Final Report

Grant Activity No. 582-5-64593-FY07-20
Analysis of TexAQS II Meteorological Data

A report to the Texas Commission on Environmental Quality

John W. Nielsen-Gammon
Department of Atmospheric Sciences
Texas A&M University
College Station, Texas

May 16, 2008

Contact Information:
John W. Nielsen-Gammon, Texas State Climatologist
Dept. of Atmospheric Sciences
3150 TAMUS
College Station, TX 77843-3150
n-g@tamu.edu
979-862-2248
979-862-4466

Executive Summary

The purpose of Project 582-5-75693-FY07-20 was to improve the quality of Texas Commission on Environmental Quality (TCEQ) numerical modeling by collecting, quality assuring, evaluating, preparing for data assimilation, and analyzing meteorological information collected during the TexAQS-II field program.

Numerous surface and upper air data sets were collected and made available for TCEQ and public access (with documentation) on our TexAQS-II web site, http://www.met.tamu.edu/texaqs2 . In many cases, primary quality assurance was performed by federal agencies or principal investigators, in which case further visual inspection of the data was used to identify outliers. Extensive quality assurance and mixing height estimation was performed on the radar wind profiler network by subcontractor Sonoma Technology, Inc. A related Texas Environmental Research Consortium (TERC) project supported generation of mixing heights from NASA lidar data and collection of other meteorological data into a complete, web-based data set.

A series of nudging files were created for meteorological modeling with the MM5 model using the radar wind profiler data and the NOAA shipborne doppler lidar data. The use of nudging in MM5 during a particular test episode (July 30-August 2, 2005) indicated that nudging directly improved the wind fields and indirectly improved the clouds and precipitation. Temperatures were improved in coastal areas where precipitation had a large impact. The use of the Grell cumulus parameterization on the inner 4km grid caused further reductions in clouds and precipitation, with the amounts similar to observations, and it is possible that Grell will be useful in other simulations as well.

The possibility of using dual-Doppler wind fields, derived from National Weather Service and mobile SMART radars, in data assimilation was investigated. Artifacts were found in the dual-Doppler wind fields that were attributable to radially-varying errors in the SMART radar velocity fields. Thus, using the SMART radar data for data assimilation is not recommended, though the SMART radar data would be useful for model validation in identifying the inland progression of the Bay Breeze front and the behavior of other meteorological features.

Climatological analysis of winds and mixing heights during TexAQS-II revealed that high mixing heights and light winds were generally associated with high ozone. All other things being equal, low mixing heights should be associated with high ozone, but in the cases of Houston, Dallas, and San Antonio, high mixing heights tended to occur when stagnation was prevalent. Low values of the ventilation index, a product of wind speed and mixing height, were associated with high ozone, indicating that the concentrating effect of low wind speeds was more powerful than the diluting effect of high mixing heights. Analysis of individual days showed a range of meteorological conditions, with no single mechanism responsible for high ozone on all the days examined.

1. Introduction

Project 582-5-75693-FY07-20 (henceforth, the Project), sponsored by the Texas Commission on Environmental Quality (TCEQ) was initiated in spring 2007 and completed in spring 2008. This summary of the background and objectives of the Project draws heavily upon the original Proposal for Grant Activities and its subsequent amendments.

The 2005-2006 Texas Air Quality Study (TexAQS-II) was a major field study conducted in eastern Texas between June 2005 and October 2006, with an emphasis on the Houston and Dallas areas. TexAQS-II attracted scientists from universities and government laboratories across the country. Intensive measurements during selected periods were supplemented by a broader, continuous array of meteorological and chemical measurements. One benefit of this field program is the availability of numerous special observations for assimilation, testing, and evaluation of meteorological models for State Implementation Plan (SIP) development.

The purpose of this Project was to improve the quality of TCEQ meteorological modeling by evaluating and providing meteorological data for data assimilation, model verification, and episode characterization. These improvements were in support of the 2005-2006 Texas Air Quality Study (TexAQS-II), the Rules of the TCEQ, the Clean Air Act, SIP development, computer modeling studies, and other goals and requirements.

