3B.6

INVESTIGATION OF NUMERICAL ERROR SOURCES IN COUPLED MODEL PREDICTIONS OF ATMOSPHERIC TRANSPORT AND DISPERSION

Nelson L. Seaman*, David R. Stauffer, A. Deng
Penn State University, University Park, PA
and
Ian Sykes
L-3 Titan Corp., P.O. Box 2229, Princeton, NJ

1. INTRODUCTION

Numerical modeling is an essential tool to study, understand and predict behavior of complex environmental systems. Ideally, one might envision a single seamless model that includes all relevant components and scales of the earth-atmosphere-ocean system. A limited example is the "online" atmospheric chemistry system of Grell et al. (2005) that fully couples the RADM2 chemical mechanism within the WRF-ARW numerical weather prediction (NWP) model. Intended advantages of online chemistry include direct two-way coupling of the dynamics and chemistry solvers at each time step, elimination of interpolations necessary for models having different grids and assurance that the same atmospheric physics drives both chemistry and weather predictions.

Exchange of certain sub-model components in fully integrated online modeling systems, such as a physics scheme, should be fairly straightforward. However, a possible disadvantage of the online approach is that installing major components, such as a new ocean model, chemistry mechanism or plume solver, may be comparatively difficult. Furthermore, fully integrated multi-disciplinary models require either a single expert developer with extensive knowledge of each discipline or stable long-term collaborations among developers from different fields. While possible, reliance on such rare qualifications would be likely to impede overall inter-disciplinary model development needed to advance environmental sciences.

An alternative approach is offered by the Earth System Modeling Framework (ESMF) now under development (DeLuca, 2006). ESMF will allow disparate "stovepipe" models of the atmosphere, oceans, chemistry, biosphere, etc., to be more easily and efficiently coupled with minimal modification to their original codes through the use of standardized interfaces, drivers and couplers. However, just because new software may allow easier standardized coupling of existing models does not mean such coupling will have no impact on model skill. In fact, horizontal and vertical interpolations of fields passed between different model grids, use of different parameterizations for essentially similar processes in the coupled models and temporal interpolations of outputs exchanged between models all represent potential sources of error.

At present ESMF lacks capability to support grid nesting and several other tools often used in mesoscale model applications. In the interim to better understand how a coupling framework such as ESMF might be best exploited in future interdisciplinary environmental modeling, this paper reports on efforts to explore errors associated with a fairly simple coupling of two atmospheric models designed for separate applications – mesoscale NWP and plume dispersion. The study has direct applicability to such problems as rapid response to disasters involving hazardous or toxic airborne releases, either accidental or intentional.

2. MODEL DESCRIPTION

The coupled system evaluated in this study consists of the mesoscale meteorological model MM5v3.6 (Grell et al. 1995) and the SCIPUFF plume model (Sykes et al. 1996). MM5 predicts 3-D atmospheric fields (wind, temperature, mixing ratio, perturbation pressure, and cloud water/ice), plus 2-D fields such as precipitation. These fields are periodically saved and provided to SCIPUFF via separate MEDOC interface software. In essence MEDOC handles the one-way model interface role, but without the generalizations under development for ESMF. SCIPUFF uses second-order turbulence equations to calculate dispersion rates for a series of Lagrangian puffs emitted from a source and advected downwind by the MM5-supplied wind field. Plumes can be simulated on scales from meters to thousands of km. Mixing rates are dependent on the shear and stability predicted by the mesoscale model, plus internal parameters determined by SCIPUFF.

For the present study SCIPUFFvG:5.0-T:5.0.001-S:2.303 (released April 2007) has been used. This updated version of SCIPUFF still supports the model's standard grid (sigma-z vertical coordinate and non-

Corresponding author address: Nelson Seaman, Penn State Univ., Dept. of Meteor., University Park, PA, 16802; email: seaman@ems.psu.edu

staggered horizontal grid) and constant surface roughness. It also contains new options to provide tighter coupling with the MM5 grid and its land-surface physics. Thus, the latest SCIPUFF supports MM5's sigma-p coordinate, Arakawa B-grid staggering and variable surface roughness. When run at identical resolutions and invoking the MM5 grid options, SCIPUFF requires no horizontal or vertical interpolations of meteorological datasets passed between the mesoscale and plume-scale models.

