Results of Chemical Data Assimilation in the BAMS/MCNC Numerical Air Quality Prediction System

Carlie J. Coats, Jr., John N McHenry,
Don Olerud, and Robert Imhoff
Baron Advanced Meteorological Systems, LLC,
Research Triangle Park, NC 27709-5064

1. Abstract

The traditional air quality model initialization approach, in which one “cycles” the model, using one forecast to initialize the next, suffers from potential model drift, in which errors in one forecast simulation (possibly due to errors in the meteorological, emissions, or air quality forecasts) propagate into the next. For the MCNC/BAMS Real-Time Ozone Forecast System, we have attempted to ameliorate these errors by assimilating chemical observational data from EPA's AIRNOW project. Several algorithm s have been suggested in the literature for this; however, we have incorporated a new algorithm that attempts to improve upon prior algorithms in the following ways:

1) It incorporates as much of the monitor data as possible (rather than looking only one hour’s data, or being overly influenced by the last data reported by a given monitor).
2) It accounts for the three-dimensional mixing in the daytime atmosphere, and has an appropriate three-dimensional influence.
3) It accounts for wind-driven transport from the times of the various monitor-observations up to the time of model-initialization.
4) It incorporates multiple adjacent observations “fairly.”
5) It does not attempt to extrapolate into large data-void areas.
6) It is inherently positive definite (does not create negative concentrations).

We have employed our new algorithm in our summer-season daily MM5/SMOKE/MAQSIP-RT air quality forecasts. Here we give a description of the algorithm and some caveats about its limitations.

2. The Algorithm

The algorithm we have developed for incorporating monitor data may be summarized as follows: For each daytime hour during the day prior to model initialization (over which the atmosphere up through the PBL height may reasonably be thought of as coupled to the surface), and for each "good" monitor observation during that period, construct a ratio event located at the monitor site that gives the ratio of monitor ozone to model ozone at that location at that time. Advect all the ratio-events forward in time to model-initialization, construct the weighted geometric mean of these ratios, and multiply the modeled ozone field up through the daytime PBL by it, using weights that "go away" as you move away from some "radius of influence" for these monitor events, and as you go vertically from the daytime minimum PBL height to its maximum.

Data is available in sufficiently timely form to be used in this fashion from the AIRNOW project of EPA's Office of Air Quality Planning and Standards. The AIRNOW data set has a coverage including most states east of the Rockies. During typical daytime hours, there are 800-950 monitors reporting to AIRNOW. Note that we exclude data sufficiently different from the modeled ozone; typically such outliers are the result of local conditions, especially depressed PBL heights not correctly modeled by the met model.

```
1.  Calculate the daytime-period
    maximum and minimum gridded
    PBL heights.

2.  For each hour of the daytime
    period and each valid monitor
    observation in that period for
    which the PBL is elevated ,
    construct a monitor-event as
    follows:
```

2.a. Interpolate the layer-1 air quality modeled ozone to the monitor site.

2.b. Compute the logarithm ***ln(monitor-ozone/model-ozone)***, excluding data for which this ratio is very large (> 2.0) or small (< 0.5)

2.c. Register the time and event-location in model coordinates.

3. Advect all monitor events forward in time to the model-initialization time, using Lagrangian advection.

4. Compute the PBL-depth/advection-period-averaged wind vector V "experienced" by the monitor-events during the process of advecting them forward to the model initialization time.

5. Compute the two-dimensional gridded log-ratio G(c,r) using modified inverse-square r2/(r2+dist2) weights that are "stretched" in the V-direction by the ratio R/(R+1800|V|) (The stretching corresponds to both ends of a one-hour period centered at the event's location.)

6. Use the same stretched weights to construct a piecewise-differentiable two dimensional gridded mask function M(c,r) that has values between 0 and 1, with zeros beyond the "stretched" radius of influence of all the advected monitor events, and 1's inside half the radius of influence of any monitor.

7. Compute the two-dimensional geometric-mean-ratio factor field F(c,r)=exp(M(c,r)G(c,r))

8. This factor-field is the weighted gridded geometric mean of the advected monitor event ratios near the locations of those events, and being identically 1 more than the radius of influence away from those advected monitor events.