The Project included the following subcomponents:

(a) Collection, quality assurance, and reformatting of meteorological data for use in data assimilation (observational nudging) with the MM5 meteorological model, for use in model evaluation, and for use in meteorological analysis (Tasks 3.1 and 3.2);

(b) Evaluation of the potential value of Shared Mobile Atmospheric Research and Teaching (SMART) radar data for use in data assimilation with MM5 (Task 3.3);

(c) Evaluation of observational nudging files produced by this Project through analysis of MM5 model performance with nudging for a sample episode, and determination of appropriate MM5 base state constant (Tasks 3.4 and 3.5);

(d) Characterization of the relationship between upper air meteorological observations and pollution transport during the TexAQS-II field study (Task 3.6).

The Project was overseen by John Nielsen-Gammon at Texas A&M University (TAMU). A major portion of component (a) involved quality-control of radar wind profiler and sodar data, together with mixing height estimation. This portion of the work was conducted by Sonoma Technology, Inc., under subcontract to TAMU. Sonoma Technology also carried out the work under component (d).

The Project was carried out by TAMU in parallel with a related project funded by the Texas Environmental Research Consortium (TERC) and administered by the Houston Advanced Research Center (HARC). This TERC/HARC project also involved collection and quality-control of meteorological observations during TexAQS-II. The emphasis of

the TERC/HARC project was on data collected during the 2006 field intensive while various other TERC-funded field measurements were being carried out. The TERC/HARC project also had a greater emphasis on broad availability and dissemination of data through development and population of a publicly-available web site.

Resources from both projects were used synergistically. For example, all radar wind profiler quality control was conducted under subcontract through this project, while the generation of National Aeronautic and Space Administration (NASA) lidar aerosol mixing heights was conducted under subcontract through the TERC/HARC project. Project personnel at TAMU were assigned to various categories of observations in order that the overall data compilation might be conducted in the most efficient way possible.

While this Project is not directly related to other studies, the quality-controlled data sets are likely to receive wide use among scientists involved with analysis of results from the TexAQS-II field program. Thus, many future studies are being assisted indirectly by this Project.

2. Data Collection, Quality Assurance, and Reformatting

Many types of data were collected, organized, and in many cases quality-assured in support of this Project and the TERC/HARC project. The complete set of data is listed here. Those data for which collection or quality assurance was funded by TCEQ are indicated by an asterisk. Further descriptions of the data and formatting, as well as the data itself, are available directly on our web site: http://www.met.tamu.edu/texaqs2, and in our final report for TERC Project H84, Development of Databases and Characterization of Meteorological Conditions for the TexAQS II Intensive Period.

Surface data:

- *NCDC (National Climatic Data Center) hourly surface data, mostly from ASOS (METAR) stations, with some hourly coastal observations added. Common parameters include air temperature and dew point temperature, sea level pressure, wind speed and direction, sky cover, and precipitation.
- *TCEQ surface monitoring data for meteorological variables, ozone, and solar radiation. Hourly data was collected for the entire field program, while 5-minute data was collected for the 2006 field intensive and other episodes of interest to TCEQ.
- *Buoy and coastal platform data from the National Data Buoy Center (NDBC). The data include wind direction, speed, and peak gusts, sea level pressure, air temperature, dew point temperature, sea surface temperature, visibility, and wave information.
- *NOAA (National Oceanic and Atmospheric Administration) Ports data, including five stations in the vicinity of Galveston Bay, reporting wind speed, direction, and gusts, air temperature, water temperature, air pressure, and relative humidity. Some of the data is hourly, some is at 6-minute intervals.

Upper air data:

• NCDC upper air data for the United States, Canada, Mexico, and the Caribbean. Most observations are twice a day, including pressure, geopotential height, temperature, dew point, and wind speed and direction.

• *Quality-controlled radar wind profiler data from NOAA, TCEQ, and STI. This data set and its quality control procedures are described in detail below. Mixing heights were estimated from this data set.

• *NOAA High-Resolution Doppler Lidar (HRDL) wind profiles from the R.V. Ron Brown. Available data includes wind profiles from individual scans, 15-minute aggregate wind profiles, and 1-hour aggregate wind profiles. The latter were created for use in data assimilation. Data is available from early August 2006 to mid September 2006.