3. EXPERIMENT DESIGN

For this paper the coupled MM5-SCIPUFF system (Deng et al. 2004), updated as in Section 2, was used to study a 24-h period of the September 18-19, 1983, case from the Cross Appalachian Tracer Experiment (CAPTEX). In this case, 208 kg of inert tracer (C_7F_{14}) were released from Dayton, OH, in a 3-h period during the late morning of 18 September in the southwesterly flow to the north of a large subtropical anticyclone. Figure 1 shows the release site, a cyclonic frontal system crossing the lower Great Lakes and observed concentrations logged at CAPTEX surface monitors at 6-h intervals over the succeeding 24 h relative to the end of the tracer-release period (1700 UTC).

MM5 was configured with 32 vertical layers and 4 nested grids of 108-, 36-, 12-, and 4-km. SCIPUFF was configured with 4-km horizontal resolution, 32 layers, and no active chemistry (passive tracer only). This allows more direct assessment of error sources associated with model coupling. The lowest layer was ~30 m in depth, with 16 layers below 850 hPa. Both models were configured with the Lambert conformal map projection. MM5 was initialized at 1200 UTC, September 18 and run for 36 h. Some SCIPUFF runs used MM5 meteorology created with multi-scale nudging FDDA (Stauffer and Seaman 1994) to reduce growth of regional-scale meteorological errors during the simulation, while others used MM5 with no FDDA (Table 1). MM5 output was provided to SCIPUFF at intervals of 10, 60 or 180 minutes. Table 1 also shows in which experiments SCIPUFF used either the standard SCIPUFF grid or MM5's grid and constant surface roughness length versus variable roughness.

4. MODELING RESULTS

4.1 *Maximum Surface Concentrations*

Examination of model results begins with Table 2 which compares maximum observed concentrations to the model-simulated concentrations, anywhere on the grid, every 6 h during the experiments. It is evident immediately that MM5-SCIPUFF predicts maxima that exceed the observed values by ratios as great as 17:1.

Table 1. Summary of the MM5-SCIPUFF experiment design.

Exp No.	MM5 freq. (min)	MM5 FDDA	MM5 time - avg'd met.	SCIP. grid	SCIP. Rough.
1	60	No	No	Std	Z_o const
2	60	No	No	Std	$Z_o(x,y,t)$
3	60	Yes	No	Std	$Z_o(x,y,t)$
4	180	Yes	No	Std	$Z_o(x,y,t)$
5	10	Yes	No	Std	$Z_o(x,y,t)$
6	60	Yes	Yes	Std	$Z_o(x,y,t)$
7	60	Yes	Yes	MM5	$Z_o(x,y,t)$

There are several factors that contribute significantly to this over-prediction. First the spacing of the CAPTEX monitors greatly exceeds the 4-km model grid spacing, so it is unlikely that the true maximum surface concentrations coincide with any of the monitors. This is evidenced by the extreme tracer gradient observed between the closely spaced monitors in northern OH (spacing ~32 km) at +5h after the end of release period (Fig. 1a), when the central portion of the plume is concentrated in central OH. The effect of this spatial sampling error is also likely to be large near +11 h when the plume is passing over Lakes Erie and Ontario which imposes large gaps on the monitor network. Second, the observed values represent averages over 20-min. sampling intervals during each 6-h period, while the model output shows instantaneous concentrations at the end of those intervals. This sampling difference could easily lead to underestimates of the actual peak concentration of the surface plume. Third, Deng et al. (2004) showed that for most of the study period, the highest predicted concentrations occur aloft. Limitations in the models' parameter-ization of stable nocturnal conditions are likely to cause excessive mixing from this elevated plume to the surface following sunset (about +7h). This would be

Table 2. Observed (bold) maximum surface tracer concentrations (fl l^{-1}) at 6-h intervals during the 18-19 September 1983 CAPTEX case versus simulated maxima anywhere on the 4-km MM5-SCIPUFF domain.

