based on the amount of stagnation. Maximum pollutant buildup occurs when air moves slowly for a long period of time. To represent this, stagnation over a given interval is computed as the displacement of an air parcel (in km) from the beginning to the end of an interval divided by the length (in hours) of that interval. So, for example, if air moves 20 km in five hours, the stagnation is 4 km/hr. This has units of velocity, except that it is calculated from the net displacement of an air parcel rather than the (often circuitous) track of an air parcel. Smaller numerical values correspond to greater stagnation.

By using net displacement, the stagnation value does not distinguish between the air that moves very slowly over an interval and air that moves farther but in a loop and returns to near its starting point. Thus, the stagnation value also captures the recirculation of air that has been shown to produce double-doses of pollutants and which is the main focus of this meteorological analysis.

For clarity of presentation, the stagnation value is linearly transformed to a scale whereby 10 represents maximum stagnation and 0 represents weak stagnation unlikely to lead to significantly enhanced pollutant buildup. The weak stagnation threshold is set to 5 km/hr. So stagnation of 0 km/hr yields a 10 on the stagnation scale, stagnation of 1 km/hr yields an 8 on the stagnation scale, etc.

The procedure for calculating the stagnation scale value is as follows. First, wind runs are extracted from the interpolated profiler data at 250 m, 300 m, and 350 m. These levels are the lowest levels present in all three profiler data sets. The use of three levels avoids erratic index behavior as recirculating air may or may not happen to return exactly to an earlier location at a discrete hourly interval. Next, the stagnation value is calculated for all possible starting and ending times within a day and averaged across the three levels. The lowest average value becomes the stagnation value for that day. Lastly, the stagnation value is converted to the 0-10 stagnation scale.

Figures 2.14 through 2.17 show examples of back-trajectories at LPT. The back-trajectories are created by assuming air remains at a constant altitude throughout the day and experiences the winds observed by the LPT profiler. The termination point of the back-trajectories (at the origin) is at 1800 LST. These four examples were chosen because they had similar average winds at 400 m but drastically different stagnation scale values: 8.6, 7.6, 0.0, and 0.0 respectively. The low-level trajectories in Figure 2.14 show a near-stagnation point in the early morning, while those in Figure 2.15 show recirculation and slow air motion. Meanwhile, there is no sign of enhanced stagnation in the other two figures, consistent with the stagnation scale values.

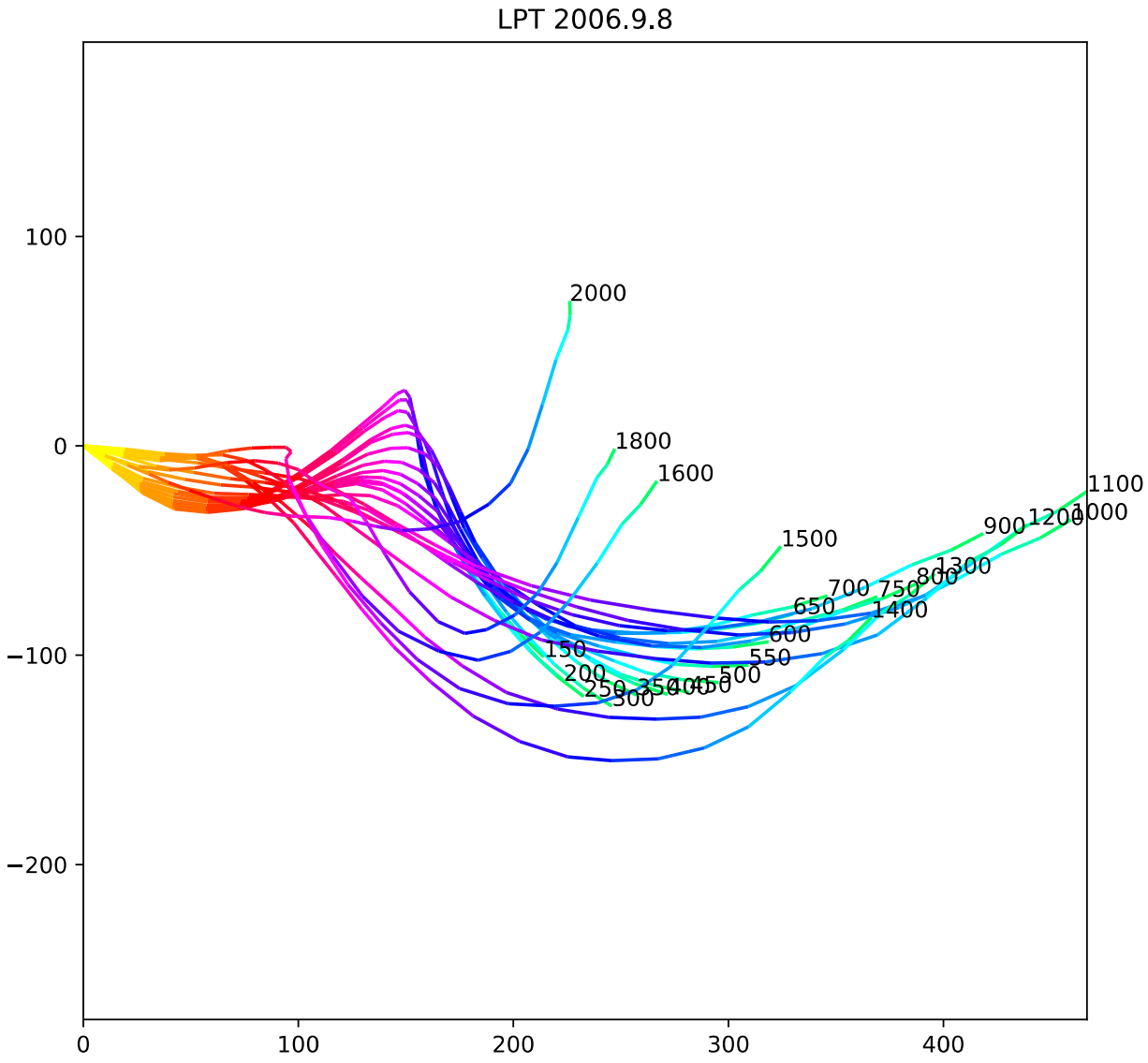


Figure 2.14 Air parcel back-trajectories from LPT, Sept. 8, 2006. Each hour-long trajectory segment is plotted in a different color, with greens corresponding to evening, blues to night, reds to morning, and oranges to afternoon. All trajectories end at 1800 LST. This case featured a large value of the stagnation index, as low-level air came to a near-halt in the early morning. Low-level air initially moved westward in the evening, then curved northward overnight. Stagnation followed around dawn, followed by a renewed motion toward the southwest during morning and northwest during the afternoon. The x and y coordinates give the location in km.

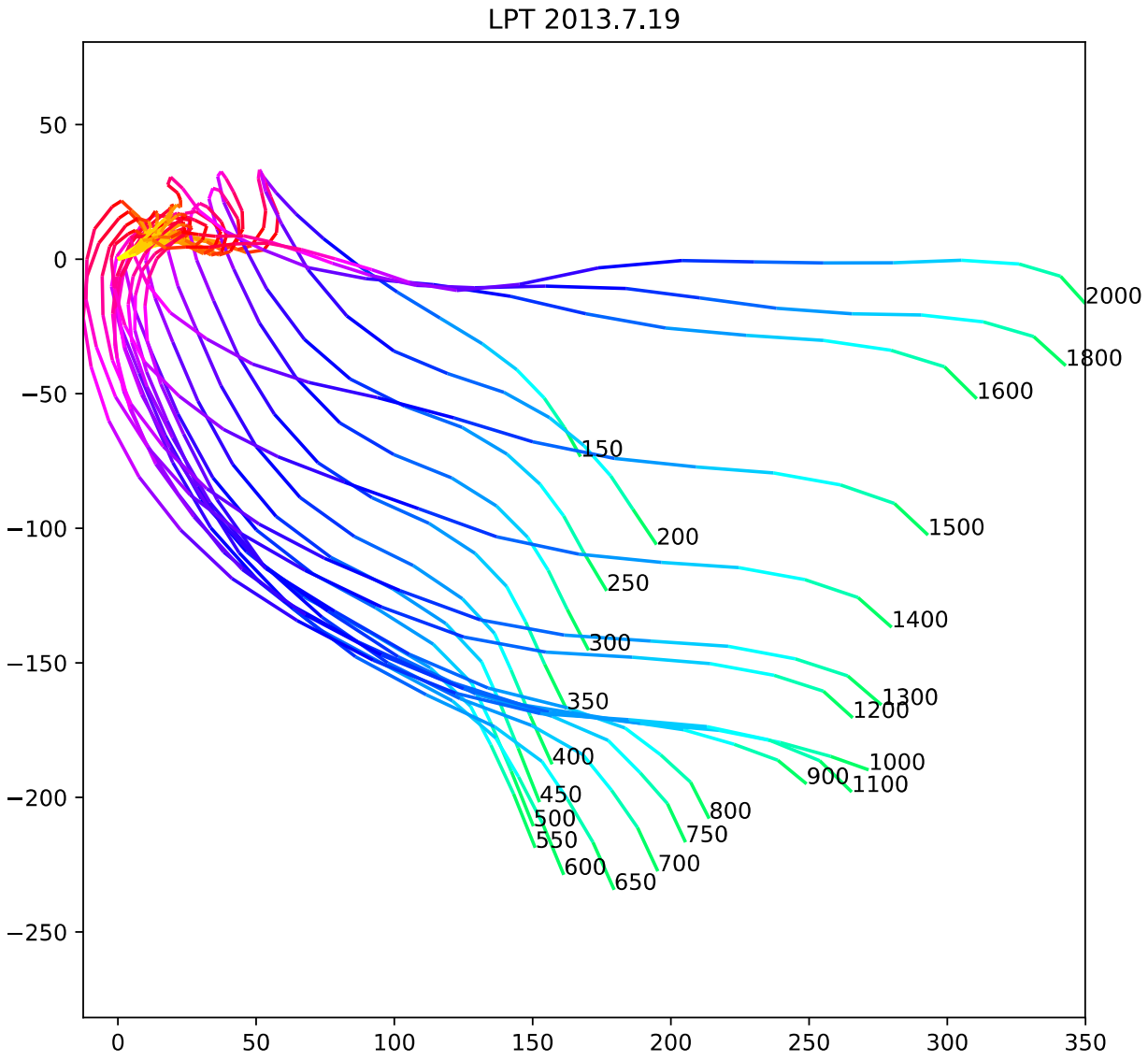


Figure 2.15 July 19, 2013 is another case with high values of the stagnation index. Air moved toward the northwest overnight before executing a complete loop during the early morning and drifting slowly westward during the remainder of the day.

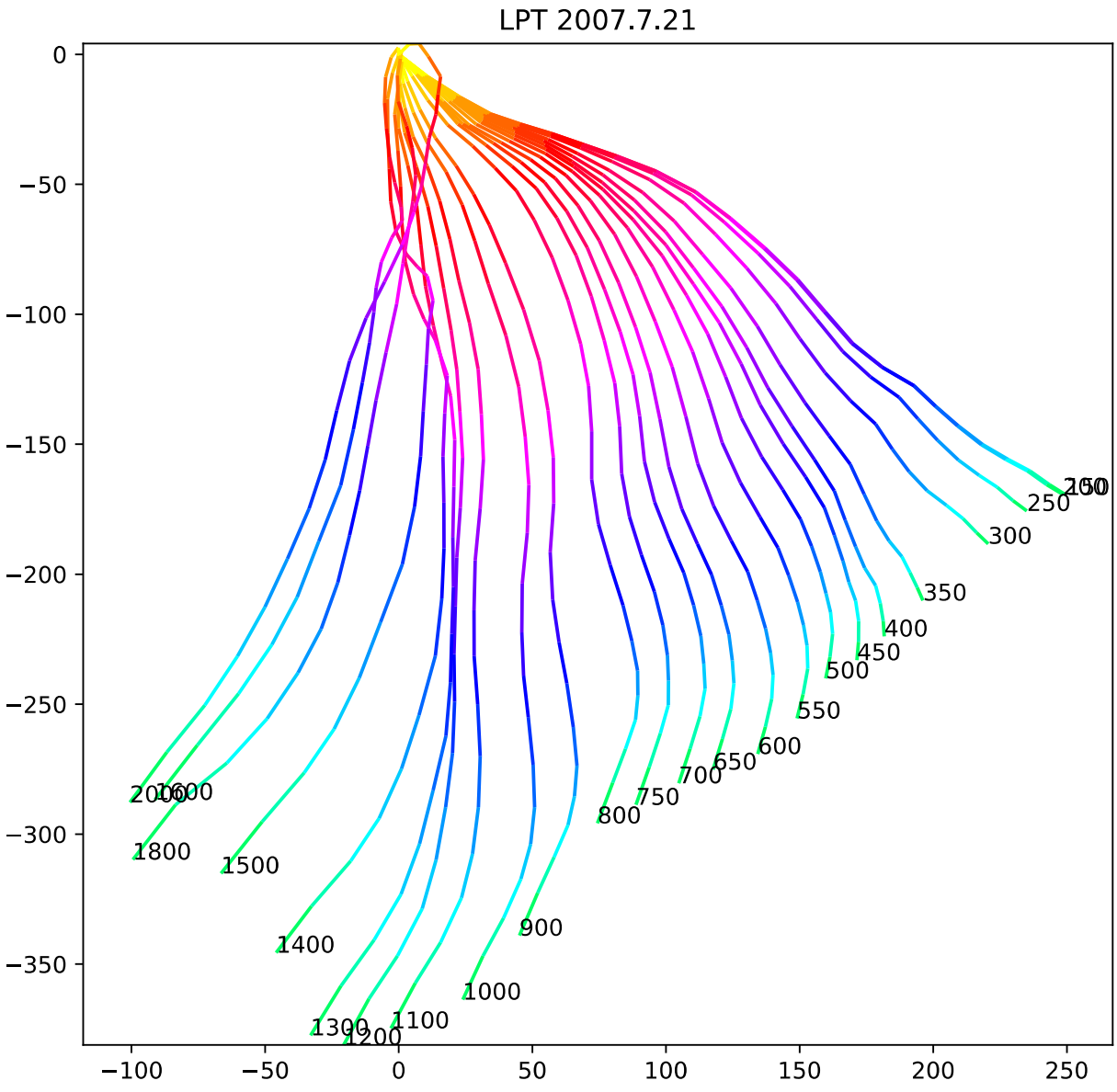


Figure 2.16 Back-trajectories for LPT on July 21, 2007, showing an absence of recirculation or stagnation. Low-level air progressed steadily toward the northwest throughout the day.

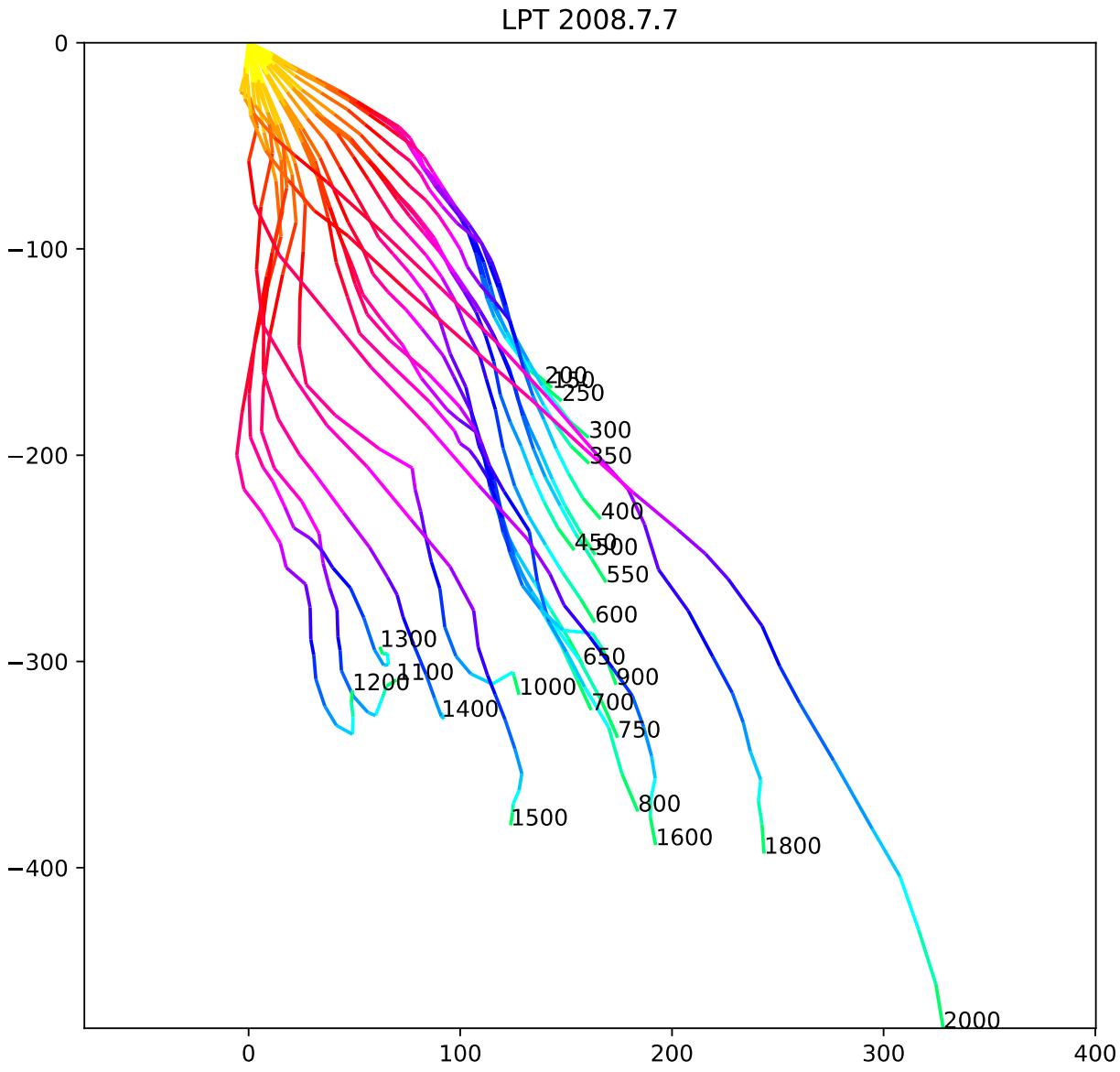


Figure 2.17: Back-trajectories for LPT on July 7, 2008, also showing an absence of recirculation or stagnation. Low-level air progressed steadily toward the northwest throughout the day.

2.4.2. Dependence of Stagnation on Large-Scale Wind

Figure 2.18 plots the stagnation scale values as a function of large-scale wind speed at LPT for all nearly-steady large-scale wind days during the warm season (April-September). As expected, stagnation is favored when wind speeds are weak. However, the coastal inertial cycle permits stagnation for almost all cases with a large-scale wind speed at 400 m less than about 2.5 m/s. Beyond this, there is a fairly sudden drop-off in stagnation potential, particularly for the most common winds from the southeast. The enhanced local winds found with moderate winds from the southwest account for the tendency for stagnation or recirculation to occur at large-scale wind speeds of up to 5 m/s from the southwest, while such stagnation did not happen at all at similar large-scale wind speeds from the southeast or northeast.

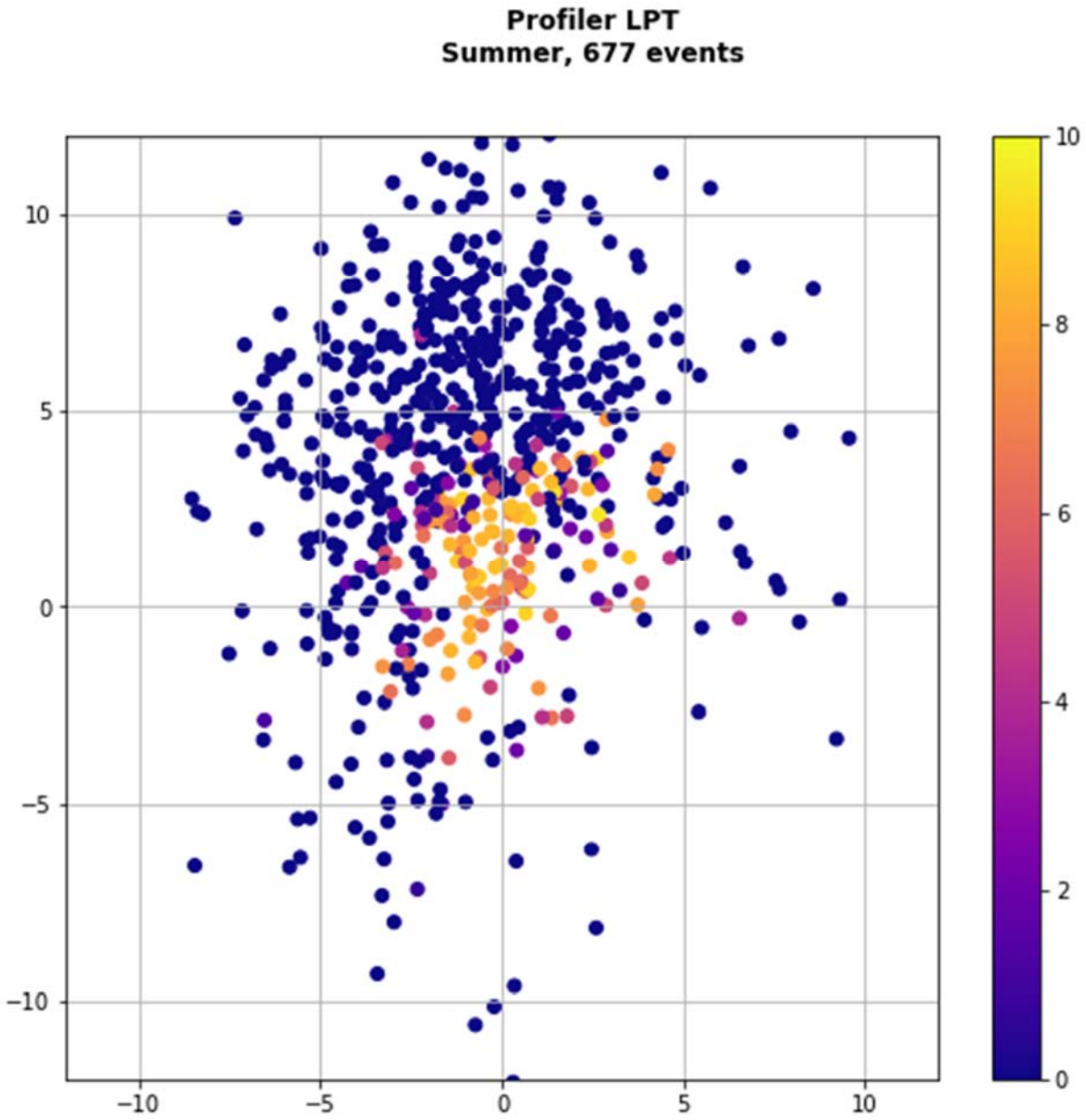


Figure 2.17 Stagnation scale values as a function of mean wind velocity (m/s) during the warm season (April-September) at LPT on all days with nearly steady large-scale winds. Each dot represents a different day and is plotted at a location corresponding to the mean 400 m wind for that day. For example, the dot in the upper left represents a day with a northward wind component of 10 m/s and a westward wind component of 7.5 m/s, which combine to yield a very strong wind toward the northwest. The dots are color-coded according to the stagnation scale. Most events have zero or near-zero stagnation scale values. The high stagnation scale values are clustered near the origin, corresponding to very light large-scale winds. Further details are found in the text.

The four back-trajectory examples (Figures 2.14-2.17) were chosen because they all had similar large-scale winds: from the southeast at about 2.5 m/s. Figure 2.18 shows that there is a wide variety of stagnation values at that particular large-scale wind speed, and Figures 2.14-2.17 explore that variety of stagnation behavior. The figures show that the coastal inertial cycle is

sometimes large enough to cause stagnation at that large-scale wind speed, but when the coastal inertial cycle is weaker no stagnation or recirculation is produced.

Examples of stagnation and non-stagnation at moderate large-scale wind speeds (6 m/s) from the southwest are now shown. Figure 2.19 is a stagnation/recirculation example. It shows that air in the 250-500 m layer traveled rapidly toward the northeast overnight, then slowed and executed a clockwise loop during the early afternoon. Figure 2.20 features similar initial trajectories, but the afternoon low-level trajectories never reverse themselves and stagnation does not occur. As with the cases with large-scale winds from the southeast, the large-scale winds are marginal for recirculation: it happens when the coastal inertial cycle is large enough but not when it is too weak.

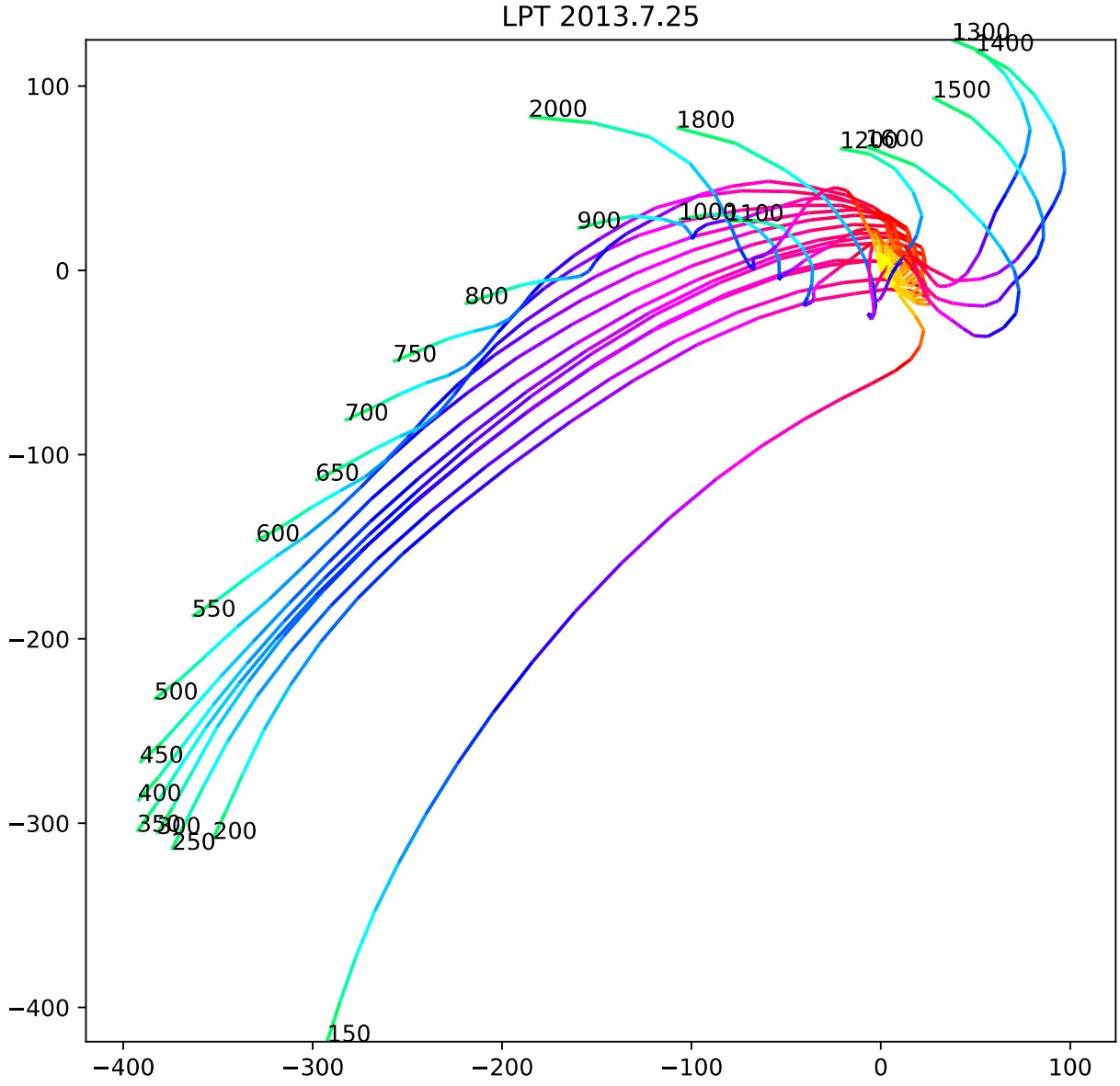


Figure 2.18 Back-trajectories at LPT on July 25, 2013. Low-level trajectories are initially fast-moving from the southwest, but during late morning and afternoon they slow and execute a clockwise loop.

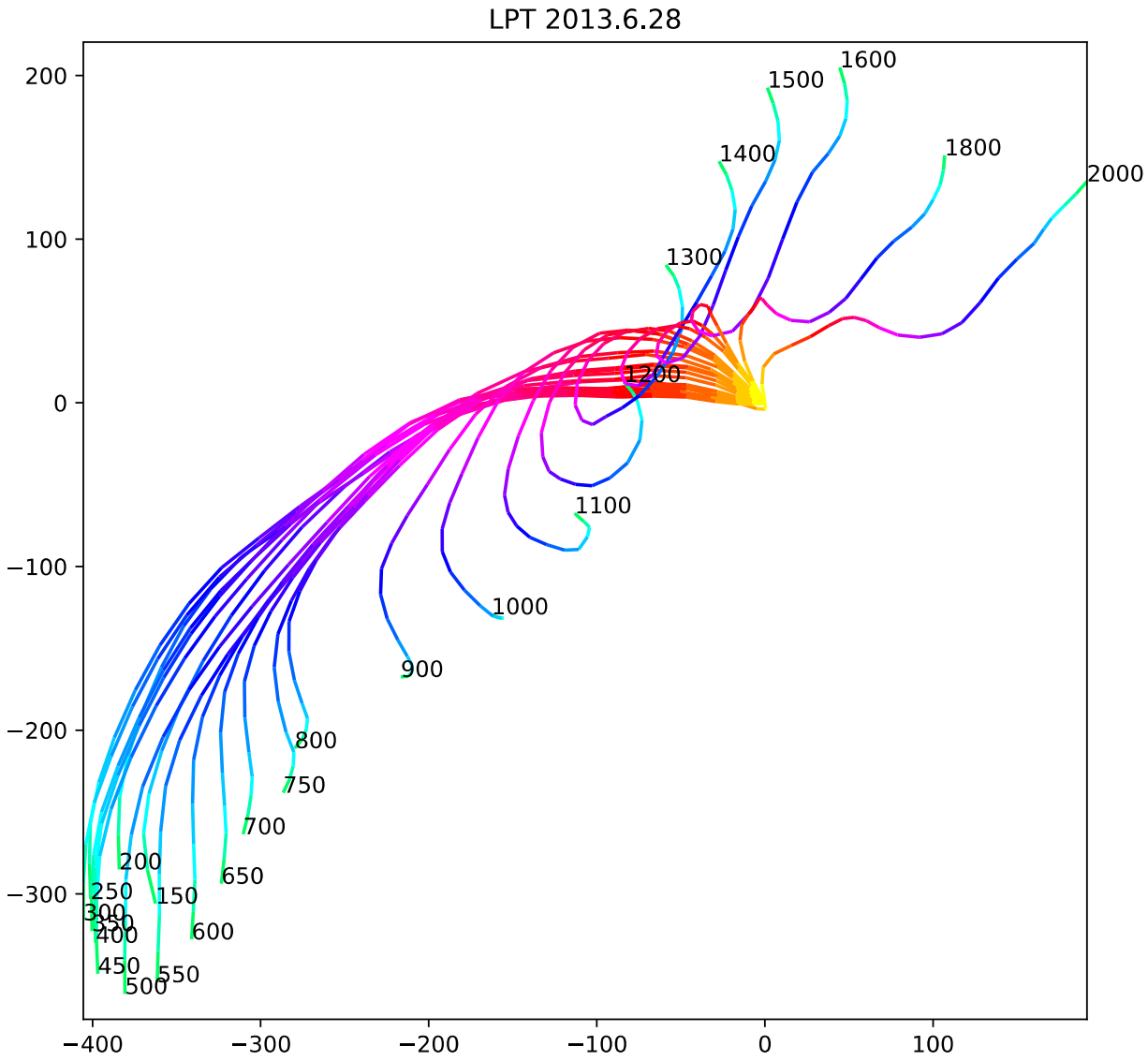


Figure 2.19 Back-trajectories from LPT for June 28, 2013. Despite similar large-scale winds as in Fig. 2.19, the daytime wind remains blowing from land to sea all afternoon rather than reversing course and recirculating.

When the large-scale wind averages near zero, simple circular wind rotation might be expected. This expectation is clouded slightly by large-scale wind shear (such that wind averages are only typically near zero for a limited range of altitudes) and other factors. A sampling of such cases is shown in Figures 2.21-2.24.

Figure 2.21 shows an extremely clear example of the coastal inertial cycle, amplified somewhat by low-level winds being from the southwest. The air around 350-400 m executes a complete circle, returning after 24 hours to the location from which it started. Also worth noting are the trajectories of air at 1600-2000 m. There, the large-scale wind is from the north, but a local-scale wind variation is also apparent, almost exactly out of phase with the local-scale winds at low-levels. At the beginning and end of the trajectories, the low-level winds are moving northwestward while the upper-level winds are moving southeastward. At the midpoint, the low-level air reaches its farthest east point while the upper-level air reaches its farthest west point. This is

consistent with the idea that the coastal inertial cycle at 1600-2000 m is being triggered by the return flow aloft associated with the sea breeze vertical circulation.

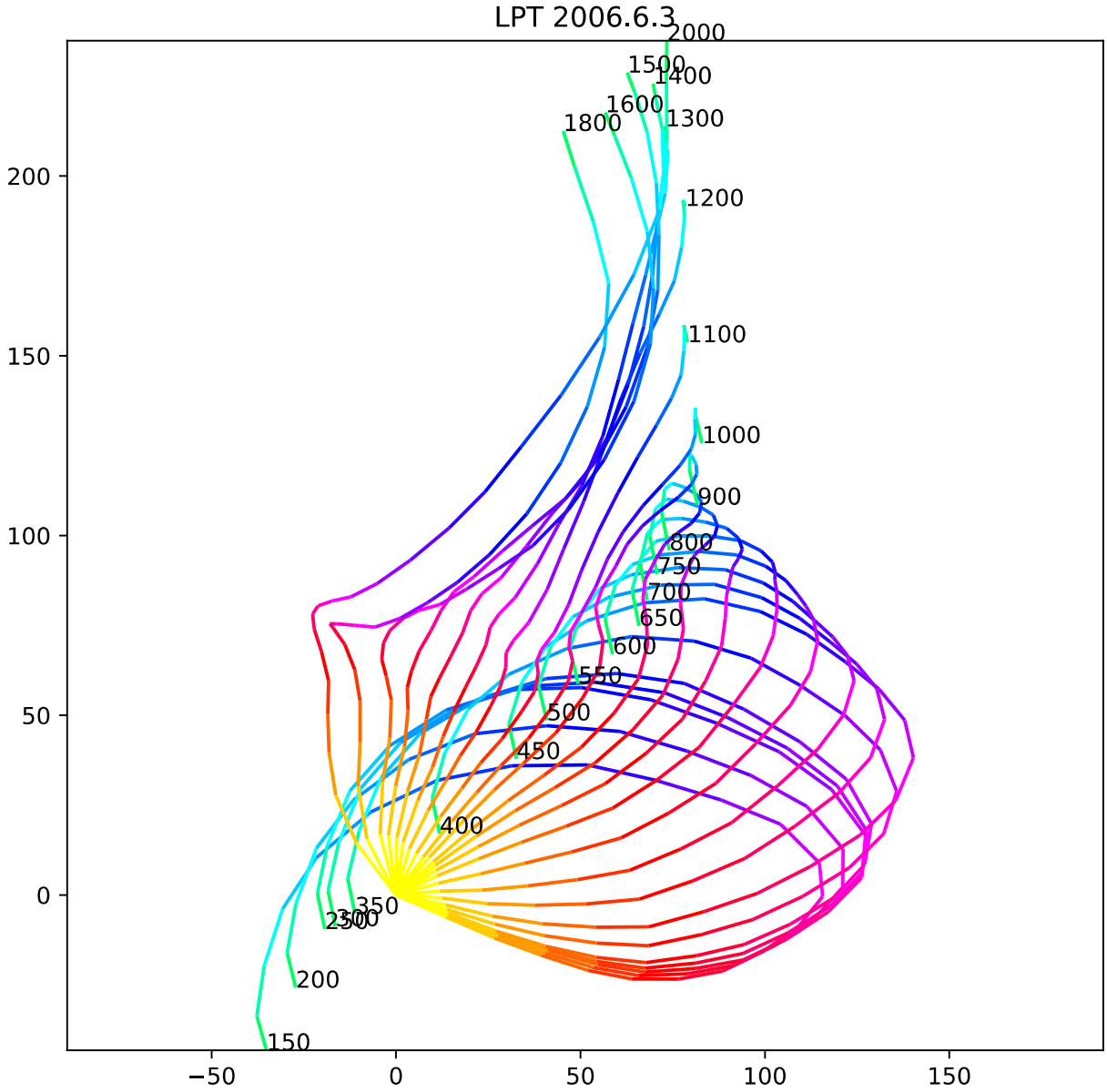


Figure 2.20 Back-trajectories at LPT during a day with weak net transport, June 3, 2006. Air at low levels made a complete 360-degree loop over the course of the day, starting out moving northward, then eastward, then southwestward, and finally southeastward. Air parcels above 1000 m executed a local wind oscillation about 12 hours out of phase while moving southwestward.

