Report 2
Data Assimilation/Ensemble Techniques

TERC Project H-45-S-B-2004-T2
Modeling Strategy in Support of TexAQS-II and 8 Hour Ozone Assessment

John Nielsen-Gammon and Gerry Creager
The Center for Atmospheric Chemistry and the Environment
Texas A&M University

November 6, 2005

Contact information:
John W. Nielsen-Gammon
Texas A&M University
3150 TAMUS
College Station, TX 77843-3150
n-g@tamu.edu
979-862-2248

Note: The preparation of this report is based on work supported by the State of Texas through a Contract from the Houston Advanced Research Center, Texas Environmental Research Consortium and the Texas Commission on Environmental Quality.

Purpose

The primary purpose of this report is to describe results from the following four tasks:

1. Task B1: Observational Data: Documentation of observational data used in support of modeling, data assimilation, verification, and analysis.
2. Task B2: Prepare for Real-Time Data Assimilation: Documentation of data assimilation procedures.
3. Task D1: Data Assimilation: Summary of progress on use of EnKF technique.
4. Task D3: Additional Models and Ensembles: Summary of met-chem ensemble modeling experiments.

Since the project itself has as its primary operational goal the support of TexAQS-II field activities, a summary of the forecasting support under this project will also be provided.

Details of the operational numerical model configuration and performance are provided in the revised version of Report 1.

Organization

This report is organized as follows. First, an overview of the project and its activities in support of field operations will be given. Next will be a combined discussion of the observational data (Task B1) and the data assimilation procedures to be used for real-time data assimilation (Task B2). This is then followed by an overview and discussion of separate reports submitted under Tasks D1 and D3. These reports were written by other scientists under this project, are being submitted simultaneously with this report, and are to be considered incorporated into this report.