Exp. No.	+5 h 2200 Z	+11 h 0400 Z	+17 h 1000 Z	+23 h 1600 Z
OBS'D.	***1608***	***121***	***89***	***47***
1	2225	1176	450	182
2	2406	1689	513	349
3	2648	1638	322	163
4	2815	1354	447	163
5	2791	2061	311	95
6	2744	1474	398	135
7	2209	1519	371	119

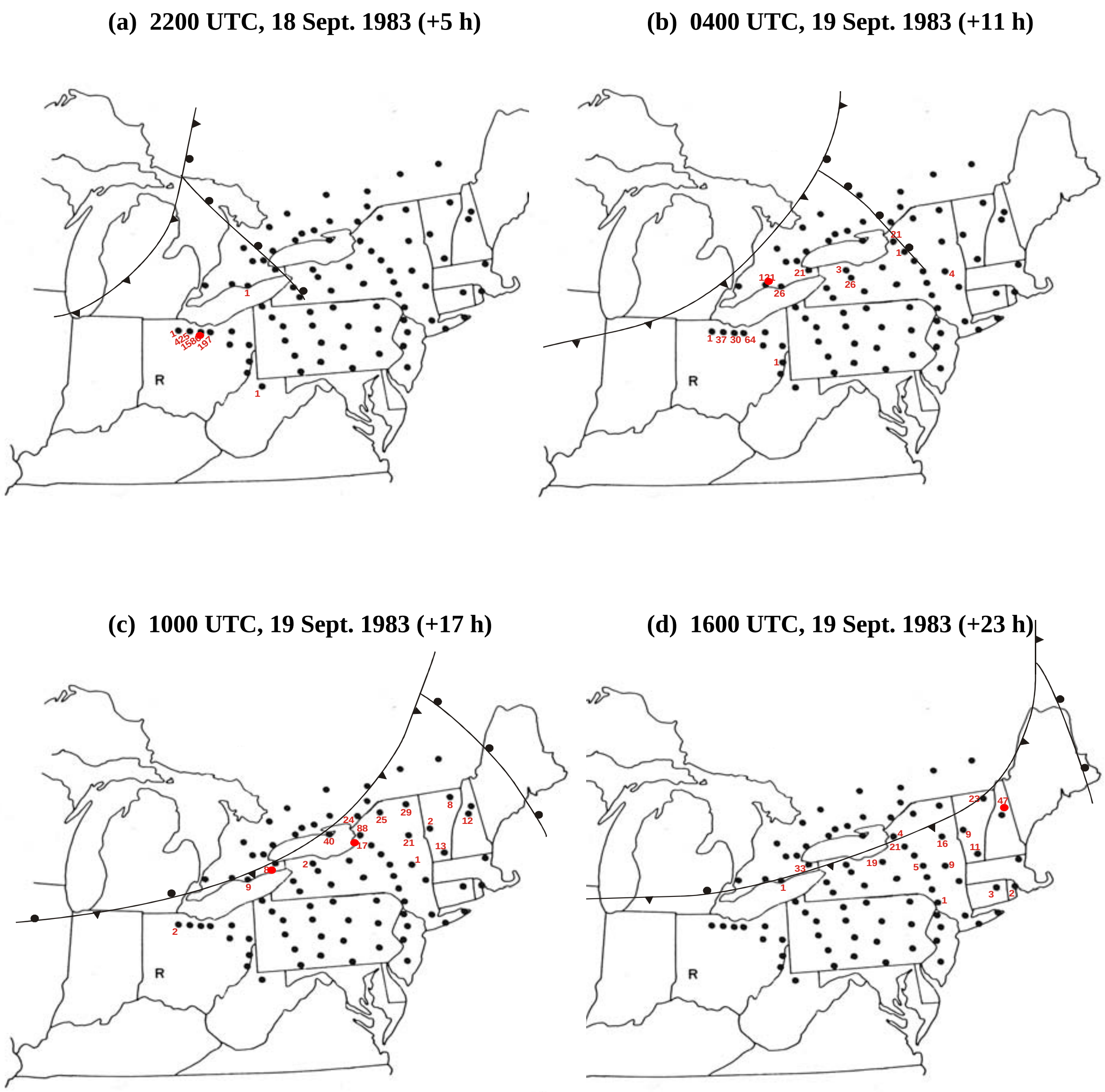

Figure 1. CAPTEX tracer network sites (black dots) and observed surface tracer concentrations (red, fl l^{-1}) at (a) 5 h, (b) 11 h, (c) 17 h, and (d) 23 h following release of inert tracer gas from Dayton (shown as "R"). Frontal positions are superimposed. Times (h) shown in parentheses are relative to end of the tracer release period (1700 UTC, 18 Sept., which is 5 hours after the beginning of the MM5 simulations).