• NASA airborne aerosol lidar data and mixing height estimates, produced by NASA under subcontract to TAMU. The data includes mixing heights along the flight track of the NASA King Air, which was active from mid August 2006 to late September 2006.

• ACARS data consisting of temperature and wind speed and direction was obtained from NOAA's Earth System Research Laboratory. Data were collected in both sounding and standard form.

Graphics:

• Radar mosaics, satellite imagery, surface maps, satellite-surface composites, and constant altitude profiler maps were retrieved from the National Center for Atmospheric Research image archive (http://www.mmm.ucar.edu/imagearchive).

• Vertical profiles from rawinsondes and ACARS soundings were generated.

• *Time-height radar wind profiler plots, with mixing heights, were generated.

• *High-resolution surface maps of Houston and Dallas were generated. Note that as of this writing, the winds from TCEQ stations are erroneously plotted at twice their correct wind speed.

• Graphics from the NASA lidar were posted.

Most of the above data had previously undergone quality control either by NOAA or by the principal investigators. For those data, our additional quality control consisted of extreme outlier spot-checks and inspection of visual depictions of the data. However, extensive quality control of radar wind profiler data was conducted under this Project by STI. STI performed quality control on the profiler data, deleting observations that were deemed to be unreliable based on examination of wind profiles, vertical velocity data, signal-to-noise ratio data, and surface meteorological wind data. This quality control was designed to eliminate known sources of error, including nighttime migrating bird wind contamination. For more details, see final report STI-907100-3215-FR.

The radar wind profiler quality control was conducted in two phases, with fifteen profilers completed in 2007 and four more profilers completed in 2008. Our interim report described the procedures, which are summarized here with updated information incorporated.

A variety of radar wind profilers were deployed during TexAQS-II. Some profilers were 404 MHz profilers, part of the Wind Profiler Demonstration Network operated by the National Oceanic and Atmospheric Administration (NOAA), and are

semi-permanent profilers designed to measure winds throughout the troposphere; these have coarser vertical resolution and are not well suited for monitoring low-level jets. The rest of the profilers were 915 MHz profilers, which provide better resolution at low levels but have a typical maximum vertical range of 3 km to 5 km. Some of the latter profilers were operated by TCEQ or under contract from TCEQ or local government entities, while others were operated by NOAA specifically for the TexAQS-II field program.

The locations of profilers included in this data set are given in Table 1. Note that the profiler named "Moody" is located at the NOAA Moody flux tower in north-central Texas rather than the Moody Tower observing site on the University of Houston campus.

Table 1: Names and locations of Texas radar wind profilers included in the quality-controlled data set. An * indicates a 404 MHz tropospheric profiler.

Station ID	Station Location (TX)	Latitude (°N)	Longitude (°W)	Elevation (m)
ACL	Arcola	29.51	95.48	21
BHM	Brenham	30.22	96.37	94
BRZ	Brazos A19	28.18	95.59	24
BVL	Beeville	28.37	97.79	75
CLE	Cleburne	32.35	97.44	250
DQU*	Dequeen, AR	34.11	94.29	195
HKL*	Haskell, OK	35.68	95.86	219
HVE	Huntsville	30.72	95.64	101
JEF	Jefferson Co.	29.94	94.10	5
JTN*	Jayton	33.01	100.98	707
LDB*	Ledbetter	30.09	96.78	122
LPT	La Porte	29.67	95.06	8
LVW	Longview	32.38	94.71	106
MDY	Moody	31.34	97.37	248
NBF	New Braunfels	29.70	98.03	195
PAT*	Palestine	31.77	95.71	119
PRC*	Purcell, OK	34.97	97.51	331
SNR	Sonora	30.26	100.65	697
WNF*	Winnfield, LA	31.89	92.78	93

The times given in the data files are the times corresponding to the start of the period of observation (typically one hour) for which wind profiles are estimated. The appropriate time to assign to these observations, however, is the time at the midpoint of the period of observation. Thus, for all observations other than CLE, JEF, and LPT, 30 minutes is added to the time of observation. CLE, JEF, and LPT typically reported wind observations at 30-minute intervals, so 15 minutes is added to the time of observation at those stations.