Figure 2.22 shows a similar low-level pattern, except the scale is much smaller due to the strong northerly winds aloft. These winds appear to cause a modification to the low-level winds near the end of the day when the convective boundary layer would be deepest, with mixing of high-velocity air from above causing low-level winds to turn back toward the southwest.

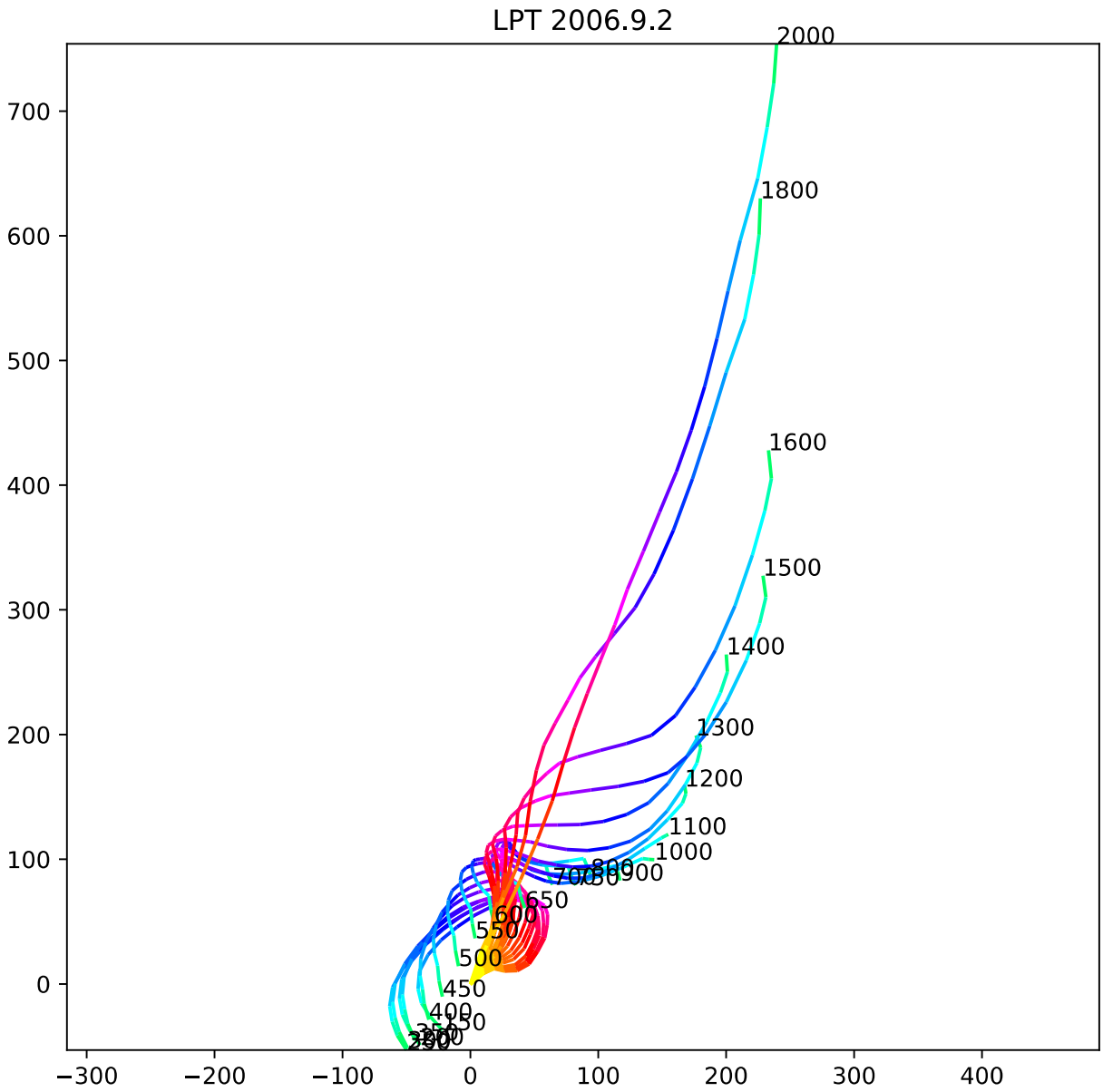


Figure 2.21 Back-trajectories at LPT during a day with weak net transport, September 2, 2006. Low-level air executes a similar loop as in Fig. 2.21, but the transition to a wind oscillation 12 hours out of phase begins lower, around 800 m, and the strong southward winds aloft appear to cause the low-level winds to change back to southwestward near the end of the period due to momentum mixing within the boundary layer.

In Figure 2.23, the coastal inertial cycle dominates the wind pattern, with air at many levels staying within a 100x100 km box while remaining in constant motion. By contrast, Figure 2.24 shows an instance where the lack of overall large-scale wind is a consequence of northwestward winds shifting to southeastward. This wind shift destroys any pure coastal inertial cycle signal that might otherwise have been present had large-scale winds remained steady.

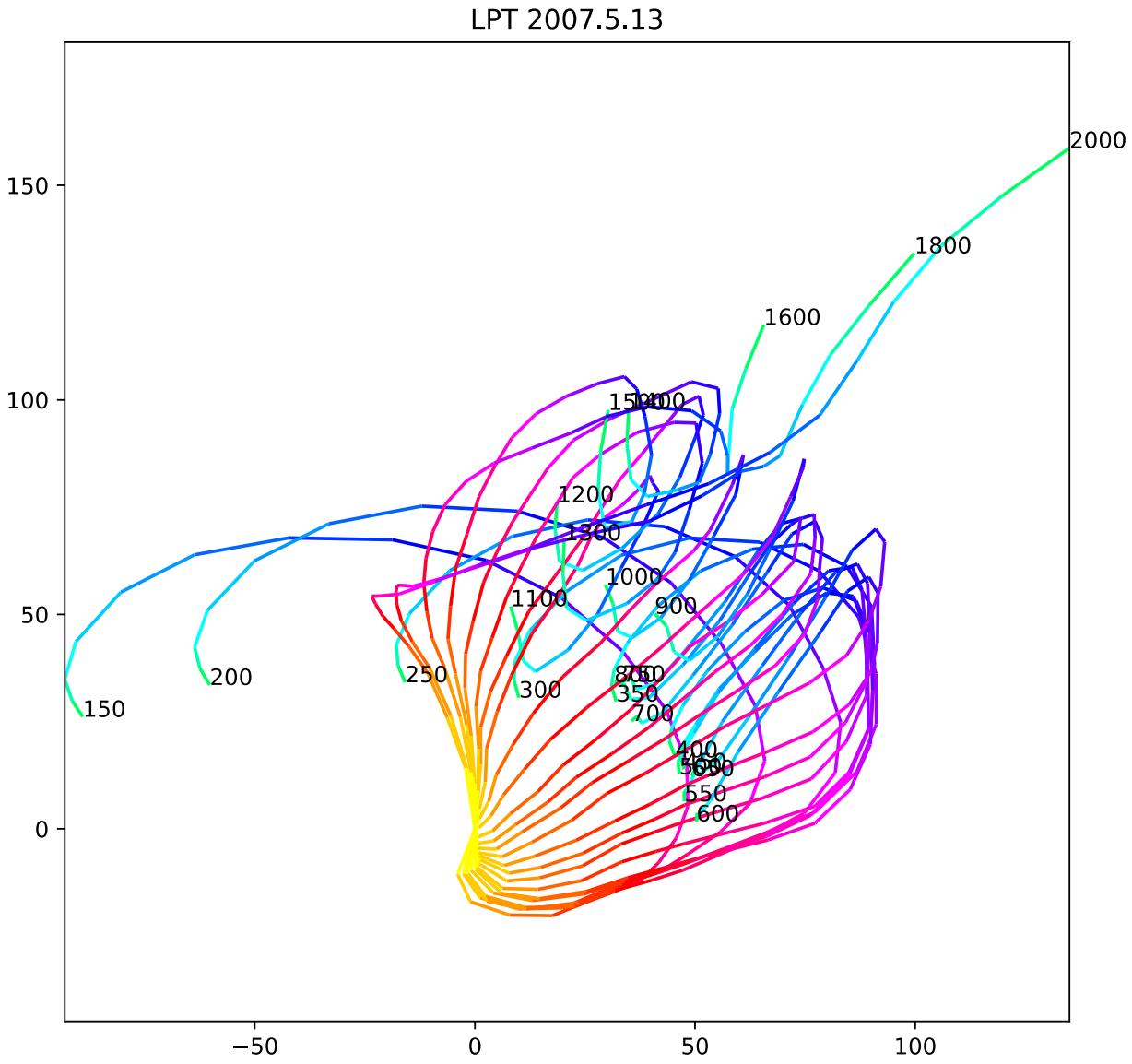


Figure 2.22 Back-trajectories at LPT during a day with weak net transport, May 13, 2007. Trajectories rotate clock-wise overtime at most levels but don't quite complete a circle. Nonetheless, background winds are weak enough that there is overall stagnation: air between 250 m and 1500 m remains within a 100x100 km box for the entire day.

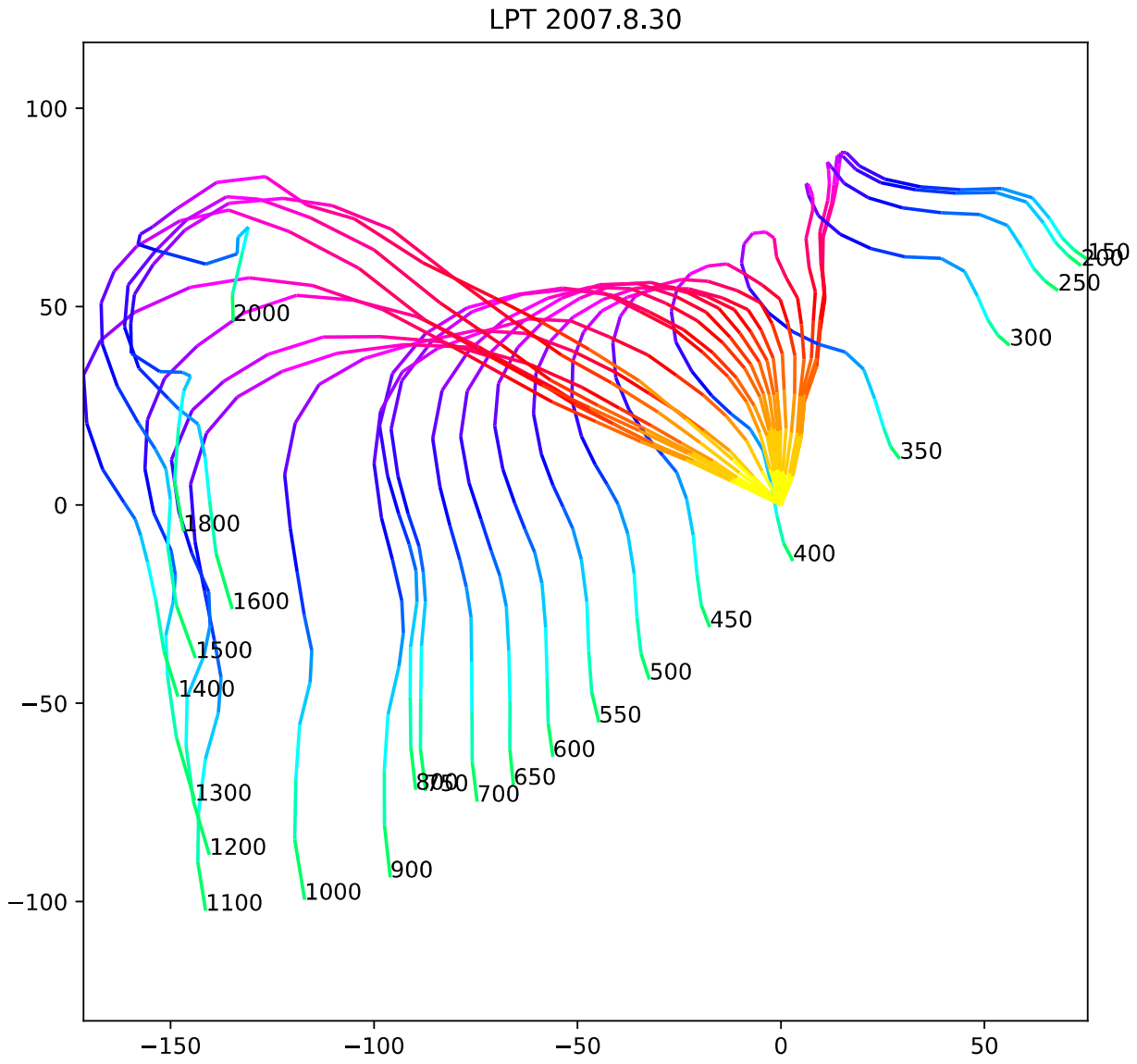


Figure 2.23 Back-trajectories at LPT during a day with weak net transport, August 30, 2007. Trajectories were toward land until early morning and then switched to be directed toward the coast. With the apparent change in large-scale winds from onshore to offshore, the classic coastal inertial cycle does not clearly appear.

2.4.3. Differences by Season and Location

During the cool season, October-March, winds are not typically weak from the southeast like they are during the other half of the year (Fig. 2.25). Thus, there are fewer opportunities for wind stagnation. Furthermore, with weaker sunlight, the driving of an inertial cycle is necessarily weaker on average, and the weaker cycle should translate into less opportunity for stagnation.

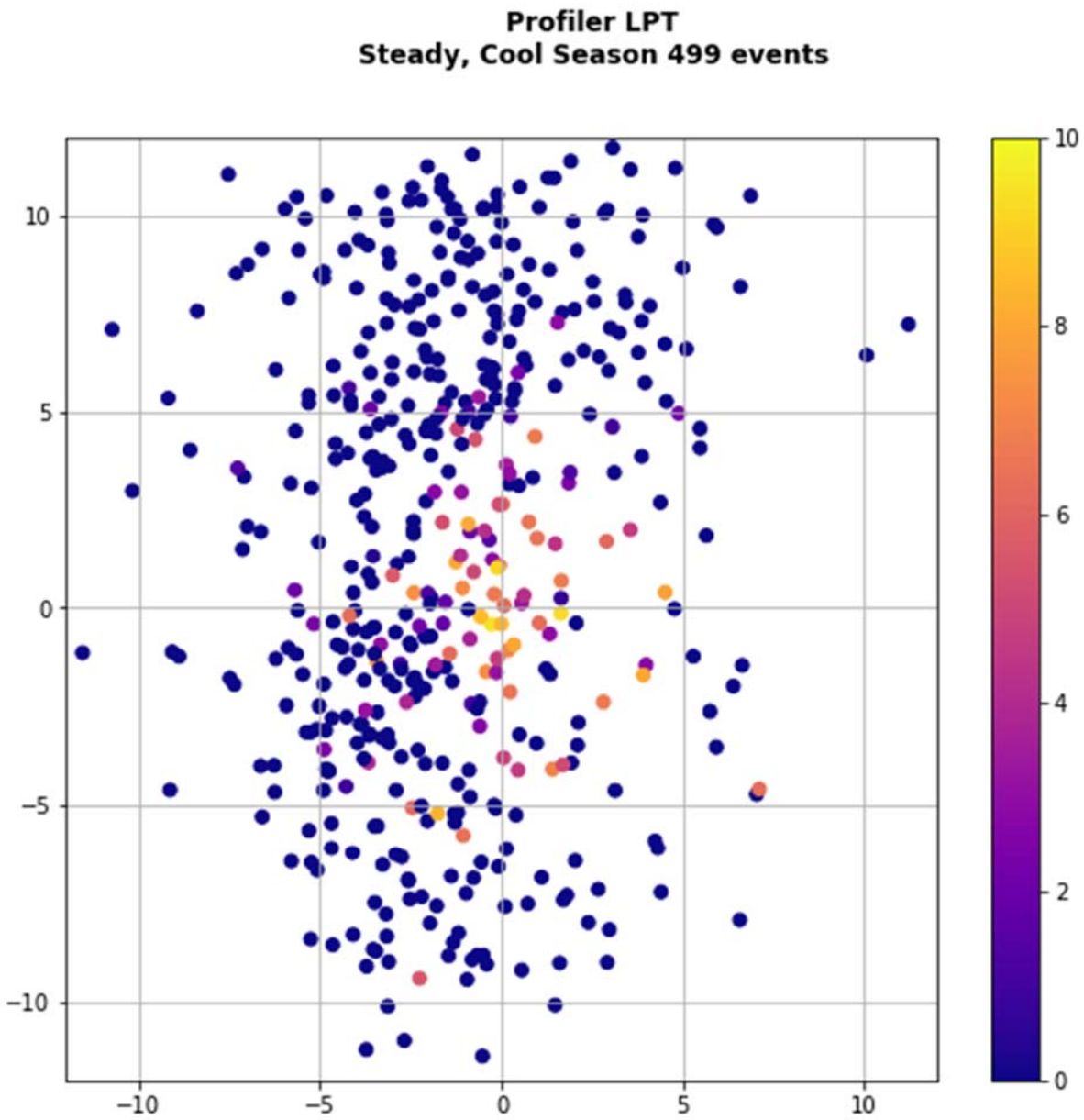


Figure 2.24 Stagnation scale values for LPT, as in Fig. 2.18, but for October-March. In general, except for extremely light wind conditions, there tends to be less stagnation or recirculation compared to the other half of the year. Winds are predominantly from the north and south, with large-scale wind speeds less than 4 m/s being typically associated with moderate stagnation.

At HSN (Figures 2.26-2.27), differences with LPT are minor. Overall, the largest differences arise with weak to moderate winds from the southwest during the warm season. At LPT, there was a clear preference for stagnation even at moderate wind speeds, but at HSN stagnation doesn't seem any more likely with that direction than with any other. This distinction might not be statistically robust, or perhaps the substantial difference in overwater fetch under southwesterly wind conditions at LPT and HSN is enough to mitigate enhancement of the diurnal cycle at HSN.

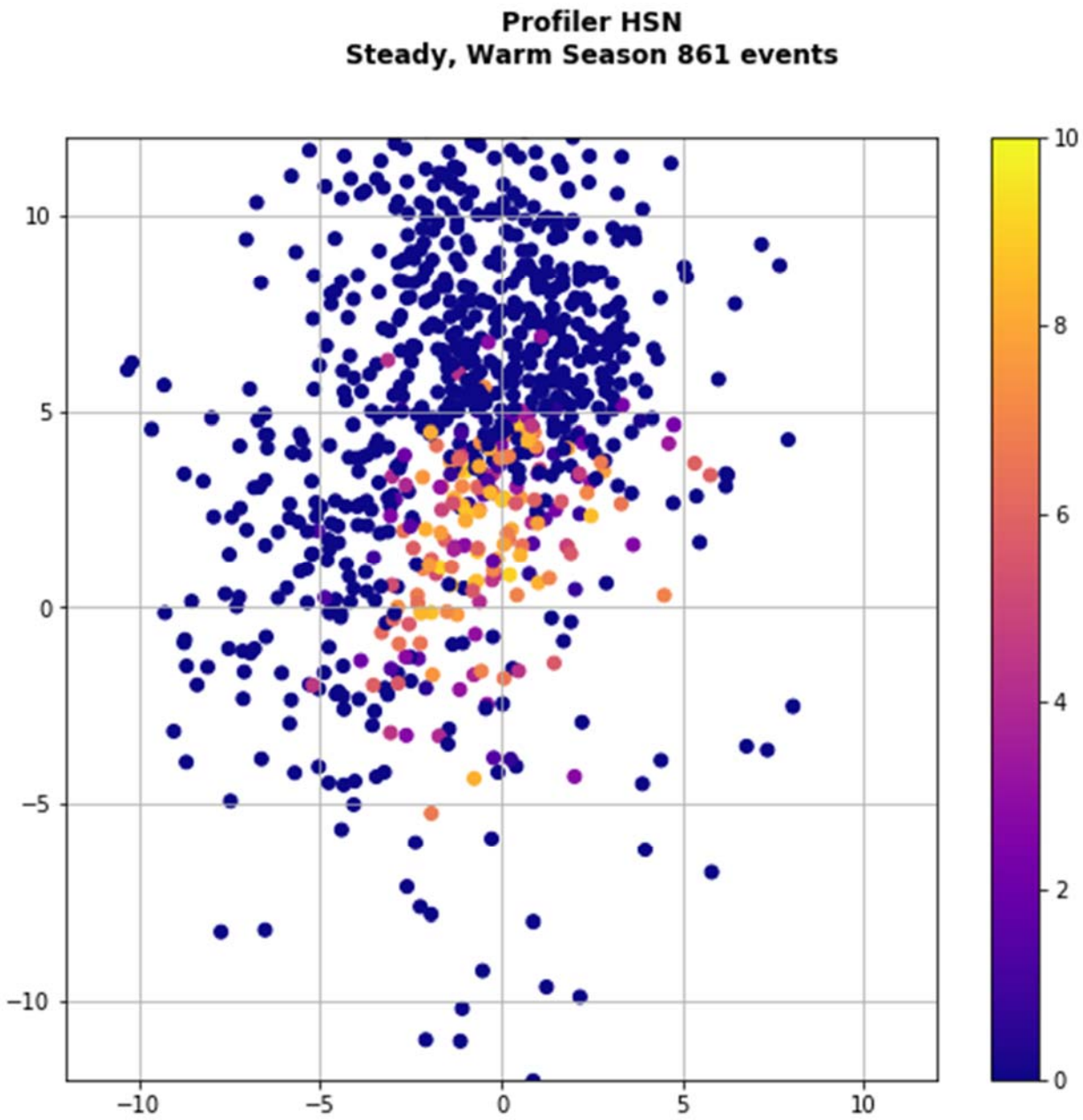


Figure 2.25 Stagnation scale values for HSN during the warm season, as in Fig. 2.18. See text for a discussion of similarities and differences with Fig. 2.18.

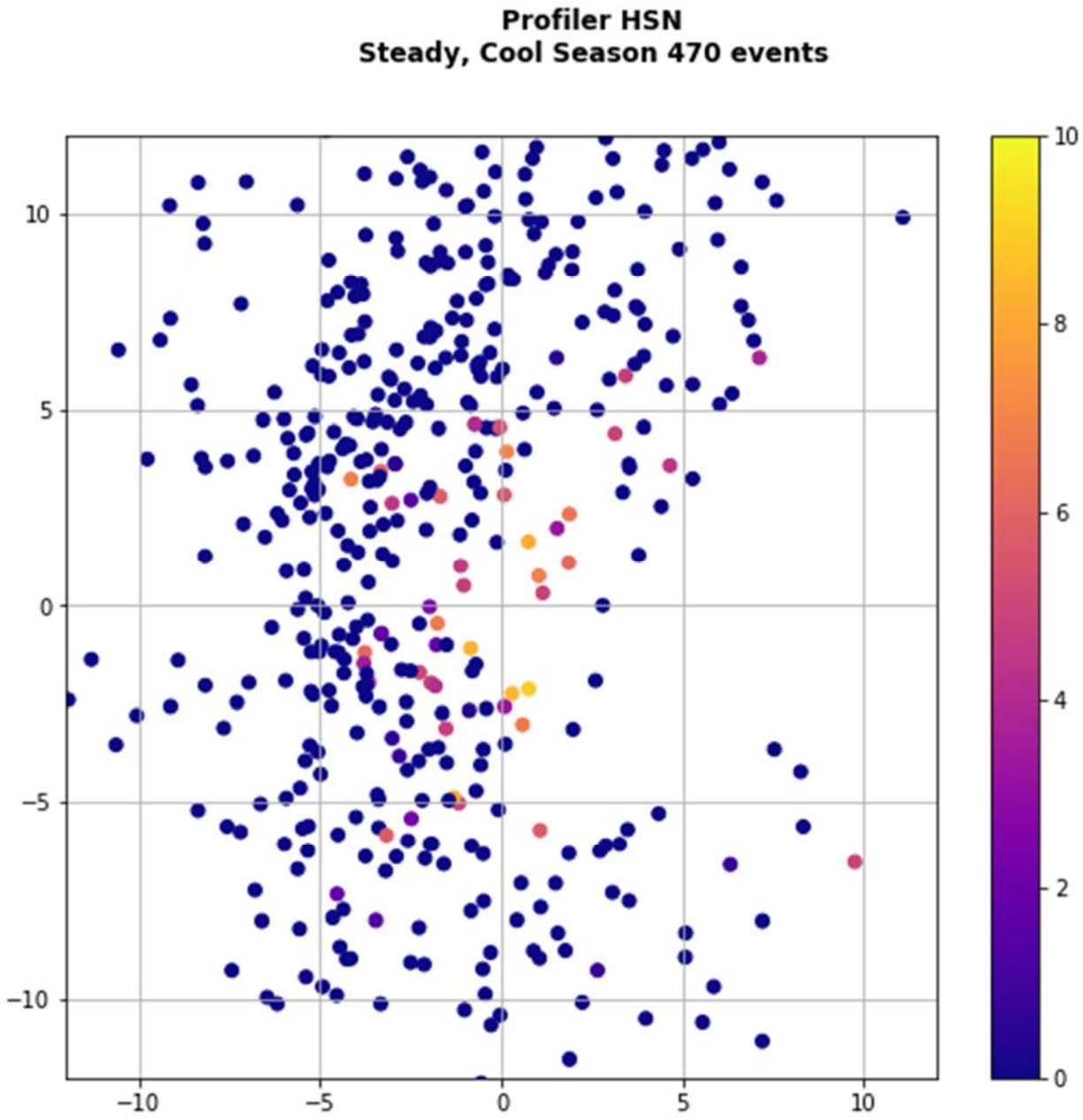


Figure 2.26 Stagnation scale values for HSN during the cool season. The overall pattern is very similar to Figure 2.25, LPT during the cool season.

At CLE (Figures 2.28-2.29), warm-season stagnation seems more likely than at HSN, despite the absence of a coastal temperature contrast to drive an inertial cycle. If anything, there's a stagnation-based preference for winds from the west during the warm season, but such a preference is not present in the cool season. Inspection of stagnation-inducing trajectories for CLE (not shown) reveals that the mechanism for stagnation is often basic adverse shear: if surface winds are weak from one direction while winds aloft are strong and from the opposite direction, the growth of the convective boundary layer will cause the low-level winds to reverse direction.

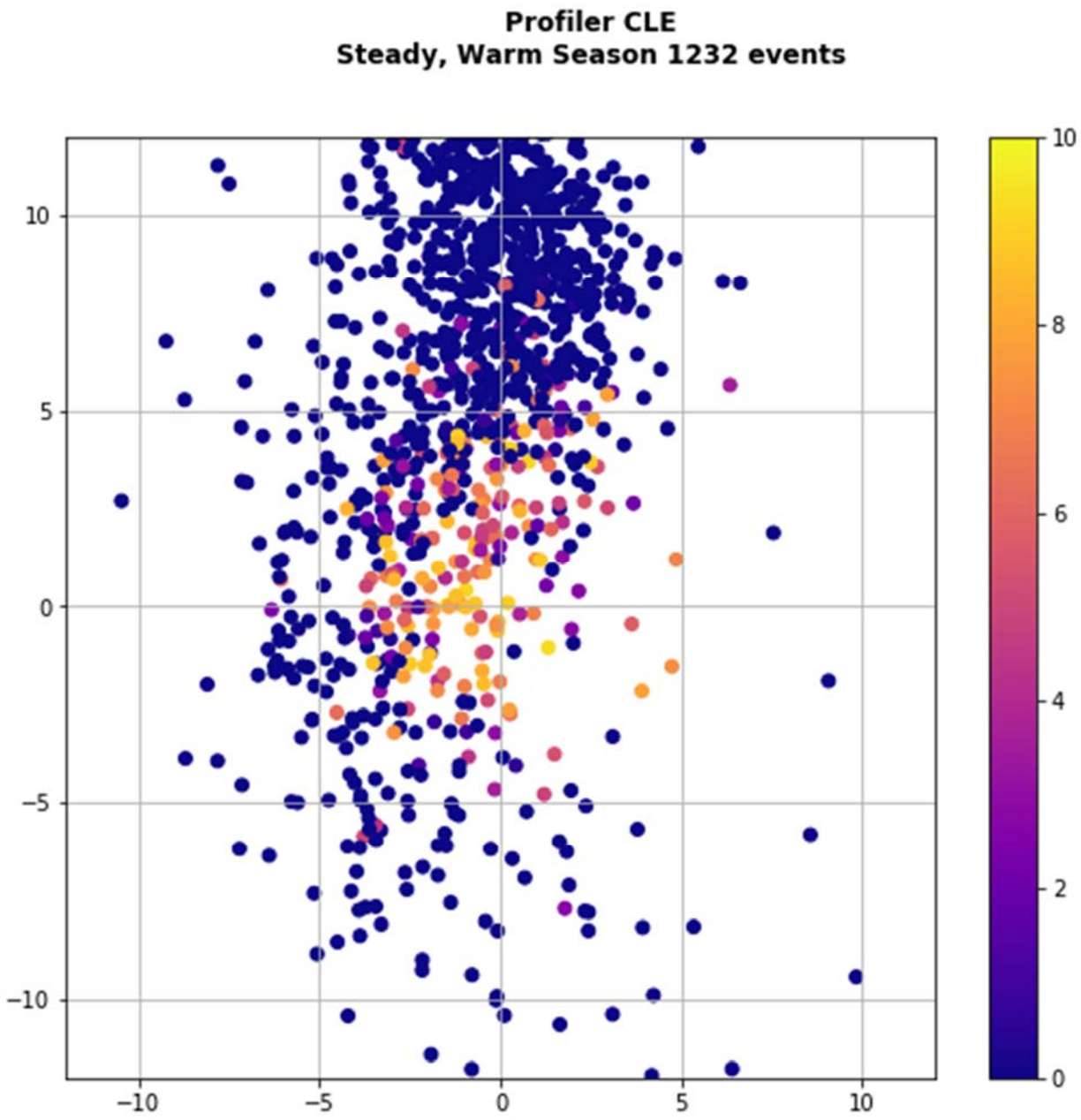


Figure 2.27 Stagnation scale values for CLE during the warm season. Most days have strong winds from the south with very little chance for stagnation. High stagnation and recirculation values are mostly confined to large-scale winds less than 2.5 m/s, but the division is much less clear than at LPT and there are some apparent exceptions at higher wind speeds.

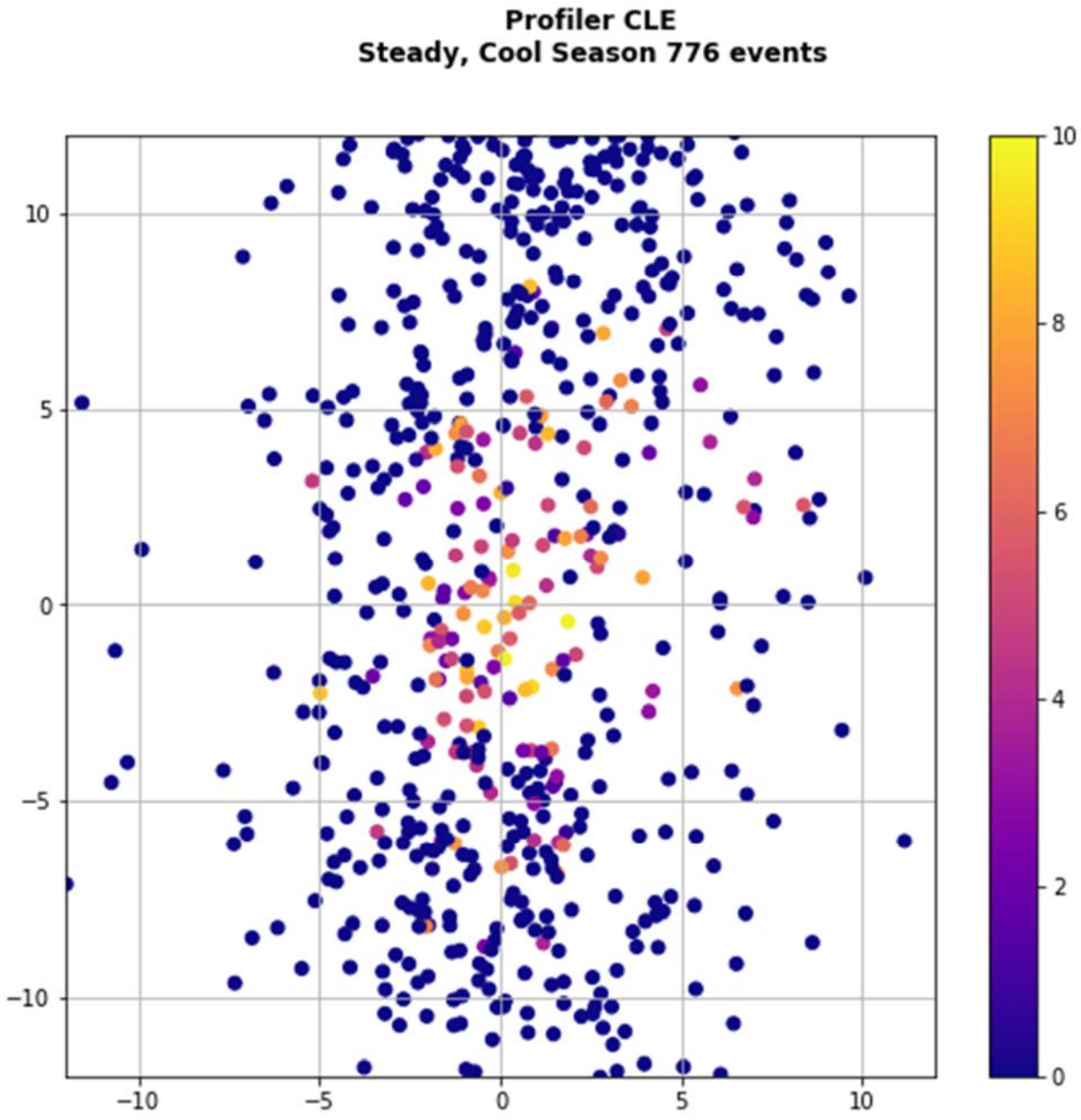


Figure 2.28 Stagnation scale values for CLE during the cool season. Large-scale light winds are rare, and when they do occur the likelihood of stagnation seems a bit weaker during the cool season as during the warm season.

2.5. Nighttime shear

While vertical wind shear can be important in determining the magnitude of the inertial oscillation and the behavior of daytime winds, it is also crucial for the nighttime transport of pollutants. If winds at different levels are moving at different directions and speeds, the nighttime Houston plume will not remain intact but will instead be spread out over a large horizontal area, thereby literally diluting its impact when it mixes down to the ground the following day.

Figure 2.30 shows the difference between an air parcel at 400m and one at 1000m after 12 hours of transport between 6 PM local standard time (LST) and 6 AM LST. Some of the large-scale wind conditions that most favor the coastal inertial cycle, such as weak to moderate winds from the southwest, are associated with a large dispersion of the nighttime pollution plume. This is consistent with the sample back-trajectories for those large-scale winds (Figures 2.19-2.20), and with the expectation that winds parallel to the coastline maximize the low-level vertical wind shear. A similar sort of effect may be causing large dispersion of the nighttime pollution plume when winds are from the east, also parallel to the coast.

