

## The Operational Mesogamma-Scale Analysis and Forecast System of the U.S. Army Test and Evaluation Command. Part III: Forecasting with Secondary-Applications Models

ROBERT D. SHARMAN,\* YUBAO LIU,\* RONG-SHYANG SHEU,\* THOMAS T. WARNER,\*<sup>+</sup> DARAN L. RIFE,\*  
JAMES F. BOWERS,<sup>#</sup> CHARLES A. CLOUGH,<sup>@</sup> AND EDWARD E. ELLISON<sup>&</sup>

*\*National Center for Atmospheric Research,\*\* Boulder, Colorado*

*+ Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, Colorado*

*#U.S. Army, Dugway Proving Ground, Dugway, Utah*

*@U.S. Army, Aberdeen Proving Ground, Aberdeen, Maryland*

*& U.S. Army, White Sands Missile Range, White Sands, New Mexico*

(Manuscript received 27 November 2006, in final form 17 October 2007)

### ABSTRACT

Output from the Army Test and Evaluation Command's Four-Dimensional Weather System's mesoscale model is used to drive secondary-applications models to produce forecasts of quantities of importance for daily decision making at U.S. Army test ranges. Examples of three specific applications—a sound propagation model, a missile trajectory model, and a transport and diffusion model—are given, along with accuracy assessments using cases in which observational data are available for verification. Ensembles of application model forecasts are used to derive probabilities of exceedance of quantities that can be used to help range test directors to make test go–no-go decisions. The ensembles can be based on multiple meteorological forecast runs or on spatial ensembles derived from different soundings extracted from a single meteorological forecast. In most cases, the accuracies of the secondary-application forecasts are sufficient to meet operational needs at the test ranges.

### 1. Introduction

Part I of this series of papers (Liu et al. 2008a) provides an overview of an operational mesogamma-scale forecast model, called the Real-Time Four-Dimensional Data Assimilation (RTFD DA) system, that is in use at the U.S. Army Test and Evaluation Command (ATEC) test ranges. The forecast component of RTFD DA is based on the fifth-generation Pennsylvania State University–National Center for Atmospheric Research Mesoscale Model (MM5; Grell et al. 1995). The forecast model, which is multiscale with nested domains, provides very high horizontal resolution (1.1–3.3-km grid increment, depending on the test range) over the test areas. To briefly review, the heart of the RTFD DA system is a data assimilation engine that continuously ingests meteorological observations as they

become available, producing model-assimilated datasets that both define the current conditions on the ranges and serve as initial conditions for the model forecast component. The Newtonian relaxation approach to data assimilation uses nonphysical nudging terms in the model predictive equations. These terms force the model solution at each grid point to observations, in proportion to the difference between the model solution and the observation (Stauffer and Seaman 1994). Each observation is ingested into the model at its observed time and location, with proper space and time weights, and the model spreads the information in time and space according to the model dynamics. The data utilized by the assimilation system include the standard hourly surface reports and twice-daily radiosondes; data from various special mesoscale networks; wind-profiler data; hourly cloud-track winds derived from infrared, visible, and water-vapor imagery; aircraft reports; and data from various observation platforms at the test ranges, including surface meteorological stations, boundary layer profilers, and rawinsondes.

Given the needs of the ranges for specialized forecast products, the RTFD DA system output is used to drive secondary-applications models to support test operations on a daily basis. This third in our series of papers

\*\* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Robert Sharman, NCAR/RAL, P.O. Box 3000, Boulder, CO 80307-3000.  
E-mail: sharman@ucar.edu

discusses three examples of the use of such secondary-applications models:

- 1) sound propagation models for range neighborhood annoyance mitigation,
- 2) ballistic trajectory and missile-debris-drift models, to ensure that range boundaries are not violated, and
- 3) transport and diffusion models for smoke, obscurants, and simulant releases.

In most cases, these application models are used to estimate the impact of the atmosphere on operationally relevant physical processes. If that prediction shows that an unsatisfactory test outcome or impact is probable, the test may be cancelled. Traditionally, these application models have used, as input, meteorological surface observations and soundings taken near the time and location of a scheduled test. Thus, the output of the application model provided a “nowcast” of the quantities. In some cases, this nowcast actually is fairly old. For example, the decision about whether to conduct an afternoon test was often based on the results of the application model that used an early morning sounding. In contrast, the use of the RTFDFA system at the ATEC ranges introduces the capability to actually *forecast* expected test conditions at the exact time and location of a test.

Application model codes that were developed for single-station, rawinsonde data input can also be used with a single vertical sounding extracted from the mesoscale model at a particular horizontal grid point. We will refer to this derived sounding as a “model sounding.” In this approach, an application model that uses only a single sounding cannot consider the effects of horizontal spatial variability in the atmosphere. Other applications, for example, some transport and diffusion models, however, do accept spatially and temporally varying mesoscale-model output fields, allowing full use of the three-dimensional plus time output from the RTFDFA system.