For ease of use, seventeen separate nudging periods were designated, each with its own nudging files. These files correspond to the eight episodes of interest identified by TCEQ in the Proposal for Grant Activities for Work Order 21, the periods between the eight episodes, and the period from the end of the final episode to the end of the quality-controlled data set. One of these periods spanned the beginning of the TexAQS-II field intensive, so was divided in two (before and during) for the convenience of those who might wish to simulate the entire TexAQS-II field intensive. Any two or more nudging files from different times may be combined into a single nudging file through concatenation in chronological order.

Table 2: Characteristics of time periods for which nudging files were created. Times are given in UTC; subtract six hours for LST.

Period	Begin mm/dd/yy	Begin Julian Day	Begin Hour	End mm/dd/yy	End Julian Day	End Hour
ep05A	05/19/05	139	0000	06/04/05	155	1200
ep05ib	06/04/05	155	1201	06/14/05	165	2359
ep05B	06/15/05	166	0000	07/01/05	182	1200
ep05ic	07/01/05	182	1201	07/25/05	206	2359
ep05C	07/26/05	207	0000	08/09/05	221	1200
ep05i06	08/09/05	221	1201	05/30/06	150	2359
ep06A	05/31/06	151	0000	06/17/06	168	1200
ep06ia	06/17/06	168	1201	07/31/06	212	2359
ep06ib	08/01/06	213	0000	08/13/06	225	2359
ep06B	08/14/06	226	0000	08/23/06	235	1200
ep06ic	08/23/06	235	1201	08/28/06	240	2359
ep06C	08/29/06	241	0000	09/09/06	252	1200
ep06id	09/09/06	252	1201	09/10/06	253	2359
ep06D	09/11/06	254	0000	09/15/06	258	1200
ep06ie	09/15/06	258	1201	09/30/06	273	2359
ep06E	10/01/06	274	0000	10/11/06	284	1200
ep06if	10/11/06	284	1201	10/17/06	290	2359

Table 3. Characteristics of nudging file data sets.

Version Number	Domain	Texas Profilers	Out-of-State Profilers	404 mHz below 1 km	HRDL Lidar
v2.0	4 km	yes	no	yes	no
v2.1	4 km	yes	no	no	no
v2.2	4 km	yes	yes	no	no
v2.3	4 km	yes	yes	no	yes
v2.4p.4km	4 km	yes	yes	yes	no
v2.4p.12km	12 km	yes	yes	yes	no
v2.4.4km	4 km	yes	yes	yes	yes
v2.4.12km	12 km	yes	yes	yes	yes

Table 3 shows the characteristics of the various nudging file data sets created under this Project. The elimination of 404 mHz data below 1 km was designed as a possible test on the (in)ability of poorly-resolved profiles at low levels to improve the model simulations. Because CLE, JEF, and LPT include observations at twice the frequency of the other profilers and lidar, their observations will effectively be weighted twice as heavily as the other observations during the MM5 nudging process. A few of the profilers were located outside the 4 km domain, so a 12 km domain nudging file was created so that data from these profilers could be assimilated into the MM5 simulations. The use of observational nudging on the 12 km grid (even if gridded nudging is also used) should improve the consistency of the boundary conditions for the 4 km simulations.

Filenames of nudging files are given by *nudg.*[version number].[period].*dat*, where [version number] is given in Table 3 and [period] is given in Table 2. We recommend that versions 2.4.4km and 2.4.12km be used for data assimilation, with the other nudging files available for sensitivity studies.

Not all data sets listed in the Proposal for Grant Activities under Tasks 3.1 and 3.2 have been collected and processed. Missing data sets include the UH rawinsonde and tethersonde data sets (the PIs requested to retain their data, which they had quality-controlled, so we simply posted links to their data archives), and the NOAA Twin Otter lidar data and surface flux data (processing and quality control of these data sets was conducted by NOAA, and the data is available from the respective PIs). For generation of nudging files, we focused on the kinematic information available from the radar wind profiler network and the NOAA HRDL lidar. Sodar data was only available from San Antonio, so it was not likely to have a significant data assimilation impact. The dual-Doppler data was judged not to be appropriate for data assimilation (see below).