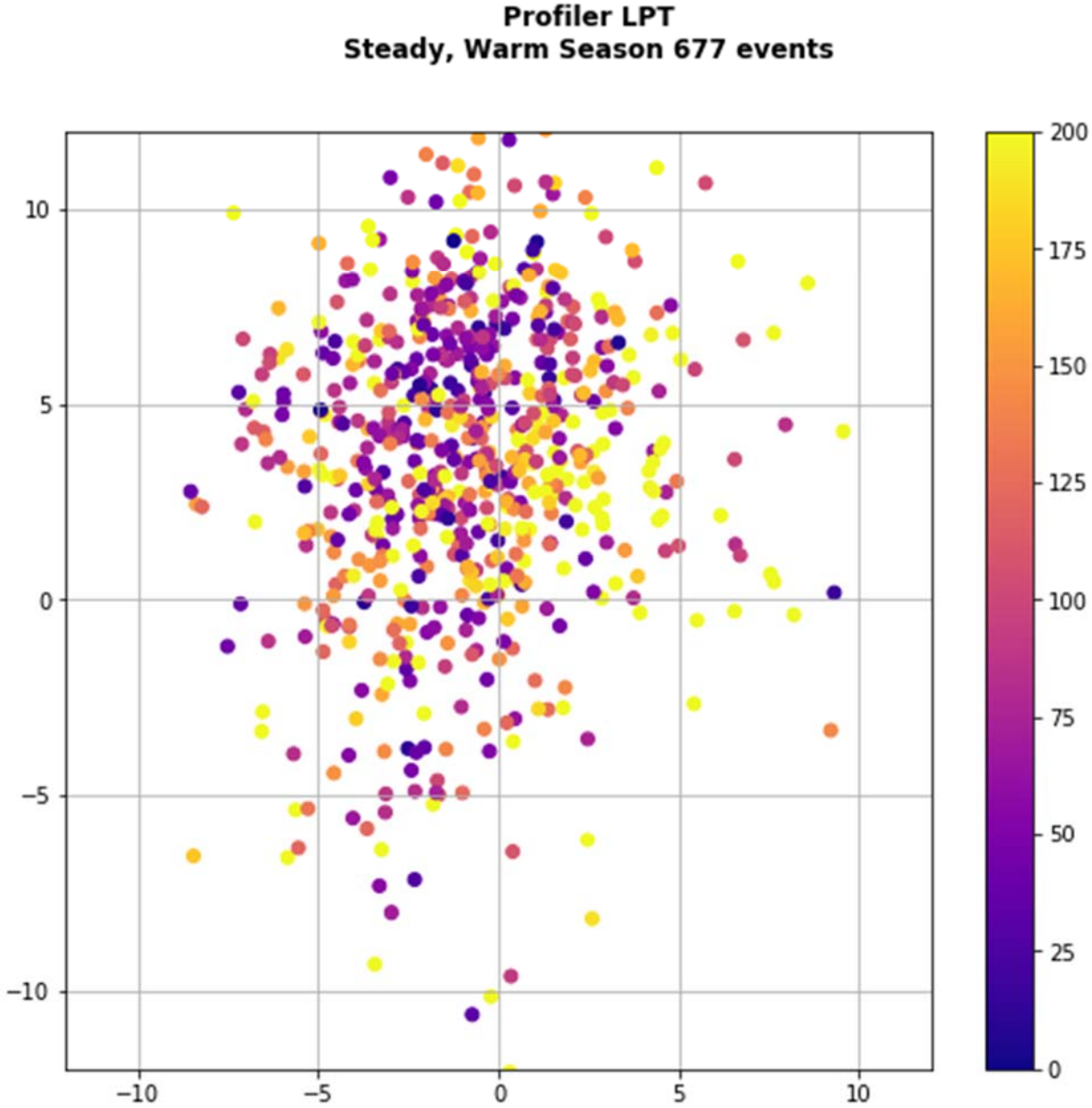


Figure 2.29 Difference (km) after 12 hours (6 PM LST to 6 AM LST) in the location of air parcels originating at 400 m and 1000 m during the warm season at LPT. Many separation differences are less than 50 km, while many more exceed 200 km. For a detailed discussion, see the text.

Conversely, there are some large-scale wind conditions where the lack of plume dispersal is favored. In particular, moderately strong (~7 m/s) winds from the south-southeast tend to have dispersal of less than 100 km between 400 m and 1000 m. The lack of low-level vertical wind shear implies a near-uniform temperature distribution, and air moving rapidly from the south-southeast would have originated over the tropical Atlantic Ocean where temperatures are nearly uniform. The large wind speed would leave minimal time for the air to interact with the heating gradients along the coastline.

Apart from these general tendencies, there appear to be cases with nearly all possible large-scale wind conditions for which limited nighttime dispersal can occur. Reliable prediction of nighttime dispersal based on winds alone doesn't appear to be possible. However, since temperature gradients and vertical wind shear are closely linked, analyses of 925 hPa temperatures may provide useful predictive information.

2.6. Summary

This portion of the study placed the local-scale wind structures observed by long-term wind profilers into the context of large-scale winds and the processes that drive those structures. Clockwise wind rotation was found to be an overwhelming presence in the lower atmosphere over Texas, due to the effects of the Earth's rotation on air movement.

While any disturbance can trigger clockwise wind rotation on a daily time scale, disturbances that are themselves tied to the time of day can operate in resonance with the inertial rotation to amplify the local-scale wind variations. In Texas, two such disturbances are the daily growth and collapse of the daytime convective boundary layer and the daytime heating of air over land relative to air over water. The latter also produces the sea breeze, and the coastal inertial oscillation can be viewed as a part of the sea breeze cycle.

Local-scale winds are strongest in North Texas under strong wind conditions. This behavior is consistent with being caused by the growth and collapse of the boundary layer. The same mechanism is responsible for the Great Plains low-level jet, which may be indirectly driving much of the local-scale winds across Texas. In the Houston area, by contrast, local winds are strong when large-scale winds are light. That fact, plus the large difference in the timing of wind variations at different altitudes, point to the coastal heating contrast as the dominant mechanism.

Stagnation and recirculation take place when the local-scale winds are strong enough to equal or exceed the large-scale wind speeds. This happens over a range of wind speeds during the warm season (April-September), mostly speeds less than 2.5 m/s. The one main exception is that southwesterly winds can produce larger local-scale wind variations, so stagnation is likely with large-scale winds up to 5 m/s when the wind is from the southwest.

Sample wind profiler back-trajectories illustrate a variety of wind patterns under light wind conditions. Near the transition zone between stagnation and non-stagnation, the coastal inertial cycle is sometimes strong enough to cause recirculation over a few hours. When large-scale winds are nearly calm, the wind rotation tends to dominate, and air can carry out large loops up to 100 km in diameter before returning to roughly the same location 24 hours later.

At night, pollutant transport is strongly affected by vertical wind shear. Some of the same conditions that favor robust recirculation also favor rapid nighttime dispersal of the Houston pollution plume. The plume is most likely to remain intact when summertime winds are from the south-southeast and moderately strong.

2.7. Quality Assurance

Programs used to download, process and generating plots have been used in previous studies and adapted for current study. All the downloaded were visually inspected by David Coates. John Nielsen-Gammon further checked ~50% of the plots generated for all the datasets. Data analysis were performed jointly by David and Dr. Nielsen-Gammon and were cross-checked to ensure the analyses were performed in a consistent way. The figures included in the Chapter were further checked by Qi Ying when the final report was assembled.

3. Development of the source and age resolved CMAQ model

3.1. Age-resolved air quality modeling

The Source- and Age-Resolved Community Multiscale Air Quality (SAR-CMAQ) model represents a further development of the source-oriented CMAQ model. It is capable of tracking emissions not only by sources/source-regions but also their age since emitted into the atmosphere. In this study, the model uses reactive tracers to track NO_x and primary VOCs emitted from different sources/source-regions at different times. The base chemical mechanism is a condensed version of SAPRC-07 (CS07A) as documented by Carter (https://intra.engr.ucr.edu/~carter/SAPRC/csaprc07.pdf). Additional non-reactive O_3 tracers are introduced to track the ozone formed at different locations (including elevation) and times. In this way, we can directly quantify the amount of locally-formed vs. regional ozone and quantitatively determine when and where the ozone affecting Houston on high ozone days are formed. To track the atmospheric age of ozone precursors, we modified the source-oriented CMAQ model so that it can be used to track the atmospheric age of the air pollutants. Conceptually, the age-resolved mechanism can be explained using the following reactions for NO and NO_2:

$$NO\_T1 + O_3 \rightarrow NO_2\_T1 + O_2 \quad \text{(R1)}$$
$$NO\_T2 + O_3 \rightarrow NO_2\_T2 + O_2 \quad \text{(R2)}$$
$$\dots \qquad \dots$$
$$NO\_Tn + O_3 \rightarrow NO_2\_Tn + O_2 \quad \text{(Rn)}$$

The NO_T[1,2…n] and NO_2_T[1,2,…,n] species are used to track NO and NO_2 with different atmospheric times from fresh to aged. In the model simulation, fresh emissions of NO and NO_2 are represented by the species with T1 tags. At the end of each predefined aging advancing time (e.g. 1 hour or 3 hours), a time bin advance operation is performed so that NO_2_Ti = NO_2_T(i-1) for i=1,2,…,n-1. For the last time bin, NO_2_Tn = NO_2_Tn + NO_2_T(n-1). The same operations are done for NO and other tagged reactive nitrogen species. This age-resolved approach is also applied to major primary VOCs (Table 3.1).

Table 3.1 List of VOC species whose atmospheric age is tracked in the modified CS07A mechanism

VOC	**Model species**	**MIR**
All alkanes with kOH less than 5 x 10^3 ppm^{-1} min^{-1}.	ALK3	2.19
All alkanes with kOH greater than 5 x 10^3 ppm^{-1} min^{-1}.	ALK4	3.11
Ethene	ETHE	5.44
Alkenes (other than ethene) with kOH < 7x10^4 ppm^{-1} min^{-1}.	OLE1	9.73
Alkenes with kOH > 7x10^4 ppm^{-1} min^{-1}.	OLE2	10.98
Aromatics with kOH < 2x10^4 ppm^{-1} min^{-1}	ARO1	6.72
Aromatics with kOH > 2x10^4 ppm^{-1} min^{-1}	ARO2	15.36
Isoprene	ISOP	12.62
Monoterpene	TERP	8.54
Formaldehyde	HCHO	6.49

Ozone is a secondary polluted formed from photochemical reactions of VOCs and NOx. In this study, the atmospheric age of ozone is determined by its residence time in the atmosphere or the age of the precursor. Thus, two sets of non-reactive ozone tracers are needed. O_3_V_T1, O_3_V_T2, …, O_3_V_Tn, are used to represent O_3 with different atmospheric ages attributed to VOCs. O_3_N_T1, O_3_N_T2, …, O_3_N_Tn, are used to represent O_3 with different atmospheric ages attributed to NOx. The total concentration of O3 for each atmospheric age group is calculated by O3_Ti = O3_V_Ti + O3_N_Ti. Thus, to accurately track the atmospheric age of ozone, it is necessary to 1) attribute the amount of ozone formed in each model time step to NOx and/or VOCs and 2) to allocate the amount of ozone attributed to NOx and/or VOCs to different age groups.

The updated ozone sensitivity indicator method developed by Wang et al. (2018) is used to attribute ozone to VOCs and/or NOx. The O_3 sensitivity indicator (R) used in this study is:

$$\mathrm{R} = \frac{P_{H_2O_2} + P_{ROOH}}{P_{HNO_3}} \tag{3-1}$$

where $P_{H_2O_2}$ is the formation rate of hydrogen peroxide; P_{ROOH} is the formation rate of organic peroxide, and P_{HNO_3} is the formation rate of nitric acid in each chemistry time step. The formation rates are calculated using the integrated process analysis (IPA) tool in the CMAQ model for each model chemistry time step. If R is less than R_{ts} (which is set to 0.045, based on Wang et al. 2018), ozone formation is considered as VOC-limited and all the ozone formed in this time step is attributed to VOCs. When R is greater than Rts, ozone formation is attributable to both NOx and VOCs and the faction of ozone attributed to NOx increases with R and asymptotically approaches one, as shown in equation (3-2).

$$F_{NOx} = 1.0 - 0.1715exp(-1.309log_{10}R) \tag{3-2}$$

Predicted F_{NOx} using equation (3-2) and the F_{NOx} based on sensitivity simulations as described by Wang et al. (2018) is shown in Figure 3.1. F_{NOx} is defined for the entire range of R but is set to zero when $R<R_{ts}$.

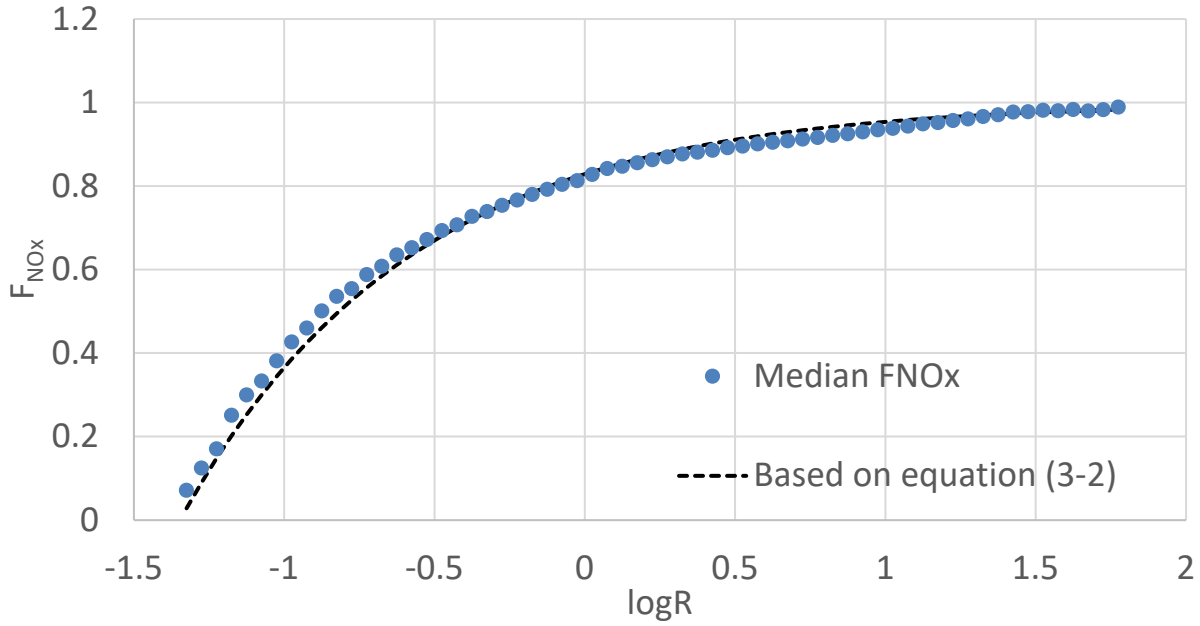


Figure 3.1 Median values of FNOx and the FNOx values predicted by equation 3-2.

After the O_3 is attributed to NO_x and/or VOCs, they are apportioned to different atmospheric age groups with respect to NO_x and VOC sources. In this study, the approach to apportion O_3 formation to different regional NO_x and VOC sources is mostly based on the equations in (Wang et al., 2018a) and (Wang et al., 2019). The approach can be summarized by the following steps.

The first step is to calculate the intermediate concentration of O_3 attributed to NO_x and VOC from the i^{th} atmospheric age group (O_3_N_Ti, O_3_V_Ti) as shown in eq (3-3) and (3-4):

$$O_3\_N\_T_i^{int} = O_3\_N\_T_i^{t-\Delta t} + P_{O_3} F_{NO_x} S_{NO_{x,i}} \qquad i = 1,2 \dots N_{TB} \qquad (3\text{-}3)$$

$$O_3\_V\_T_i^{int} = O_3\_V\_T_i^{t-\Delta t} + P_{O_3}\left(1 - F_{NO_x}\right) S_{VOC,i} \qquad i = 1,2 \dots N_{TB} \qquad (3\text{-}4)$$

where i is the atmospheric age bin index; the superscript *int* indicates intermediate concentrations, and superscript *t-Δt* denotes concentrations from the previous time step; F_{NOx} is the attribution function defined in (Wang et al., 2018a); $S_{NOx,i}$ and $S_{VOC,i}$ are the atmospheric age contribution functions to apportion incremental O_3_N and O_3_V to the i^{th} age group of NO_x and VOCs, respectively. The calculations of S_{NOx} and S_{VOC} are shown in equations (3-5) and (3-6). The MIR values needed for equation (3-6) are shown in equations (3-6).

$$S_{NOx,i} = \frac{NO_{2,i}}{\sum_i^{N_{TB}} NO_{2,i}} \qquad (3\text{-}5)$$

$$S_{VOC,i} = \frac{\sum_j^{N_c} C_{VOC,j,i}\, MIR_j}{\sum_i^{N_{TB}} \sum_j^{N_c} C_{VOC,j,i}\, MIR_j} \qquad (3\text{-}6)$$

The second step is to use the intermediate concentrations to update O_3_N_T_i and O_3_V_T_i by considering the O_3 removal terms, as shown in equations (3-7) and (3-8):

$$O_3\_N\_T_i^t = O_3\_N\_T_i^{int} - D_{O_3} \frac{O_3\_N\_T_i^{int}}{\sum_{i=1}^{N_s}\left(O_3\_N\_T_i^{int} + O_3\_V\_T_i{}^{int}\right)},\ i = 1,2 \dots N_{TB} \quad (3\text{-}7)$$

$$O_3\_V\_T_i^t = O_3\_V\_T_i^{int} - D_{O_3} \frac{O_3\_V\_T_i^{int}}{\sum_{i=1}^{N_s}\left(O_3\_N\_T_i^{int} + O_3\_V\_T_i{}^{int}\right)}, i = 1,2 \dots N_{TB} \quad (3\text{-}8)$$

where D_{O_3} is the in-situ O_3 removal rate.

The additional information from the atmospheric age distribution modeling allows more accurate modeling for the physical, chemical and optical processes related to air pollutants in the atmosphere. Different proportions of fresh and aged pollutants reflect the difference in meteorological conditions in diluting/retaining air pollutants. Atmospheric age is also directly related to the transport distance and can be useful in analyzing local vs. regional contributions to air pollution events.

3.2. Improve the computation efficiency of source- and age- resolved chemical mechanism

3.2.1. Introduction

Source-oriented chemical mechanisms have been used extensively in source apportionment modeling studies to determine source (or source region) contributions to NO_x (Zhang and Ying, 2011), VOCs (Ying and Krishnan, 2010), secondary inorganic (Ying and Kleeman, 2006) and organic aerosols (Wang et al., 2018b), and ozone. In these mechanisms, additional reactions are introduced in the gas phase chemical mechanisms to track primary emissions and their reaction products from different sources. For the source apportionment of secondary aerosol products from gas-to-particle partitioning, aerosol and cloud processes are also modified to include additional model species to represent the semi-volatile products from different sources. While this is conceptually simple, the source-oriented mechanisms are computationally expensive because the number of reactions increases almost quadratically with the number of source types due to reactions that involve two source-tagged species. For example, consider the simple reaction of NO + NO_3 = $2NO_2$, if the mechanism is expanded to track two explicit sources in addition to the vanilla type (which may represent emissions from all other sources and from initial and boundary conditions), the following nine reactions in reaction set 1 (RS1) are needed (using model species names):

NO + NO3 = NO2 + NO2
NO + NO3_X1 = NO2 + NO2_X1
NO + NO3_X2 = NO2 + NO2_X2
NO_X1 + NO3 = NO2_X1 + NO2
NO_X1 + NO3_X1 = NO2_X1 + NO2_X1
NO_X1 + NO3_X2 = NO2_X1 + NO2_X2
NO_X2 + NO3 = NO2_X2 + NO2
NO_X2 + NO3_X1 = NO2_X2 + NO2_X1
NO_X2 + NO3_X2 = NO2_X2 + NO2_X2 (RS1)

The names of the source-tagged species have an '_X' in them, followed by a number as the source-origin index. The number of equations for this simple reaction increases quadratically

with the number of sources to be tracked. For 9 explicit + one vanilla sources of NO_x (i.e. 10 sources), 100 reactions are needed instead of one reaction. As there are quite a number of such NOx + NOx reactions in the gas phase inorganic chemistry, the number of reactions needed for the chemical mechanism grows quickly. This problem has prevented broader applications of the source-oriented models for long-term source apportionment studies.

The NOx chemistry is further complicated by the necessity to deal with the double-source-tagged species, such as N_2O_5, which is generated when two reactive nitrogen species react. For example, consider again two explicit sources plus a vanilla source type, the following nine reactions are needed to describe the formation of N_2O_5 and keep track of the sources of NOx, as shown in RS2:

```
NO2 + NO3 = N2O5
NO2 + NO3_X1 = N2O5_X01
NO2 + NO3_X2 = N2O5_X02
NO2_X1 + NO3 = N2O5_X10
NO2_X1 + NO3_X1 = N2O5_X11
NO2_X1 + NO3_X2 = N2O5_X12
NO2_X2 + NO3 = N2O5_X20
NO2_X2 + NO3_X1 = N2O5_X21
NO2_X2 + NO3_X2 = N2O5_X22
```

(RS2)

In the above notation, the double source-tagged N_2O_5 has two numbers after "_X" to indicate the individual source of NO_2 and NO_3 that form the specific N_2O_5. The '0' represents NO_2 or NO_3 from the vanilla type. Nine different types of N_2O_5 species are needed to properly track sources of NO_x, as N_2O_5 dissociates to give NO_2 and NO_3, which contributes to ozone and particulate nitrate formation. In this regard, the number of N_2O_5 species also increases quadratically with the number of explicit sources, leading to near-quadratic growth of the overall number of species when the number of types to track gets higher.

In ozone source apportionment calculations, it is also necessary to track the sources of primary emitted VOCs as well as some of their reaction products in addition to the sources of NOx. Some of the unsaturated VOCs such as olefins can react with the NO_3 radical. In the source-oriented mechanism, the number of reactions needed for these VOC+NO_3 reactions also increases quadratically, as shown in RS3 below, using the ethene (ETHE) + NO_3 reaction from the SAPRC-99 mechanism as an example. Again, nine reactions are needed for three types of NOx and ETHE.

```
ETHE + NO3 = RO2_R + RCHO
ETHE + NO3_X1 = RO2_R + RCHO
ETHE + NO3_X2 = RO2_R + RCHO
ETHE_X1 + NO3 = RO2_R + RCHO
ETHE_X1 + NO3_X1 = RO2_R + RCHO
ETHE_X1 + NO3_X2 = RO2_R + RCHO
ETHE_X2 + NO3 = RO2_R + RCHO
ETHE_X2 + NO3_X1 = RO2_R + RCHO
ETHE_X2 + NO3_X2 = RO2_R + RCHO
```

(RS3)

For accurate VOC source apportionment calculations that involve reactions between two source-oriented organic radical species, such quadratic dependence of source types and reaction numbers also arises (Ying and Krishnan, 2010). For example, consider the reaction between CCO_O2 (acetyl peroxy radical, CH3C(O)O2.) and C_O2 (methyl peroxy radical, CH3O2.) that forms CCO_OH (peroxy acetic acid, CH3C(O)OH) and formaldehyde (HCHO) in SAPRC-99, the following nine reactions (RS4) are needed:

CCO_O2 + C_O2 = CCO_OH + HCHO
CCO_O2 + C_O2_X1 = CCO_OH + HCHO_X1
CCO_O2 + C_O2_X2 = CCO_OH + HCHO_X2
CCO_O2_X1 + C_O2 = CCO_OH_X1 + HCHO
CCO_O2_X1 + C_O2_X1 = CCO_OH_X1 + HCHO_X1 (RS4)
CCO_O2_X1 + C_O2_X2 = CCO_OH_X1 + HCHO_X2
CCO_O2_X2 + C_O2 = CCO_OH_X2 + HCHO
CCO_O2_X2 + C_O2_X1 = CCO_OH_X2 + HCHO_X1
CCO_O2_X2 + C_O2_X2 = CCO_OH_X2 + HCHO_X2

Due to the necessity of explicitly handling some or all of these reactions in source-oriented mechanisms, the source-oriented modeling approach becomes extremely computationally intensive so that previous applications were limited to up to 9 explicit sources for secondary nitrate in a single run (Kleeman and Cass, 2001; Ying et al., 2004; Ying et al., 2014b). In some previous work for VOC and secondary organic aerosol source apportionments, only one explicit source was tracked at a time to simplify the reactions and to reduce the computation burden (Wang et al., 2018b; Ying and Krishnan, 2010). However, multiple model runs are needed to determine the contributions from all sources. To make the source-oriented approach practical for a larger number of source types, it is necessary to improve the computation efficiency of the source-oriented approach.

3.2.2. Method

In the original source-oriented model, a reaction set that involves two source-tagged species as reactants for n source types can be written in the following more general form, with n^2 reactions:

A1 + B1 = C1 + D1 + E, k
...
A1 + Bn = C1 + Dn + E
...
...
Ai + B1 = Ci + D1 + E
... (RS5a)
Ai + Bn = Ci + Dn + E
...
...
An + B1 = Cn + D1 + E
...
An + Bn = Cn + Dn + E

where A, B, C, and D are typed model species and the numbers following the letters denote the source origin index of these species. C is the product from A and D is the product from B. E

represents a general product whose source-origin is not tracked in the model simulation. k is the second-order reaction rate coefficient, which is the same for all the reactions in this reaction set. Reaction sets RS1, RS3, and RS4 can all be expressed in this form.

Based on the above reactions, the loss rate of Ai and Bi (the i^{th} typed species of A and B, respectively) can be written as pseudo first-order reaction rate equations (3-9a) and (3-9b), respectively:

$$d[Ai]/dt = -k[Ai]([B1]+[B2]+...+[Bn])=-k[B\_tot][Ai]=k_{eff1}[Ai] \quad (3\text{-}9a)$$
$$d[Bi]/dt = -k[Bi]([A1]+[A2]+...+[An])=-k[A\_tot][Bi]=k_{eff2}[Bi] \quad (3\text{-}9b)$$

where [A_tot]= [A1]+[A2]+...+[An] and [B_tot]=[B1]+[B2]+…+[Bn] are the total concentrations of species A and B, respectively. The pseudo first-order reaction rate coefficients k_{eff1} and k_{eff2} are product of k and [B_tot] and [A_tot], respectively.

The above n^2 reactions can be equivalently represented by the following $2n$ pseudo first-order reactions, with the pseudo first-order reaction rate coefficients as defined below,

$$\begin{aligned} &A1 = C1 + E, && k_{eff1}=k[B\_tot] \\ &... \\ &An = Cn + E, && k_{eff1}=k[B\_tot] \\ &B1 = D1, && k_{eff2}=k[A\_tot] \\ &... \\ &Bn = Dn, && k_{eff2}=k[A\_tot] \end{aligned} \quad (RS5b)$$

It is easy to show that the formation rate of Ci and Di from this set of $2n$ equations is the same as the original n^2 equations, as shown in equations (3-10a) and (3-10b):

$$d[Ci]/dt = k_{eff1}[Ai] = k[B\_tot][Ai] = k[B1][Ai] + k[B2][Ai] + ... + k[Bn][Ai] \quad (3\text{-}10a)$$
$$d[Di]/dt = k_{eff2}[Bi] = k[A\_tot][Bi] = k[Bi][A1] + k[Bi][A2] - ... + k[Bi][An] \quad (3\text{-}10b)$$

For the non-typed product E, it can appear in either the Ai reactions or the Bi reactions. For example, if it appears in the Ai reactions, as shown above, the overall formation rate of E is

$$d[E]/dt = k_{eff1}[A1] + k_{eff1}[A2] + ... k_{eff1}[An] = k_{eff1}[A\_tot] = k[B\_tot][A\_tot] \quad (3\text{-}11)$$

This can also be derived based on the original equation set with n^2 reactions. It is also easy to show that if E goes with the Bi reaction, the overall formation rate of E is still the same as equation (3-11).

The double-typed N_2O_5 formation reactions can be simplified as well. For N2O5_Xij, it can be equivalently written as 0.5N2O5_Xi + 0.5N2O5_Xj, in terms of preserving the source contributions to NO_2 and NO_3, which is needed for the proper source apportionment of nitrate and ozone. With this simplification, as well as the simplification of the n^2 reactions described above, the reactions of N_2O_5 formation for n source types can be written in $2n$ reactions:

$$\begin{aligned} &NO2 = 0.5*N2O5, && k_{eff1} = k[NO3\_tot] \\ &... \\ &NO2\_Xn = 0.5*N2O5\_Xn \\ &NO3 = 0.5*N2O5, && k_{eff2} = k[NO2\_tot] \end{aligned} \quad (RS6)$$

...
NO3_Xn = 0.5*N2O5_Xn

With double-typed N_2O_5 species, a total of n^2 N_2O_5 reactions are needed with the original source-oriented representation. Now, only $2n$ reactions for n single-typed N_2O_5 species are needed. This represents a significant reduction in terms of the number of species as well as the number of reactions for the source-oriented mechanism.

3.2.3. Implementation of the improved method

To implement this approach in a kinetic modeling framework in air quality models, two approaches can be considered. The first approach uses additional species to track the total concentrations of the typed species that are needed in the calculation of the pseudo first-order reaction rates (NO2_tot, NO_tot, etc.). Since the total concentrations are included as model species, the reactions are still second-order and the reaction rate coefficients remain unchanged, as shown in RS5c below:

A1 + B_tot = C1 + E, k
...
An + B_tot = Cn + E
B1 + A_tot = D1 (RS5c)
...
Bn + A_tot = Dn

The total concentrations of the A and B species might be affected in other reactions. In any other reactions that involve with the typed species, the corresponding total concentration also needs to be included in the product list so that total concentrations are kept up-to-date. For example, reactions in (RS7),

X + A1 = Y + B1
... (RS7)
X + An = Y + Bn

need to be written as:

X + A1 = Y - A_tot + B_tot (RS7')
...
X + An = Y - A_tot + B_tot

The advantage of this approach is that no special modifications to the air quality model function that calculates the reaction rates are needed. The pseudo first-order reaction rates (i.e. the k_{eff}'s mentioned above) are automatically handled by the inclusion of total concentrations in the calculations.

The second approach is to modify the reaction rate coefficient function/subroutine in the kinetic modeling framework to calculate the total concentrations of the typed species without additional explicit model species. These total concentrations calculated on-the-fly are then used to calculate the pseudo first-order rate coefficients for the typed reactions shown above. The function that calculates the reaction rates needs to be modified to recognize these special pseudo first-order reactions. Usually, modification of the mechanism pre-processor, which

converts the reaction mechanisms into modules or include files for the kinetic solver is needed as well. Code modification is usually more cumbersome, but the advantage of this approach is that it is computationally more efficient as the total concentrations do not need to be tracked explicitly.

The chemical mechanism preprocessor (CHEMMECH) and the general ode solvers (SMVGEAR (a vectorized GEAR solver) and ROS3 (an implementation of the Rosenbrook solver) in the Community Multiscale Air Quality (CMAQ) model have built-in capability to treat "special" reaction rates, which makes them ideal for testing the second approach. List 1 shows a simplified example of the mechanism definition file compatible with the CHEMMECH preprocessor, which implements RS1 for 10 NO_x source types.

List 1. The special reaction rate section (between the keywords SPECIAL and END SPECIAL) and reactions used to implement the source-oriented NO2+NO3 reactions with 10 source types using the chemical mechanism preprocessor CHEMMECH for the CMAQ model. Due to the limitation of the current mechanism preprocessor, a dummy reaction (<10_dum>) is needed so that the original reaction rate can be included in the calculation of the special rate constants. Using the special rate expression is signaled by including the symbol '?' in the reaction rate coefficient expression.

```
SPECIAL =
RNO_NO3 = K<10_dum>*C<NO3> + K<10_dum>*C<NO3_X1> + <10_dum>*C<NO3_X2>
        + K<10_dum>*C<NO3_X3> + K<10_dum>*C<NO3_X4> +
K<10_dum>*C<NO3_X5>
        + K<10_dum>*C<NO3_X6> + K<10_dum>*C<NO3_X7> +
K<10_dum>*C<NO3_X8>
        + K<10_dum>*C<NO3_X9>;
RNO3_NO = K<10_dum>*C<NO> + K<10_dum>*C<NO_X1> + K<10_dum>*C<NO_X2> +
          K<10_dum>*C<NO_X3> + K<10_dum>*C<NO_X4> +
K<10_dum>*C<NO_X5> +
          K<10_dum>*C<NO_X6> + K<10_dum>*C<NO_X7> +
K<10_dum>*C<NO_X8> +
          K<10_dum>*C<NO_X9>;
...
END SPECIAL
...
<10_dum> dummy1 + dummy1 = dummy1 + dummy1 #1.80e-11@-110;
<10_aX0> NO = NO2 #1.0?RNO_NO3;
<10_aX1> NO_X1 = NO2_X1 #1.0?RNO_NO3;
<10_aX2> NO_X2 = NO2_X2 #1.0?RNO_NO3;
...
<10_aX9> NO_X9 = NO2_X9 #1.0?RNO_NO3;
<10_bX0> NO3 = NO2 #1.0?RNO3_NO;
<10_bX1> NO3_X1 = NO2_X1 #1.0?RNO3_NO;
<10_bX2> NO3_X2 = NO2_X2 #1.0?RNO3_NO;
...
<10_bX9> NO3_X9 = NO2_X9 #1.0?RNO3_NO;
```

3.2.4. Results

3.2.4.1 Test Mechanism

To evaluate how much improvement in computation efficiency can be achieved by using this method, a series of source-oriented mechanisms for simultaneous attribution of ozone and secondary inorganic aerosol were constructed based on the SAPRC-99 photochemical mechanism. The SAPRC-99 mechanism is chosen instead of the more recent versions of SAPRC because it is faster with fewer species and reactions, and thus is more suitable for operational air quality forecasting. The source-oriented mechanism based on this will be applied in a future air quality forecasting model that also forecasts source and source-region contributions to air pollution. Since the main purpose of this paper is to test the algorithm, aerosol and cloud processes were disabled in the analyses described below.

The original SAPRC-99 mechanism used in this study contains a total of 90 species and 226 reactions. Among these species, 14 species are reactive nitrogen species. In addition, SO_2 and sulfuric acid (SULF) were also expanded in the source-oriented mechanism. To evaluate source contributions to ozone, 13 primary VOC species were also treated as source-oriented species. As formaldehyde (HCHO) is an important oxidation product from several parent VOCs, sources of secondary HCHO from the first-generation oxidation of parent VOCs are also tracked. The original mechanism and the source-oriented mechanisms are solved using the vectorized GEAR (SMVGEAR) solver in CMAQ, with a relative tolerance (RTOL) of 1e-3 and an absolute tolerance (ATOL) of 1e-9. Although the SAPRC-99 has a more efficient Euler Backward Iterative (EBI) solver for the stock SAPRC-99 mechanism, it is not used because it cannot be used for the source-oriented mechanisms without modifications. Another limitation of the EBI solver in CMAQ is that it does not incorporate the ability for integrated reaction rate (IRR) analysis, an important feature that is needed for the ozone source apportionment algorithm.

3.2.4.2 Solution Accuracy

To test the solution accuracy, a source-oriented mechanism using the improve source-oriented algorithm described in the previous section with special pseudo first-order reaction treatment was constructed for 10 different source types (one vanilla and 9 explicit sources). This version of the source-oriented mechanism contains 399 species and 1372 reactions (see Table 3-2).

Table 3.2 Number of model species, reactions and gas-phase chemistry time for source-oriented mechanisms with the increasing number of simultaneous sources

# of sources	Species	Ratio	Reactions	Ratio	Time (s)	Ratio
1	90	1.000	226	1.000	466.12	1.000
5	229	2.544	752	3.327	1560.43	3.348
10	399	4.433	1372	6.071	3207.34	6.881
16	603	6.700	2116	9.363	5268.02	11.302
Old 10*	331	3.678	1772	7.841	4644.00	9.963

* An old source apportionment scheme that tracks 10 sources of NOx and their products for secondary nitrate source apportionment. Primary VOCs, HCHO and ozone were not tracked in that mechanism.

The following simulations were conducted for a two-day period, with (1) the original SAPRC-99 mechanism and (2) source-oriented mechanism with emissions evenly distributed into 9 explicit types (initial and boundary conditions mapped to the vanilla type). The following graph shows that total ozone, NO, NO2 and HCHO on day 2 at 1400-1500 from all three cases

are nearly identical. Thus, the improved source-oriented method is shown to yield the same overall concentrations as the non-source-oriented mechanism for all practical purposes.