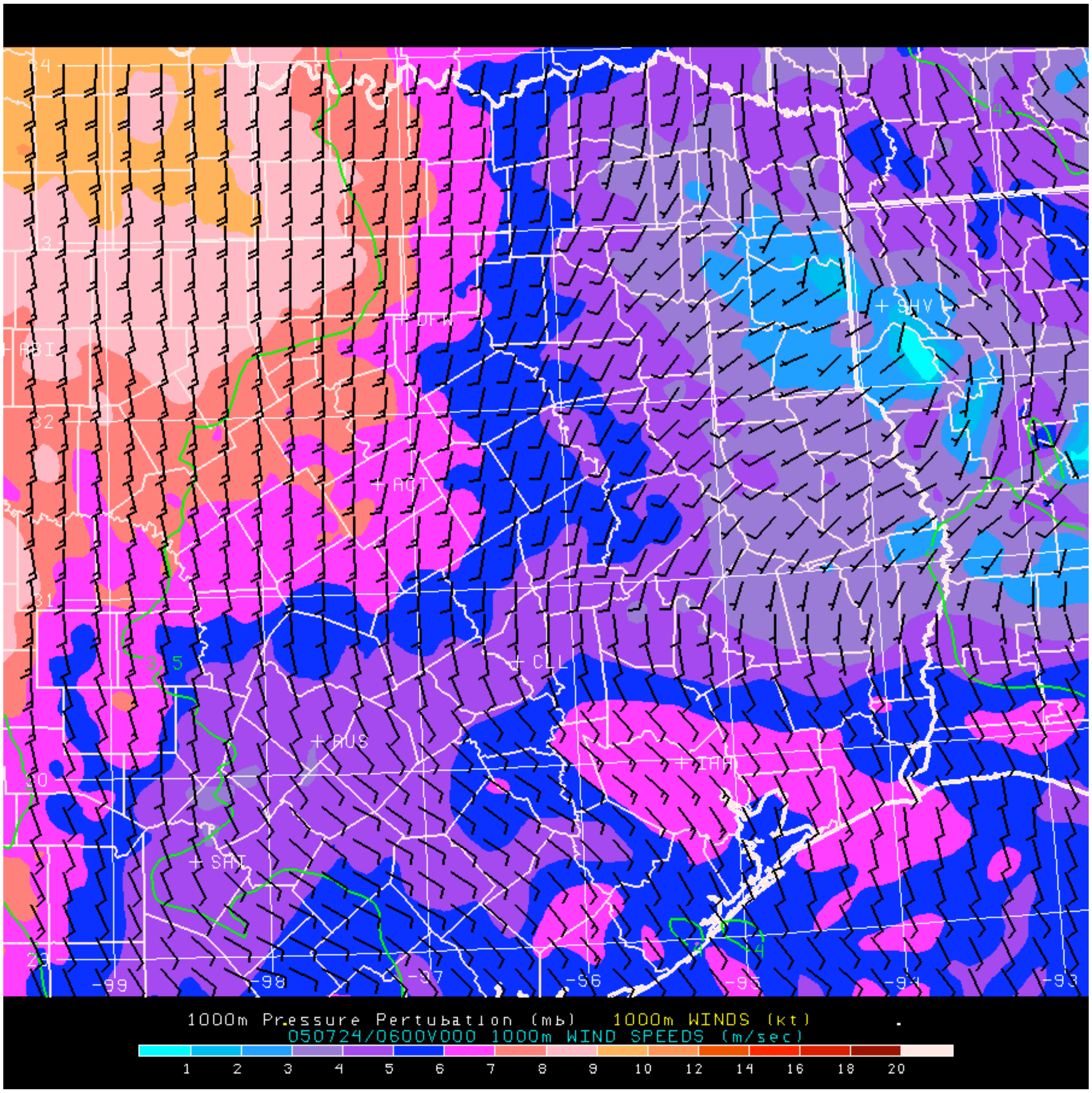
Field Support Overview

Field support was requested for two field projects: the Southeast Texas Transport Study (SETTS) and the Northeast Texas Plume Study (NETPS, sometimes known as the Northeast Texas Transport Study). The field support in both cases consisted of high-resolution meteorological modeling and subjective forecasting guidance, but the nature of the support depended on the specific field projects.

SETTS was scheduled for July 16-31, 2005, with the operations base in Conroe, Texas. The primary contact person was Carl Berkowitz of the Pacific Northwest National Laboratory. The objective of the study was to conduct controlled-altitude balloon flights downwind of Houston during both daytime and nighttime conditions. By request of SETTS scientists, a climatological study was made of transport winds in East Texas during the summer season, as an aid in field activity planning. This report was not originally specified as one of the deliverables of the project. A copy of this report was provided to TERC in July 2005.

Modeling support was provided by our operational MM5 model runs. The model was run once per day, initialized with data from 0000 UTC. The model would typically run from 10:30 PM to 6:30 AM, completing a 56-hour forecast that, by the time of output availability early in the morning, provided forecast guidance for that day, that night, the following day, and the following evening. Output products designed specifically for the SETTS include hourly maps of horizontal wind distributions at various altitudes above ground level, including 100 m, 300 m, 500 m, 1000 m, 850 mb (approximately 1500 m), and 2000 m. Maps were also generated of hourly cloud and precipitation distribution. The innermost (4 km) forecast domain included the entire area of expected tetroon travel.

Model output was transmitted, as soon as it was generated, to the University of Alabama at Huntsville (UAH) and the University of Houston (UH). UAH ingested the data and used it to drive their plume trajectory tool, which was modified to provide forecasts of tetroon drift. We provided additional quality control for the UAH plume product, calling attention to an inconsistency between the MM5 forecasted winds and the UAH plume forecast, which turned out to be caused by an error in the UAH processing of the MM5 fields. The UH data feed was intended for UH to use it to drive a photochemical model for additional forecasting guidance.



A sample operational MM5 forecast from the model run initialized July 23, 2005 at 0000 UTC. This particular image is of 1000 m winds at 30 hours into the model forecast, or 0600 UTC July 24, 2005. The wind barbs represent the direction and speed of the winds. The colors also indicate wind speed, in meters per second, with the scale given at the bottom of the figure. This image, like all others generated by the operational model, is accessible via the web from the model image archive. This particular image is at http://mesonet.tamu.edu/TexAQS/images/050723/tatxam.30.gif.

In preparation for subjective forecasting support, mock forecasts were made of daytime and nighttime winds and expected ozone levels in and downstream of Houston beginning on July 1, 2005. Participating in the mock forecasts were John Nielsen-Gammon and graduate students James Tobin and Brent McRoberts. These mock forecasts continued until July 16, 2005, at which point they became operational.

The SETTS did not request on-site forecasting support. Instead, the SETTS staff included an undergraduate meteorology student who was responsible for the local forecasting. Our forecasting support consisted of three classes of activities: operational forecasts, conference calls, and access to meteorological information.