Incorporation of rawinsonde/tethersonde data was not done, because the data is not entirely compatible with the profiler data. Because the profiler data is at high frequency, its characteristics determine the nudging parameters. The rawinsonde data is much sparser in time, so assimilation of this data will be sub-optimal. It was beyond the scope of this project to conduct the modeling tests necessary to determine the proper compromise between profiler-appropriate nudging coefficients and rawinsonde-appropriate nudging coefficients. In any case, our future data assimilation work will involve the Ensemble Kalman Filter, within which each data set can be handled independently.

3. Evaluation of SMART Radar data

During summer 2005, mobile Doppler radars (known as SMART radars) were operated in the Dallas and Houston areas to provide dual-Doppler wind coverage over key areas of interest. The report entitled Analysis of TexAQS II Meteorological Data describes our examination of the Houston data for possible use in data assimilation; a summary is provided here.

Doppler radar measurements in clear air are generally confined to the boundary layer or residual layer and detect velocities through radar back-scatter from insects and

refractive index variations. Under the observed conditions, the range at which Doppler velocities can be detected is about 50-75 km.

The principle of dual-Doppler analysis is that since each radar measures one component of velocity (along the radial direction, directly toward or away from the radar), two radars viewing a given segment of the atmosphere at different angles provide two components of velocity. If vertical motion can be assumed to be zero or estimated through vertical integration of the continuity equation, vector horizontal winds can be recovered within the region where wind observations from both radars are available.

To carry out this analysis, radar data is placed in a three-dimensional Cartesian grid. We experimented with various parameters controlling the amount of vertical and horizontal interpolation and smoothing, and developed optimal parameter settings.

Unfortunately, the resulting dual-Doppler wind fields had spatially-invariant artifacts symptomatic of systematic errors in one or both of the radars. Because of the pattern of the artifacts, we concluded that the artifacts were most likely due to errors in the velocities measured by the SMART radar. These artifacts consisted of wind variations as large or larger than real wind variations. We therefore concluded that assimilation of these dual-Doppler winds into a model simulation would be more likely to degrade the model simulation than improve it.

Although the SMART data does not appear to be suitable for dual-Doppler analysis, it can still be used for model validation, particularly through analysis of the spatial progression of the Galveston Bay breeze front and comparison with model simulations.

4. Evaluation of Observational Nudging Files and Base State Constants

Five MM5 model runs were conducted to get a sense of the performance benefit to observational nudging and to evaluate the quality of the observational nudging dadta sets. These experiments and their results are described in our earlier report, entitled Model Configuration and Performance with Wind Profiler Nudging, and are summarized here.

Model setup and initial condition files were provided by TCEQ, and all simulations were performed on TCEQ computers. Only the inner (4 km) domain was simulated. Visual inspection of the profiler observations for the four-day period chosen for the simulation (July 30 through August 2, 2005) showed no obviously erroneous values that needed elimination.

The five model runs included a control run, a run without nudging, a run with different nudging parameters, a run with data from the HVE profiler and data from below 1000 m from the tropospheric profilers withheld, and a run with inclusion of the Grell cumulus parameterization.

The model runs were evaluated for their ability to reproduce the observed clouds and convection. All model runs produced too much cloud cover and too much convection. The run without any nudging was worst in that respect, while the run with the Grell cumulus parameterization was best. The simulations did not seem to be generating spurious convection because of nudging.

The model runs were evaluated for their ability to reproduce observed mixing heights. Because of the erroneous cloud cover, too much of the domain had mixing heights suppressed by cloudiness. Where clear skies were simulated, the mixing heights were spatially variable but generally of the about the right magnitude. Random variability was large enough within the model simulations that no clearly superior model run was identifiable.

The sea breeze was generally simulated better with nudging than without, because the nudging produced more accurate large-scale winds to which the sea breeze responds, and more accurate spatial patterns of convection which can disrupt sea breeze formation. Simulation of the nighttime low-level jet required nudging, and even with nudging it is not clear that the proper spatial pattern of nighttime winds was reproduced, since withholding of a single profiler had a large effect. Nighttime wind patterns did tend to agree with surface observations where winds were strong enough to be felt at the surface.