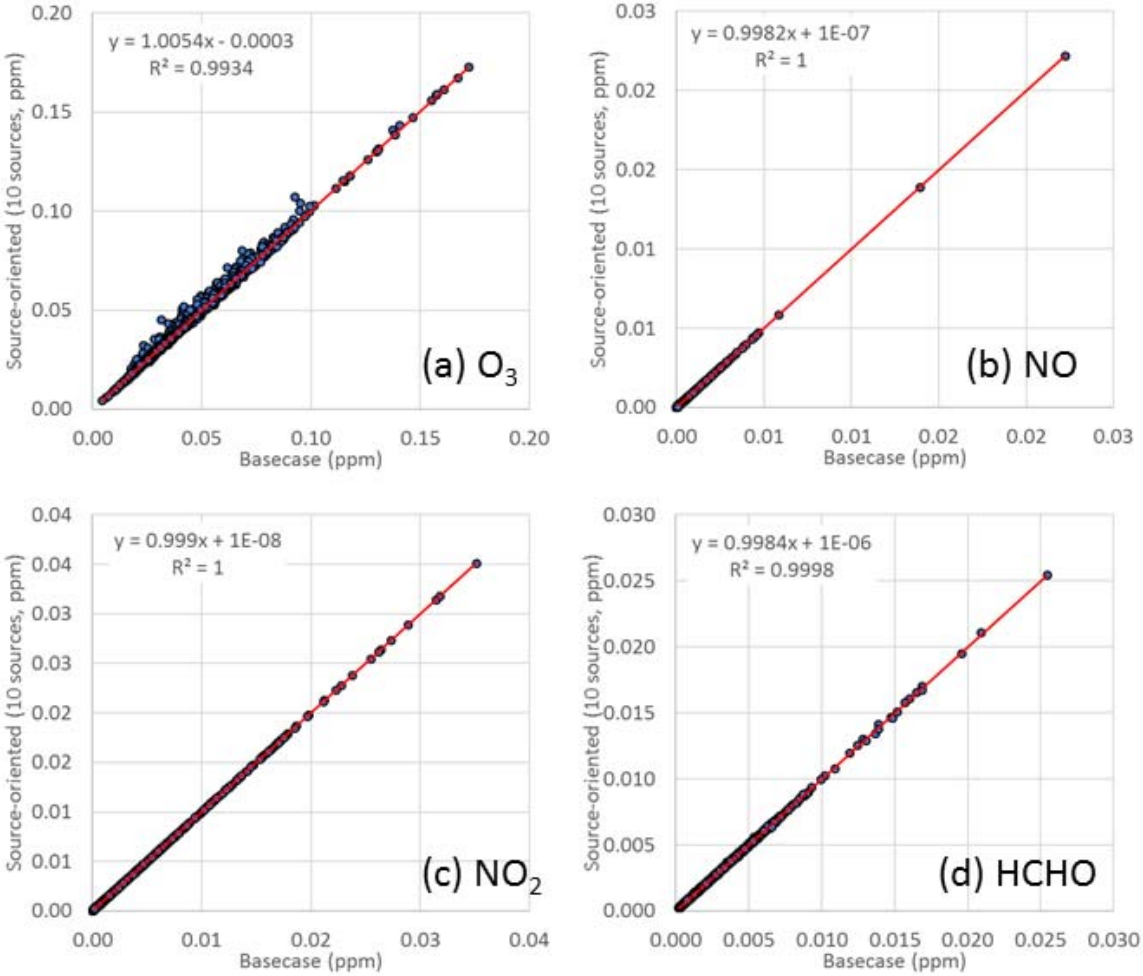


Figure 3.1 Comparison of predicted (a) O_3, (b) NO, (c) NO_2, (d) HCHO, (e) HO and (f) HO_2 concentrations using the original SAPRC-99 and the 10-source (vanilla + 9 sources) SAPRC-99 with all emissions evenly distributed among 10 source types.

3.2.4.3 Computation efficiency

A 2-day simulation for August 28-29, 2006 for the eastern United States was used in the analyses. Overall emissions of all source-oriented primary emission species were evenly divided into N sources for a source-oriented mechanism that has N explicit sources. The number of species and reactions of these source-oriented mechanisms are listed in Table 1. The mechanisms for 5, 10 and 16 source types only need about 2.5, 4.4 and 6.7 times of the original species and 3.3, 6.1 and 9.4 times of the original reactions. The number of species and reactions for a previous version of the source apportionment scheme (also based on SAPRC-99) that tracks 10 sources of NOx and their products for secondary nitrate source apportionment (Qiao et al., 2018) are also included for comparison. Note that this old mechanism only tracks NOx and SOx species and does not track VOCs and no ozone tracers are included in the mechanism either. The number of reactions for this mechanism is 1772, or 7.8 times that of the original SAPRC-99. In comparison, the mechanism with 10 source types using the efficient approach described in this paper only needs 1372 reactions, which is only 6.1 times of the original SAPRC-99, even though the new scheme includes source-oriented reactions for primary VOCs, HCHO, and ozone tracers

(approximately 400 additional reactions are needed for these). If these reactions were excluded, the new scheme would only require 972 reactions, or ~45% fewer reactions than the old approach to obtain the same source apportionment results for secondary inorganic aerosols.

This set of simulations were conducted on a 10-core/20-thread E5-2670 v2. Only 8 cores were requested for these simulations. The wall-clock time for the gas-phase mechanism as well as the total run time was recorded using the time function in MPI. Figure 2 shows the details of the computation time for the gas-phase chemistry. Figure 2(a) shows the time spent in gas-phase chemistry at each time step for four sets of simulations with 1, 5, 10 and 16 sources, respectively, for the entire two-day simulation period. Figure 2(b) shows the total amount of time in the gas-phase chemistry for the two-day simulations. The new scheme that includes source apportionment of NOx, SO2, NH3, for 10 sources needs only 70% of the time of the old scheme yet generates much more information.

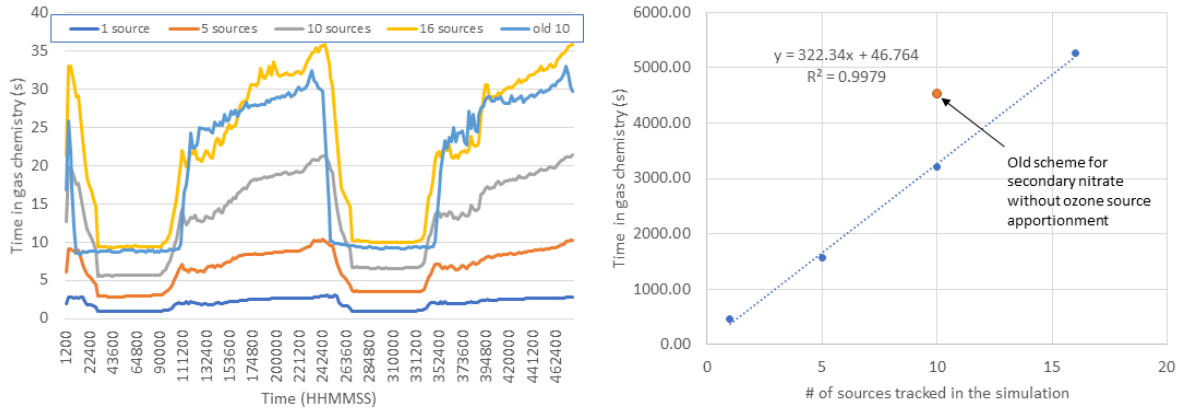


Figure 3.2(a) Wall-clock time for gas phase chemistry at each time step during the two-day simulation (August 28-29, 2006) for simulations with a different number of sources (including the vanilla type) and (b) total wall-clock time for the gas phase chemistry part of the two-day simulation. Units are seconds. All simulations were performed on a 10-core/20-thread E5-2670 v2. Only 8 cores were requested for these simulations.

3.2.5. Conclusion

In this study, the computation efficiency and thus scalability of the source-oriented approach is greatly improved with a new approach of dealing with these two-tagged-species reactions. The new approach is based on tracking the total concentration of the source-tagged species and reduce the n^2 number of second-order reactions for n sources into 2n pseudo first-order reactions. The production and removal rate of individual species remains unchanged in the new approach. The number of reactions and number of model species increases linearly with the number of source types, thus greatly improved the computation efficiency. Test cases based on the Texas Air Quality Study 2006 ozone episode showed that a source-oriented SAPRC-99 mechanism that simultaneously performs the source apportionment of NOx, SO2, primary VOCs, HCHO, and ozone for 16 sources needs only 11 times of the computation time of the original non-source-oriented mechanism. While efficient source-oriented approach for primary particles are already available to track a large number of sources simultaneously, the efficient approach developed in this study has the potential to track a large number of sources to evaluate their impact on secondary pollutant formation, and has the potential to be applied in air quality forecasting models that provide source or source-region contribution information for policymakers for better emission regulations under adverse meteorological conditions.

3.3. Test of the source and age-resolved model

To test the source and age-resolved CMAQ model based on the CS07A chemical mechanism, a previous summer ozone episode (August 15 to September 7, 2000) during the Texas Air Quality Study 2000 was studied. The most vigorous wind rotation took place between August 25 and September 6, 2000. The model domains for this study is shown in Figure 3.3. The 36-km resolution domain covers the Eastern US and the 12-km domain covers East Texas, which includes the Houston, Galveston and Beaumont and Port Arthur (HGBPA) areas. Since this is a test run, 4-km and 1-km domains are not used in the study.

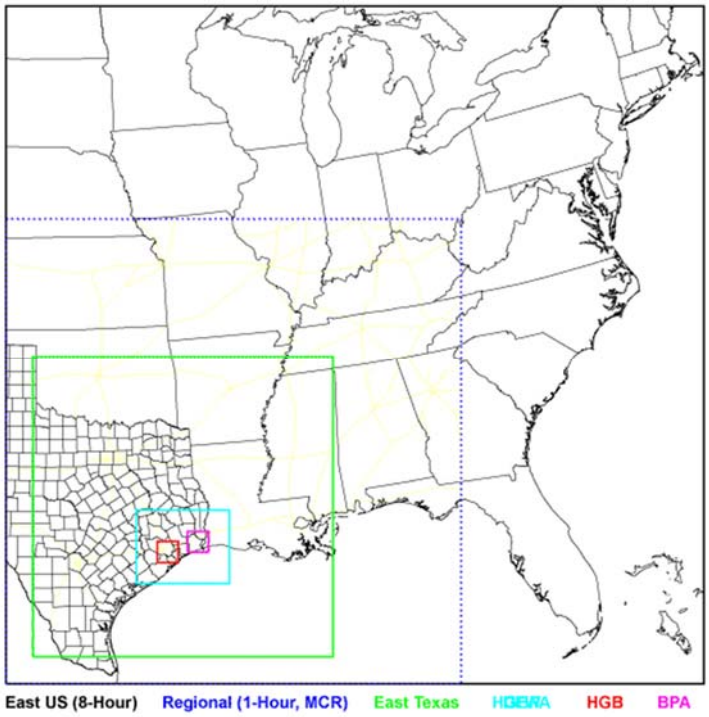


Figure 3.3 CMAQ model domains used in this study: East US (36-km horizontal resolution) and East Texas (12-km).

Emissions of anthropogenic emissions were based on the National Emission Inventory 2002 and biogenic emissions were generated using the Biogenic Emission Inventory System (BEIS) v3.12. Both biogenic and anthropogenic emissions are processed using the SMOKE emission processor. Details of the emission processing for this modeling episode has been reported in Zhang and Ying (2010). The meteorological inputs to drive the model simulation were prepared using outputs from the MM5 model provided by the Texas Commission on Environmental Quality (TCEQ). Several sets of simulations were conducted: (1) Base case simulation without age and source tracking; (2) Source-region apportionment simulation – the age distribution capability was turned off; (3) Atmospheric age distribution simulation with 1-hour time-bins; and (4) Atmospheric age distribution simulation with 3-hour time-bins; Simulations were conducted from August 15, 2000 to September 7, 2000. The first five days were used as spin-up days and the results were not used in the analysis.

Figure 3.4 compares the predicted ozone concentrations at Galveston (a coastal site) and Aldine (an urban site north of downtown Houston). The model predictions agree better with observation in the second half of the simulation, between August 27 and September 6, 2000, at both locations. At Aldine, a significant amount of non-background ozone with peak concentration around 30-40 ppb was predicted on top of the background ozone. The non-background ozone is mostly attributable to NOx, although on a few days (e.g. August 28 and September 1) more ozone is attributed to VOCs. At Galveston, on most of the days, the model predicts low concentrations of non-background ozone. However, high concentrations of ozone were predicted on a number of days in late August and early September. These high concentrations of ozone were attributed almost entirely to NOx. The atmospheric age distribution shows that at peak ozone hours in Aldine, approximately 10-12 ppb of ozone was formed within one hour and ozone with atmospheric age of 1-2 hours have higher concentrations than fresh ozone. This is due to additional ozone formation from aged precursors. A small amount of ozone with atmospheric age of 3-4 hours also contributed to the total ozone concentrations at both locations. It is also obvious that on higher ozone concentration days there are more contributions due to ozone with older atmospheric ages.

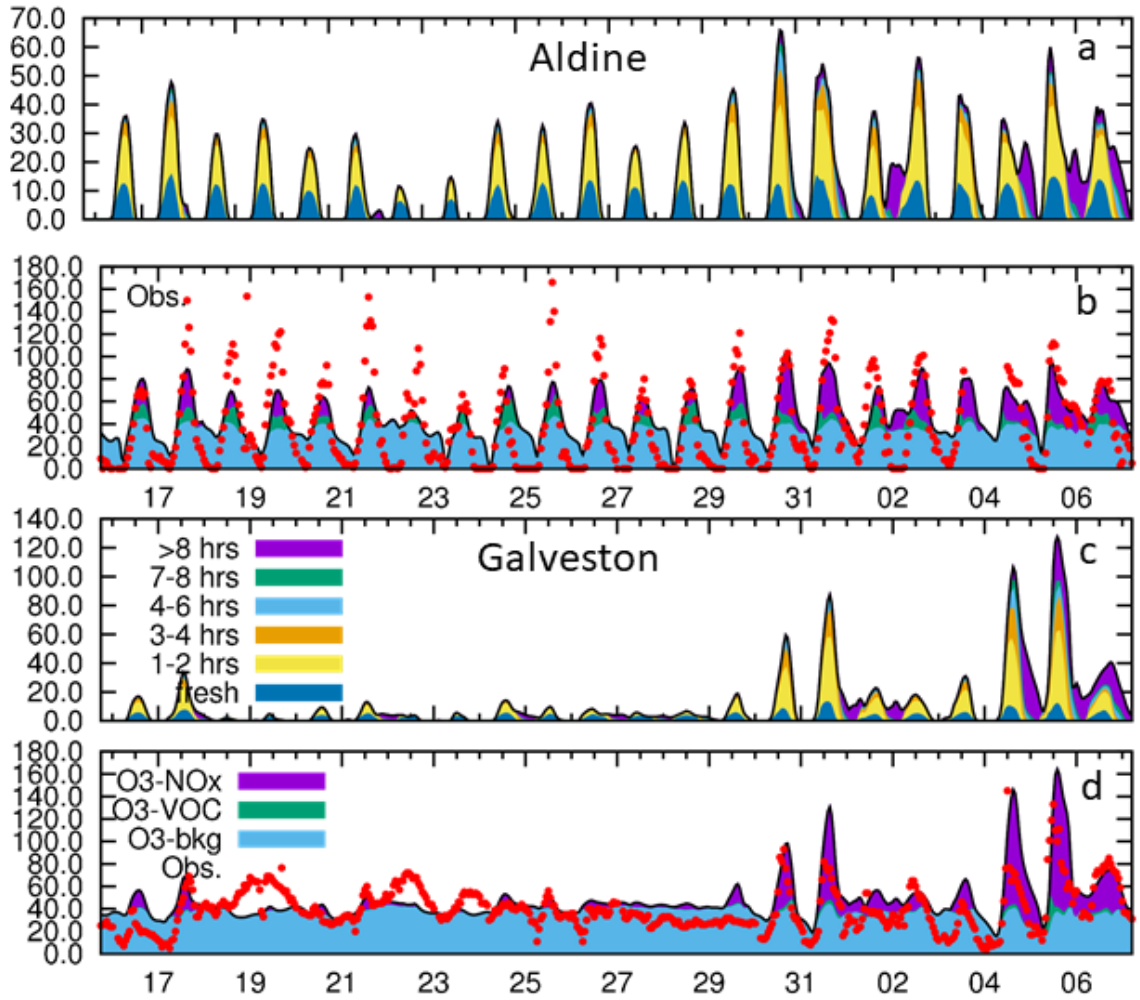


Figure 3.4 Predicted atmospheric age distribution of non-background ozone (i.e. ozone attributed to NOx and VOCs) at Aldine (a) and Galveston (c) and the breakdown of predicted total ozone to NOx, VOC and background

contributions (b and d). Observed and predicted concentrations are in units of ppb. The age distribution results are based on 1-hr time-bin and a few age bins are combined to make it easier to visualize the results.

High contributions of highly aged ozone (>8 hours old) occurred at both Aldine and Galveston on the nights of September 1-2, 4-5, 5-6 and 6-7. On 5-6 and 6-7, the aged ozone was originated from Louisiana but on the other two nights, it was due to emissions originated in Texas (Figure 3.5). Banta et al. (2005) focused on August 30 and presented evidence for the bulk of the ozone being multi-hour aged pollution. Figure 3.6 shows the simulated O_3 concentrations with different atmospheric ages on 1500-1600 CST, August 2000. The results clearly show that aged ozone (or ozone from aged precursors) accounts for a majority of the non-background ozone in the south-east portion of the modeling area, along the coastline of the Gulf of Mexico.

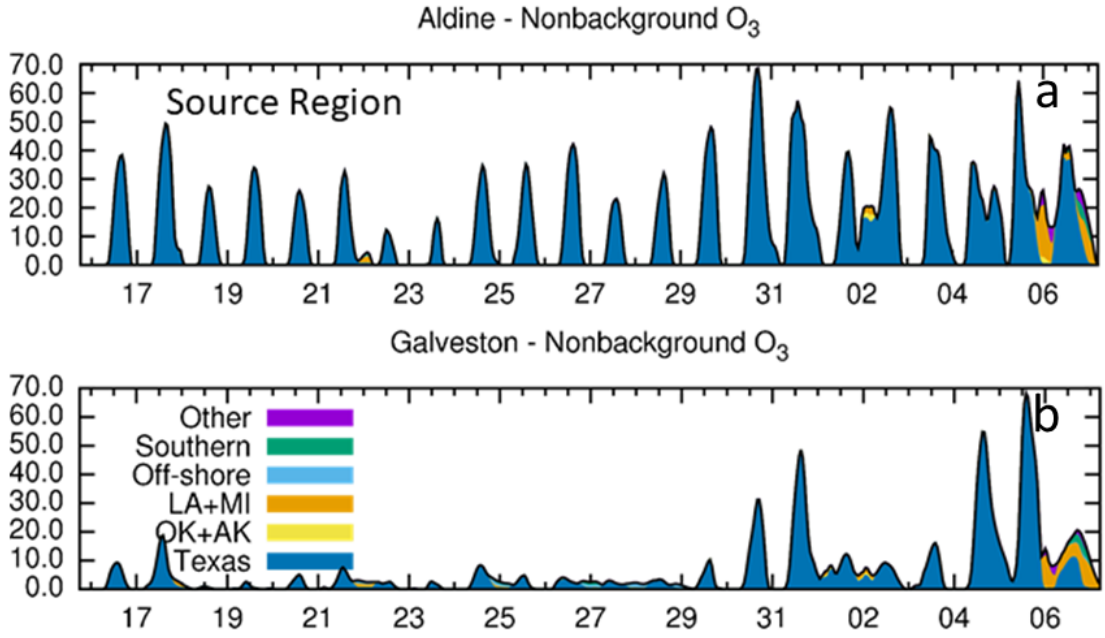


Figure 3.5 Source-region contributions to predicted non-background ozone concentrations at Aldine (a) and Galveston (b).

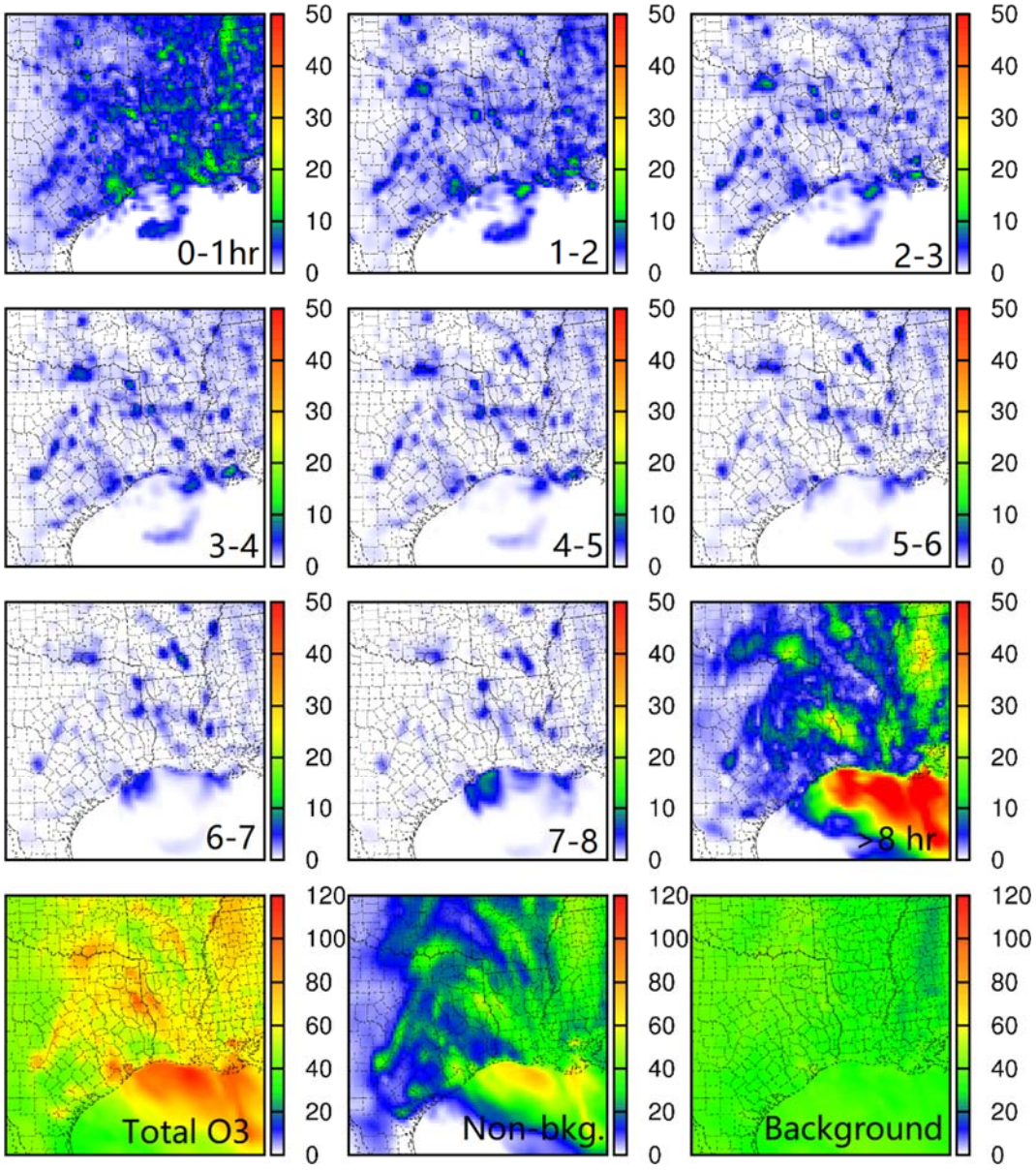


Figure 3.6 Atmospheric age distribution of non-background ozone (top three rows); total ozone, non-background ozone and background ozone (last row) concentrations on 1500-1600 CST, August 30, 2000. Units are ppb.

Figure 3.7 shows that the vertical distributions of ozone at Galveston and Aldine have distinctive patterns. At Galveston on high surface ozone days (August 30-31, and September 4-5), high ozone is limited below 500 m due to the shallower boundary layer. From September 1-3, high ozone concentrations occur at higher elevations between 500-1000 m, and some downward transport can be observed in the afternoons. At Aldine, ozone concentrations always peak near the surface and high ozone concentrations extend to 1500 m or more above the surface in the afternoon.

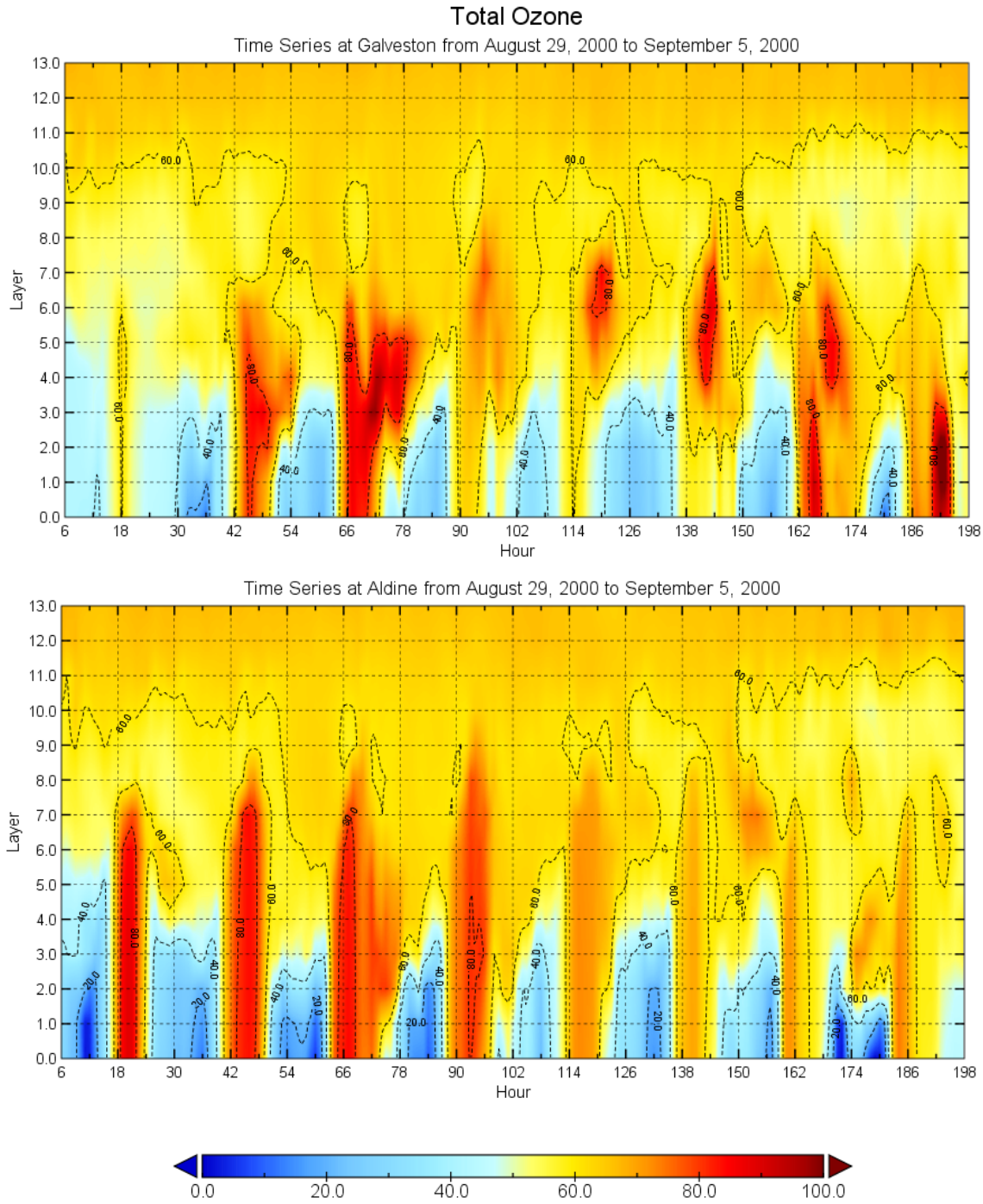


Figure 3.7 Time series of the vertical distribution of ozone at Galveston (top) and Aldine (bottom) from August 29, 2000, to September 5, 2000. Hour 6 is the 0000 hours local time on August 29, 2000. The mid-layer evaluations for each layer from the lowest are 22, 68, 135, 275, 460, 700, 1090, 1640, 2280, 3140, 4330, 6180, 9550, and 16200 m.

Figure 3.8 shows that aged ozone (> 8 hours) at Galveston can be a significant contributor to total ozone at both surface and higher elevations. On August 30 and 31, higher aged ozone occurred at higher elevations about 250-500 m above the surface with concentrations exceed 30 ppb. In the afternoon and evening hours of September 4, a high concentration of aged ozone

occurred throughout the boundary layer and lasted for more than 12 hours. Fractional contributions of aged ozone to total ozone exceeds 50% on these hours.

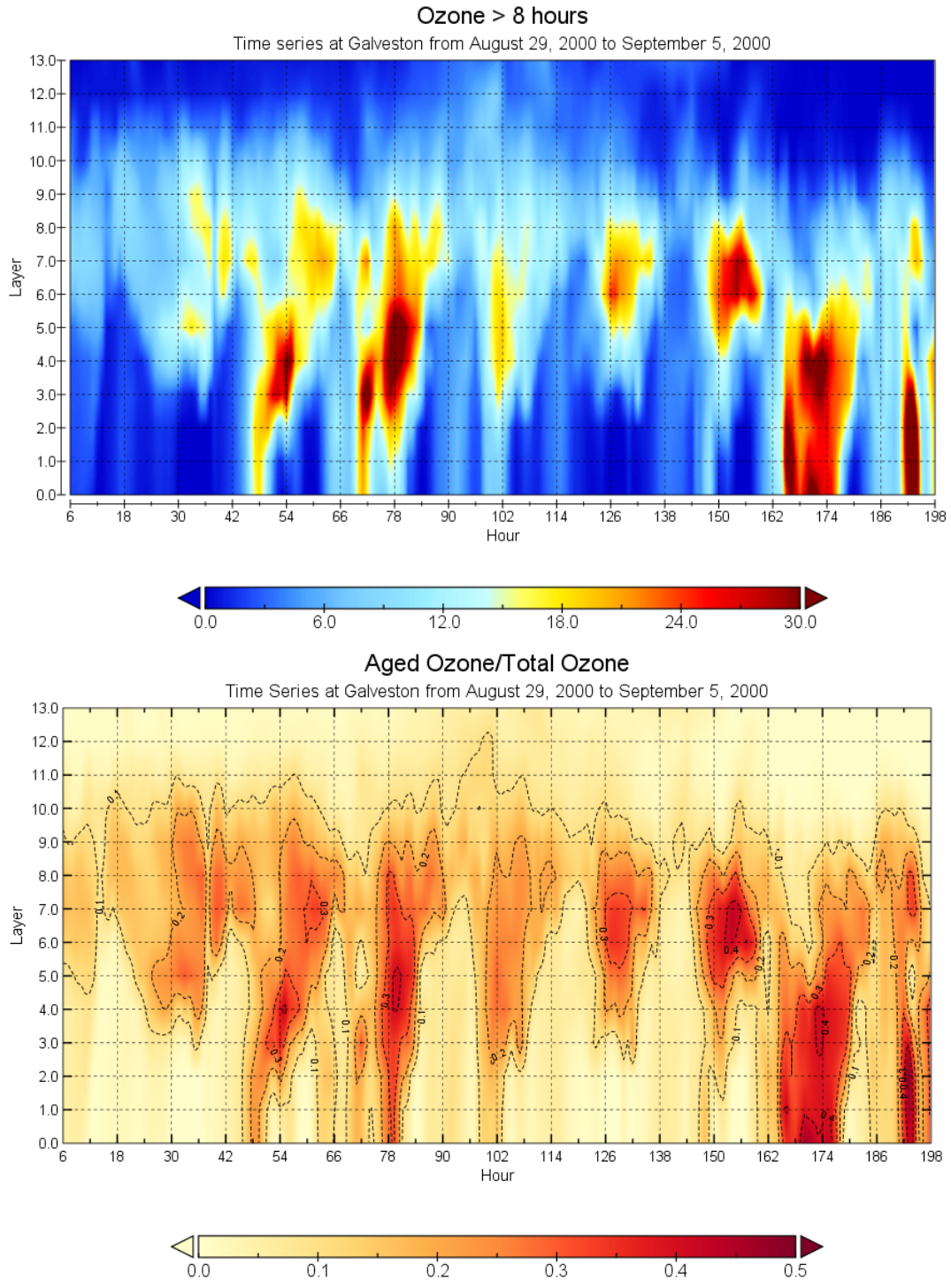


Figure 3.8 Time series of the vertical distribution of aged ozone concentrations (top) and aged ozone fraction (bottom) at Galveston from August 29, 2000, to September 5, 2000. Hour 6 is the 0000 hours local time on August 29, 2000.

Figure 3.9 shows that at Aldine, aged ozone is limited at higher evaluations near the top of the boundary layer. Concentrations of aged ozone near the surface are low, with contributions to total ozone less than 10% in most times.

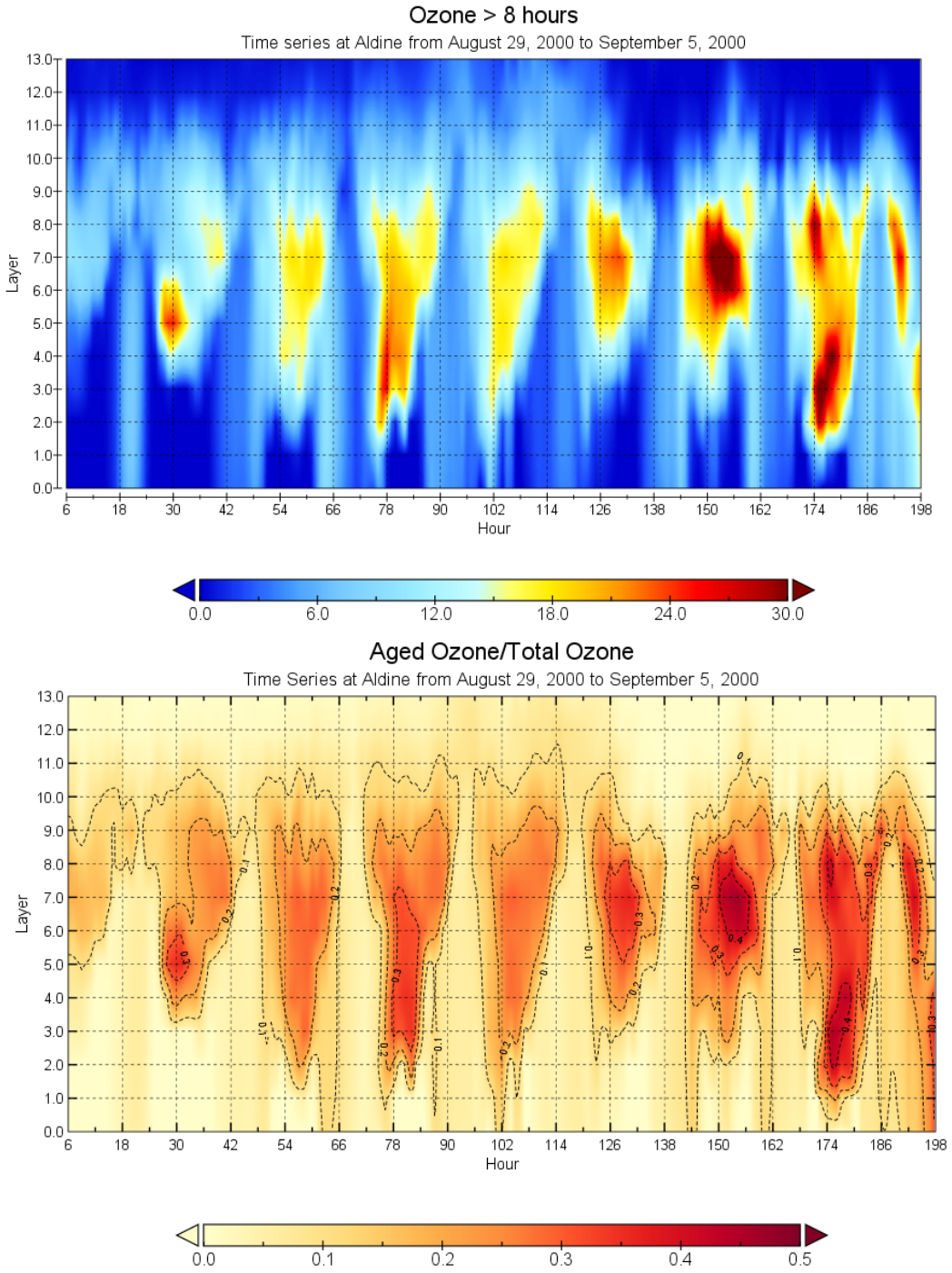


Figure 3.9 Time series of the vertical distribution of aged ozone concentrations and aged ozone fractions at Aldine from August 29, 2000, to September 5, 2000. Hour 6 is the 0000 hours local time on August 29, 2000.

Contributions of regional emissions to ozone at Galveston are shown in Figure 3.10 and Figure 3.11. Throughout the entire boundary layer, most of the ozone at Galveston is due to Texas

emissions. All other regions have small contributions to non-background ozone. Ozone from emissions in Louisiana and Mississippi contribute to 10-20 ppb of ozone on the night of September 4 and early morning hours of September 5. This long-range transport of ozone appears to occur within 0-500 m above the surface.