9. Read in the previous forecast cycle's initialization time three-dimensional gridded ozone field.

10. For each column, row, and layer in the model grid, multiply the gridded ozone field by F(c,r)P, where P is a piecewise linear function of the vertical sigma coordinate, with value 1 up to the minimum PBL layer and value 0 one layer above the maximum PBL. This has the effect of "tapering out" the influence of the ratio-field F(c,r) from its full effect up to the minimum PBL to no effect above the maximum PBL.

11. Output the resulting 3-dimensional ozone concentration field for use in initializing the air quality model.

Plots of the "mask function" and layer-1 ozone ratio field for September 30, 2000, may be found in Figures 1 and 2.

3. Air Quality Modeling Configuration

The BAMS/MCNC numerical air quality prediction system is a 45/15/5KM-resolution multi-nested compound modeling system whose components include the MM5 meteorology model (as modified by US EPA and by BAMS, the SMOKE emissions modeling system, the MAQSIP-RT chemistry/transport model, and various analysis and visualization tools, all tied together into an operational real time forecast system running twice daily on the 0Z and 12Z forecast cycles. MM5 is modified by the inclusion of the MCPL output module [Coats et al 1998]to produce model-ready output meteorology files that contain all the additional derived variables needed for SMOKE and MAQSIP-RT. It is run in single-execution one-way-nested mode. SMOKE is a chemical emissions model that uses sparse-matrix technology to organize and optimize the "data processing" parts of emissions modeling, making it hundreds of times faster than conventional data processing approaches. SMOKE uses the meteorology data to compute the meteorology-driven parts of the emissions—biogenic emissions, mobile source evaporative emissions, and point-source plume rise—in real time. There are independent SMOKE executions for each air quality modeling domain. MAQSIP-RT is a descendant of the prototypes that also led to the EPA's CMAQ air quality model. It has been parallelized and highly optimized for microprocessor based systems, and contains enhanced cloud and photolysis treatments. MAQSIP-RT runs one grid at a time, with the outer nests providing time stepped chemical boundary conditions for the inner.

4. Operational Considerations

The operational schedule is constrained by the available hardware (a 24-processor SGI Origin), and by the 1:00 PM EST product-delivery times required by the state and regional air quality forecasters. On the 0Z cycle, the 45-km MM5 and MAQSIP-RT both initialize at 0Z (7:00 PM EST) and run for 60 hours; 15KM nests are initiated at 6Z and run for 30 hours; 5KM nests begin 8Z and run for 20 hours. At 12Z, MM5 and MAQSIP-RT initialize at 12Z; 15-km nests are initiated at 18Z, and 5-km nests at 20Z. Assimilation is only performed on the 0Z cycle 45-km runs, since that initialization only is shortly after the daytime period and the assimilation will not have been disrupted by such effects as stable night-time stratification, transport, and dry deposition scavenging. Other 45-km and 15-km nests are initialized by "cycling" data from the most recent forecast on that nest. The 5-km nests are initialized by interpolating from the enclosing 15-km nest. We chose 50 kilometers for the radius of influence, and 08:00-17:00 EST (13Z-23Z) for the daytime assimilation period. All monitors from thjis period meeting the PBL-height and ratio criteria will go into the calculation of the ozone correction factor field.

5. Application Episode

The ozone assimilation algorithm was evaluated over the period of July 15 through August 3, 2002. For the time series analysis in this paper, we chose to focus upon those 45-km grid cells most affected. Figures 3 has the time-series plot of the results for this grid cell of both a "base-case" model (brown) and the "data-assimilated" model (light blue) for the (29,44) cell in the Chicago area, together with observation statistics (max, min, mean) of all the monitors in this grid cell. This plots fits the typical pattern, with an initial difference on the order of 10-20 PPB, decaying to differences of less than two PPB within 24 hours. For only two forecasts and sub-regions was there a difference larger than 5 ppb extending past 24 hours.