Statistical validation was performed separately with the TCEQ surface observations and the NCDC surface observations. The TCEQ observations are clustered in areas of greatest interest, but the NCDC surface observations tend to be more representative since they are sited primarily for meteorological purposes. Evaluations were performed separately for subdomains centered on North Texas and Southeast Texas.

The impact of nudging was generally larger in Southeast Texas, where model errors (due partly to erroneously widespread convection) were greater. Data assimilation was almost uniformly beneficial. Nudging improved temperatures by reducing erroneous cloud cover and convection, and improved winds directly.

The base state constants were found to be appropriate for the simulations at hand. The model top was set sufficiently high to avoid spurious convection interaction with the upper boundary condition. Other parameters were set to give a base state temperature profile that corresponded reasonably well to expected mean conditions during 2005 and 2006 ozone episodes in eastern Texas.

5. Characterization of Upper-Air Meteorological Observations and Pollution Transport

Under this project, STI conducted an examination of mixing heights and transport during the TexAQS-II field program and three episodes within the TexAQS-II period. The analysis included a climatological component, looking at statistical behavior over the entire field program, and a case study analysis of the three episodes. The full report is STI-907101-3344-FR, accompanying this report.

The climatological analysis of mixing heights quantified the variations in mean mixing heights as a function of time of day, season, and location. In most areas, mixing heights tended to be highest in the summer and lowest in the winter, but in Houston mixing heights were highest in the spring. Mixing heights also tended to peak earlier in Houston, 2 PM rather than 3 PM. The synoptic pattern leading to the highest mixing heights was also different: a ridge in Houston, but a broad high elsewhere.

Despite these differences, all locations experienced higher mixing heights during the ozone season for episode days (with high ozone) than for non-episode days, although Houston did tend to have lower morning mixing heights on episode days. Other independent analyses of TexAQS-II data have found similar results. All other things

being equal, high mixing heights should lead to low ozone, but all other things are not equal. High mixing heights tend to be coincident with sunny skies and rapid photochemistry. Also, high mixing heights are associated with high pressure, which tends to occur with light winds.

Wind speeds and transport distances were found to be strongly related to ozone concentration, with the predominant transport distance much smaller for episode days than non-episode days. Houston, in particular, had an average episode transport distance of about 270 km, while Dallas and Austin/San Antonio had episode transport distances of 450 km and 330 km respectively.

Multiplying the wind speed and mixing height together gives the ventilation index, a measure of the volume of air in which pollutants will accumulate. The ventilation coefficient was found to be much lower at all three locations during episode days than during non-episode days. This demonstrates that the increased mixing heights during episode days are being more than offset by the decreased wind speeds during episode days.

The meteorology on specific days were examined. This examination expands upon the number of days described on our daily field program pages at http://www.met.tamu.edu/texaqs2 (June 22-25, 2005; July 31-August 2, 2005) and adds a second perspective to some key high ozone days during the field intensive (August 31-September 2, 2006). A variety of factors were found to be important in individual days, such as light winds, sea breeze recirculation, mixing height variations, and cloud cover. Inherent in this variety is the unfortunate fact that no single factor or combination of factors accounts for all high ozone days in Houston. Different processes are important on different days.

Statistical analysis of mixing heights and other factors through statistical regression techniques is ongoing and will be expanded and included as part of TERC/HARC project H107.

6. Recommendations for Future Work

The newest project by TAMU funded by TCEQ, Work Order #24, involves used of advanced data assimilation with the goal of improving on nudging both through the assimilation technique itself and by direct incorporation of clear-air Doppler radial velocity data. This is an appropriate next step to take full advantage of the observations collected during TexAQS-II to improve the modeling of ozone episodes during that period.

In addition to the mixing heights derived from profiler data by STI, NOAA derived profiler data for the field intensive period. It may be interesting to compare the two mixing height data sets to determine the consistency and relative accuracy of the mixing height estimates. Also, if mixing heights have been estimated from the NOAA airborne lidar, the data would be useful as an incremental additional source of information to evaluate numerical modeling, if extensive modeling will be done during the period in which the NOAA airborne lidar flew.