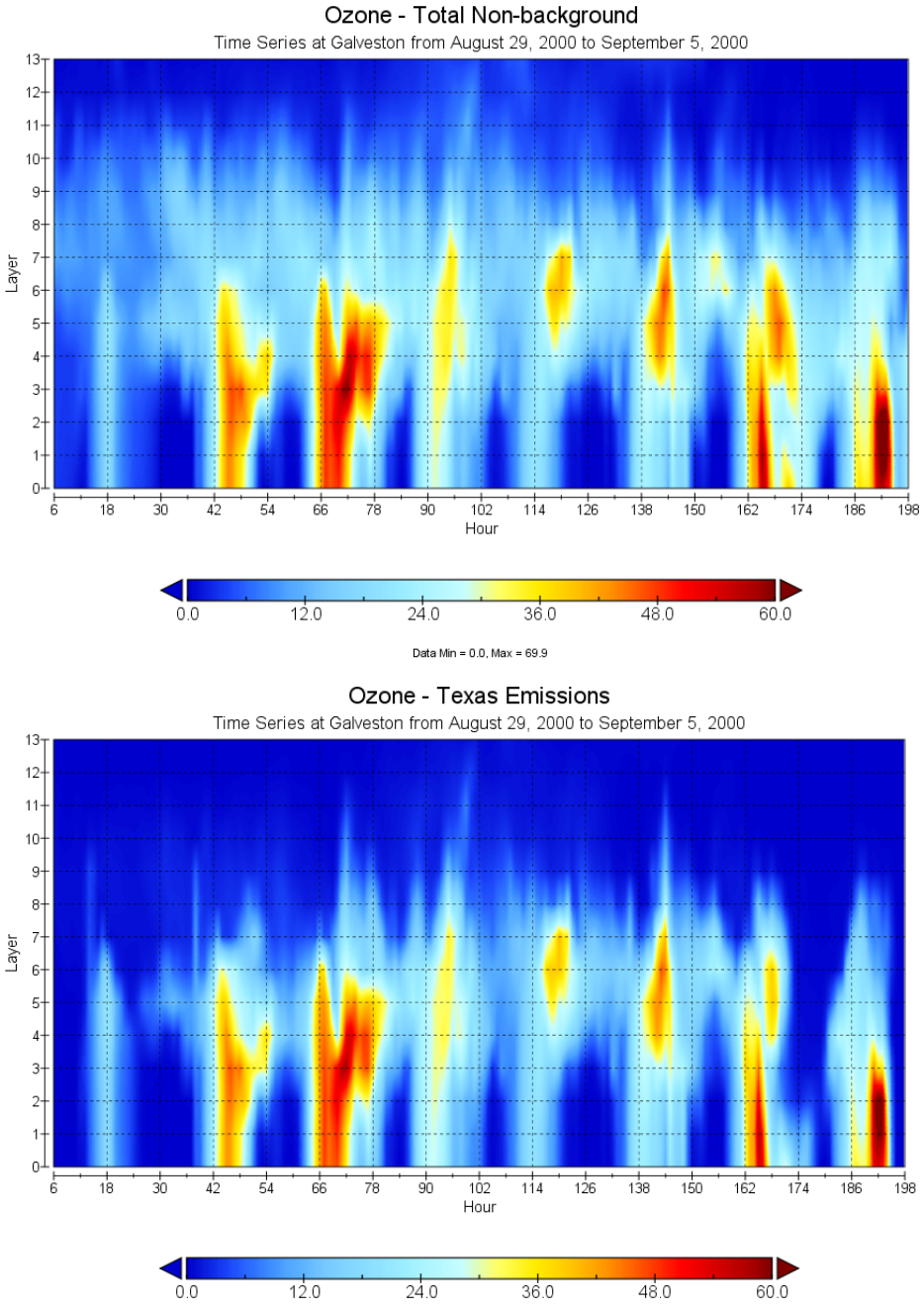


Figure 3.10 Time series of the vertical distribution of total non-background ozone and ozone at Galveston due to Texas emissions from August 29, 2000, to September 5, 2000. Hour 6 is the 0000 hours local time on August 29, 2000.

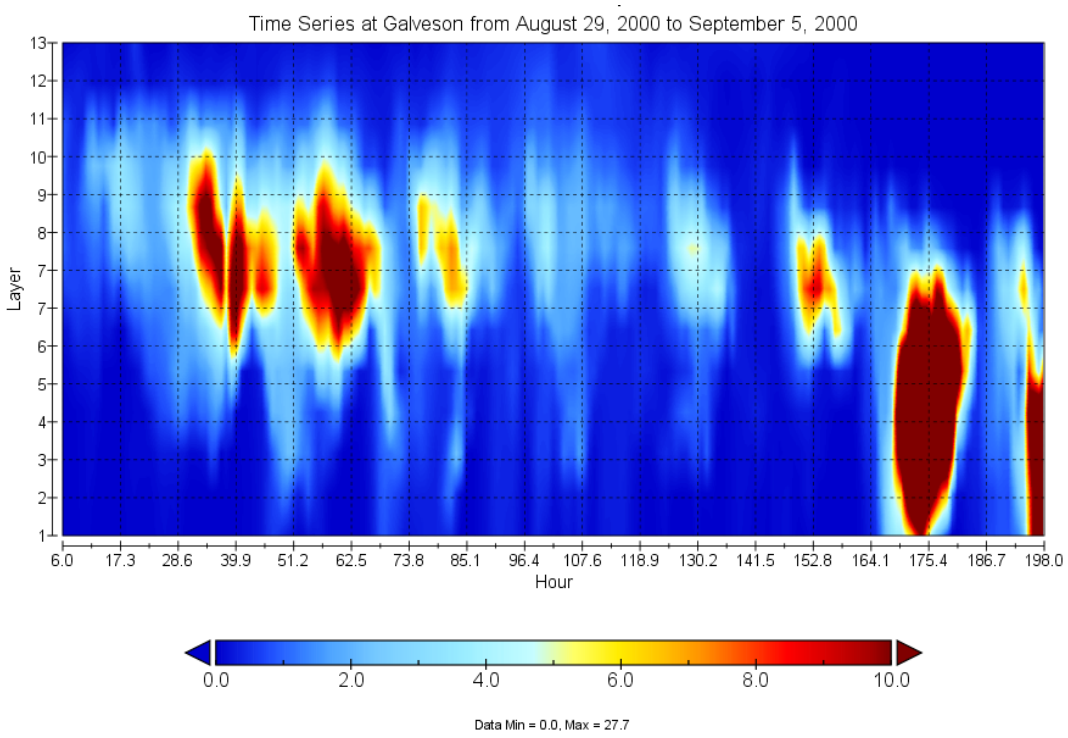


Figure 3.11 Time series of the vertical distribution of ozone at Galveston due to Louisiana and Mississippi emissions from August 29, 2000, to September 5, 2000. Hour 6 is the 0000 hours local time on August 29, 2000.

3.4. Quality Assurance

The source-oriented CMAQ code was mostly developed by Dr. Ying. Jie Zhang subsequently studied all (i.e. 100%) of the modified code and inspected for potential errors. The accuracy of the code has also been studied by comparing the predicted concentrations of common species with the original code that does not explicitly track the sources/ages of air pollutants. The total concentrations from the source and age resolved CMAQ model matches very well with the unmodified code, providing evidence that the code is work correctly. The input data used in this chapter, including meteorology and emission inputs, were taken from a previous study and thus their qualities have been evaluated. The predicted model predictions of ozone are compared at monitoring sites (see Figure 3.4) and show agreement that is comparable to previous studies. The data extraction programs were mostly developed in previous studies. The figures were generated jointly by Dr. Ying and Jie Zhang, and all are cross-checked to avoid mislabeling.

4. Analysis of the interaction of mesoscale winds and ozone formation during key episodes

4.1. Case Selections

In order to investigate model simulations of wind rotation and resultant pollutant evolution, four specific cases were chosen. The suite of cases was selected to include a wide variety of wind rotation behaviors, all of which were associated with ozone exceedances, during years of high interest by TCEQ in which radar wind profiler measurements were available.

The first case study was August 25-31, 2000. This case, during the original TexAQS Field Program, includes several days that have been the subject of intense investigation, with substantial aircraft-based pollutant measurements. More importantly for the purposes of this study, the episode included some of the largest wind rotation events ever documented in the Houston metropolitan area. With strong southwesterly wind conditions, the ideal setup was in place for high-amplitude wind rotation, similar to what was discussed in Section 2 earlier. Such a high-amplitude event should facilitate the diagnosis of the role of wind rotation and recirculation in the context of the ozone tracer module.

The second case study was September 25-26, 2013. This event took place during the Deriving Information on Surface conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) field program. The first day featured winds from the north, while the second day featured winds from the south. At low levels, the wind executed a complete clockwise loop during the day on September 25, while on September 26 the wind followed a classic moderate southerly wind pattern with variations in wind direction but no complete rotation. While this case doesn't offer a pure wind rotation example, it is the most relevant example from the DISCOVER-AQ field program.

The third case study, May 6-7, 2016, was a high vertical wind shear case, with a low-level flow from the northeast and flow at 1800 m from the southwest. This case will be interesting because of that mix of different trajectories. Also, the overall wind speed was unusually light on May 6, so the sea breeze circulation should be easier to detect. Indeed, winds above 1200 m were from the northeast in the late afternoon on May 6 due to the sea breeze return flow, and that return flow persisted until nearly midnight on May 6. This case will test the ability of the model to properly represent pure thermally-forced boundary layer circulations.

The fourth case study was July 21-22, 2016. This event was a near-classic light southeast wind event, except that trajectories above 1000 m were mainly from the east where they might be expected to contain aged continental background air. The winds were lighter on July 21, and low-level winds executed a complete loop in the early morning. Winds at 400-800 m did not loop but instead experienced a few hours of stagnation around sunrise. The following day was similar except that winds were slightly stronger and no trajectories exhibited looping behavior. This case should be interesting for the mix of fresh vs. aged ozone and ozone precursors as well as the critical sea breeze behavior that yields recirculation below a large-scale wind threshold and simpler trajectories above that threshold.

4.2. WRF Model Configuration

As mentioned in Chapter 2, part of predicting and simulating local-scale wind patterns that contribute to pollutant buildup requires accurate modeling of the events in question. Sea breeze recirculation is already a wind pattern that is sensitive to the large-scale flow conditions, and

modeling it requires a high-resolution model configuration that itself is highly sensitive to a number of different factors beyond the large-scale flow. Part of accurate modeling is performing sensitivity studies that characterize not only the model's ability to produce the features in question but the model's ability to produce a realistic simulation of the atmosphere in the regions of interest in general and what factors native to the model environment can affect the accuracy of that simulation.

The model used for this project is the Advanced Research Weather Research and Forecasting Model (ARWRF, hereafter simply WRF). It is primarily a mesoscale model, with a wide variety of physics and dynamics parameterizations and data assimilation techniques built-in giving it flexibility in simulation. The most recently updated version of WRF, 4.1, contains nearly a dozen planetary boundary layer (PBL) parameterization schemes, allowing for several test environments to be created to determine how accurately WRF simulates sea breeze circulation, recirculation, and stagnation that is common to Houston and coastal Texas.

This subsection and the ones that follow the overall WRF meteorological model environment, changes to parameterization and domain initialization that comprises the sensitivity test suite, and the results from each simulation, from how each reproduces observed wind profiles on days of interest to the overarching local and nonlocal dynamics that produce each model outcome. This section discusses the baseline model configuration: the domains, the cases, and the input data, the baseline parameterizations, what PBL schemes are used, and how they are implemented in the model. Section 4.3 compares observed wind profiles near Houston to model simulated wind profiles for the various test environments. Section 4.4 discusses the local wind field for the innermost model domain for the whole boundary layer at height levels of interest. The PBL test completed to date focus on the September 2013 and May 2016 episodes.

4.2.1. Domain Layout

The WRF model domain for the sensitivity portion of the study is configured according to the TCEQ domain standards for air quality modeling (Figure 4.1). It is a Lambert-Conformal grid with true-latitudes of 33°N and 45°N that is centered at 97°W, 40°N. The outer domain is a 36km mesh grid of 163 by 129 grid points, covering CONUS, much of southern Canada, northern Mexico, and the Caribbean Islands. The first child domain is a 12km mesh grid of 175 by 139 grid points covering the Southern Plains, including New Mexico and the Lower Mississippi River Basin. The innermost child domain is a 4km mesh grid of 217 by 289 grid points that covers central and east Texas and includes southeastern Oklahoma and western Arkansas and Louisiana. The outermost domain time step is 60 seconds, scaling down by a factor of 3 to the innermost domain. Each model domain has the same vertical structure, consisting of 43 sigma levels. Of those 43, 20 are contained within the bottom 3km of the atmosphere, 16 fully within 2km, and 11 fully within the bottom 1km (Table 4.1).

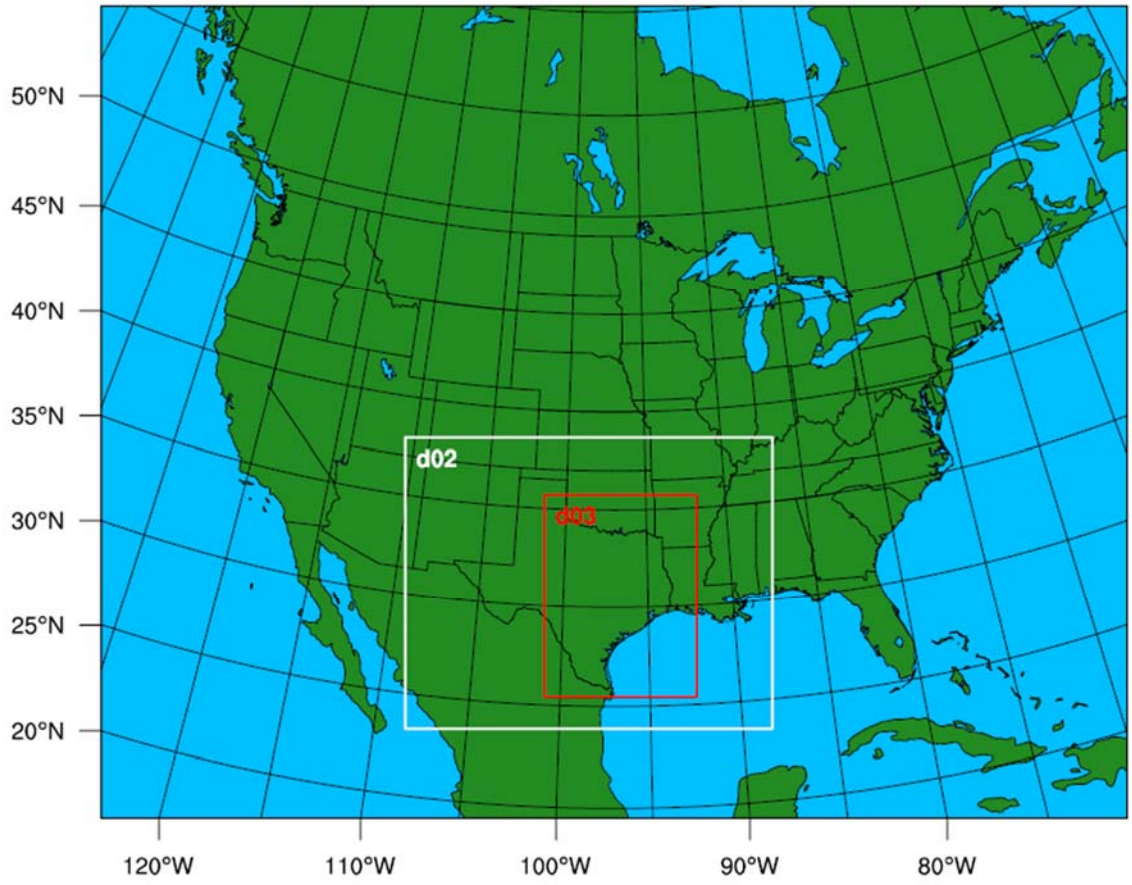


Figure 4.1 WRF domain configuration for domains 1 through 3. Domain d01 encompasses the entire plotted region: all of CONUS, much of southern Canada, northern Mexico, and the Caribbean Islands. Domain d02 covers the Southern Plains, New Mexico, and the lower Mississippi Basin. Domain d03 covers most of central and east Texas, the Lower Rio Grande Valley, and parts of Oklahoma, Arkansas, and Louisiana.

4.2.2. Input Data and Initialization

Input data for the simulations were the Global Forecast System Final Analysis (GFS FLN) reanalysis data set (NCEP et al. 2015). The North American Regional Reanalysis (NARR) data set was also considered, but ultimately the GFS FNL data was chosen because the GFS model compares favorably to the Eta model, which the NARR data set is derived from. The temporal and spatial resolutions of the GFS data is every 6 hours at 1° by 1° for the September 2013 case, and every 3 hours at 0.25° by 0.25° for both 2016 cases; the GFS data set does not have the high-resolution coverage for cases prior to 8 July 2015, so the lower resolution data set was used instead (NCEP et al. 2000). Analysis nudging is used for all three cases, occurring every boundary condition update time. Because of the contribution of the Galveston Bay to the sea breeze recirculation near Houston, the geographical parameterization makes use of the surface lake model, which accounts for surface water depth, as this is considered essential to the model set-up.

There are 3 cases for the model sensitivity portion of the study, selected because they each were observed to have high ozone concentrations above the primary United States Environmental Protection Agency (US EPA) standard and were observed to experience a complete rotation in their circulation (hereafter recirculation). Each case experienced above standard ozone

concentrations for two days: September 25-26, 2013; May 6-7, 2016; and July 21-22, 2016. Accounting for model spin-up time and spin-down for the emissions processing, the total model simulation was 15 days for each case, with 10 days of spin-up and 3 days of spin-down. For the emissions testing, each model domain was initialized at the same time and run for the entirety of the simulation time.

Table 4.1 WRF vertical layer configuration

WRF Level	Sigma Level	Top (m AGL)	Center (m AGL)		WRF Level	Sigma Level	Top (m AGL)	Center (m AGL)
0	1.000	0	0		22	0.63	3733	3553
1	0.996	34	17		23	0.6	4106	3919
2	0.99	85	59		24	0.57	4494	4300
3	0.98	170	128		25	0.54	4898	4696
4	0.97	257	214		26	0.51	5320	5109
5	0.96	344	300		27	0.475	5836	5578
6	0.95	432	388		28	0.44	6382	6109
7	0.94	520	476		29	0.405	6961	6671
8	0.93	610	565		30	0.37	7578	7270
9	0.92	700	655		31	0.33	8338	7958
10	0.91	790	745		32	0.29	9167	8752
11	0.895	928	859		33	0.25	10079	9623
12	0.88	1068	998		34	0.21	11097	10588
13	0.865	1210	1139		35	0.175	12097	11597
14	0.85	1353	1281		36	0.145	13058	12578
15	0.825	1597	1475		37	0.115	14146	13602
16	0.8	1847	1722		38	0.09	15179	14663
17	0.775	2103	1975		39	0.065	16367	15773
18	0.75	2366	2234		40	0.045	17470	16919
19	0.72	2690	2528		41	0.025	18766	18118
20	0.69	3026	2858		42	0.01	19917	19341
21	0.66	3373	3199		43	0.0	20807	20362

As a sensitivity test, the inner two domains were initialized at different times to determine how strongly the spin-up time affects the inner domain wind profiles. Three initialization experiment runs were performed: one where d01 and d02 were initialized at the start time and d03 was initialized on the fifth simulated day; one where only d01 was initialized at the start time, d02 was initialized on the fifth simulation day and d03 was initialized on the sixth simulation day; and one where d01 was initialized at the start time, d02 was initialized on the eighth simulation day, and d03 was initialized on the ninth simulation day. The WRF-derived wind profiles in the boundary layer at La Porte (LPT) and the University of Houston Coastal Research Center (HSN) on the first high ozone day were compared between the three simulations (not shown), and it was observed that the initialization time did not strongly impact the wind profiles at these locations: at this point in each simulation, the instantaneous winds had converged to similar profiles as that

of the simulation where all domains were initialized at the same time with only minor differences in wind speed and direction.

4.2.3. Model parameterizations

Because the recirculation of the sea breeze is strongly impacted by the background flow through turbulent mixing, the planetary boundary layer scheme parameterization is a key component to how well WRF resolves the recirculation. The goal is not only for WRF to accurately simulate the observed wind features on the case days, but also to produce realistic boundary layer conditions beyond simply matching with observations.

To test the sensitivity, all other parameterizations are kept consistent for each simulation except when absolutely impossible. For instance, the Yonsei University Scheme (YSU), the Mellor-Yamada-Janjic Scheme (MYJ), and the Quasi-normal Scale Elimination Scheme (QNSE) each make use of a specific surface clay physics model, so adjusting the surface clay physics model between simulations is unavoidable. For many other boundary layer schemes, multiple surface clay physics schemes are compatible and thus the surface clay physics is not changed. Situations, where multiple parameterizations must be changed concurrently, are indicated during the discussion of that run's results, and situations where this is not the case, will not be indicated, so it can be assumed that *only* the boundary layer scheme changes between simulations unless otherwise specified. The complete list of WRF parameters utilized in this portion of the study are listed in Table 4.2

Table 4.2 WRF model parameters

Option	**Parameterization**
Microphysics	New Thompson et al. for high-resolution simulations
Longwave Radiation	Rapid Radiative Transfer Model
Shortwave Radiation	Goddard Shortwave Scheme
Land Surface Model	Noah LSM
Surface Layer	Modified Monin-Obukhov (unless otherwise specified)
Lake Physics	Community Land Model 4.5
Cumulus Physics	Grell-Devenyi Ensemble Scheme
Urban Surface	Urban Canopy Model

4.3. Back Trajectories

Figure 4.2 depicts the back trajectories for September 25-26, 2013 for LPT. On 25 September, particularly in the lower boundary layer, trajectory lengths tend to shorten during the afternoon to late evening during which a wind rotation of almost 360° below 350m AGL and around 180° below 100m AGL. While no complete loop forms in the trajectories, the short hourly length, and large rotation indicate a high degree of stagnation occurring near the end of the day. Compare this to 26 September, where the low-level winds tend to be more strongly southerly; the large rotation occurs above 1000m AGL, with the 2000m AGL trajectory forming a complete loop. The profiler trajectories at HSN at the same time (Figure 4.3) show similar patterns—noticeable rotation in wind trajectory over the course of the day with progressively shortening segments—with weaker (stronger) rotation on the first (second) day at the same levels as LPT.

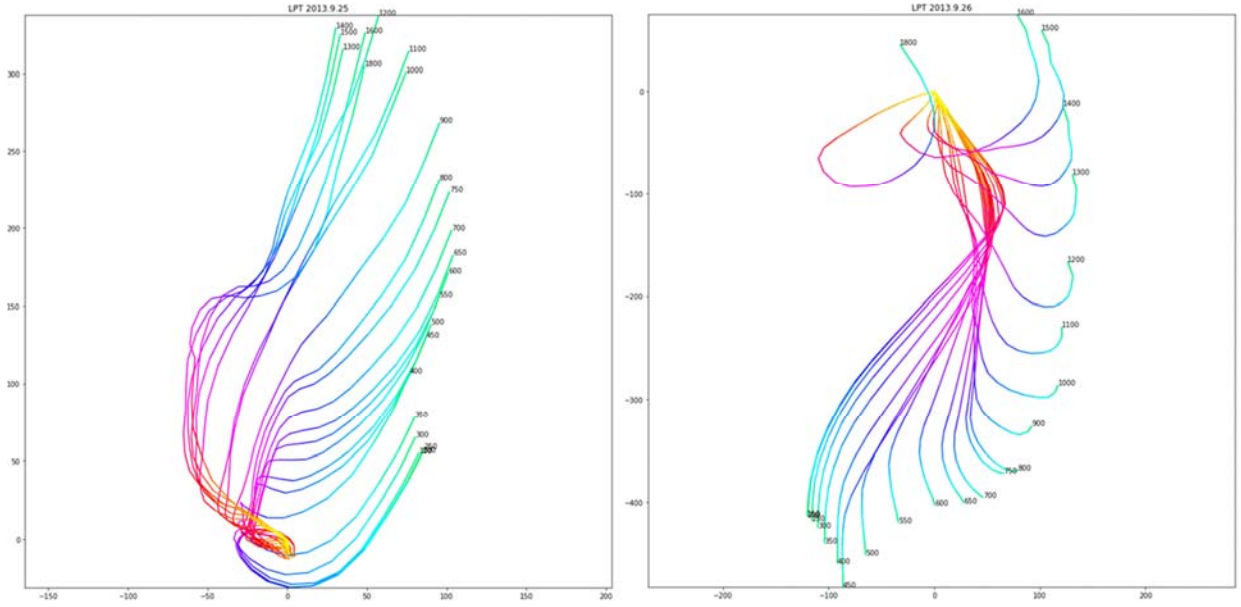


Figure 4.2 Back trajectories calculated from the La Porte (LPT) profiler observations on 25-26 September 2013. Observations are of a low-level recirculation near the LPT profiler location near the end of the day on 25 September, with a stagnation of winds in the middle boundary layer at the same time. Recirculation or near-recirculation is observed in the upper boundary layer on 26 September, with a gradual phase shift in winds above 1000m AGL from north to south.

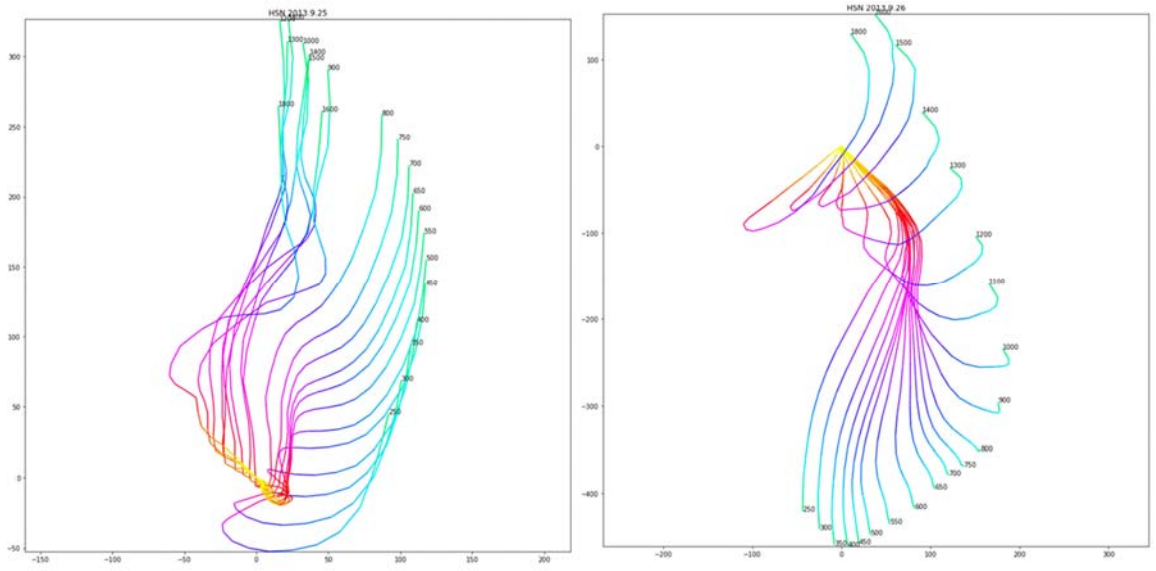


Figure 4.3 Back trajectories calculated from the University of Houston Coastal Research Center (HSN) profiler observations on 25-26 September 2013. Winds at HSN are similar to those at LPT, with low-level recirculation after a complete wind rotation near the end of the day, with a uniform middle and upper layer entering the profiler location from the north and northwest. Winds on 26 September again show recirculation in the upper boundary layer and

winds entering the profiler location from the south and southeast, and there is less uniformity in the middle and lower boundary layer than at LPT.

Figures 4.4, 4.5, and 4.6 depict the WRF simulated back trajectories at LPT for the simulations using the YSU, MYJ, and QNSE PBL schemes, respectively; note that each of these schemes requires a concurrent adjustment to the surface layer physics scheme. The WRF simulated trajectories at both LPT and HSN on September 25 do not show the same stagnation pattern as observed; the general pattern of winds is similar, as the trajectories show parcel flow in to the profiler locations from the north with a westward deflection in the middle of the boundary layer, but the parcels all flow into the profiler location from the north uniformly. It is clear that the trajectory segment lengths for the WRF simulations are much longer: observations show the incoming parcels to be sourced no farther than ~300m away from the profiler location, but the simulation has no parcel sourced *closer* than 300m. There is some reduction in segment length for the lower half of the boundary layer, but there is almost no trajectory rotation, indicating little to no rotation in the wind direction, occurring at the same time. On 26 September, the simulated back trajectories are more varied but not much more accurate. The YSU scheme does show some degree of rotation in the upper boundary layer, but the 180° rotation is the most prominent feature in the mid to lower boundary layer, which matches poorly to the observations. The MYJ scheme produces winds with a greater length of southward flow in the mid to upper PBL, though with a stronger-than-observed shift to westward flow in the late afternoon; incidentally, there is one trajectory that shows recirculation, but it is at 3km, likely outside of the boundary layer. The QNSE scheme produced trajectories that show almost no northward flow at all. Interestingly this scheme produces flow approaching what can almost be called recirculation and stagnation when the trajectory segment lengths are short in the lower boundary layer in the afternoon, but this still does not match observations well.

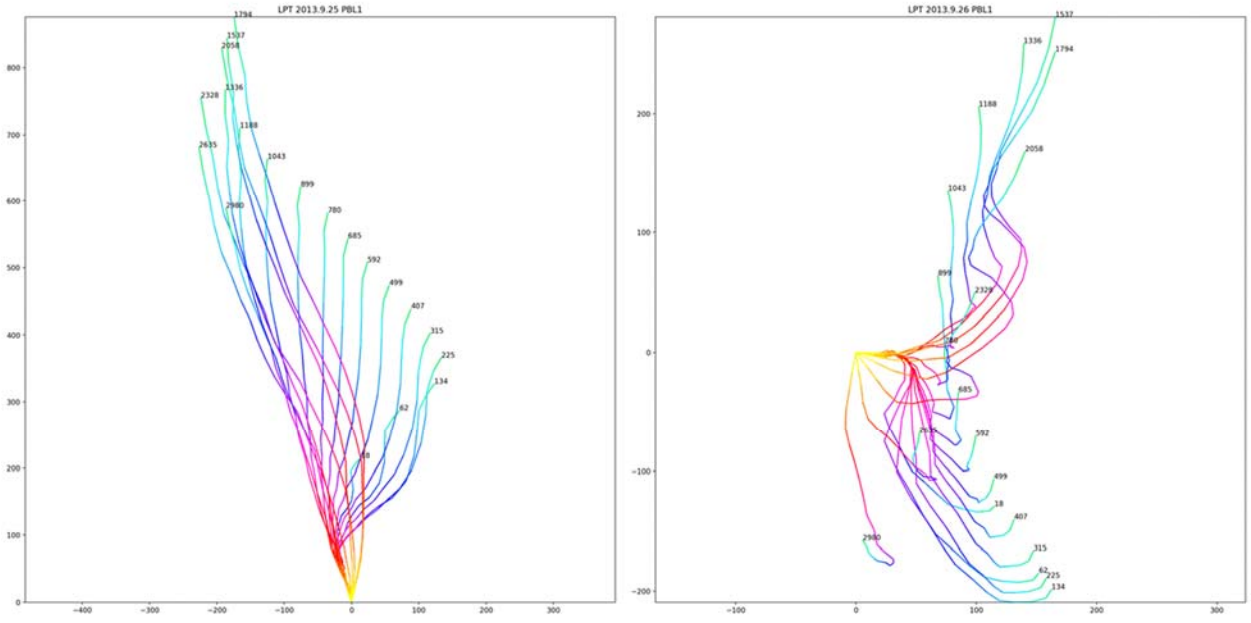


Figure 4.4 Back trajectories calculated from WRF model output at LPT from the simulation using the YSU PBL scheme. The YSU WRF simulation produces no recirculation on 25 September at low levels. Mid-level winds do show some reduction in speed, but rotation in wind direction poorly matches observation. No recirculation is

exhibited on 26 September either, and the rotation in winds seen in the upper boundary layer is instead seen in the lower boundary layer, with stronger wind speed reduction near the end of the day than observed.

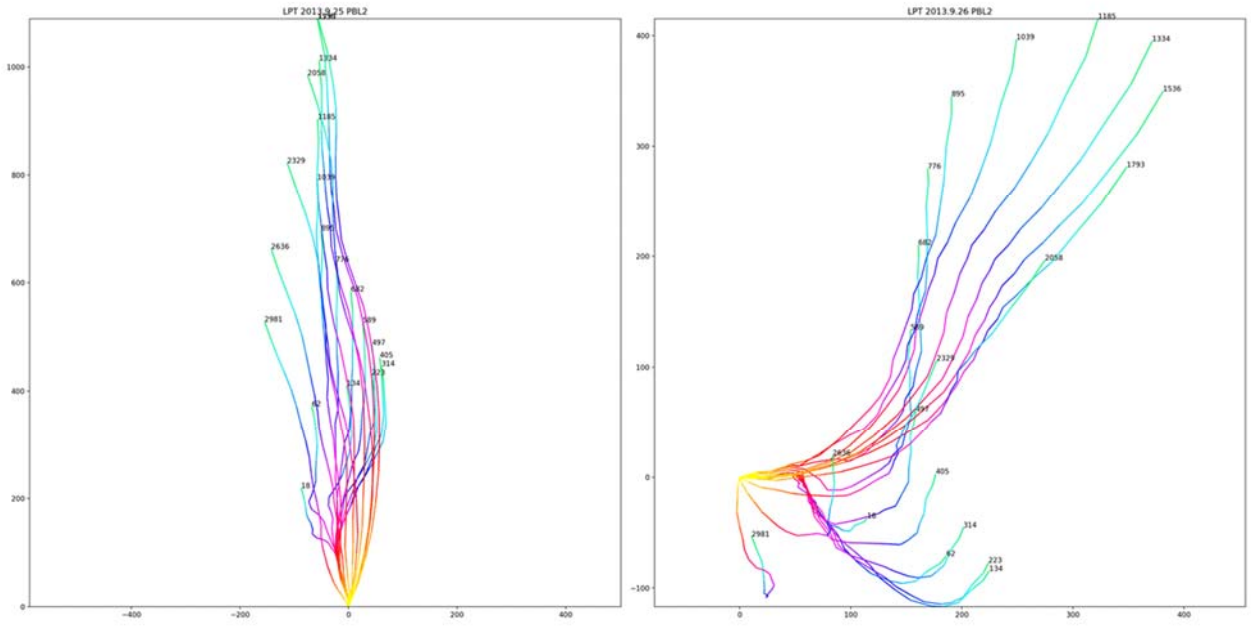


Figure 4.5 Back trajectories calculated from WRF model output at LPT from the simulation using the Mellor-Yamada-Janjic PBL scheme. The MYJ scheme produces no recirculation, the entire column is more homogenous in wind direction than observed, and wind speeds, as indicated by trajectory segment length, as much higher than observed. Recirculation is seen on 26 September, but outside of the boundary layer, and the wind rotation is seen in the lower boundary layer like the YSU simulation.

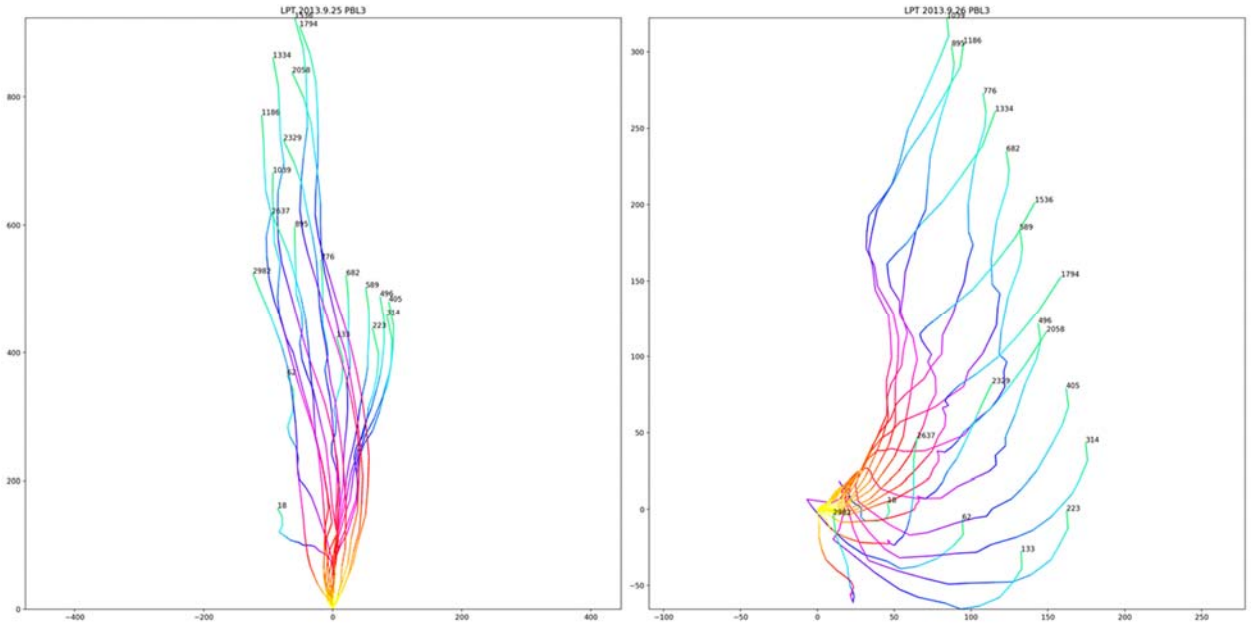


Figure 4.6 Back trajectories calculated from WRF model output at LPT from the simulation using the QNSE PBL scheme. The QNSE simulation produces wind patterns like that of the MYJ scheme on 25 September: stronger than observed and consistent throughout the column, but with a less defined phase shift over the course of the day. Wind direction is far from observation on 26 September, and winds are in general too strong.