Our operational forecasts were produced regularly, posted on the web (http://mesonet.tamu.edu/TexAQS/trac/cgi-bin/trac.cgi/wiki/ForecastDiscussion), and verified on the web. That web site is still up and will remain up indefinitely. An example of a Houston forecast and verification for July 23, 2005, is included here:

Forecast by John Nielsen-Gammon, Brent McRoberts, and James Tobin
Houston: Light northerly winds in Houston will shift to the southeast this afternoon after becoming calm, and increase to near 10 kts. There will be scattered thunderstorm activity during the day but organization is doubtful and the activity should die away by 6PM. Winds overnight should be southerly, veering to southwesterly, at 5-10 knots at 1000 m and 10-15 knots at around 400 m. Because the MCS destroyed the possibility of a westerly wind overnight, the best launch site will remain to the north. Tomorrow, as low pressure develops over the Great Plains, winds will be light in the morning but pick up sooner from the SE during the day and be 10-20 kt around sundown, again with a low-level jet developing overnight. Chance of precipitation slightly higher tomorrow: 80% area, 25% point. Sunday maximum ozone: 1-hr: 80 ppb, 8-hr: 50 ppb.

Verification
Houston: Winds were light and variable during the morning, mostly from the northeast, with some scattered showers between Houston and the coast. The showers spread inland but only affected about half of Houston; except for western Harris County they had died away by 3 PM. Winds picked up from the southeast around 2 PM and blew at 7-10 kt for the rest of the day. A low-level jet developed at the Huntsville profiler with winds from the S approaching 20 kt at 300m-500m. 1-h ozone max = 107 ppb, 8-h = 77 ppb (3) (5)

The verifications serve both as a measure of the accuracy of the forecasts and as a basic summary of the air pollution meteorology on each day during field operations.

The conference calls, conducted upon request from SETTS scientists, involved a discussion of the upcoming weather situation in the context of field operations. The independent forecasts made by us and by SETTS scientists were compared and a consensus was reached. We also provided extensive insight into the summertime meteorological environment, such as the diurnal nature of the precipitation and the unusual character of the southeast Texas sea breeze and low-level jet.

The access to meteorological information included not only the model output from our own MM5 modeling, but also a clearinghouse web site designed to provide direct links to all useful meteorological and ozone observations and forecasts. This web page, which was also heavily utilized during the NETPS, is accessible at (http://mesonet.tamu.edu/TexAQS/fcstlinks.htm).

The combined forecasting efforts of SETTS project scientists and ourselves were successful. SETTS conducted flight operations on three evenings, deciding to avoid daytime flights because of the danger of daytime convective activity. The nighttime flights all traveled northward, away from Houston, as forecasted.

The NETPS (lead scientist: Noor Gillani) requested on-site support for its field operations in Waxahachie, Texas. The on-site support we provided included two forecasters for the first week (for training and consistency) and one forecaster for each of the remaining three weeks. The forecasters stayed at the same hotel as the NETPS scientists, conduced daily weather briefings in the early afternoon, provided evening forecast updates when requested, and provided early morning go/no-go decision support prior to all daytime flight activities.

The forecasters continued the practice of posting daily forecasts on the web. A sample set of forecasts for Thursday, August 25, 2005, are shown here.

Daily forecast by Brent McRoberts
Dallas/Fort Worth: Old outflow boundaries and strong heating today give the Dallas area a slight chance of afternoon storms (40% area, 20% point), but rainfall should be out of the picture for the rest of the week. Southerlies around 10 kt at noon should diminish to around 5 kt by mid-afternoon. Winds will increase in intensity to 15-20 kt and shift to the SSE or SE by late evening. Overnight winds should stay mainly southerly at 15-20 kt with a slight shift to SSW or SW expected by sunrise. Winds should decrease to 6-8 kts during mid-morning. A shift in the winds from the SW to the SE looks to occur, but the models conflict on when this will take place. For now, will go with the ETA and MM5 models with a shift occurring between 11am and 1pm. Strong ridging should prevent any convection on Thursday afternoon, so will not include any POPS. Thursday evening winds will intensify to 10-15 kt out of the southeast. Winds overnight Thursday into Friday look very similar to the forecast for the night before (southerly at 15-20 kt). Also, the Friday wind forecast is much like the Thursday forecast. Maximum 8-hour ozone: today 85 ppb, tomorrow 90 ppb, Friday 90 ppb

6:00 AM Update by Brent McRoberts in support of Thursday's flight
Winds at DFW at 6AM were out of the SSW at 20 kt. Expect winds to turn to SW by mid-morning, decreasing in intensity to roughly 10 kt by noon tou of the SSW. Winds will be southerly in the afternoon,

possibly turning slightly to the SSE and should decrease to about 5 kt, although yesterday winds never made it much below 10 kt. Convection should be non-existent in the metroplex all of today.