In the talk, we will show several more such time series. Also, if PowerPoint behaves, we will display animated-GIF "movies" showing the difference between assimilated-model and base-case-model layer 1 gridded ozone for various forecast periods.

6. Conclusions

Whereas initially the assimilated ozone data have a moderately large effect—10-20%—we see that this effect "decays" over the next 12-24 hours so that in only one case was there a lingering effect of more than 5% 24 hours deep into the forecast. This is probably because we were able to assimilate ozone data only but not the rest of the chemistry. If one were able to make correlations between the ozone and the rest of the chemical state of the atmosphere (e.g., O3+NO2), it might be possible to infer corrections to be applied to the other chemical state variables. If possible, this would lead to a more chemically balanced assimilated atmosphere, and therefore better resulting forecast.

References

Coats, C. J, 1996: High Performance Algorithms in the Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System. J Ninth AMS Joint Conference on Applications of Air Pollution Meteorology with A&WMA, 584-588, 1996

Coats, C.J., J. N. McHenry, A. Lario-Gibbs, and C. D. Peters-Lidard, 1998: MCPL(): A drop-in MM5-V2 module suitable for coupling MM5 to parallel environmental models; with lessons learned for the design of the weather research and forecasting (WRF) model. The Eighth PSU/NCAR Mesoscale Model Users' Workshop. MMM Division, NCAR, Boulder, Co., 117-120.

Coats, C. J, A. Hanna, D.Hwang, and D.W. Byun, 1993: Model Engineering concepts for Air Quality Models in an Integrated Environmental Modeling System. Trans. AWMA Specialty Conf. on Regional Photochemical Measurement and Modeling Studies, San Diego, Nov. 1993.

Grell, G.A., J. Dudhia, and D.R. Stauffer, 1994: A description of the fifth-generation

Penn State/NCAR mesoscale model (MM5). NCAR Technical Note TN -398, Boulder CO.

J. N. McHenry, W.F. Ryan, N.L. Seaman, C. J. Coats, S. Arunachalam, and J. M. Vukovich, A Real-Time Eulerian Photochemical Model Forecast System and Initial Ozone Forecast Performance in the NE US Corridor, accepted for publication in the Bulletin of the American Meteorological Society.

McHenry, J. N., N.L. Seaman, C. J.˜ Coats, A Lario-Gibbs, J.Vukovich, N.Wheeler, and E.Hayes 1999a: Real-Time Nested Mesoscale Forecasts of Lower Tropospheric Ozone Using a Highly Optimized Coupled Model Numerical Prediction System, Preprints, AMS Symposium on Interdisciplinary Issues in Atmolspheric Chemistry, Dallas TX 1999.

MCNC Environmental Programs, 1996: The Multiscale Air Quality Simulation Platform. http://envpro.ncsc.org/products/maqsip/

MCNC Environmental Programs, 1998: Numerical Air Quality Prediction. http://envpro.ncsc.org/NAQP/

Xiu, A., R. Mathur, C. J. Coats 1999: On the Development and Testing of an Air Quality Modeling System with Integrated Meteorology, Chemistry, and Emissions. Preprints, AMS Symposium on Interdisciplinary Issues in Atmolspheric Chemistry, Dallas TX 1999