Simulated trajectories at HSN for 25-26 September 2016 are shown in Figures 4.7, 4.8, and 4.9 for the YSU, MYJ, and QNSE schemes, respectively. Like for LPT, the simulated HSN profiles do a poor job reproducing the observed profiles. One 25 September, wind direction is near uniformly southward with very little rotation at any point and at all times stronger than in observation. For 26 September, the YSU scheme produces a weaker rotation in the lower boundary layer, which is closer to the observed winds than the same simulation produced at LPT, but the winds weaken more noticeably at the end of the day at HSN than observed, and the rotation is more extreme and uniform across the bottom half of the boundary layer. The MYJ scheme at HSN is near identical to LPT: erroneous recirculation at a level above the boundary layer, strong 180° rotation of wind direction in the lower boundary layer, and weak rotation with strong winds in the mid- to upper boundary layer. The QNSE scheme fares better at HSN than LPT in that it is able to produce northerly flow in the lower boundary layer, but those same levels have more stagnation than observed and still exhibit a near 180° rotation in their direction over the course of the day. The upper boundary level winds still exhibit little of the rotation seen in the observations.

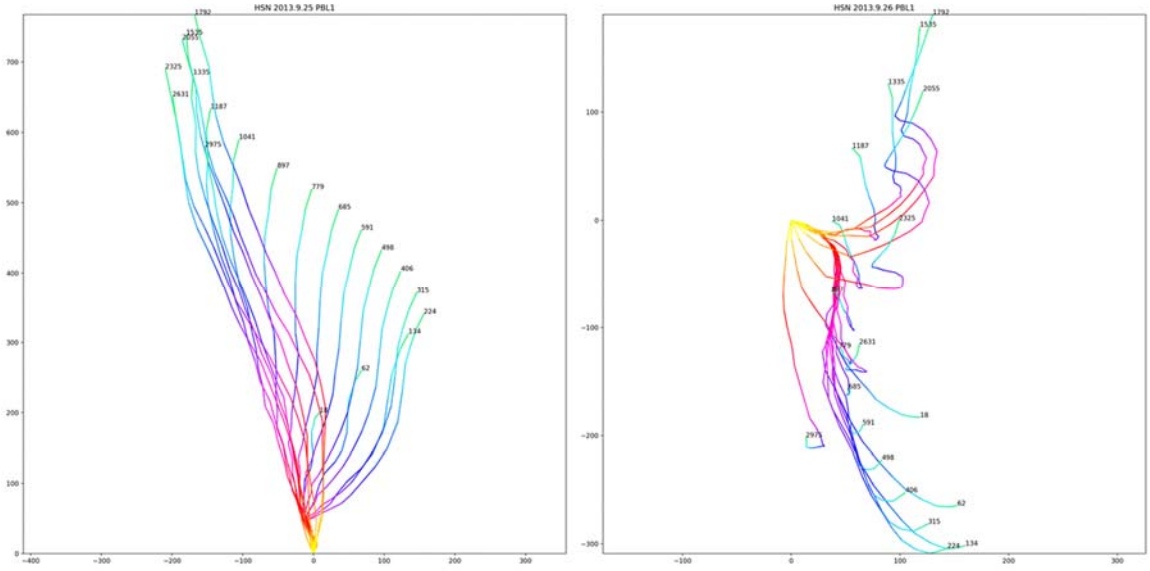


Figure 4.7 Back trajectories calculated from WRF model output at HSN for the simulation using the YSU PBL scheme. The YSU scheme produces a wind profile at HSN nearly identical to at LPT on 25 September: no recirculation, weak wind rotation, and uniformity throughout the column. Winds on 26 September are more uniform in the lower and middle boundary layer than observed, though wind speeds are closer at HSN than they were at LPT. The upper boundary layer shows a stronger rotation over the day than the simulated LPT profile, which the timing of the rotation poorly matches observation.

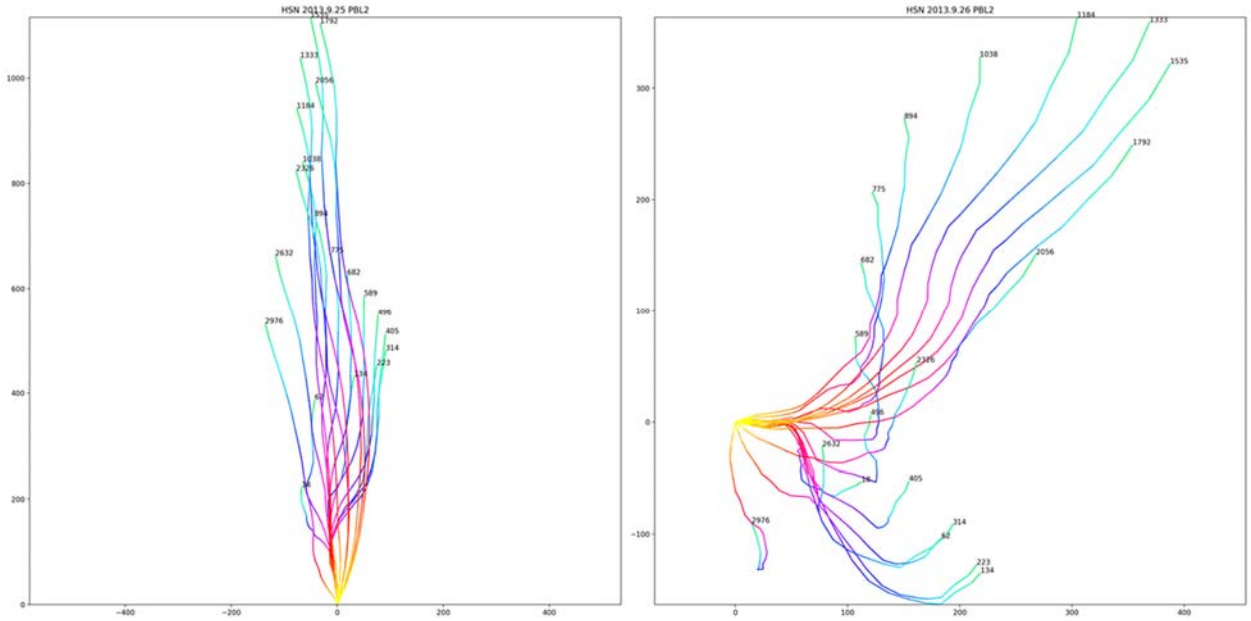


Figure 4.8 Back trajectories calculated from WRF model output at HSN for the simulation using the Mellor-Yamada-Janjic PBL scheme. The profiles at HSN in the MYJ simulation are nearly identical to those at LPT and exhibit the same mismatch between modeled wind speed and direction and observation. Winds on both days are too strong, feature little to no rotation on 25 September and rotation in the wrong layers on 26 September with too much uniformity.

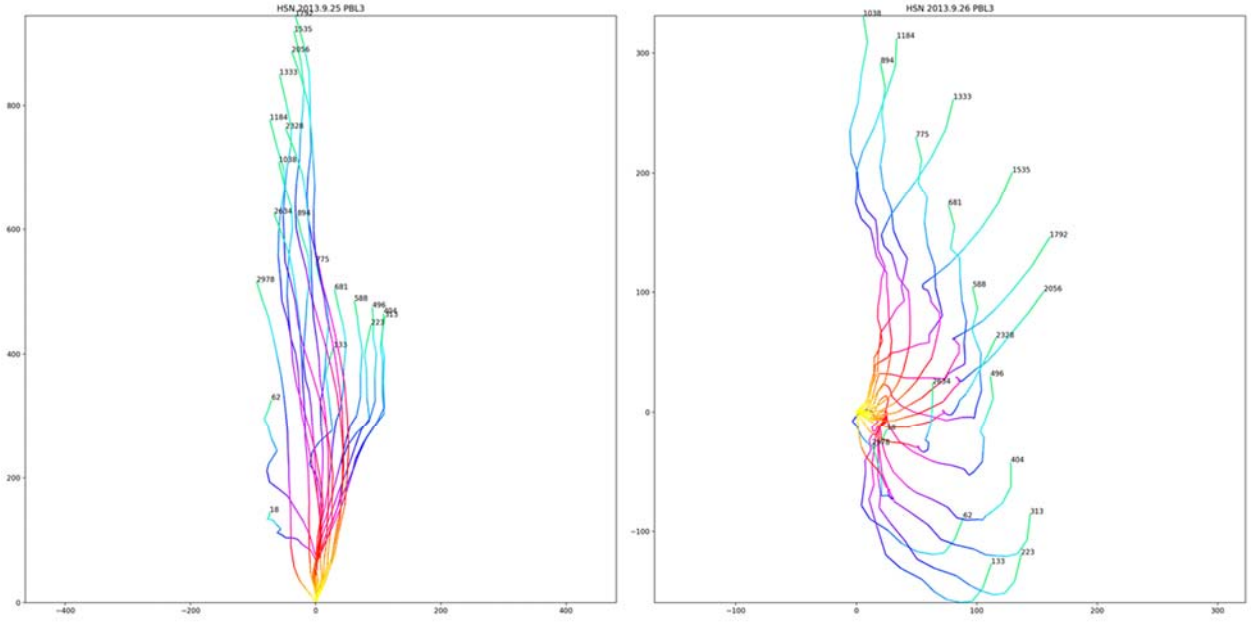


Figure 4.9 Back trajectories calculated from WRF model output at HSN for the simulation using the QNSE PBL scheme. The HSN profile from the QNSE simulation is more accurate to observation than the LPT profile but is still different in many of the same ways that the MYJ scheme is. The QNSE on 26 September exhibits less uniformity than the YSU and MYJ schemes, which is good as the observations are not uniform through the PBL, but the wind directions are southerly at the end of the day in the simulated profile while the observations are northerly.

For the two 2016 cases, the LPT profiler was not operational, so the analyses between observations and the WRF simulations are for the HSN profiler only. Figure 4.10 depicts the profiler-derived back-trajectories for 21-22 July 2016. There is prominent recirculation at the lower levels of the boundary layer, which gradually give-way to weakly rotating and decelerating wind patterns in the middle and upper boundary layer on 21 July. 22 July experienced weakly rotating winds in the upper boundary layer in the morning and strongly rotating winds in the lower boundary layer in the evening, but no recirculation and little to no wind speed reduction.

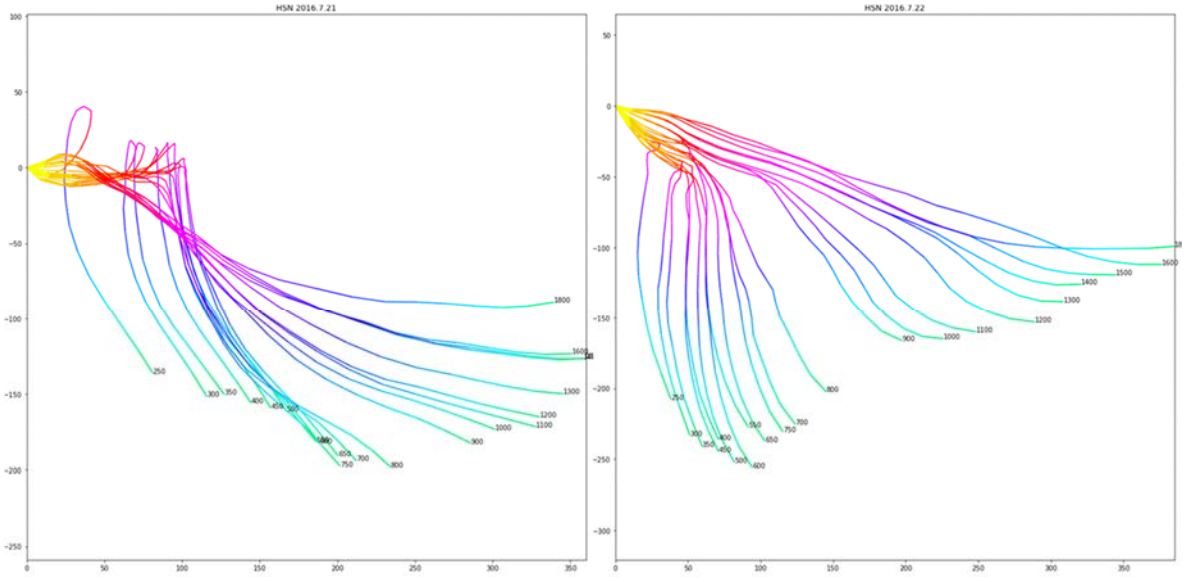


Figure 4.10 Back trajectories calculated from the HSN profiler for 21-22 July 2016. Wind directions are predominately westward, with observed recirculation in the lower boundary layer at midday entering the profiler location from the east at night on 21 July. On 22 July, the wind direction is largely the same, with wind speed reduction in the lower boundary layer and a quick wind rotation from north/northeast to northwest/west.

Figures 4.11, 4.12, and 4.13 show the WRF simulation trajectories for the YSU, MYJ, and QNSE schemes, respectively, which all fail to reproduce the recirculation at the lower portions of the boundary layer in the observations on 21 July. The simulations do reproduce a rotation in the lower and middle boundary layer of 90°, but unlike the observations, the simulations depict winds from the west shifting to the south rather than from the south to the east. The same pattern is observed in the models for 22 July to varying degrees, with the exception of the QNSE scheme, which shows no strong rotation at any level but rather a gradual 90° rotation in the lower boundary layer. The latter is no more accurate in spite of that, as the bottom half of the boundary layer shows wind trajectories coming to the profiler location from the wrong direction. In all of the simulated trajectories, the total length of the trajectories is again much longer than observed.

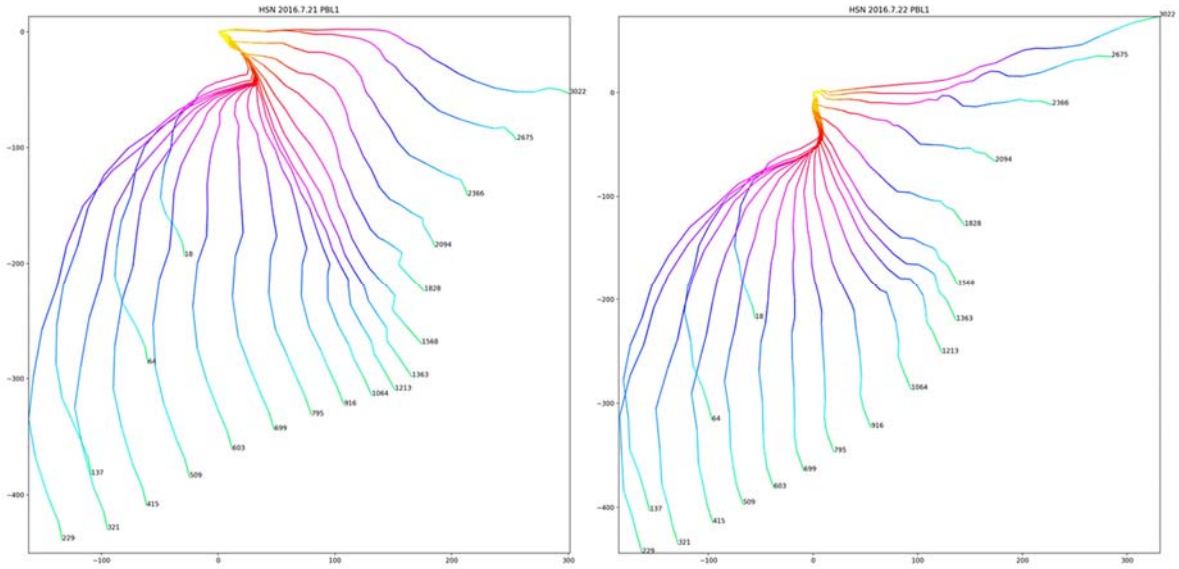


Figure 4.11 Back trajectories calculated from WRF model output at HSN for the simulation using the YSU PBL scheme. The YSU scheme exhibits wind direction uniformly shifting from the west to the south-southeast at the end of the day in the lower boundary layer, with a more gradual shift in wind directions in the morning and early afternoon from northward to westward through the depth of the column on 21 July with no recirculation. On 22 July, back trajectories are largely the same: predominantly northward with a strong, uniform shift from east-northeast to northward near the end of the day, with lower boundary trajectories poorly matching observation.

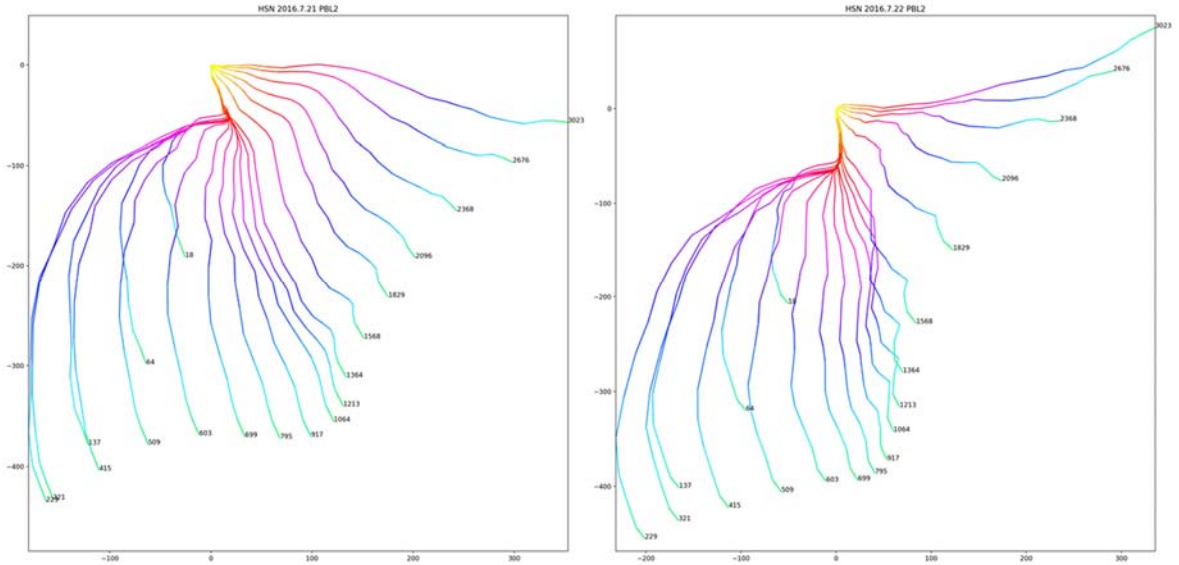


Figure 4.12 Back trajectories calculated from WRF model output at HSN from the simulation using the Mellor-Yamada-Janjic PBL scheme. The MYJ simulation exhibits the same wind patterns as the YSU simulation on 21 and 22 July: northward winds in the lower boundary layer that gradually rotate eastward before strongly rotating northward at the end of the day with no simulated recirculation.

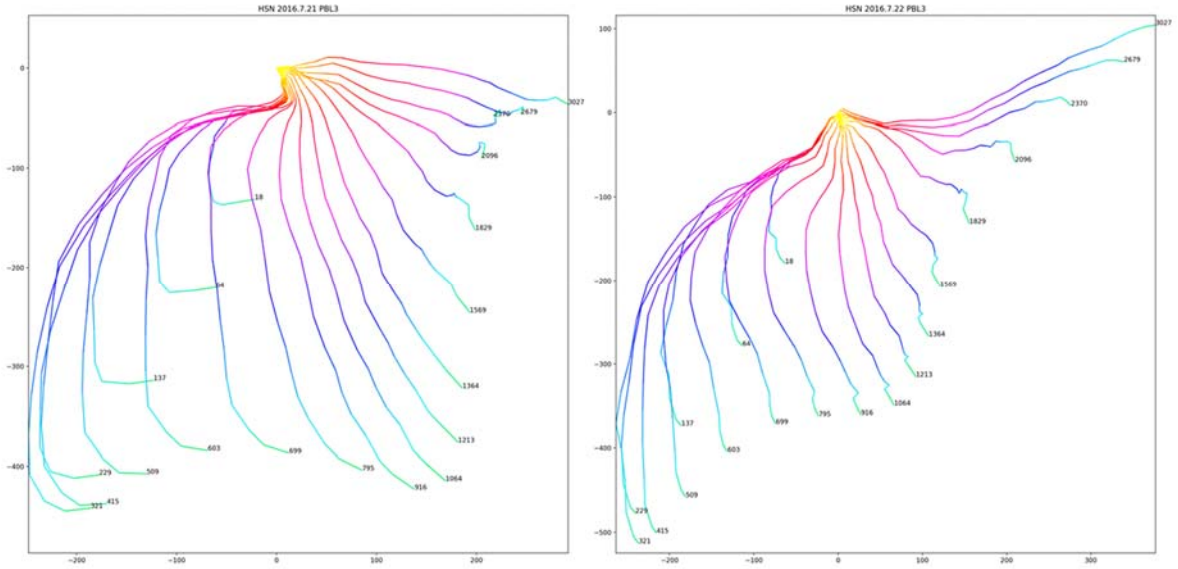


Figure 4.13 Back trajectories calculate from WRF model output at HSN from the simulation using the QNSE PBL scheme.

On 21 July, the QNSE simulated profile does not strongly differ from the YSU and MYJ profile, though there is a 90° shift in wind direction in the early morning hours in the lower boundary layer. On 22 July, there is no simulated rotation from eastward to northward at the end of the day, but rather the trajectories in the lower boundary layer experience a gradual eastward rotation and enter the profiler location from the west, while the middle and upper boundary layer winds approach the profiler from the south and east, respectively.

Figure 4.14 shows the profiler-derived back-trajectories for HSN on 6-7 May 2016. On 6 May, the lower boundary layer winds form recirculation patterns to the east of the profiler location and the total boundary layer wind column experiences a 270° shift in wind direction over the course of the day. Trajectory segments do not change significantly during the course of the day, so much of the stagnation comes from the large rotation in the wind field during the day. On 7 May, the strongest rotation is seen in the upper boundary layer, which exhibits and almost 180° doubling-back at several layers above 1km; the lower layer shows little to no rotation until the evening, and even then it is weak. Figure 4.15 shows the WRF simulation using the YSU scheme, and again the WRF simulation fails to reproduce the rotation of the wind over the course of 6 May, showing about a 180° rotation at the lowest levels and almost no rotation at the top of the boundary layer. For 7 May, the model more accurately produces the flow direction for the lower boundary layer but does not produce the rotation from 1-2km at all.

Generally speaking, the WRF models as configured struggle to show any recirculation in the boundary layer near Houston at all. It was often seen that rotation was weaker than observed or more uniform across the depth of the boundary layer than observed, and more often the winds were higher speed than observed, with trajectory segments and total lengths much longer than those of the observations.

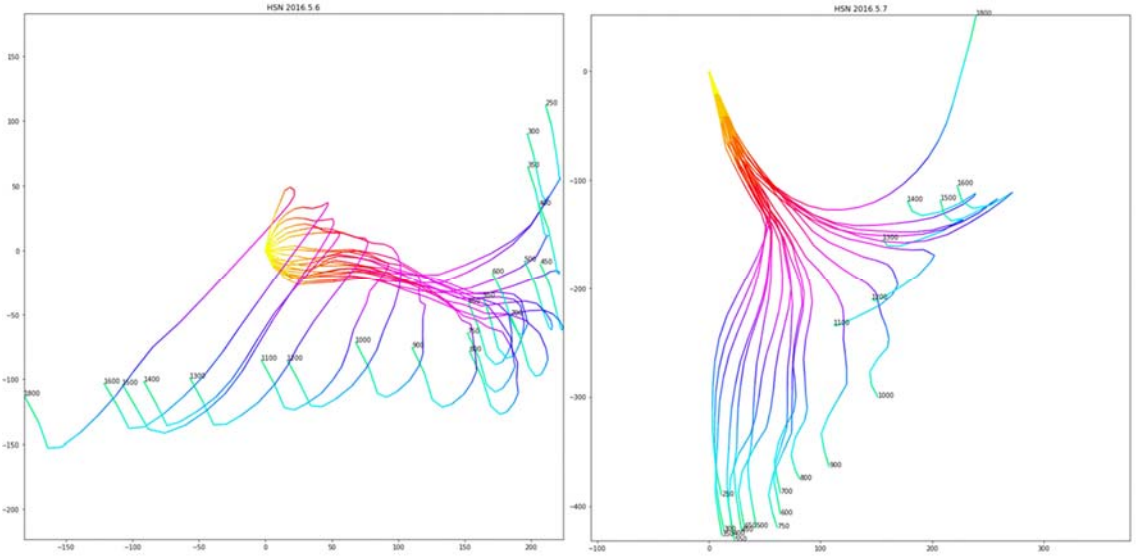


Figure 4.14 Back trajectories calculated from the HSN profiler observations for 6-7 May 2016. Observed back trajectories show a uniform rotation in wind direction throughout the entire column from the north to the south in the early morning hours of 6 May, which results in recirculation in the lower boundary layer. Winds uniformly approach the profiler location from the east at the end of the day, with the upper boundary layer wind experiencing a 90° rotation at that time from the southwest to the east. On 7 May, there is an observed recirculation in the upper boundary layer in the early morning hours at winds shift from eastward 180° to westward, with the middle and lower boundary layer trajectories showing little shift in wind direction from the south into the profiler location.

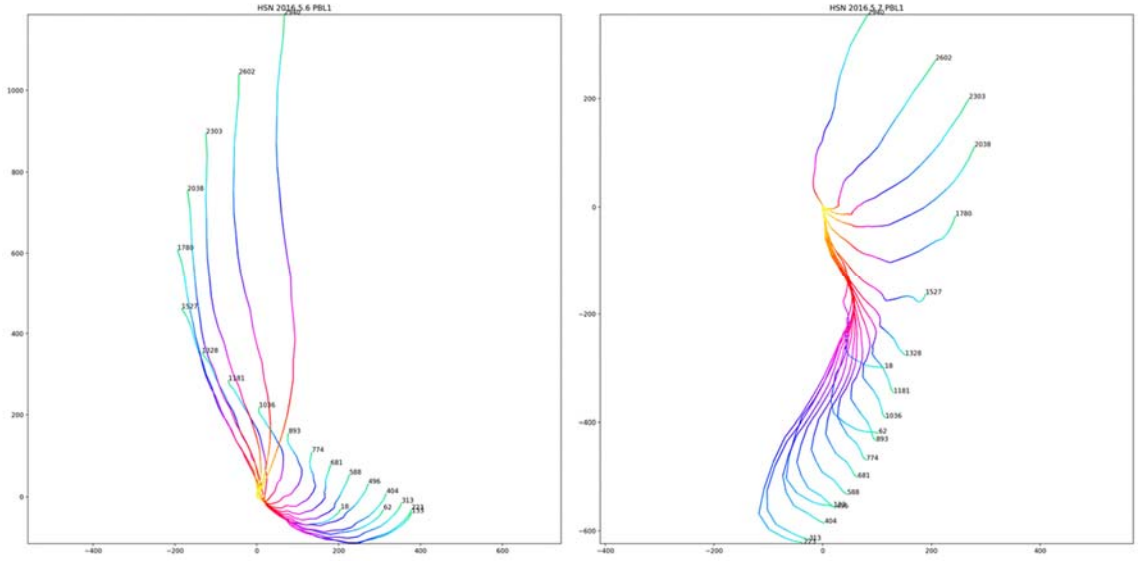


Figure 4.15 Back trajectories calculated from WRF model output at HSN from the simulation using the YSU PBL scheme. The YSU simulation poorly matches observation, with no early morning wind rotation and no resulting recirculation on 6 May; instead, winds gradually rotate from the north to the east in the lower boundary layer, and the upper boundary layer does not experience a wind shift and instead flow from the north throughout the day. On 7 May, the entire boundary layer column exhibits a gradual rotation from the southeast to the northeast throughout the day but does match observations uniformly approaching the profiler location from the south.

4.4. Time-mean perturbations

One major factor in the development of a diurnal sea breeze circulation is the background mean flow: as described in Chapter 2, one of the necessary conditions for diurnal rotation is weak boundary layer mixing. Strong instantaneous surface layer winds relative to the temporal averaged mean are indicative of a more well-mixed boundary layer, and are a prohibiting factor in sea breeze rotation. For this section, time perturbation wind vectors are plotted at a variety of boundary layer heights at 02Z, 08Z, 14Z, and 20Z (21CDT, 03CDT, 09CDT, and 15CDT, respectively), corresponding to maxima in insolation outgoing radiative cooling and transitions between those phases. Averaging is done via centered differencing a 25-hour time slice of the WRF model output, as shown in equation (4-1):

$$\bar{U} = \frac{1}{T}\sum_{n-12}^{n+12} U_n \tag{4-1}$$

where n is the time for which we are determining the mean, U is the total wind, and T is the total duration of the period. Because the sea breeze rotation takes place over the course of the diurnal period, successful resolving of the sea breeze rotation would result in a nonzero perturbation wind whose phase tilts eastward with increasing height at a given time. Locally large perturbation winds are not the sole indicator of recirculation, however, as strong perturbation winds also produce strong turbulent mixing, which is more indicative of low-level jet formation in the Southern Plains than sea breeze recirculation. Ideally, the perturbation wind at a given point would be nonzero but not strongly nonzero.

Figure 4.16 shows the perturbation wind vectors for the WRF model using the MYJ scheme for 25 September 2013. In the lower levels of the boundary layer near Houston, the instantaneous perturbation wind is weakly different from the mean flow—about 2 ms^{-1} between 250m AGL and 1000m AGL. At these levels, there is not much sign of a sea breeze wind at all, as the perturbation wind at the time of peak insolation at along the coast, and differs only slightly from the wind direction at grid points away from the coast. In the upper boundary layer, there does seem to be a return flow, with the perturbation wind phase differing by 180° between the lower and upper boundary layers, but like the lower boundary layer, it is not meaningfully different from the grid points away from the coast. Compare this with the simulation using the YSU scheme (Figure 4.17), which shows a clearer rotation in the lower boundary layer over the course of the day with more pronounced perturbation winds. However, the perturbation wind direction is not in the same phase as the expected sea breeze circulation for the majority of the Gulf Coast, and the boundary layer columns are roughly in phase with each other, indicating the whole column is experiencing phase shift together for all other times. The QNSE scheme (Figure 4.18) produces weak perturbation wind in the lower boundary layer with weak return flow in the upper boundary layer like the YSU scheme, but unlike the YSU scheme, the middle boundary layer exhibits a strong offshore perturbation wind, possibly indicative of erroneously high turbulent mixing near the coast.

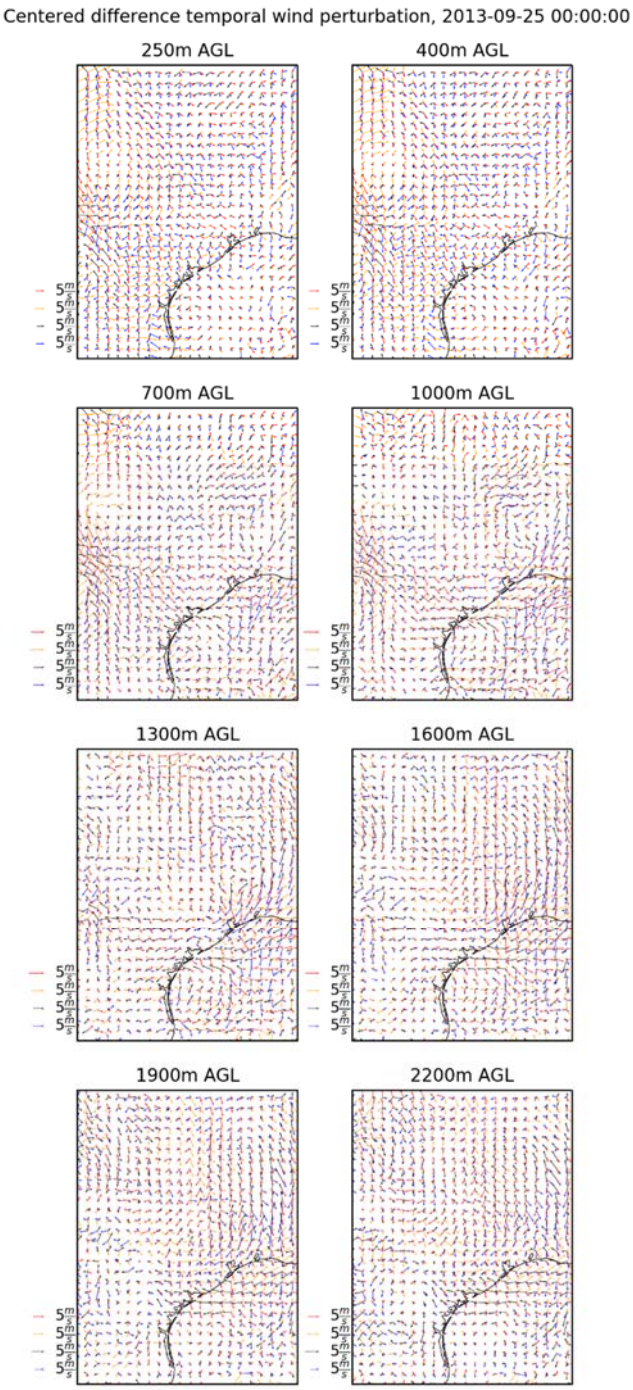


Figure 4.16 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

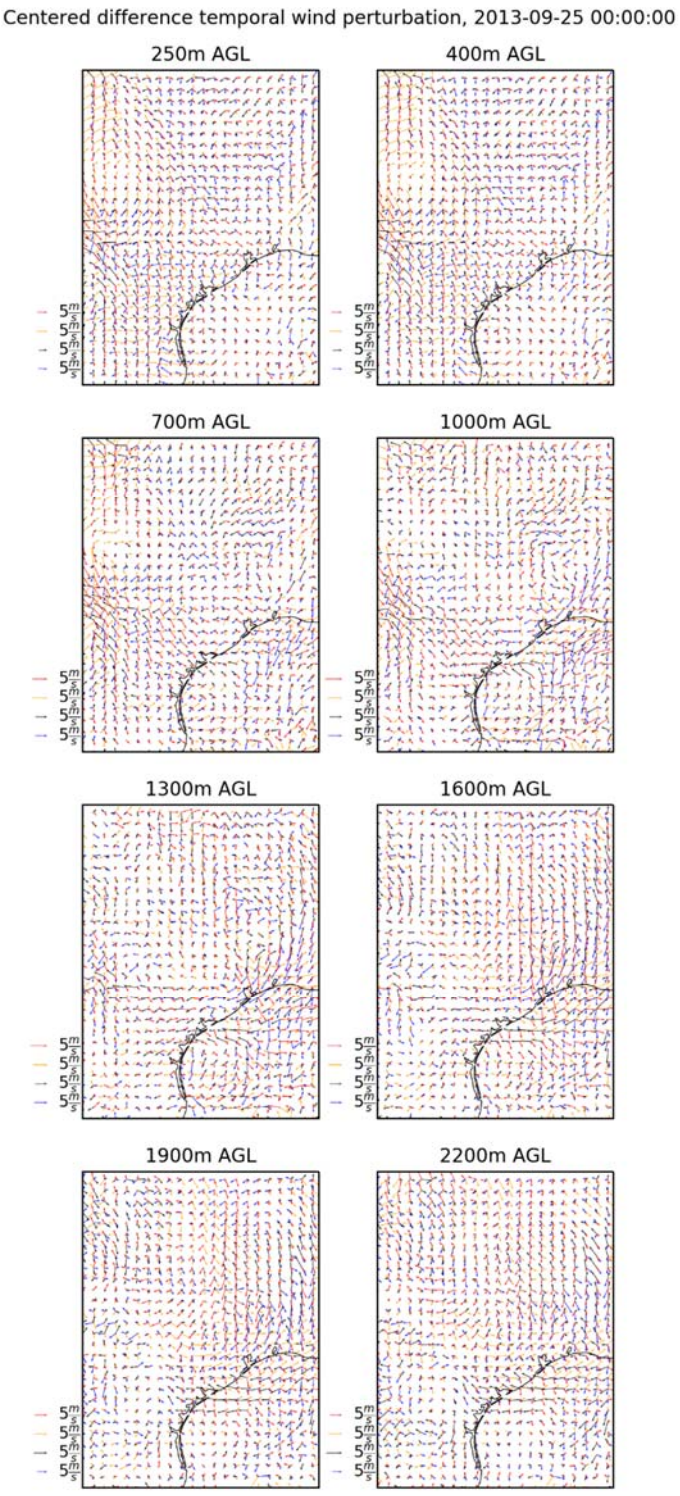


Figure 4.17 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

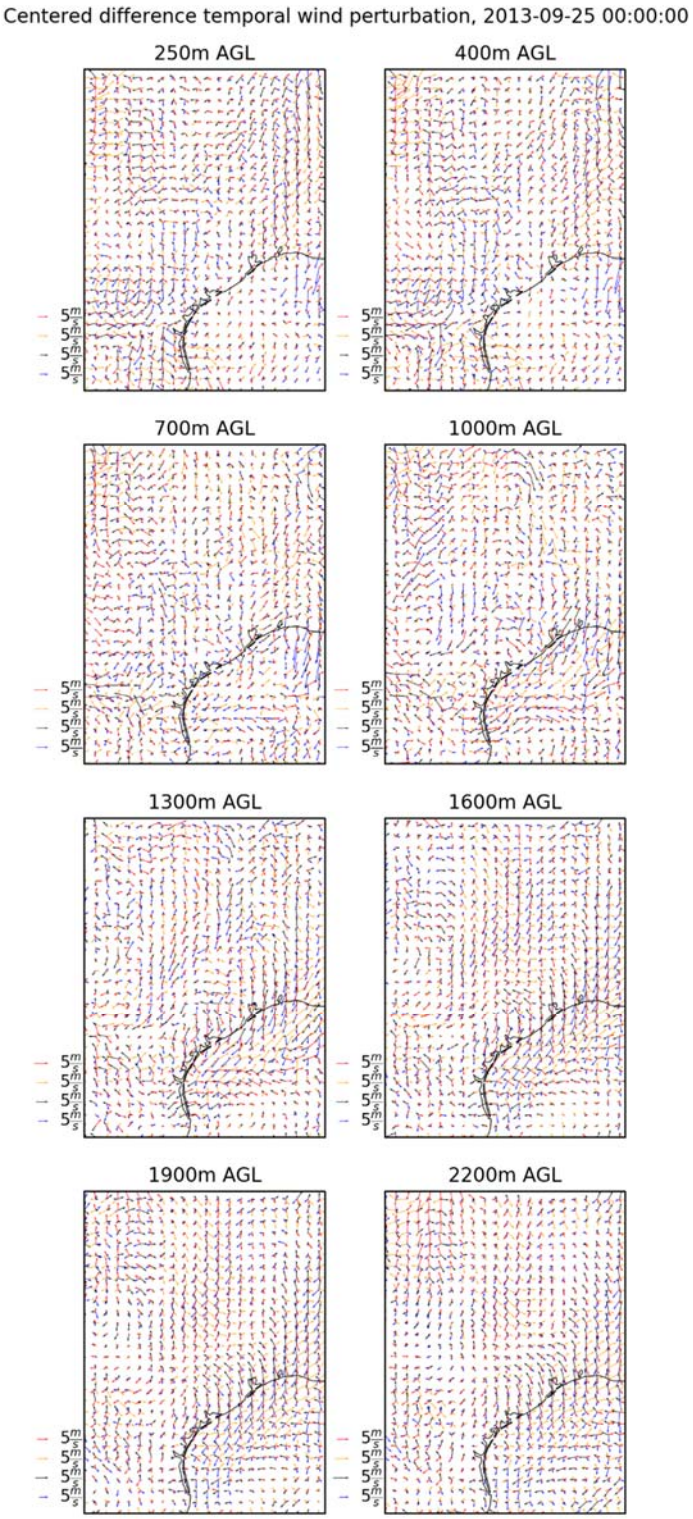


Figure 4.18 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

The following day, the MYJ scheme (Figure 4.19) produces a weak offshore flow at 20Z. The lower boundary layer exhibits a better rotation during this day than the previous day, with a roughly 90° phase shift per hour, but the middle and upper boundary layer is nearly in phase with the lower boundary layer. The YSU scheme (Figure 4.20) produces a weaker rotation in the lower boundary layer than the MYJ scheme, and the upper boundary layer is not in phase with the lower portions; however, the upper boundary layer is out of phase by a clockwise rotation of 90°, not 180, and there is not much evidence of a complete sea breeze circulation. The QNSE (Figure 4.21) exhibits similar offshore flow in the perturbation wind field at 20Z to the YSU and MYJ simulations. Like the previous day, the QNSE perturbation wind field immediately offshore near Houston is larger than the land winds in the middle boundary layer, but the wind magnitude is weaker than the previous day at the same location.