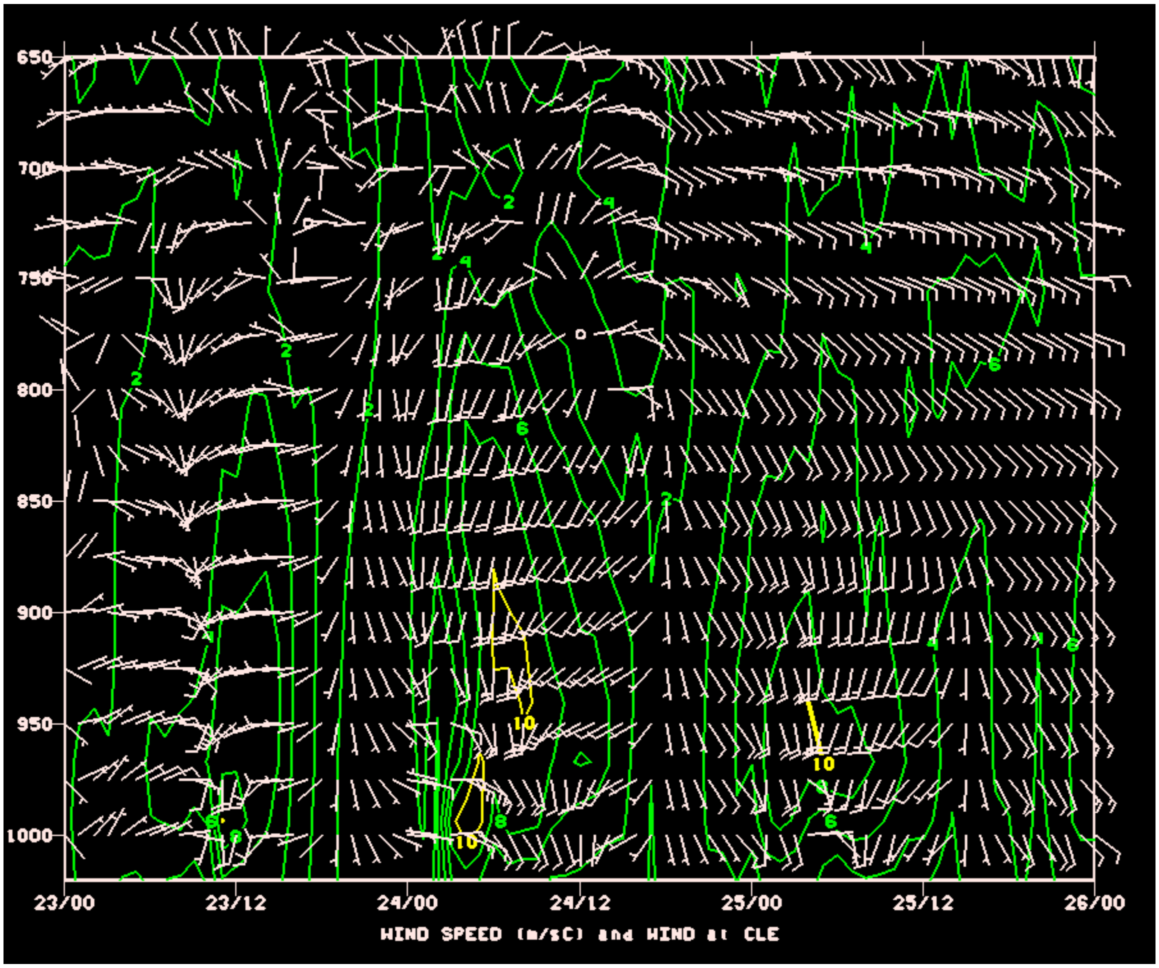
Verification
Dallas/Fort Worth: Winds overnight were out of the S at about 20 kt, veering slightly to the SW at 15 kt by 9am. By noon, winds were out of the S at 10 kt and SSE by 3pm. Winds remained out of the SSE for the rest of the afternoon veering to the S by late evening, increasing to 20 kt by midnight. Again, the only convection in NE Texas was confined to the Tyler/Longview area, lasting from 3-9pm. Maximum 8-hr ozone: 88 ppb. (5) (5) (5)

Note that the forecast discussion made Wednesday includes forecasts for Thursday and Friday. It became clear early on in the field operations that the NETPS scientists wished to have operations guidance two days in advance. To support that requirement, the daily MM5 forecast runs were extended from 56 hours to 72 hours. The full 72-hour model output was used to generate graphics, and the full model output was sent to UAH and UH as well.