consistent with the comparatively large ratios of obs-to-simulated maximum concentrations found at night (+11 and +17 h), but only modest ratios during the day when unstable mixing in a deep boundary layer is expected (+5 and +23 h). Thus, even without addressing issues related to coupling, errors in tracer sampling and inaccuracies in meteorological physics are likely to cause over-predictions versus the data.

To gain some further insight into the role played by the inherently incomplete sampling associated with any practical monitoring network, Table 3 shows maximum concentrations predicted in each experiment *at the sites of the CAPTEX monitors*. Table 3 reveals much lower maxima, compared to Table 2, due to failure of the monitors to coincide with actual locations of the SCIPUFF-predicted maxima on the 4-km grid mesh. For example, at +5 h the model-predicted maxima lie southwest of the first arc of monitors, resulting in large under-estimates of the plume maximum, even though Table 2 showed an over-prediction. Based on observed surface and boundary layer winds at this time (not shown), it is very likely that the true plume maximum is also to the southwest, so the plume is just starting to impact the monitors. However, by +11 h and +17 h, the plume is crossing the heart of the network and Table 3 shows the predicted maxima at the monitors consistently underestimate the predicted grid maxima by up to 300 percent. At the final time, +23 h, the plume is beginning to leave the monitoring domain and the ratio of "actual to monitored" maxima predicted by the modeling system again rises dramatically.

Table 3. Observed maximum surface tracer concentrations (bold, fl l^{-1}) at 6-h intervals during the 18-19 September 1983 CAPTEX case versus simulated maxima simulated by MM5-SCIPUFF *at any monitoring site*. Ratios in parenthesis compare actual predicted maxima (Table 2) to the maxima simulated at the CAPTEX monitors.

Exp. No.	+5 h 2200 Z	+11 h 0400 Z	+17 h 1000 Z	+23 h 1600 Z
OBS'D.	***1608***	***121***	***89***	***47***
1	282 (7.9 : 1)	647 (1.8 : 1)	225 (2.0 : 1)	20 (9.1 : 1)
2	380 (6.3 : 1)	881 (1.9 : 1)	379 (1.4 : 1)	12 (29.1 : 1)
3	92 (29.3 : 1)	637 (2.6 : 1)	288 (1.1 : 1)	40 (4.1 : 1)
4	163 (17.3 : 1)	817 (1.7 : 1)	263 (1.7 : 1)	27 (6.0 : 1)
5	215 (13.0 : 1)	673 (3.1 : 1)	278 (1.1 : 1)	73 (1.3 : 1)
6	244 (11.2 : 1)	475 (3.1 : 1)	294 (1.4 : 1)	30 (4.5 : 1)
7	133 (16.6 : 1)	501 (3.0 : 1)	208 (1.8 : 1)	48 (2.5 : 1)

Clearly, we are most interested in the actual maxima predicted by the coupled model, since these represent the highest risk when simulating hazardous plumes. Nevertheless, this exercise confirms that even well-designed monitoring networks are likely to under-estimate actual maximum plume concentrations by a substantial amount, significantly complicating the use of such data for model evaluations. Thus, if a meteorological-dispersion system were "tuned" to match observed maximum concentrations for a case, its predictions quite probably would be misleading.

4.2 *Plume Advection and Spread*

Along with prediction of maximum surface concentrations in a simulated plume, perhaps the most important other characteristics to be evaluated are the plume's advection by the mesoscale wind and its spread due to turbulence. The resulting spread of the surface plume and its progress downwind are referred to collectively as its "footprints," defined here as the area having the minimum observable concentration (1 fl l^{-1}). In this section we simultaneously evaluate maximum concentrations and the accuracy of plume "footprint" predictions for each experiment using standard statistical measures of accuracy (Table 4).