Each of the utilized PBL schemes does not seem to reliably produce the expected sea breeze circulation. The magnitudes of the instantaneous perturbation winds are around the magnitudes one would expect to see for the sea breeze circulation, i.e. nonzero wind speed at or below order 10^1 meters per second. The various PBL schemes, when they demonstrate the ability to produce perturbation winds that veer in time, do not reliably produce a boundary layer column that shows a phase shift with height associated with a low-level complete circulation, but rather the upper and lower boundary layers frequently are in-phase with each other. Often the simulated wind direction does not match the theoretically expected wind direction for a sea breeze circulation, and the wind direction along the coast does not meaningfully differ from the wind direction at grid points away from the coast in inland Texas.

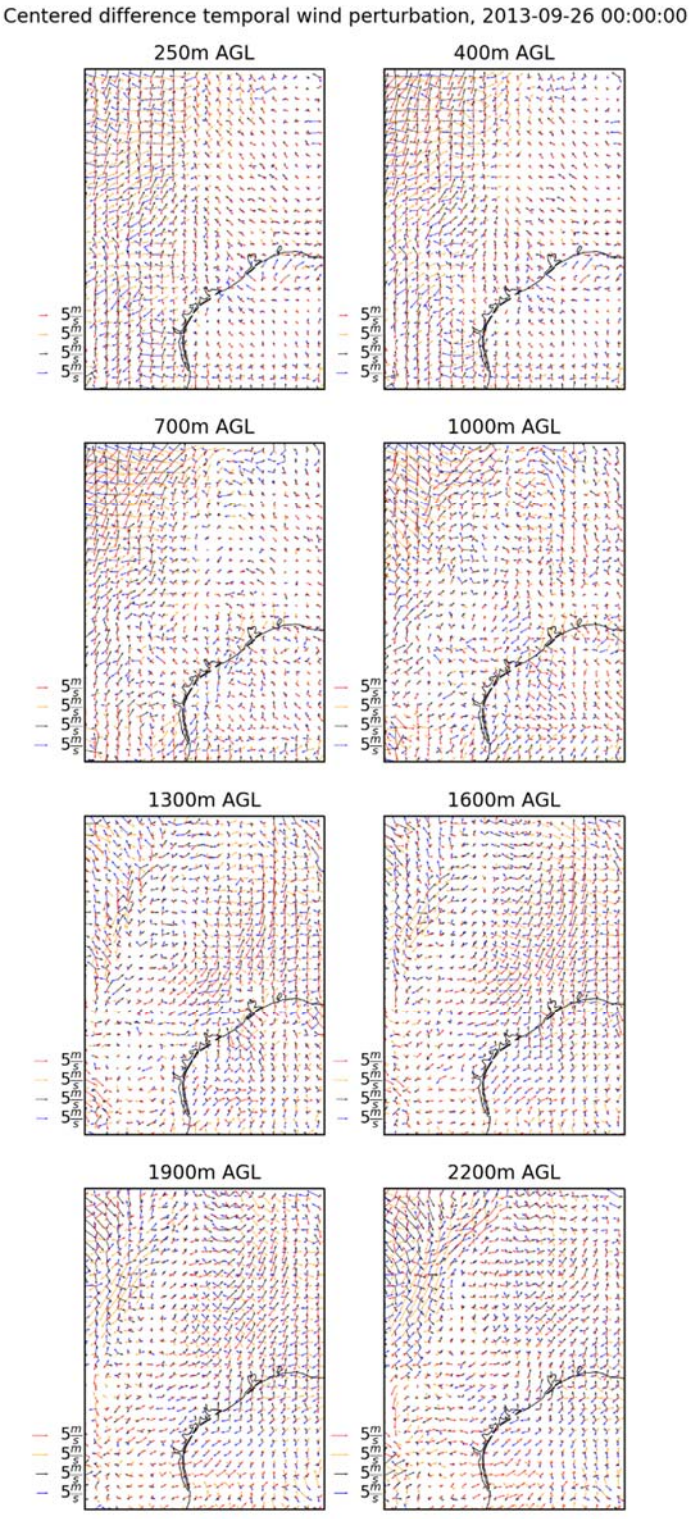


Figure 4.19 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

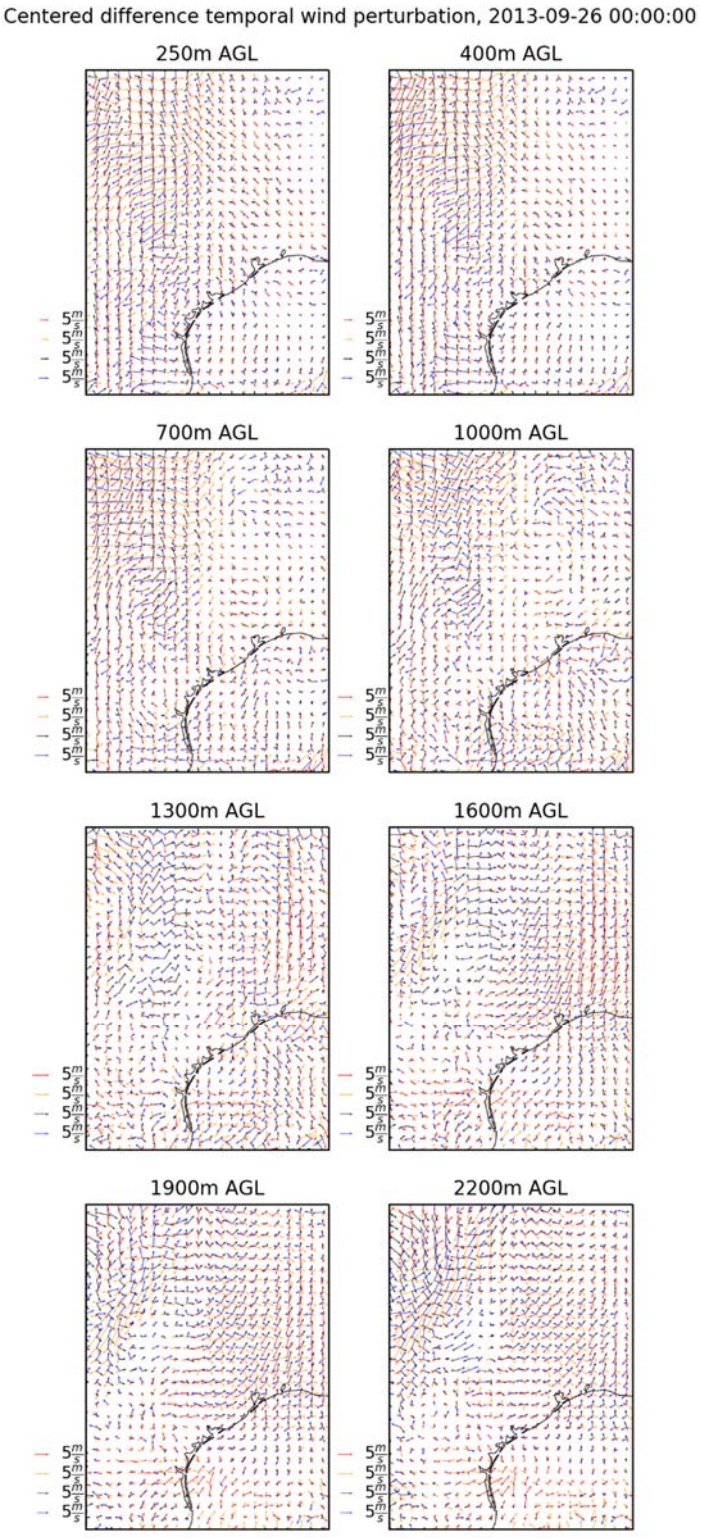


Figure 4.20 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

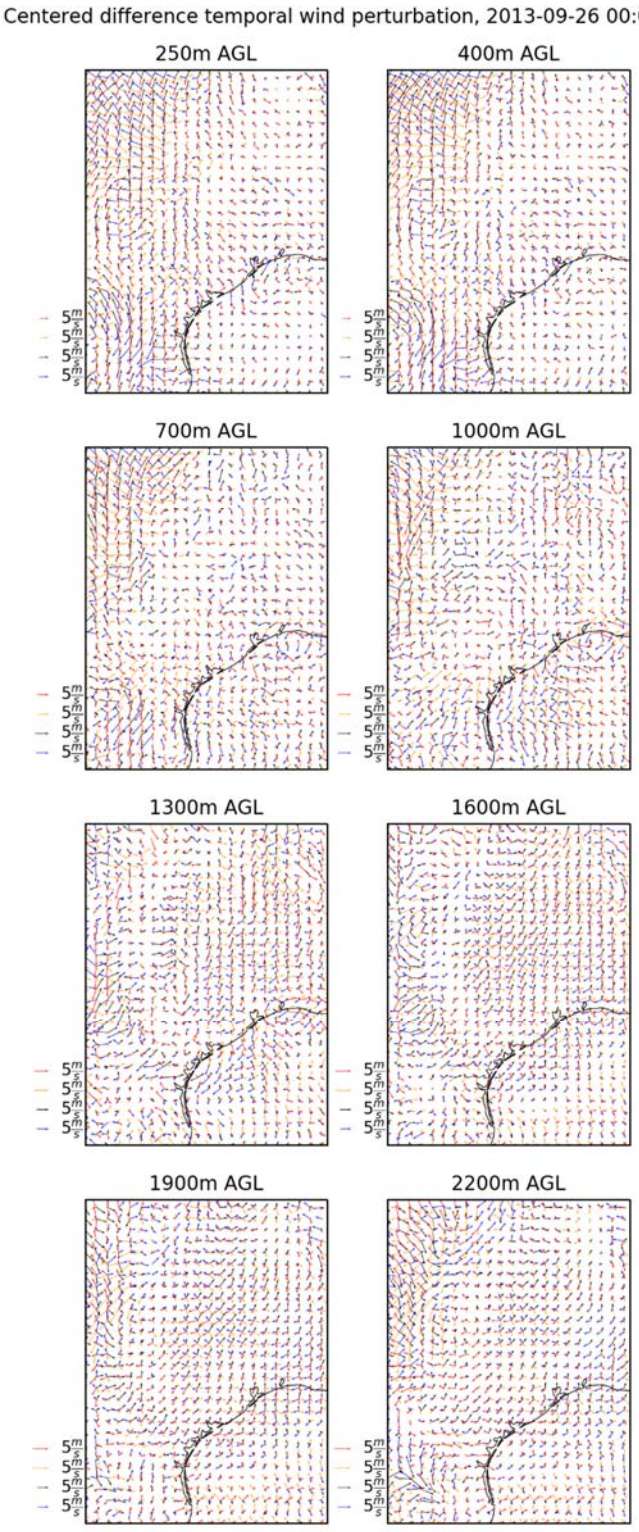


Figure 4.21 Simulated perturbation wind on the 250m, 400m, 700m, 1000m, 1300m, 1600m, 1900m, and 2200m isoheights at 02Z (blue), 08Z (black), 14Z (orange), and 20Z (red).

4.5. CMAQ modeling

4.5.1. Air quality modeling domain and model setup

A three-level nested domain is used in this study (See Figure 4.22), following the RPO Comprehensive Air Model with Extensions (CAMx) domains used by the TCEQ for ozone air quality modeling. Map projection parameters, and other details such as vertical domain structures, are described in detail in http://www.tceq.texas.gov/airquality/airmod/rider8/modeling/domain. In summary, the 36-km, 12-km, and 4-km resolution domains have sizes of 148x112, 149x110, and 191x218 grid cells. All have the same vertical layer structure, with 28 stretching layers up to approximately 15 km above the surface. The first layer thickness is approximately 34 m.

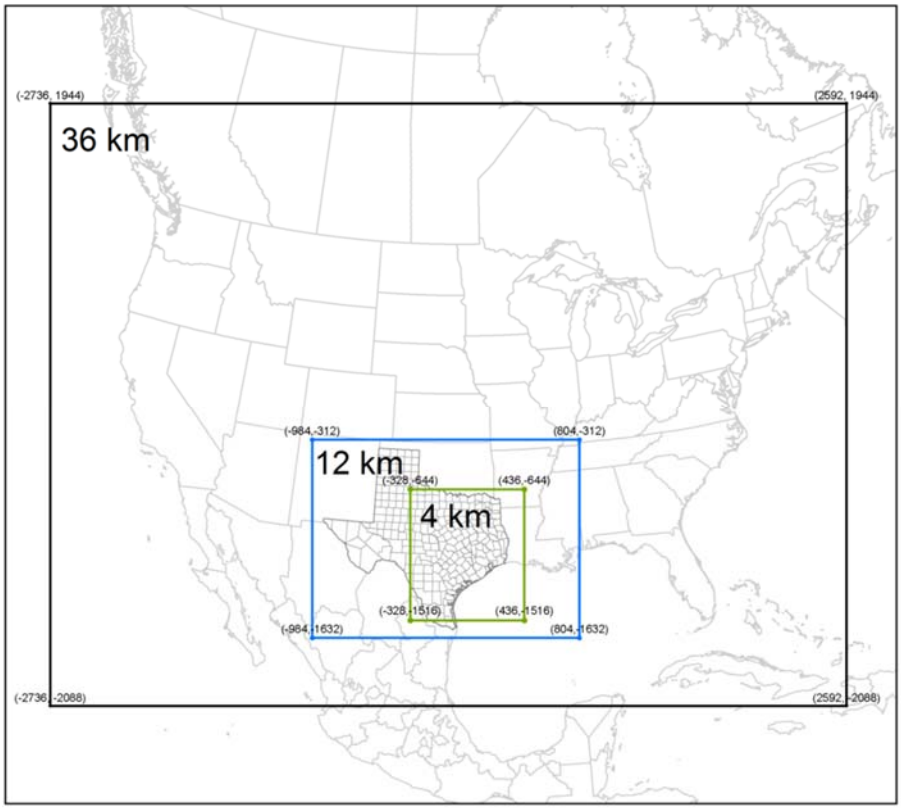


Figure 4.22 Three-level nested CMAQ model domain.

The CMAQ (v5.0.2) model configuration is listed in Table 4.3. Previously, online dust emission calculation depends on the USGS land use information used in the inline biogenic emission module. The CMAQ model was modified to allow online windblown dust emission simulation using the 20-category Moderate Resolution Imaging Spectroradiometer (MODIS) land use classification data. The equation to estimate the vertical dust flux was modified to follow that of Shaw et al. (2008) and the PM_{10} fraction in total PM emission is estimated based on Choi and Fernando (2008). The same equations were used in an offline dust module in previous applications of CMAQ in China (Ying et al., 2014a; Zhang et al., 2012). Photolysis rates are also calculated inline to correctly account for the reduction of actinic flux due to aerosol loading.

Table 4.3 Configuration of CMAQ

Options	Value	Notes
Mechanism	CS07A	Custom version modified from the original CS07A to allowed source and age tracking
Aerosol	AERO6	Version 6 of the aerosol mechanism - treatment of trace metals; aging of primary organics
Solver	SMVGEAR	
Plume rise	Calculated offline in SMOKE	
Dry deposition	Inline	
Dust emissions	Inline	Modified to use MODIS land-use type.
Photolysis	Inline	
Vertical diffusion	ACM2	
Lighting NOx	Not included	
Surface HONO	Enabled	
Biogenic emission	Offline BEIS 3.14	

4.5.2. Anthropogenic and biogenic emission processing

Due to the time limitation of the project, anthropogenic emissions for the year 2011 are used in the final report to simulate air quality for the three identified cases in 2013 and 2016 (see section 4.1) without further adjustment. We will continue to update the analysis using more recent emission inventories. A summary of the emission processing for NEI 2011 is given below. The National Emission Inventory (NEI) 2011 source sectors, as shown in Table 4.3, were processed using SMOKE v3.5.1. Details of the NEI 2011 as used in the EPA's 2011v6 platform can be found in ftp://ftp.epa.gov/EmisInventory/2011v6/v1platform/README_2011v6_package.txt. A short summary regarding point and on-road mobile source sectors is included in the following. In NEI 2011, emissions from electrical generating units (EGUs) are divided into three sectors: *ptegu*, *ptegu_pk* and *ptnonipm*. In older NEIs, the *ptegu* sector was called "*ptipm*" or "Integrated Planning Model". This sector incorporates Continuous Emissions Monitoring (CEM) hourly emissions for a majority of sources. The *ptegu_pk* sector includes units that only operate during times of peak demand, rather than for most or all of the year, as defined by EPA's Clean Air Markets Division (CAMD). Peaking units are kept in a separate sector by the EPA for the purposes of source apportionment in future modeling applications. This sector incorporates CEM hourly emissions for all sources. The *ptnonipm* sector includes emissions from all other industrial point sources. The run scripts provided with the 2011v6 platform were modified so that emissions from all three CMAQ modeling domains can be generated. For the 4-km domain, spatial allocation surrogates for the United States were provided by the US EPA. However, spatial allocation surrogates for Mexico are not available but the 4-km domain does contain a small fraction of Mexico in the lower-left corner. The Spatial Allocator program developed by the US EPA was used to re-grid the 12-km resolution emissions (othar and othon, see Table 4.4) into 4-km resolution emissions. Biogenic emissions were generated using the BEIS 3.14 emission processor in the SMOKE model.

Table 4.4 Source sectors processed using SMOKE 3.5.1 for CMAQ modeling

Source sectors	Type	Notes
afdust	nonpoint	Area fugitive dust
ag	nonpoint	Agriculture ammonia sector
c1c2rail	nonroad	Class 1/Class 2 commercial marine vessels and locomotives
c3marine	nonroad	treated as point sources; Class 3 commercial marine vessels
nonpoint	nonpoint	Other non-point sources
nonroad	nonroad	Non-road mobile equipment sources
np_oilgas	nonpoint	Oil and gas extraction-related emissions
othar	nonpoint/nonr oad	Area and nonroad mobile sources from Canada and Mexico
othon	onroad	Onroad mobile sources from Canada and Mexico
othpt	point	Offshore Class 3 CMV; drilling platforms; Canada and Mexico point sources
ptegu	point	Electrical generating unit; non-peaking units
ptegu_pk	point	Electrical generating unit; peaking units
ptfire	point	Wildfire and prescribed burning
ptnonipm	point	Other industrial point sources
pt_oilgas	point	Oil and gas extraction-related emissions
rateperdistance_cat x	onroad, RPD	California and Texas on-road emissions[1]; on-network emissions[2]
rateperdistance_no RFL	onroad, RPD	On-road emissions for other states; on-network emissions
rateperdistance_Rf only	onroad, RPD	Refuling emissions[3]; all states; on-network emissions
rateperprofile_catx	onroad, RPP	California and Texas on-road emissions; off-network emissions, fuel vapor venting
rateperprofile	onroad, RPP	On-road emissions for other states; off-network emissions, fuel vapor venting
ratepervehicle_catx	onroad, RPV	California and Texas on-road emissions; off-network emissions, non-venting
ratepervehicle_noR FL	onroad, RPV	On-road emissions for other states; off-network emissions, non-venting
ratepervehicle_RF Lonly	onroad, RPV	On-road emissions for other states; off-network emissions, non-venting; refuel only
rwc	nonpoint	Residential wood combustion

[1] Total of the California and Texas emissions were adjusted to match the States' reported totals.
[2] On-network emissions include running emissions from rural and urban roads.
[3] Off-network emissions include start, evaporative and extended idle emissions.

4.5.3. CMAQ modeling of the age distribution of O_3

CMAQ simulations were conducted for three cases described in Section 4.1: September 15 to September 29, 2013, April 25 to May 10, 2016, and July 10 to July 25, 2016. Initial and boundary conditions in these simulations were based on the MOZART-4 global simulations. The first five days in each case were considered as spin-up and were excluded in the analysis. For the September 2013 and May 2016 cases, the nested simulation completed successfully for all three levels. For the July 2016 case, the 4-km resolution simulation was not able to complete due to an unidentified error in the CMAQ source code. Thus, the results from the 12-km simulations were used in analyzing the July 2016 case. Predicted O_3 generally agrees with observations at the stations within the 4-km domain. In the following, we focus on analyzing the O_3 age distribution during the target days (September 25-26, 2013; May 6-7, 2016 and July 21-22, 2016) in the three

selected cases. Two stations, Aldine, a suburban site northeast of downtown Houston and Galveston, a coastal site on the Galveston Island on the Gulf Coast are selected to demonstrate the difference in O_3 age distributions between a suburban site and a coastal site that is under the influence of polluted air from the urban Houston and the Houston Ship Channel and the relatively clean air from the Gulf of Mexico.

Figure 4.23 shows that non-background O_3 at Galveston reaches 30-35 ppb on September 25 and ~25 ppb on September 26, 2013. The two days show very distinct O_3 age distributions. On September 25, the highest O_3 of 78 ppb (~35 ppb non-background O_3) occur at 1400 CST (here we ignore the peak at 1800 and 1900 CST). At that time, the O3 is relatively fresh with 32% of the non-background O_3 has an age of fewer than 2 hours and 80% of the non-background O3 less than 5 hours old. Aged O_3 (> 8 hours old) accounts for approximately 8%. It should be noted that while the peak O_3 on September 25 is captured well (78 ppb vs 70 ppb), observed O_3 concentrations in the afternoon and evening hours are higher than predicted. The imperfect wind field prediction by the WRF model can be a cause of this under-prediction. This will be further explored in Figure 4.24. On September 26, new O_3 formation appears to be low, with 60-80% of the non-background O_3 attributed to aged O_3. The overprediction in the total O_3 Is believed to be related to over-prediction of the background O_3, which is related to high O_3 at the southern boundary of the 36-km domain. At Aldine, the total concentration of O_3 is well predicted on that day.

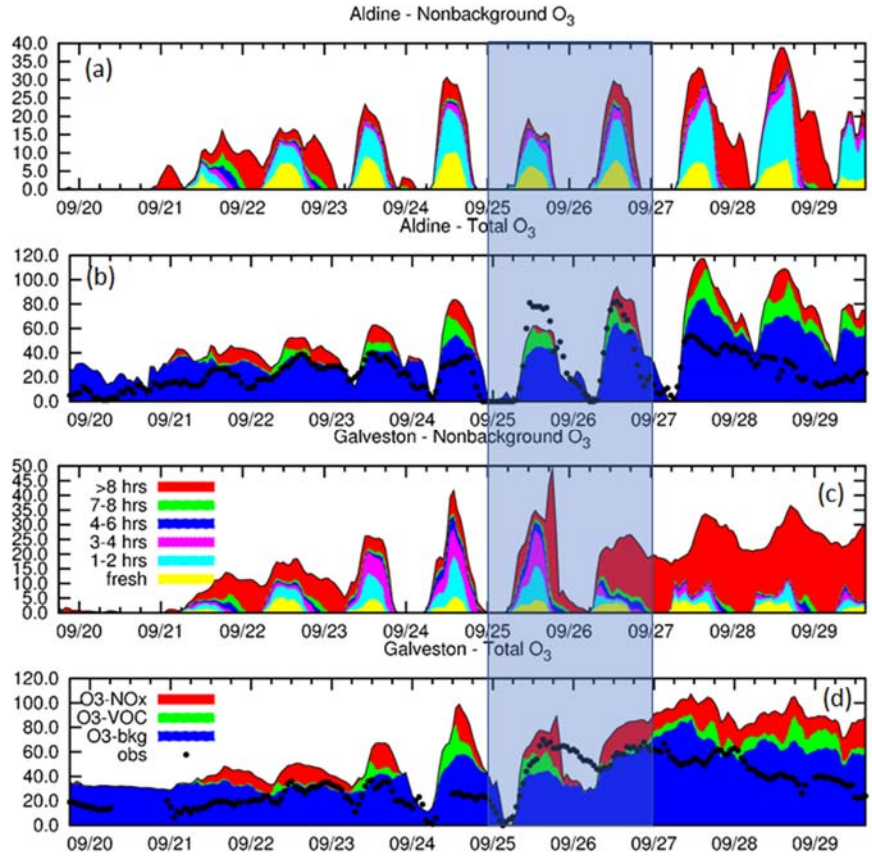


Figure 4.23 Predicted atmospheric age distribution of non-background ozone (i.e. ozone attributed to NOx and VOCs) at Aldine (a) and Galveston (c) and the breakdown of predicted total ozone to NOx, VOC and background contributions (b and d) from September 20 to 29, 2013. Observed and predicted concentrations are in units of ppb. The age distribution results are based on 1-hr time-bin and a few age bins are combined to make it easier to visualize the results.

Figure 4.24 shows the temporal evolution of the regional distribution of aged O_3 on September 25, 2013. At 1100 CST, the surface wind is northwest along the Gulf coast in Texas, which limits the aged O_3 slightly off-coast. At 1500 CST, surface wind near the coast becomes generally stagnant. Concentrations of aged O_3 increased to as much as 50 ppb in some areas over the ocean. These "aged" O_3 is recently formed from aged precursors of NOx and VOCs emitted in the morning and transported to the ocean during the day. At 1800 CST onshore flow develops and pushes the aged O_3 towards inland regions and by 2000 CST the aged O_3 is moved further inland in certain areas. Galveston is on the edge of the high aged O_3 region. A slight difference in the wind field can push the high aged O_3 plume towards it and lead to higher overall O_3 at night time.

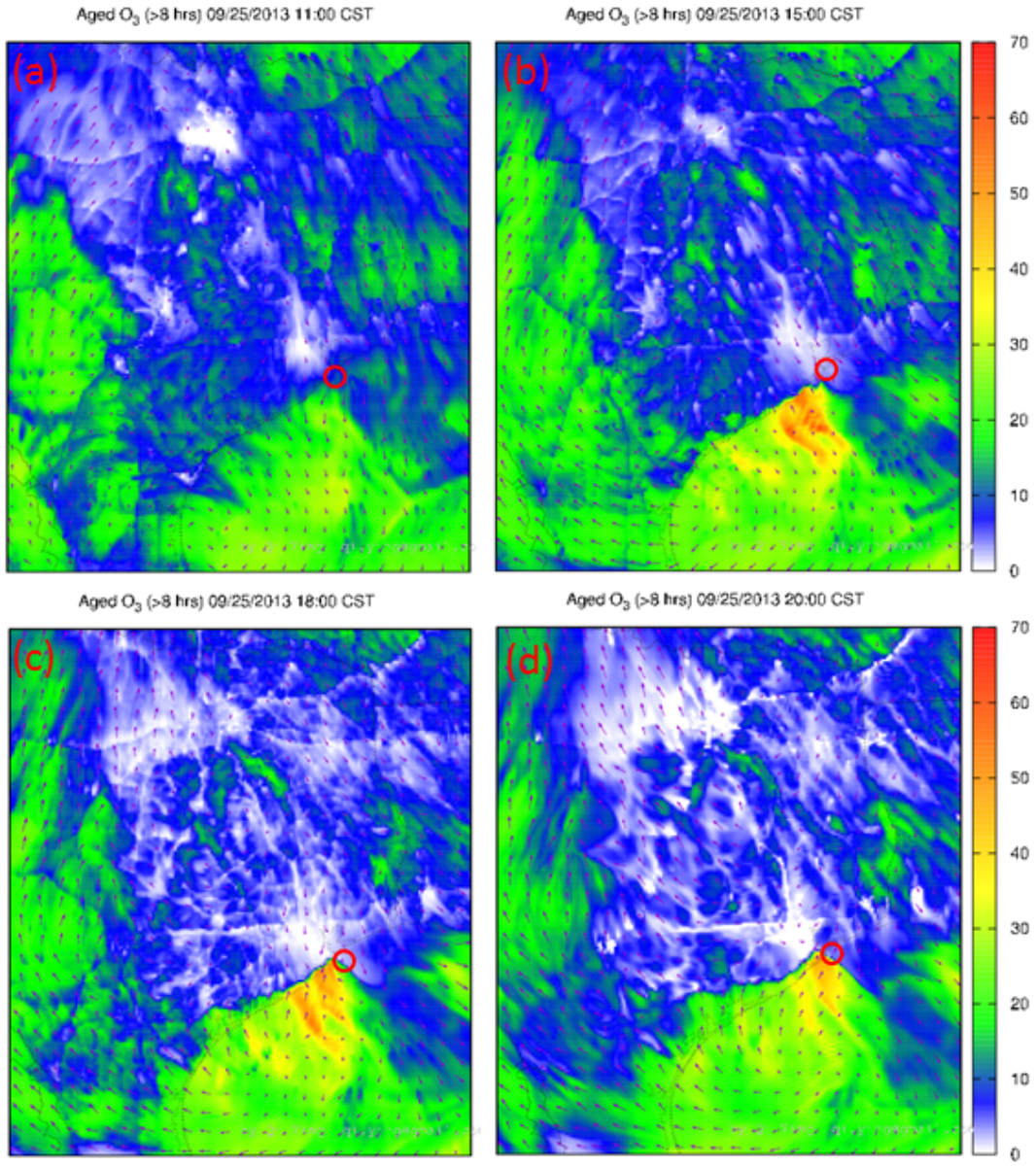


Figure 4.24 Regional distribution of ground-level aged O_3 (> 8 hours) on September 25 at 1100 CST (a), 1500 CST (b), 1800 CST (c) and 2000 CST (d). Units are ppb. The red open circle shows the location of Galveston.

Figure 4.25 shows that the CMAQ model captures the O_3 concentrations at Aldine and Galveston from May 4 to May 9, 2016. Before May 3, the model predicts little formation of background O_3 at both sites and the observed 3 concentrations in these two days are low and without the typical diurnal variations of O_3. Higher O_3 concentrations are observed and predicted at Galveston on May 4 and 5, and at Aldine on May 4-7. O_3 formation at Galveston is mostly attributed to NOx while O_3 formation at Aldine in this episode is mostly attributed to VOCs. At Aldine, non-background O_3 is always fresh with most of the O_3 has an age of 2 hours or less. At Galveston, O_3 is comparatively older with much higher contributions from aged O_3.

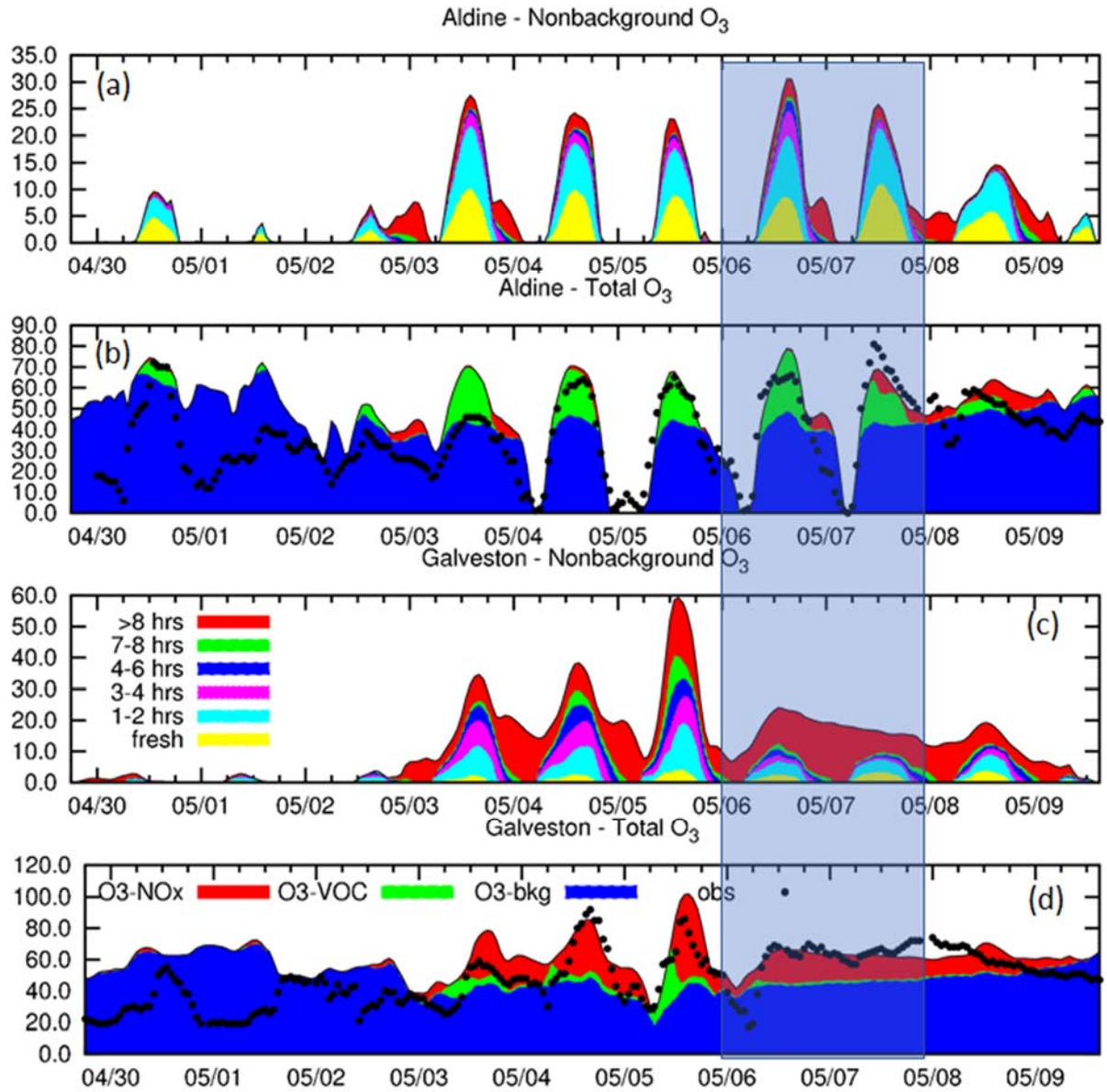


Figure 4.25 Predicted atmospheric age distribution of non-background ozone (i.e. ozone attributed to NOx and VOCs) at Aldine (a) and Galveston (c) and the breakdown of predicted total ozone to NOx, VOC and background contributions (b and d) from April 30, 2016, to May 9, 2016. Observed and predicted concentrations are in units of ppb. The age distribution results are based on 1-hr time-bin and a few age bins are combined to make it easier to visualize the results.