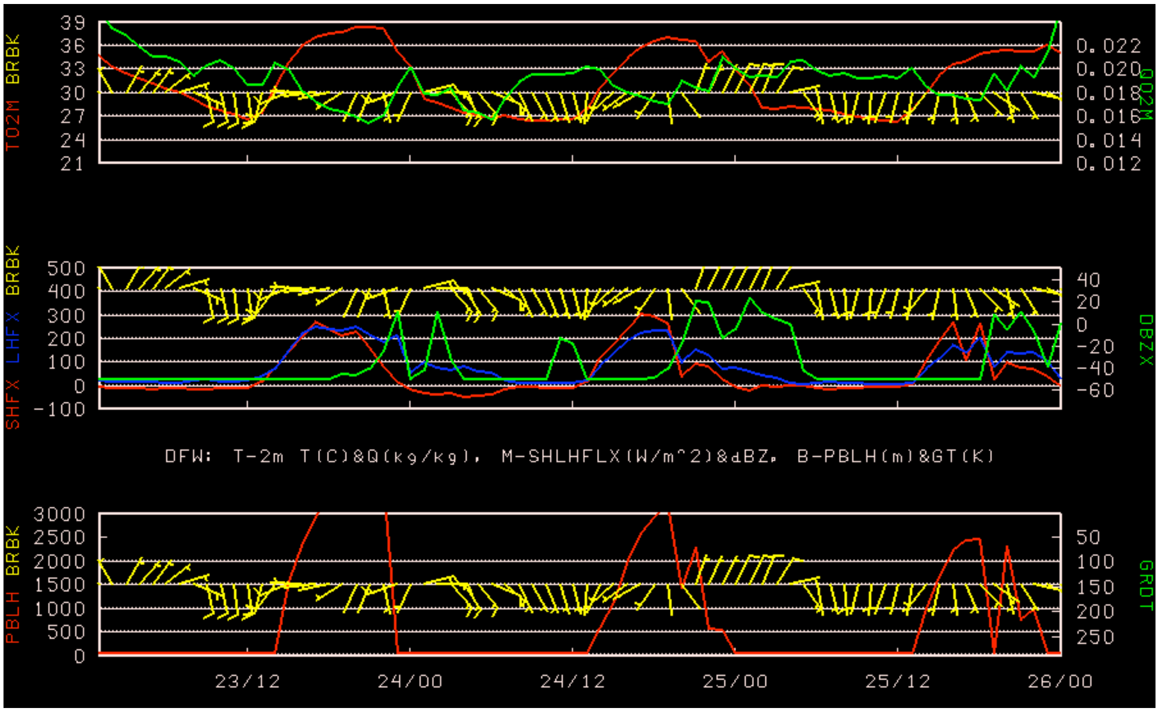
Forecast discussions were sometimes of interest to scientists not physically present at the briefing center. To enable such scientists to fully participate by telephone, links to the images used in the forecast briefings were posted prior to each briefing on yet another web page (http://mesonet.tamu.edu/TexAQS/trac/cgi-bin/trac.cgi/wiki/DiscussionImages). Such telephone-based forecast discussions were particularly valuable when considering McNider-type dead zone flights. The posted images were also useful to NETPS scientists for planning purposes.

The MM5 model forecast product output was expanded to accommodate the particular needs of the NETPS. One particularly valuable forecast product was the forecasted wind profiler plot. Wind profiles were generated on an hourly basis at each existing profiler site, and the forecasted images included the full three-day period. Also of value was the meteogram product, or time series of the forecasted weather evolution at key airports. Both products were valuable for determining the expected evolution of power plant plumes in the Dallas area on an hour-by-hour basis.

One additional support activity was conducted by us during the NETPS. Because Dr. McNider was unable to travel to Waxahachie conveniently, Dr. Nielsen-Gammon served as the airborne mission scientist for one dead zone flight. Because the frontal zone was forecasted to be close to the Red River, with westerly winds in Texas and easterly winds in Oklahoma, the flight took the form of a low-level northward traverse of the frontal zone followed by a southward higher-altitude traverse along the same track with embedded ascents and descents. High ozone was observed downwind from Dallas during both the outbound leg and the return leg.



An example of a time-height section of wind at a profiler site, generated from MM5 model output. The model was initialized at 0000 UTC August 23, 2005, and the initial conditions at Cleburne are plotted along the left margin as 23/00 with pressure (in mb) as the vertical coordinate. Successive hourly model forecasts are plotted toward the right, ending at 0000 UTC August 26, 2005 (26/00). The winds are plotted in conventional meteorological notation, similar to their appearance on a profiler plot, with a wind from the north indicated with a vertically-oriented barb. Wind speed (in m/sec) are contoured as well, with different colors indicating higher wind speeds. The tendency of strong, veering winds to form during the night is apparent on both August 24 and August 25.