The first two experiments listed in Table 1 examine the impact of using constant surface roughness length in SCIPUFF (Exp. 1) versus variable roughness passed from MM5 (Exp. 2). Constant roughness (Z_o=0.50 m) in Exp. 1 reflects the dominant forests of the Northeast U.S., but it ignores extensive agriculture, water, and urban land types (0.15, 0.0001, 0.50 m, respectively). Table 2 reveals greater predicted surface concentrations (8 - 92% higher) with variable surface roughness in Exp. 2. Accounting for reduced roughness over

Table 4. Time-averaged statistical evaluation of MM5-SCIPUFF sfc. tracer concentrations: Threat Score (TS), Bias Score (BS), Probability of Detection (POD), False Alarm Rate (FAR), Hanssen-Kuiper score (HS); perfect score shown in parentheses. The experiment having the best score for each statistic is shown in bold.

Exp No.	TS (1.0)	BS (1.0)	POD (1.0)	FAR (0.0)	HS (1.0)
1	**0.407**	0.694	**0.490**	**0.294**	**0.427**
2	0.355	**0.714**	0.449	0.371	0.367
3	0.339	0.612	0.408	0.333	0.345
4	0.339	0.612	0.408	0.333	0.345
5	0.356	0.633	0.429	0.323	0.365
6	0.328	0.571	0.388	0.321	0.331
7	0.344	0.674	0.429	0.364	0.353

water and agricultural land leads to less vertical and horizontal mixing near the surface in the model's Lagrangian puffs, so plume dilution by clean air proceeds more slowly. Figure 2 compares surface concentration patterns in Exps. 1 and 2 at +11 h as the plume crosses the lakes. Sparse monitoring near the lakes very likely contributes to the large over-prediction of concentrations at this time. However, the statistics in Table 4 reveal that Exp. 1 produces the overall best prediction of the observed footprint, with Exp. 2 having the next-best statistical skill. This result demonstrates the sensitivity of predictions to the treatment of physical parameters affecting mixing. However, it also shows that efforts to correct one set of problems in a complex modeling system may allow other formerly undetected problems to have greater impact, resulting in worse apparent skill. This type of result is widely known as the "compensating error" problem and can be quite difficult to overcome.

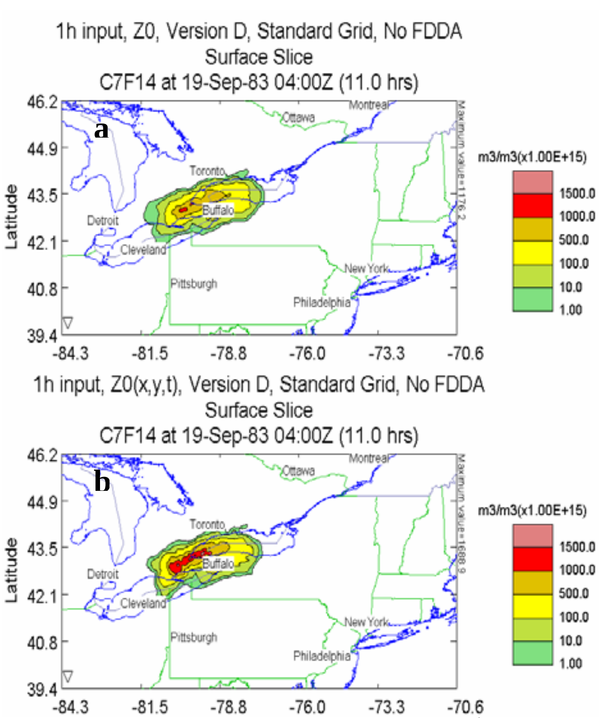

Figure 2. Surface tracer concentrations (fl l^{-1}) predicted by MM5-SCIPUFF for Sept. 19, 1983, CAPTEX case at 0400 UTC, +11 h following the tracer release. (a) Exp. 1, (b) Exp. 2.