From May 6-8, O_3 concentrations at Galveston remain relatively high throughout the days, with non-background O_3 at ~20 ppb. The predictions agree very well with observations. The age-resolved model predicts that the relatively fresh O3 still shows a clear diurnal variation with peak concentrations of ~10 ppb in the early afternoon. On top of the fresh O3, the aged ozone accounts for a significant fraction, especially at nighttime hours.

Figure 4.26 illustrates the change in the surface wind from 0000 CST on May 5 to 0500 CST on May 6, 2016, and the regional distribution of aged O3. This is the day that a significant amount of O_3 formation is predicted at Galveston and is confirmed by the observations. Panel (a) shows that at midnight of May 5, the strong surface wind near the coast is generally from the southwest and is parallel to the coastline. This confines the high concentrations of aged O_3 (~20-30 ppb) to the ocean. The offshore flow gradually develops during the day, and by 1000 CST has blown the aged O_3 away from the coast. Corresponding, O_3 concentration at Galveston reaches the lowest concentrations as the O_3 depleted air from urban Houston arrives. As the day develops, offshore flow near the coast becomes weak and allows the pollutants to accumulate. Significant O_3 formation is predicted as shown in Figure 4.25. By midnight of May 6, flow is reversed to onshore completely, which the strongest southeast wind to the west of Galveston. At 0600 CST on May 6, a plume of high concentrations of aged O_3 (up to 30 ppb) has reached far inland between San Antonio and Houston.

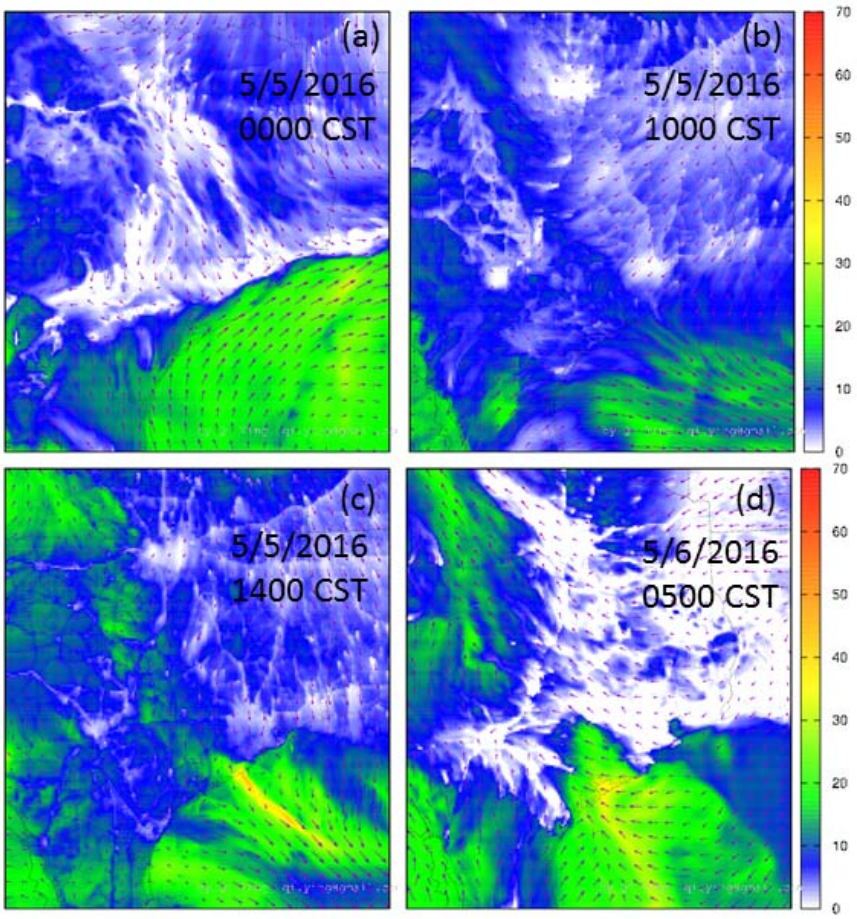


Figure 4.26 Regional distribution of ground-level aged O_3 (> 8 hours) at 0000 CST (a), 1000 CST (b), 1400 CST (c) on May 5, and at 0500 CST (d) on May 6, 2016.

Figure 4.27 compares non-background O_3 at 1400 CST on May 5 and May 6, 2016. The surface wind patterns on these two days are drastically different. As discussed previously, the surface wind is relatively stagnant in the Houston area, which allows the non-background O_3 to reach ~20 ppb at Aldine, and as much as 45-50 ppb of non-background O_3 is predicted in areas where the wind is the slowest. Winds off-coast is mostly from the northwest, and plumes of high non-background O_3 over the ocean can be observed downwind of major emissions sources. On May 6, the situation is quite different. A clear onshore flow is predicted and the non-background O_3 is generally 20-30 ppb along the coastline. Highest non-background O_3 occurs further inland, northwest of the major urban areas in the model domain.

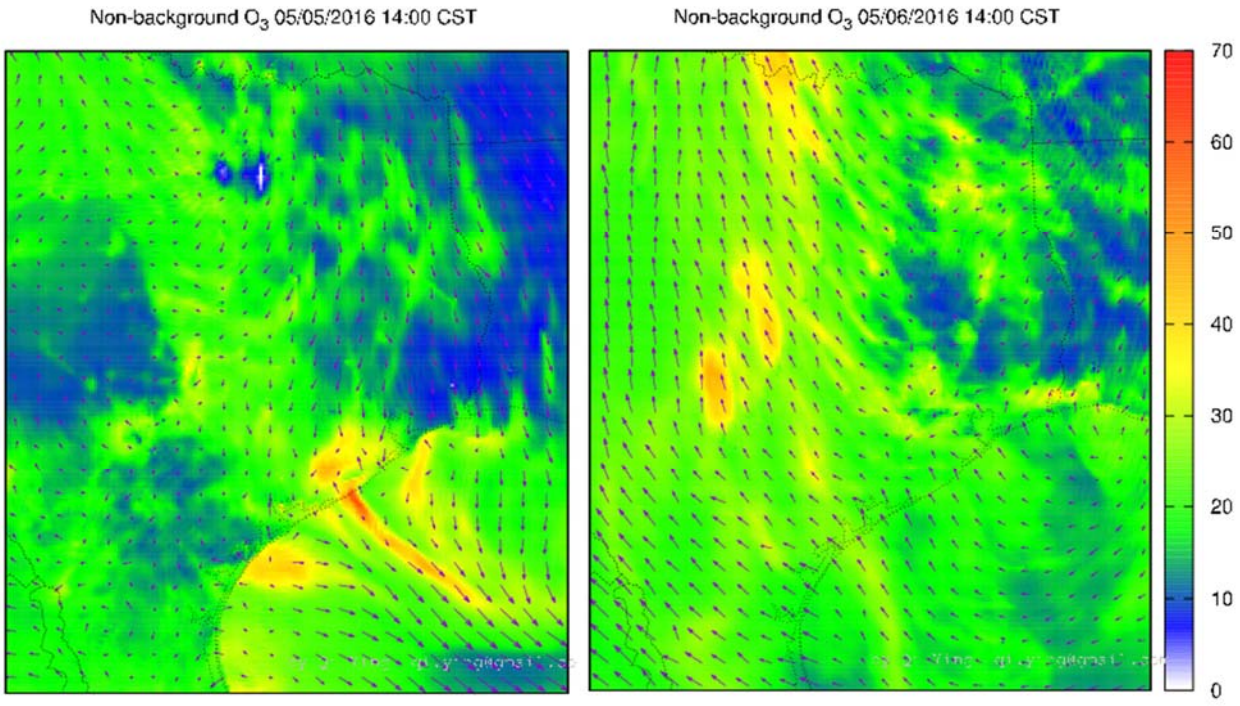


Figure 4.27 A comparison of non-background O_3 at 1400 CST on May 5, 2016, and May 6, 2016. Units are ppb.

For the July 21-22, 2016 case, the age-resolved model stopped due to an unresolved float overflow error for the 4-km domain so the results from the 12-km domain are used in the analysis. As shown in Figure 4.28, predicted O_3 concentrations agree very well with observations at Aldine for almost all days. Predicted O_3 at Galveston is also good although there is a clear over prediction of O_3 on July 23-24. At both Aldine and Galveston, O_3 concentrations are highest during two selected days. At Aldine, O_3 formation is more VOC-limited but during most of the daytime hours, O_3 formation is NOx-limited, as can be seen from Panel (b). At Galveston, a morning time VOC-limited O_3 formation is predicted for these two days and daytime O_3 is still attributed to NOx, suggesting that O_3 formation is NOx-limited. High mixing during the day and less vehicular emissions, which is high in NOx, explains this variation in the O_3 formation sensitivity regime. At Aldine, non-background reaches 40 ppb on July 21 and 75 ppb on July 22, which are higher than non-background O_3 predicted in any other cases investigated in this study. On July 22, aged O_3 (> 8 hours) is approximately 20 ppb at peak hour (1400 CST), which accounts for 26% of the total non-background O_3 at that time. The amount of fresh O_3 (i.e. O_3 < 2 hours old)

is approximately the same for all the days (~20-30 ppb) at Aldine during this period and the difference in the total non-background O_3 is due to the amount of more aged O_3. At Galveston, non-background O_3 on the two high O_3 days are approximately 30-35 ppb, which is similar to the modeled non-background O_3 in high O_3 days in other episodes. As before, aged O_3 accounts for a higher fraction of the non-background O_3 on these two days. For example, at 1300 CST on July 22, the concentration of aged O_3 is 19 ppb, which is 50% of the total non-background O_3 of 38 ppb.

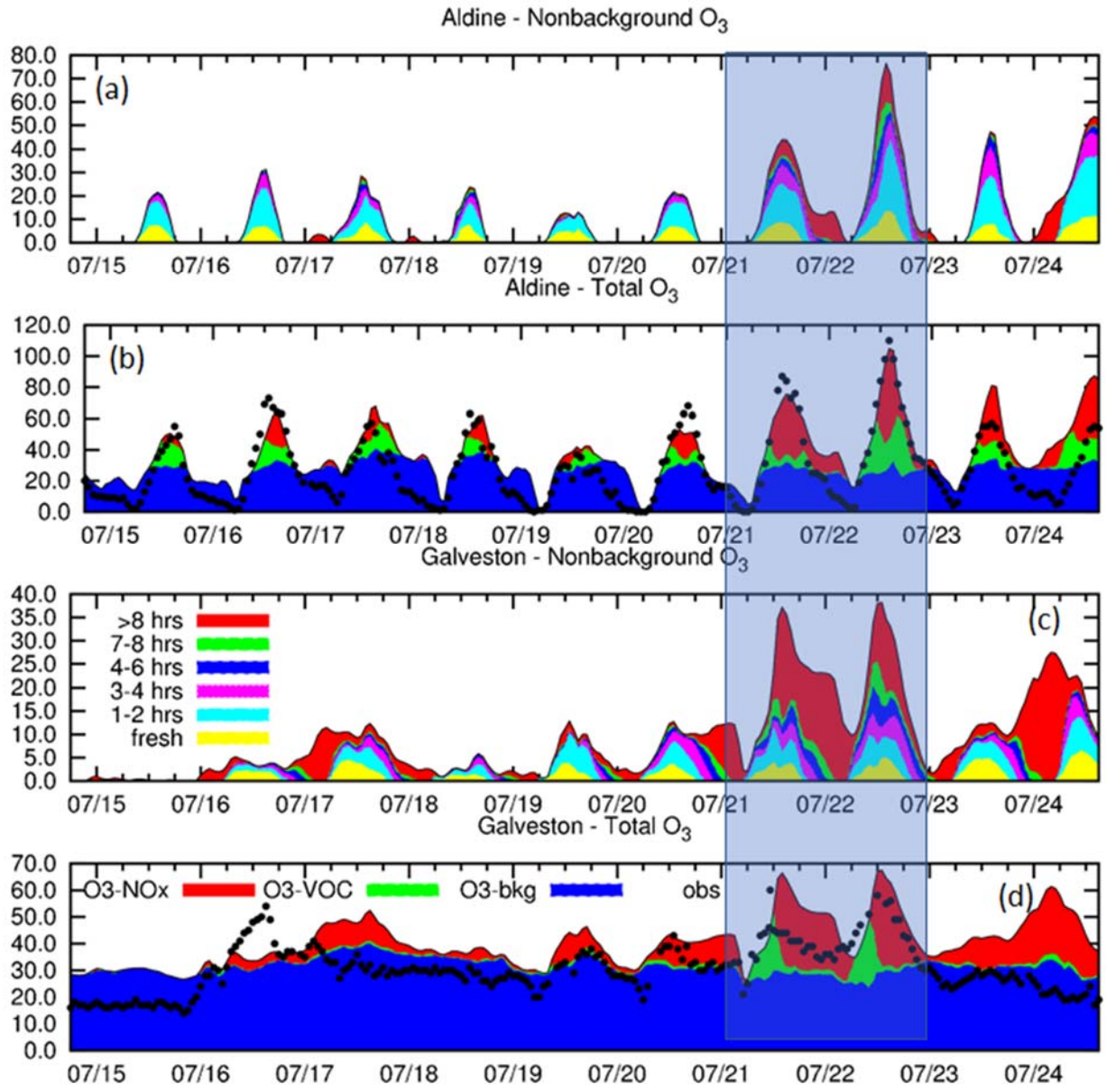


Figure 4.28 Predicted atmospheric age distribution of non-background ozone (i.e. ozone attributed to NOx and VOCs) at Aldine (a) and Galveston (c) and the breakdown of predicted total ozone to NOx, VOC and background contributions (b and d) from July 15 to July 24, 2016. Observed and predicted concentrations are in units of ppb. The age distribution results are based on 1-hr time-bin and a few age bins are combined to make it easier to visualize the results.

The surface wind on these two days shows an interesting persistent circular pattern throughout the day, as shown in Figure 4.29 below. The center of the circular wind pattern appears to be near the coastal area where Texas and Louisiana state lines meet at 0600 CST on July 21. It moves slightly to the west to reach an area between the Galveston Bay and the Matagorda Bay at 1000 CST, and north of the Matagorda Bay at 1400 CST. A blob of aged NOx with concentrations of 2-4 ppb is to the northeast of Houston at 0800 CST. This blob of aged NOx then circulates clockwise to reach the coastal area on the Texas side of the Texas/Louisiana border at 1000 CST. As the boundary layer height increases during the day, the concentration decreases but some of the aged NOx reaches and influences the Galveston area at 1200 CST and remains in the region until 1600 CST. The aged NOx along with O_3 formed previously, contribute to the high concentrations of aged O_3 seen in Galveston. A similar pattern repeats on July 22, leads to even higher concentrations of non-background O_3 and aged O_3 in the Houston, Dallas/Fort Worth and San Antonio areas, as shown in Figures 4.30 and 4.31, respectively.

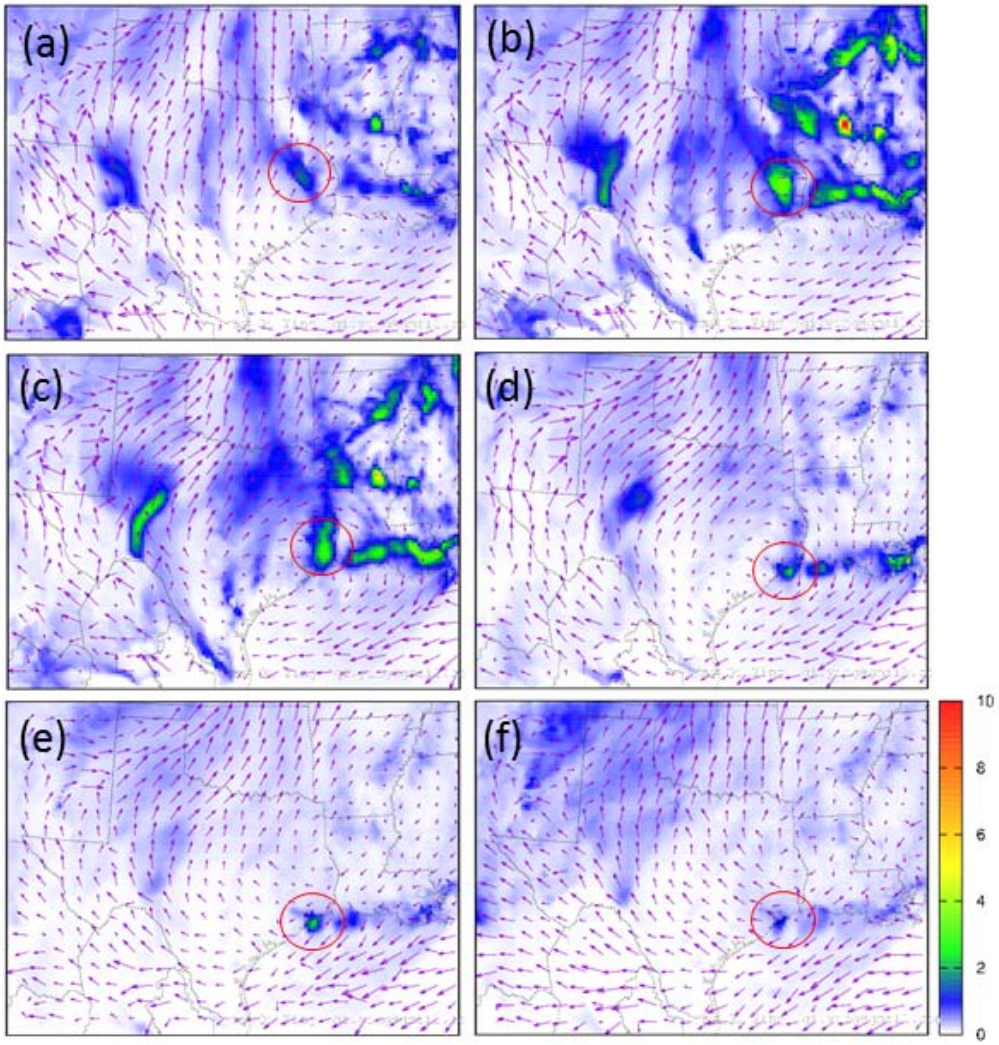


Figure 4.29 Recirculation of aged NOx (> 8 hours) to Galveston Bay. Panels (a) – (f) are 0600, 0800, 1000, 1200, 1400 and 1600 CST, July 21, 2016. Units are ppb. The circles follow the movement of aged NOx as it moves to the Galveston Bay.

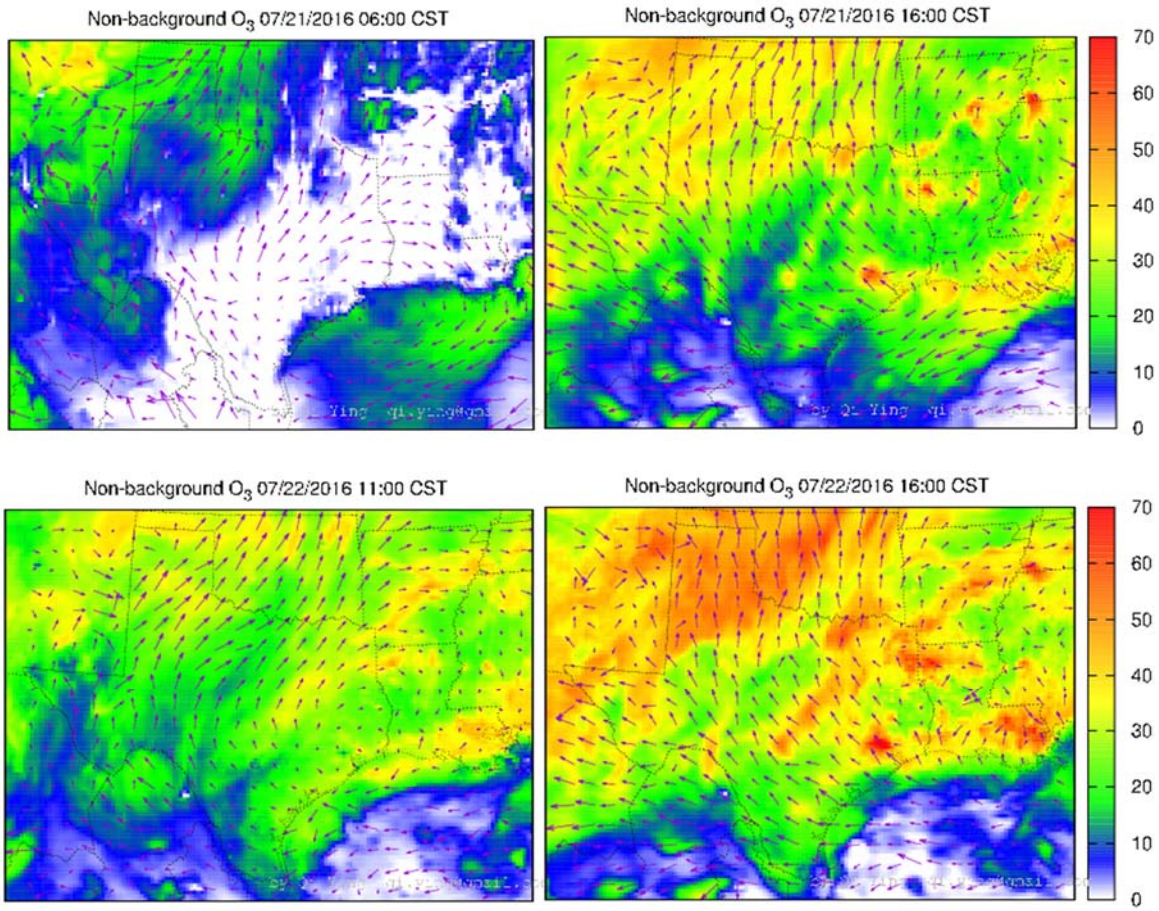


Figure 4.30 Non-background O_3 on 0600 July 21, 1600, July 21, 1100, July 22 and 1600 July 22, 2016. All times are CST and units are ppb.

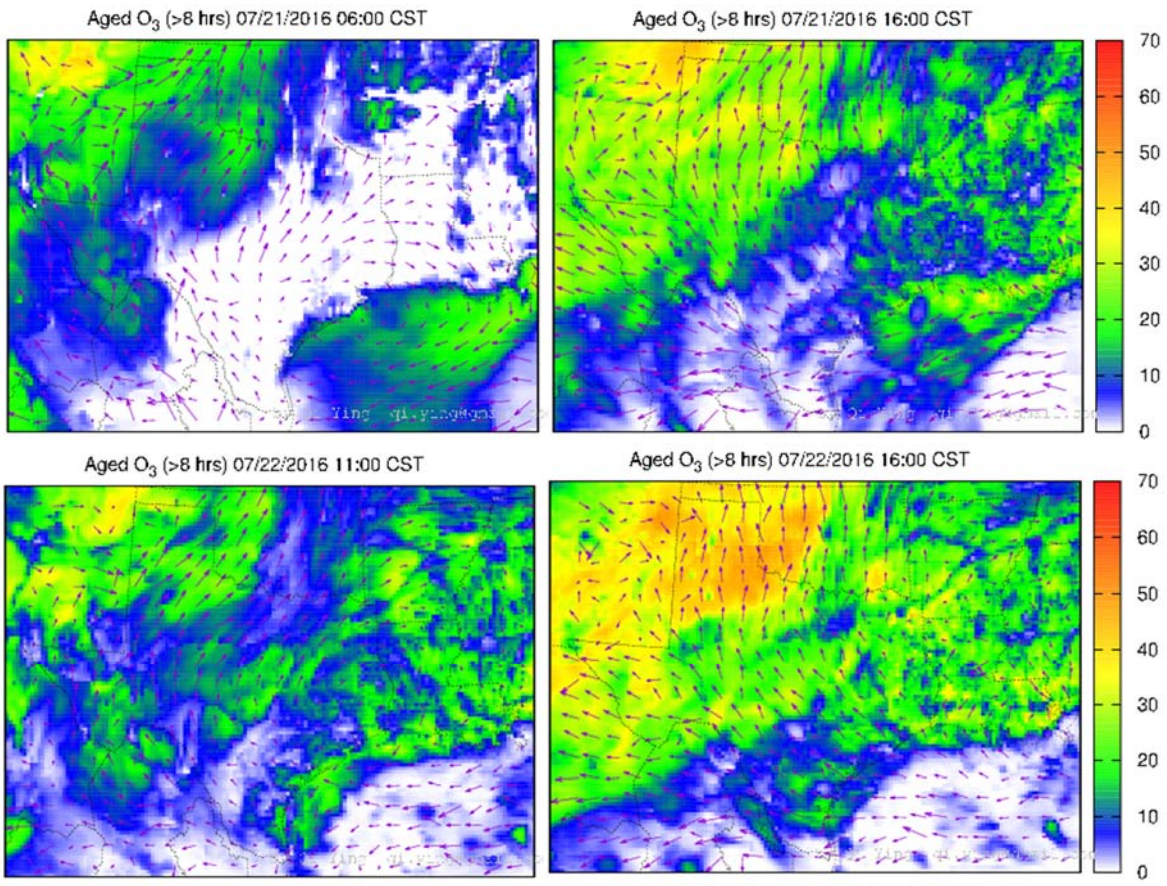


Figure 4.31 Aged O_3 (>8 hours) on 0600 July 21, 1600, July 21, 1100, July 22 and 1600 July 22, 2016. All times are CST and units are ppb. Results are based on the 12-km domain simulation.

4.5.4. Discussion

The age-resolved CMAQ model developed in this study has been demonstrated through the case studies to be a useful tool to help elucidating the cause of the high O_3 concentrations in the study area. The coastal area in Texas is prone to be influenced by high O_3 over the Gulf of Mexico. High O_3 can accumulate over the water as there are much fewer NOx emissions to titrate it. The high O_3 can then contribute to elevated O_3 concentrations onshore. The cases studies in this project show that the contributions of aged O_3 can be as high as 50% (or ~20 ppb) of the peak time non-background O_3 at Galveston. This level of aged non-background O_3 is almost as high as the fresh O_3 predicted in the vicinity of the urban Houston area on high O_3 days. While the WRF/CMAQ system applied in this study gives an acceptable performance of O_3 at Galveston, it is obvious that correct assessment of the impact of aged O_3 to local O_3 events depends on the correct prediction of the wind patterns in the coastal area. As a sharp land/sea gradient of O_3 often exist, particularly at nighttime hours when O_3 over the urban areas are titrated by high NOx emissions, small errors in the wind (esp. the wind directions) can lead to significant differences in predicted O3 concentrations at coastal areas. The current study demonstrates that the WRF model does not correctly reproduce the observed wind circulation patterns derived from the wind profiler data. Further studies are needed to better understand the cause of the errors in the WRF model in the Gulf of Mexico coastal area. This will likely lead to improved assessment of O3 attainment and emission control strategies in the region.

4.6. Quality Assurance

The WRF results described in this chapter has been extensively compared with all available wind profiler measurements and inspected visually. Back-trajectories generated from the WRF simulated fields are compared with back-trajectories computed using the wind profiler measurements. WRF simulations were conducted by David and all (100%) input data, model configurations and model outputs were doubled-check or inspected by Jie Zhang to ensure that no mistakes were made. Standard procedures to evaluate WRF model performance were carried (model performance statistics) by Jie Zhang for 50% of the modeled periods and no significant deviations from past model performances were noticed. Regional plots of the wind vectors were also generated to further ensure the results are reasonable. The CMAQ simulations were conducted jointly by Jie Zhang and Qi Ying. Input files and run scripts were cross-check to make sure that correct inputs were used in the simulations. The performance of the CMAQ model were further studies using time series. Data extract and plotting procedures were mostly done by Jie Zhang. Qi Ying checked about 25% of the work by independently carried out the data extraction and plotting procedures. The generated figures are compared and no difference were ever noticed.

References

Banta, R.M., Senff, C.J., Nielsen-Gammon, J., Darby, L.S., Ryerson, T.B., Alvarez, R.J., Sandberg, S.R., Williams, E.J., Trainer, M., 2005. A bad air day in Houston. Bulletin of the American Meteorological Society 86, 657-+.

Choi, Y.J., Fernando, H.J.S., 2008. Implementation of a windblown dust parameterization into MODELS-3/CMAQ: Application to episodic PM events in the US/Mexico border. Atmospheric Environment 42, 6039-6046.

Darby, L.S., 2005. Cluster analysis of surface winds in Houston, Texas, and the impact of wind patterns on ozone. J. Appl. Meteorol. 44, 1788-1806.

Daum, P.H., Kleinman, L.I., Springston, S.R., Nunnermacker, L.J., Lee, Y.N., Weinstein-Lloyd, J., Zheng, J., Berkowitz, C.M., 2003. A comparative study of O-3 formation in the Houston urban and industrial plumes during the 2000 Texas Air Quality Study. J. Geophys. Res.-Atmos. 108.

Daum, P.H., Kleinman, L.I., Springston, S.R., Nunnermacker, L.J., Lee, Y.N., Weinstein-Lloyd, J., Zheng, J., Berkowitz, C.M., 2004. Origin and properties of plumes of high ozone observed during the Texas 2000 Air Quality Study (TexAQS 2000). J. Geophys. Res.-Atmos. 109.

Davis, J.M., Eder, B.K., Nychka, D., Yang, Q., 1998. Modeling the effects of meteorology on ozone in Houston using cluster analysis and generalized additive models. Atmospheric Environment 32, 2505-2520.

Day, B.M., Rappengluck, B., Clements, C.B., Tucker, S.C., Brewer, W.A., 2010. Nocturnal boundary layer characteristics and land breeze development in Houston, Texas during TexAQS II. Atmospheric Environment 44, 4014-4023.

ENVIRON, 2015. The Comprehensive Air quality Model with extensions (CAMx), version 6.20.

Kleeman, M.J., Cass, G.R., 2001. A 3D Eulerian source-oriented model for an externally mixed aerosol. Environmental Science & Technology 35, 4834-4848.

Kwok, R.H.F., Baker, K.R., Napelenok, S.L., Tonnesen, G.S., 2015. Photochemical grid model implementation and application of VOC, NOx, and O-3 source apportionment. Geosci. Model Dev. 8, 99-114.

Langford, A.O., Senff, C.J., Banta, R.M., Hardesty, R.M., Alvarez, R.J., Sandberg, S.P., Darby, L.S., 2009. Regional and local background ozone in Houston during Texas Air Quality Study 2006. J. Geophys. Res.-Atmos. 114.

Ngan, F., Byun, D., 2011. Classification of Weather Patterns and Associated Trajectories of High-Ozone Episodes in the Houston-Galveston-Brazoria Area during the 2005/06 TexAQS-II. Journal of Applied Meteorology and Climatology 50, 485-499.

Nielsen-Gammon, J.W., 2016. Analysis of wind rotation in Houston and surrounding areas as detected by the TexAQS-II radar wind profiler network: A report to the 8-Hour Ozone SIP Coalition. Texas A&M University, College Station, TX.

Nielsen-Gammon, J.W., Tobin, J., McNeel, A., 2005. A Conceptual Model for Eight-Hour Ozone Exceedances in Houston, Texas. Part II: Eight-Hour Ozone Exceedances in the Houston-Galveston Metropolitan Area. Texas A&M University.

Parrish, D.D., Allen, D.T., Bates, T.S., Estes, M., Fehsenfeld, F.C., Feingold, G., Ferrare, R., Hardesty, R.M., Meagher, J.F., Nielsen-Gammon, J.W., Pierce, R.B., Ryerson, T.B., Seinfeld, J.H., Williams, E.J., 2009. Overview of the Second Texas Air Quality Study (TexAQS II) and the Gulf of Mexico Atmospheric Composition and Climate Study (GoMACCS). J. Geophys. Res.-Atmos. 114.

Pierce, R.B., Al-Saadi, J., Kittaka, C., Schaack, T., Lenzen, A., Bowman, K., Szykman, J., Soja, A., Ryerson, T., Thompson, A.M., Bhartia, P., Morris, G.A., 2009. Impacts of background ozone production on

Houston and Dallas, Texas, air quality during the Second Texas Air Quality Study field mission. J. Geophys. Res.-Atmos. 114, D00F09.
Qiao, X., Guo, H., Tang, Y., Wang, P., Deng, W., Zhao, X., Hu, J., Ying, Q., Zhang, H., 2018. Local and regional contributions to fine particulate matter in the 18 cities of Sichuan Basin, southwestern China. Atmos. Chem. Phys. Discuss. Under review.
Rappengluck, B., Perna, R., Zhong, S.Y., Morris, G.A., 2008. An analysis of the vertical structure of the atmosphere and the upper-level meteorology and their impact on surface ozone levels in Houston, Texas. J. Geophys. Res.-Atmos. 113.
Shaw, W.J., Allwine, K.J., Fritz, B.G., Rutz, F.C., Rishel, J.P., Chapman, E.G., 2008. Evaluation of the wind erosion module in DUSTRAN. Atmospheric Environment 42, 1907-1921.
Tucker, S.C., Banta, R.M., Langford, A.O., Senff, C.J., Brewer, W.A., Williams, E.J., Lerner, B.M., Osthoff, H.D., Hardesty, R.M., 2010. Relationships of coastal nocturnal boundary layer winds and turbulence to Houston ozone concentrations during TexAQS 2006. J. Geophys. Res.-Atmos. 115.
Wang, P., Chen, Y., Hu, J., Zhang, H., Ying, Q., 2018a. Attribution of Tropospheric Ozone to NO x and VOC Emissions: Considering Ozone Formation in the Transition Regime. Environmental science & technology 53, 1404-1412.
Wang, P., Chen, Y., Hu, J., Zhang, H., Ying, Q., 2019. Source apportionment of summertime ozone in China using a source-oriented chemical transport model. Atmospheric Environment 211, 79-90.
Wang, P., Ying, Q., Zhang, H., Hu, J., Lin, Y., Mao, H., 2018b. Source Apportionment of Secondary Organic Aerosol in China using a Regional Chemical Transport Model and Two Emission Inventories. Environ. Pollut. 237, 756-766.
Ying, Q., Kleeman, M.J., 2006. Source contributions to the regional distribution of secondary particulate matter in California. Atmospheric Environment 40, 736-752.
Ying, Q., Krishnan, A., 2010. Source contributions of volatile organic compounds to ozone formation in southeast Texas. J. Geophys. Res.-Atmos. 115, D17306.
Ying, Q., Mysliwiec, M., Kleeman, M.J., 2004. Source apportionment of visibility impairment using a three-dimensional source-oriented air quality model. Environmental Science & Technology 38, 1089-1101.
Ying, Q., Wu, L., Zhang, H., 2014a. Local and inter-regional contributions to PM2.5 nitrate and sulfate in China. Atmos. Environ. 94, 582-592.
Ying, Q., Wu, L., Zhang, H., 2014b. Local and inter-regional contributions to PM 2.5 nitrate and sulfate in China. Atmospheric Environment 94, 582-592.
Zhang, H., Li, J., Ying, Q., Yu, J.Z., Wu, D., Cheng, Y., He, K., Jiang, J., 2012. Source apportionment of PM2.5 nitrate and sulfate in China using a source-oriented chemical transport model. Atmos. Environ. 62, 228-242.
Zhang, H., Ying, Q., 2010. Source apportionment of airborne particulate matter in Southeast Texas using a source-oriented 3D air quality model. Atmospheric Environment 44, 3547-3557.
Zhang, H.L., Ying, Q., 2011. Contributions of local and regional sources of NOx to ozone concentrations in Southeast Texas. Atmospheric Environment 45, 2877-2887.