An example of meteogram output from the MM5 numerical model. As in the previous figure, time runs from left to right. This particular meteogram is for Dallas/Fort Worth airport from August 23-26, 2005. Surface wind is plotted in all three graphs. The topmost graph includes temperature (in red) and water vapor mixing ratio (in green). The middle graph includes surface sensible heat flux (red), surface latent heat flux (blue), and simulated radar reflectivity (green). The bottom graph includes the height of the planetary boundary layer (red).

The forecasting support activities described herein were tailored to the needs of the field operations. Based on feedback from SETTS and NETPS project scientists, this forecasting support was entirely satisfactory. Similar activities would be planned in support of field intensive operations in 2006.

Observational Data (B1) and Data Assimilation Procedures (B2)

One of the important tasks under this project is the collection of observational data that can be used for real-time forecasting, data assimilation, and model verification.

While model verification can be done after-the-fact, the other uses of data require that the data be obtained and available in real-time or near-real time.

We have established five real-time data streams in support of the objectives of this project. These data streams are (1) the Family of Services (FOS) conventional data stream, (2) the MADIS (Meteorological Assimilation Data Ingest System) data stream, (3) the Texas Mesonet data stream, (4) the NIDS (NEXRAD Internet Dissemination System) Level-2 radar data stream, and (5) numerical model output from the National Meteorological Center (NMC) via ftp. Each of these data streams, and their uses for model verification, are described in turn.

The FOS conventional data stream was the primary source of meteorological observations during TexAQS-2000, but now it serves mainly as a backup data feed and source of processed model output. It includes conventional hourly surface and twice-daily upper air observations, profiler observations, satellite and radar observations, aircraft observations, National Weather Service forecasts and warnings, gridded numerical model output (generally at coarse resolution), and objective point forecasts from model output. Data in this form are the most common data utilized by meteorologists. Except for the gridded numerical model output, all the data is in ASCII format.

The MADIS data stream is provided by the National Oceanic and Atmospheric Administration's (NOAA) Earth System Research Laboratory's (ESRL) Global Systems Division (GSD), formerly known as the Forecast Systems Laboratory (FSL). (We find it more convenient to refer to NOAA/ESRL/GSD by its anagram, Lea Snodgras.) Ms. Snodgras's MADIS data stream (http://www-sdd.fsl.noaa.gov/MADIS/) is designed to support data assimilation and numerical weather prediction, and includes the observational data in the FOS as well as additional non-NOAA data sources. Of particular importance to our project is its inclusion of radar wind profiler observations from the NOAA demonstration network, TCEQ, NOAA/ETL (now known as NOAA/ESRL/PSD or Rope Sandals), and other data sources. The surface data in the MADIS data feed includes a number of nonstandard networks such as the West Texas Mesonet, PORTS, and RAWS, but does not include the TCEQ surface network.

The Texas Mesonet data stream is our term for a collection of individual data streams pulled independently into Texas A&M University as part of the Texas Mesonet project (http://mesonet.tamu.edu). This data stream presently does not include anything that is not already in the MADIS data stream, but negotiations are under way to obtain the TCEQ surface 5-minute observations as part of this data stream. Such requests have been made before, but it is our understanding from TCEQ that they recognize this to be a high priority item so we are hopeful of success this time.

The NIDS Level-2 radar data stream consists of radial velocity and reflectivity data from all WSR-88D radars operated by the National Weather Service and the Department of Defense. These radars constitute the conventional national network for monitoring severe weather and precipitation. The volume of data generated by each of these radars is immense. We are a second-tier distribution node for this data set, which means we receive all the Level-2 data in a timely manner. The data is archived locally.

The numerical model output is obtained via an automated FTP program from NMC. Gridded model output is available at several grid resolutions; we choose to receive the grid that spans the greatest area, so that the MM5 model's outer grid can include much of the source region for springtime agricultural burning in Central America. Model output is obtained from the 0000 UTC model run as well as the previous 1800 UTC model run as backup in case of model or communications failures at 0000 UTC.

The wide range of data type and quality necessitate a careful and measured utilization of the data. The various uses of the data are for forecast model initial conditions, forecast model boundary conditions, data assimilation and analysis, and model verification. The data being used or to be used for each of these purposes are now discussed.

The initial conditions for a numerical model consist of a gridded analysis or a gridded short-range forecast from a previous numerical model. For the operational model runs during Summer 2005, the NAM (North America Mesoscale) Eta model analysis for 0000 UTC was used as the initial condition for the MM5 forecasts. This type of model run is known as a "cold restart", because the model essentially starts cold with no information about its previous forecasts at the scale of the innermost model grid. Long-term plans call for the use of a “hot restart” once the data assimilation cycle has been implemented and has been shown to be of suitable accuracy.