Next, the role of FDDA to reduce the growth of meteorological errors over time is shown by comparing peak concentrations for Exps. 2 and 3 (Table 2), surface footprint statistics (Table 4) and by examining the spatial plume pattern at +17 h (Fig. 3). Prior work by Deng et al. (2004) showed that for this CAPTEX case surface-layer RMS wind speed errors were reduced from 2.69 ms^{-1} without FDDA to 2.13 ms^{-1} with FDDA. Meanwhile, FDDA reduced mean absolute errors for boundary-layer wind direction from 18 degrees to 13 degrees. Consequently, all subsequent experiments here are based on MM5 with FDDA. Nevertheless, Table 4 indicates that by most statistical measures, the surface footprint in Exp. 3 was less accurate despite clearly improved winds.

Figure 3 indicates that the meteorology with FDDA produces a narrower plume at +17 h, especially along its axis in the St. Lawrence Valley. The monitors reporting positive tracer amounts at this time (Fig. 1c) imply the narrower plume in Exp. 3 is more realistic. Despite the smaller plume footprint in Exp. 3, Table 2 shows that peak concentrations are lower (closer to the peak observed values) at all times after +11h. This suggests more of the plume's tracer material remains elevated through the period.

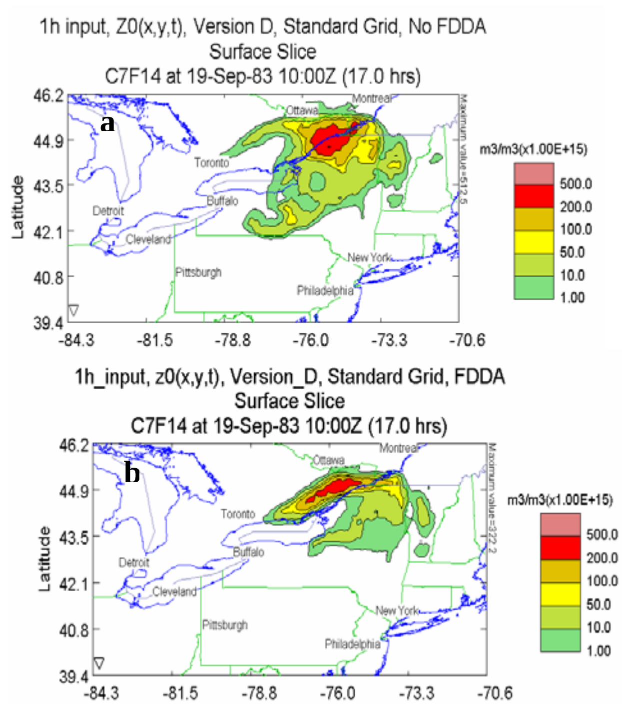

Figure 3. Surface tracer concentrations (fl l^{-1}) predicted by MM5-SCIPUFF for Sept. 19, 1983, CAPTEX case at 1000 UTC, +17 h following the tracer release. (a) Exp. 2, (b) Exp. 3.

The effect of temporal interpolations in the model coupling was studied by providing MM5 inputs to SCIPUFF at 10, 60 and 180-min intervals (Exps. 5, 3, and 4, respectively). At +5 h, peak tracer concentrations vary by only ~5% due to changes in the time sampling (Table 2), but the variation grows rapidly to 44 - 72% during the +11 h - +23 h period. Moreover, the effect can be erratic. At +11 h, temporal sampling of the meteorology at 10-min. intervals (Exp. 5) leads to the most extreme over-prediction of concentrations, but by +23 h the same experiment has the least error of all. Also, the surface footprint at +23 h in Exp. 4 (180-min. sampling, Fig. 4a) exhibits a nearly straight-line plume core, as might be expected with poor temporal resolution, while in Exp. 5 (10-min. sampling, Fig. 4b) the plume core has lower concentrations and shows evidence of meandering in a more-resolved wind field. Meanwhile, statistical scores for the plume footprint show some improvement due to the high-frequency meteorological inputs in Exp. 5, as might be expected.