Figures

Figure 1: Mask Function for Sept. 30, 2000

Adjustment fully effective where mask=1, unused where mask=0

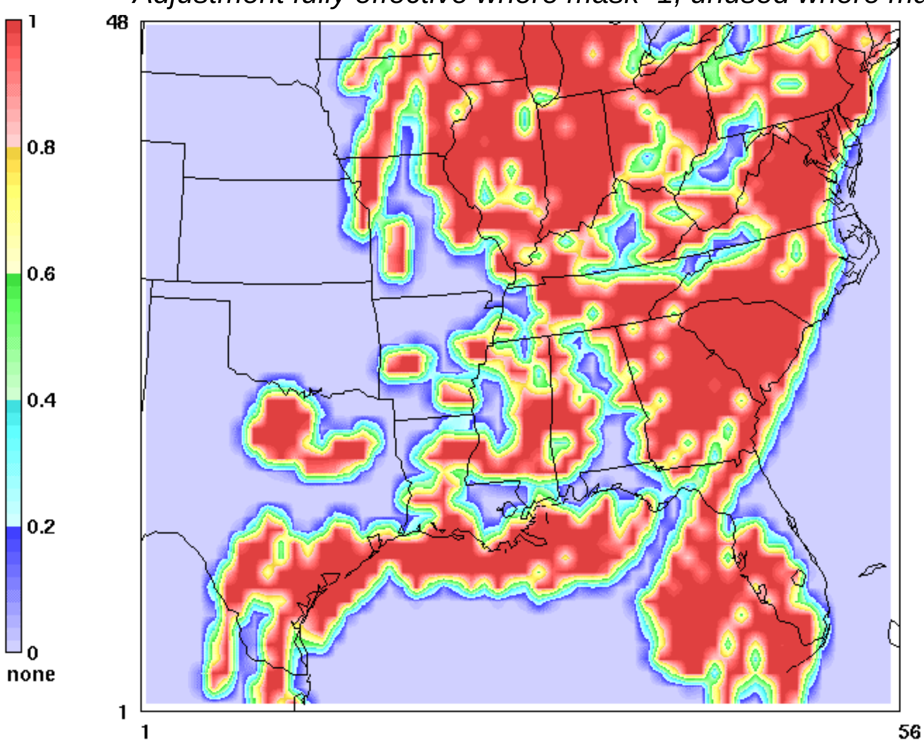

Figure 2: Ozone Adjustment Factor for Sept. 30, 2000

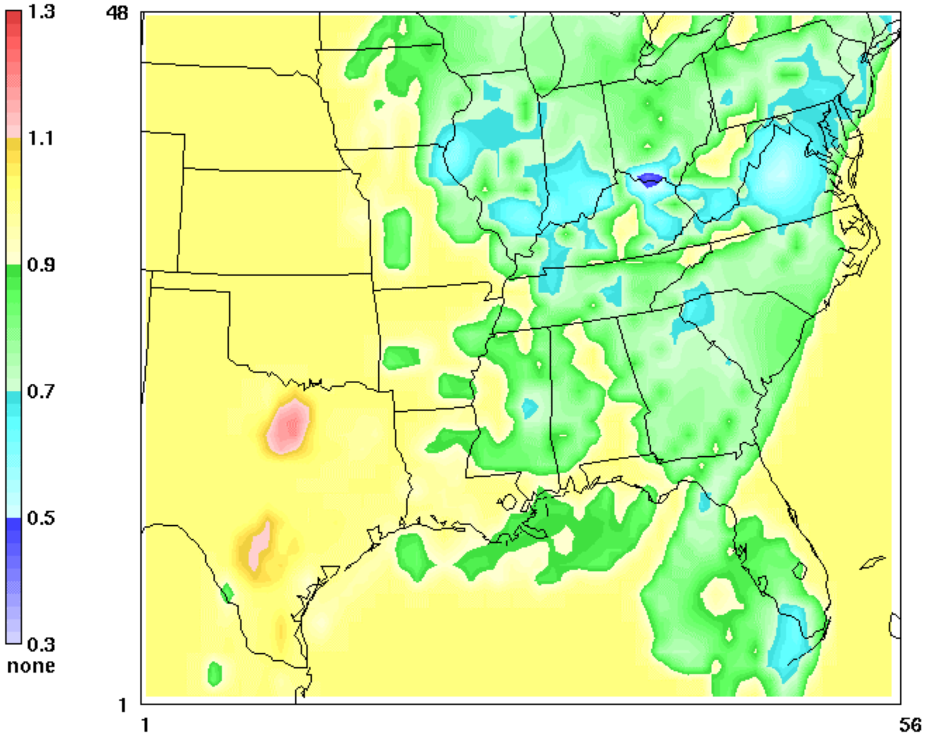

Figure 3: IC-adjusted Model vs. Base-case Model vs. Observation statistics: Time series for (29,44), July 28, 2002