The boundary conditions for a numerical model are derived from forecasts from a previous numerical model. For the operational model runs during Summer 2005, the NAM Eta model forecasts initialized at 0000 UTC were used as the boundary conditions for the MM5 forecasts. No plans exist to modify this procedure. Note that because our real-time MM5 is run with a large outer domain, the boundary conditions have comparatively little impact on the forecast.

Data assimilation in the MM5 numerical model can be conducted using analysis or observational nudging; we are also developing the capability to employ the Ensemble Kalman Filter (EnKF) for analysis. Observational data must be used with care in any data assimilation system, for fear of introducing nonphysical artifacts into the analysis or forecast that would have the effect of reducing its accuracy. We are therefore adopting a phased approach to data assimilation.

The most troublesome variables for data assimilation are thermodynamic variables (temperature and humidity) within the boundary layer. Assimilation of these variables can directly affect the stability characteristics of the atmosphere and adversely affect the mixing of pollutants. Furthermore, the temperature and humidity within the boundary layer are driven by surface fluxes, and assimilation that does not directly feed back on those surface fluxes will not have a lasting effect. Finally, temperature and humidity in the boundary layer are strongly affected by local surface characteristics and geographical aspects such as coastline proximity. If such data is assimilated, discrepancies in the model’s surface characteristics and coastline configuration will introduce biases in the temperature distribution on the scale of the assimilated data, which in general is much larger than the scale of the surface geography errors.

The first priority for data assimilation is wind profiler information. These data are obtained by Texas A&M from the MADIS data feed, and include all routine and special wind profilers in Texas and surrounding states. The data are available on an hourly basis, with a vertical resolution and vertical extent that varies from profiler to profiler. The data are generally of good quality, but are subject to known biases caused by migrating birds in spring and late summer/early fall. The New Braunfels and Sonora profilers may also be subject to biases caused by resident bats. To avoid this bias problem, data assimilation will initially occur only with data collected during daytime. Additional problems can occur when small-scale wind variations detected by the profiler during convective activity is assimilated over a larger area by the numerical model. At present, such data will be included, but future plans call for screening any wind assimilation when the observation increment (the difference between the observation and model forecast) depart substantially from the trend over several hours.

The wind data associated with the NEXRAD scanning precipitation radars will be the next priority for data assimilation. Initially, wind profiles will be constructed from the radial velocity information in the NIDS Level-2 data stream. These wind profiles will be similar in observation characteristics to the radar wind profiler data, except that they will be less subject to contamination from small-scale wind variations because they represent horizontal winds integrated over a larger area. Assimilation of these data will proceed in a similar manner as the wind profiler data, with assimilation again restricted to daylight hours.

The assimilation of both wind profiler and NEXRAD data can proceed using nudging on the 4 km domain, while tests are conducted with the EnKF. Following demonstrated reliability of the EnKF technique, a transition to operational EnKF data assimilation will take place. Because the EnKF requires at least 20 times the computer resources of nudging, the EnKF data assimilation will take place on the intermediate (12 km) grid.

Following the transition to EnKF, additional data sources will be incorporated into the assimilation cycle, pending demonstration of positive benefit. These data sources include surface wind observations, RASS virtual temperature observations from profilers, ACARS aircraft sounding and flight-level data, rawinsonde observations, and GPS precipitable water observations. The RASS and GPS data, in particular, are impossible to assimilate accurately using nudging but present no special difficulties within the EnKF. Offshore wind observations would likely be the first of this data set to be assimilated, since they present no siting issues and are somewhat similar in character to profiler observations. Except for the NIDS Level-2 data, which is received in a separate data stream, all data to be assimilated into the MM5 model are presently acquired through the MADIS data feed from Lea Snodgras.

Because of the data limitations discussed above, data assimilation will initially take place only during daylight hours, which for the purposes of this report will be taken to be the period 1200 UTC to 0000 UTC. The assimilation cycle will utilize data from this period, but to be incorporated into a hot restart, the assimilation cycle must be completed before the next model forecast is scheduled to begin. This imposes a deadline of 0230 UTC for completion of assimilation. We intend to allocate 8 hours of clock time for the assimilation cycle, implying a starting time of 1830 UTC for the assimilation. Therefore, the assimilation cycle would be able to use the 1200 UTC NAM Eta model run for nudging on the outer grid(s) while assimilation of observations takes place in the inner grid(s).