In Exp. 6 the MM5 is run exactly as in Exp. 3, except that the 60-min. meteorological outputs are created by time averaging the solutions over all time steps from -15 min. to +15 min. about the output times. This is intended to damp influences of passing internal gravity waves, convective updrafts and other transient features that can produce unrepresentative features in instantaneous model fields. In principle, one could argue that the effect of such time averaging would be fairly similar to using more frequent outputs as in Exp. 5. However, Table 2 shows that the impact of different meteorology for Exps. 5 and 6 still results in changes to peak surface concentrations, especially at +11 h. However, Table 4 shows the statistics for the plume footprint appear to be degraded somewhat in Exp. 6.

Finally, Exp. 7 combines the time-averaged 60-min. meteorological outputs (similar to Exp. 6) with the use in SCIPUFF of MM5's horizontal and vertical grid system, eliminating intermediate spatial interpolations. Table 2 indicates Exp. 7 produces the smallest surface concentration at +5 h of all tested configurations, exceeding the peak observation by just 1.4:1. As expected, the over-prediction at +11 h remains severe, most likely due to problems with the sampling network in the vicinity of the lower Great Lakes. However, the predicted maximum concentrations at +17 h and +23 h (also Fig. 5) are among the lowest produced in these experiments, with the ratio of the final simulated to observed concentration at ~2.5:1. Comparison of Figs. 5a and 3b, and Figs. 5b and 4b, reveal generally similar footprint characteristics for Exps. 3, 5 and 7. Meanwhile, Table 4 shows that the improved coupling of MM5 and SCIPUFF grids had little impact on footprint statistics. Thus, while tighter temporal and spatial coupling between models may be somewhat effective in reducing maximum plume concentration errors, many unanswered questions remain about the requirements for accurate mesoscale plume calculations, especially relating to plume (footprint) spread. We hypothesize that variance in the large-scale wind field (above the turbulence scale) may not be represented adequately in SCIPUFF, since deterministic meteorological predictions provide no information on case-dependent wind variance. In this scenario, tighter spatial and temporal coupling actually may degrade results by eliminating false dispersion sources.

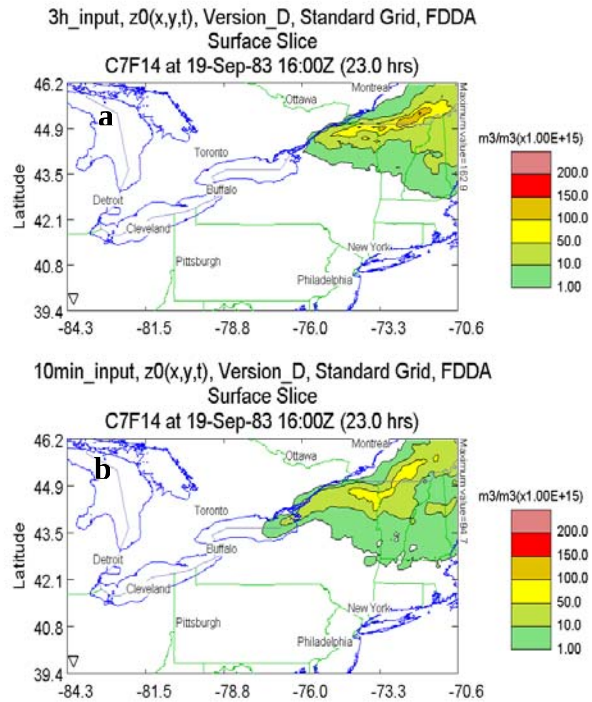

Figure 4. Surface tracer concentrations (fl l^{-1}) predicted by MM5-SCIPUFF for Sept. 19, 1983, CAPTEX case at 1600 UTC, +23 h following the tracer release. (a) Exp. 4, (b) Exp. 5.