As noted above, the assimilation would take place on the 4 km grid with nudging and on the 12 km grid with the EnKF. The 12 km gridded output would be interpolated to the 4 km grid to provide the initial conditions for the innermost grid.

Some users of the gridded meteorological information may require 4 km resolution for the analyses. To satisfy that potential need, two possibilities exist: interpolation of gridded analyses down to 4 km resolution, or very short range 4 km simulations driven by the 12 km analyses. The best choice is not obvious, and will probably require working with users to determine which procedure gives the best results.

Unlike data assimilation, verification data does not suffer from limitations associated with data consistency and representativeness. As long as the comparison between observations and model forecasts are interpreted as point comparisons subject to instrument error and representativeness error as well as model error, all forms of data can be used for model verification. But because observation error represents two of the three sources of difference between observations and models, care must be taken to assess the quality of the observations.

The procedures for objective model verification were described in an earlier report. Here we describe the data sources for verification.

Because the primary purpose of the real-time numerical modeling is to provide guidance and information for photochemical model simulations and field measurements of atmospheric photochemistry, the verification data of interest are within the atmospheric planetary boundary layer, primarily within major metropolitan areas. The primary verification data set will be ASOS and AWOS surface observations from airport sites, operated by the National Weather Service and the Federal Aviation Administration. These observing stations use a common instrument suite and are sited according to exacting standards to provide good meteorological exposure. There are approximately 100 such sites across Texas, many of which are concentrated in major metropolitan areas. Thus, the use of these sites will automatically provide greater weighting of verification statistics in the areas of greatest interest.

The TCEQ surface network was the primary verification data set for model simulations of TexAQS-II. The network provides excellent coverage within the Houston and Dallas areas and includes meteorological observations that are coincident with air quality observations. However, the station locations were generally chosen for the purposes of air quality monitoring rather than meteorological monitoring, and previous analyses by us indicate that some stations produce biased or unreliable meteorological measurements. Because verification against the TCEQ data set is likely to be of considerable interest to other parties, we will perform model verification against this data set in a manner similar to that used for the ASOS/AWOS surface observations, but the verification statistics will be kept separately from the ASOS/AWOS verification statistics.

Surface verification only tells part of the story. For upper-air verification, we will utilize the same profiler and radar observations used for data assimilation, and we will also make extensive use of the rawinsonde observations. Because the rawinsonde observations are relatively sparse, they are of limited value in data assimilation, so they will provide an independent measure of the accuracy of both the model forecasts and the model data assimilation cycle. As with the surface observations, verification information from the different data sources will be maintained separately, because each data source has unique observation error characteristics.

At the present time, the objective model verification is behind schedule because the observation archive suffered a disk failure in mid-September and the RAID mirror disk suffered a failure during the recovery procedures for the first failure. As the data archive is repopulated, verification statistics will be computed on model runs over the previous summer. Based on present data recovery rates, the first objective verifications, using surface data, are presently planned to begin during the week of November 7.

The EnKF Technique (D1) and Additional Models and Ensembles (D3)

The research in these areas is summarized in two reports. The first report is an evaluation of the accuracy of the MM5 model in recreating a 2004 ozone episode in retrospective mode. The model was run using analysis nudging on the outer grids. The episode in question was identified in our conceptual model report as a possible episode of interest for future photochemical modeling work. The report includes a summary of the air pollution meteorology in Houston and Dallas during the episode, followed by a comparative assessment of the fidelity of the MM5 model in reproducing the episode.

This report was intended to be an initial step toward photochemical modeling of this episode, but even if this particular episode is not used, the report provides a summary of the sort of model behavior that can be expected in other ozone episodes. Without local data assimilation, the MM5 produces meteorologically reasonable simulations that often do not precisely agree from day to day but capture the essence of the variety of meteorological situations during the ozone episodes. The results from this study give support to the idea of using extended episode modeling to avoid the unattainable requirement of absolute fidelity to meteorological conditions on particular days.

The second report is a detailed summary of the EnKF/Ensemble Modeling research. Our efforts in this area have focused on understanding the nature of ensemble spread in summertime East Texas meteorological situations (an essential element of EnKF data assimilation) and the role of meteorological uncertainty and ensemble spread in the performance of photochemical model simulations of an ozone episode. For the purposes of this particular study, the well-examined August 30, 2000 Houston ozone day from TexAQS-2000 is utilized. The most interesting conclusion from this report is that the photochemical simulations are sufficiently sensitive to the meteorology that, even with perfectly accurate emissions and photochemical mechanisms, an accurate photochemical simulation is not possible given the uncertainties inherent in the meteorological observations.