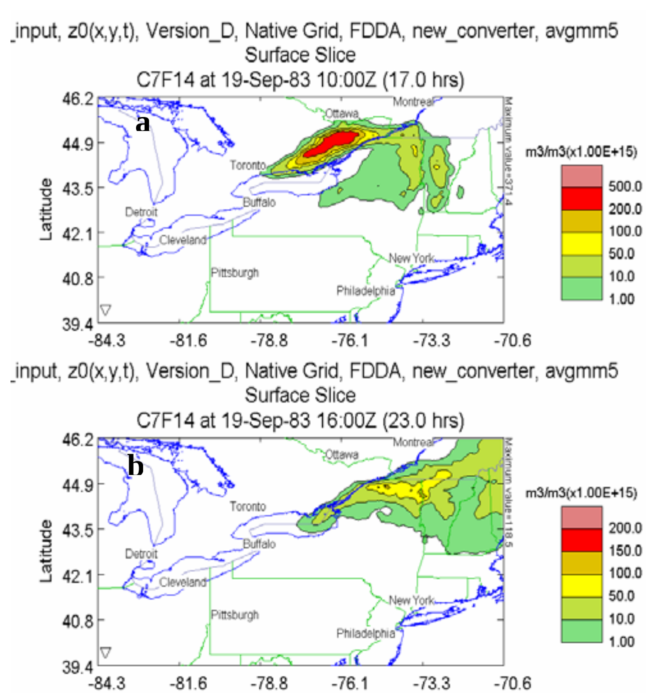

Figure 5. Surface tracer concentrations (fl l^{-1}) predicted by MM5-SCIPUFF for Sept. 19, 1983, CAPTEX case in Exp. 7. (a) 1000 UTC, +17 h following tracer release, (b) 1600 UTC, +23 h following tracer release.

5. CONCLUSIONS

The present study on numerical and physical coupling of models was conducted using the MM5-SCIPUFF system applied to the 18-19 Sept. 1983 CAPTEX case. The study shows that predicted maximum plume concentrations over a 24-h period can be considerably greater than observed due to gaps in the monitoring network, inaccuracies in coupling methodologies, and the treatment of physical processes affecting vertical mixing. Simulations of peak surface tracer concentrations and plume footprint spread can vary rather erratically depending on modifications to the roughness length and time-space interpolations. A coupling framework such as ESMF will allow component models to exchange fields every time step, eliminating effects due to poor temporal sampling. However, this study indicates that even use of data-assimilated meteorology and identical horizontal and vertical grids may be ineffective for reducing errors in standard statistics used to evaluate plume accuracy. A possible source of error unaccounted for in these experiments is the absence of case-dependent large-scale variance in the wind fields. This is being addressed in a related study conducted by a team of scientists at Penn State (Kolczynski et al. 2007).

6. ACKNOWLEDGEMENTS

We acknowledge the assistance of Brian Reen in running some of the MM5 experiments and Jeff Zielonka in the calculation of statistical scores. This work was sponsored by DTRA through contract DTRA01-03-D-0013 with L3-Titan.

7. REFERENCES

DeLuca, C., 2006: ESMF development status and new directions. ESMF 5th Community Meeting, 23 May, Baltimore, MD. http://www.esmf.ucar.edu/presentations/

Deng, A., N.L. Seaman, G.K. Hunter and D.R. Stauffer, 2004: Evaluation of interregional transport using the MM5-SCIPUFF system. *J. Appl. Meteor.*, **43**, 1864-1886.

Grell, G., J. Dudhia and D.R. Stauffer, 1995: A description of the fifth-generation Penn State/ NCAR mesoscale model (MM5). NCAR Tech. Note NCAR/TN-398+STR, 122 pp.

Grell, G., S.E. Peckham, R. Schmitz, S.A. McKeen, G. Frost, W.C. Skamarock and B. Eder 2005: Fully coupled "online" chemistry within the WRF model. *Atmos. Environ.*, **39**, 6957-6975.

Kolczynski, W.C., Jr., D.R. Stauffer, and S.E. Haupt, 2007: A simple method for calibrating ensemble variability to represent meteorological model uncertainty. Preprints, AMS 22nd Conf. on Wea. Anal. and Fcstng./18th Conf. on Num. Wea. Pred., Park City, UT, 25-29 June, 15 pp.

Stauffer, D.R. and N.L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.*, **33**, 416-434.

Sykes, R.I., D.S. Henn, S.F. Parker, and R.S. Gabruk, 1996: SCIPUFF – A generalized hazard dispersion model. Preprints, 9th Joint Conf. on Appl. of Air Poll. Meteor., Atlanta, GA, AMS, 184-188.