As mentioned in Part I, the ability to correctly anticipate meteorological conditions at the time of a test, even with a forecast of only a few hours, has tremendous implications for cost savings associated with accurate go–no-go test decisions. In addition to the expense associated with unproductive range time, there may be safety concerns as well (e.g., high winds during a missile launch could cause the missile trajectory to leave the test range boundaries). The output from a secondary application driven by a meteorological forecast model such as RTFDFA thus provides a forecast of the quantities of interest to range users that can serve as a form of a decision support system (DSS), useable by both

trained meteorologists and test directors who may have little meteorological training.

For decision support purposes, it would be most useful if the output of the secondary application could provide estimates of the probability of exceeding specified thresholds relevant to the particular application. The probability estimates should take into account uncertainties associated with weather model forecast errors at the scales resolved by the model, errors in the secondary model by itself, and errors introduced by unresolved scales (e.g., turbulent fluctuations) in both the forecast model and the secondary application.

The most obvious way of dealing with uncertainties in the forecast model is to use ensembles of different forecasts (e.g., Kalnay 2003) based on some combination of perturbations to the model initial conditions, or the use of different model parameterizations, or even of different models. Relevant to the topic of this paper, Warner et al. (2002) derived an ensemble of plume model forecasts based on 12 different MM5 executions with various initialization and boundary layer parameterization options to derive an ensemble of plume model forecasts. In addition, McKeen et al. (2007) showed that the output of seven different air quality models to predict aerosol concentrations used in an ensemble mode provided a better comparison to field observations than did a single model.

Accounting for uncertainties introduced by the effects of turbulence is more difficult. Some transport and diffusion models (e.g., Sykes and Gabruk 1997) attempt to account for turbulence effects by providing both mean and variance concentrations. For this and other applications, Frehlich (2006) describes an alternative method whereby meteorological ensembles might be produced by perturbing the input wind and temperature fields under the constraint that they satisfy a “universal” (statistically) turbulent behavior. Frehlich and Sharman (2003) used the same technique to perturb model output from a single thunderstorm simulation to produce an ensemble of model outputs. Last, uncertainties in the secondary model itself might be accounted for by perturbing other model inputs (e.g., the mass of a projectile) or internal tunable parameters.

Although an ideal DSS should account for all of these various sources of uncertainty, there are practical considerations that make their inclusion difficult for routine application. For one thing, including ensembles of forecast models is not practical in the current ATEC operational settings because of the computational burden associated with providing tens of forecast model runs to derive probabilities since the CPU usage is linearly dependent on the number of meteorological ensembles used. For the current hardware available to

compute forecasts at the ranges, only one member can be generated and still provide the model (deterministic) forecasts at reasonable times. For another, at the moment it is unclear what might be the best approach to including the effects of turbulence in either the RTFDDA input or output fields. Further, since the effects of turbulence will have different influences on different secondary models (e.g., a given turbulence level may affect sound propagation much differently than a missile trajectory), inclusion of those effects is best left to the developers of the secondary models. For the same reasons, accounting for uncertainties by perturbing parameters of the secondary model should also be prescribed by model developers.

However, in an effort to include at least some of the uncertainties inherent in the model soundings produced by the RTFDDA model, a strategy has been developed whereby several model soundings are used to drive the secondary application, each sounding derived from a different part of the range. This technique produces a spatial (or “poor person’s”) ensemble of secondary model output. The use of these ensembles helps to account for forecast errors in the timing and location of synoptic and mesoscale features (e.g., sea-breeze fronts) produced by the RTFDDA system, and also to account for regions of interest for which the secondary application spans several grid points.

In this paper, examples are provided of secondary models, some run in the spatial ensemble mode, that are in operational use at some of the ATEC ranges, and the combined model performance using RTFDDA forecasts as input is assessed. In contrast to Part II (Liu et al. 2008b), which provides evaluations of RTFDDA output, here the performance is evaluated in terms of the combined RTFDDA–secondary model output through comparisons with observations to the extent possible given the limited availability of test results. In particular, in section 2 we discuss the sound propagation modeling system used at the Aberdeen Proving Ground (APG), Aberdeen, Maryland, to predict noise levels in the surrounding communities resulting from ordinance testing, and provide some statistical verifications against available microphone data. In section 3, we describe the missile trajectory modeling system used at the White Sands Missile Range (WSMR), White Sands, New Mexico, and provide examples of comparisons of predicted and observed missile impact points. In section 4, we discuss the atmospheric transport and diffusion modeling system in operational use at the Dugway Proving Ground (DPG), Dugway, Utah, and provide example comparisons with available data. Section 5 provides a summary and conclusions.

## 2. Blast-noise forecasting

The mission at APG includes daily testing of small weapons such as grenades and mortars, and relatively large howitzers, tank guns, and bombs. Because of the proximity of APG to surrounding residential neighborhoods, blast noise associated with these tests is an undesirable side effect. Figure 1, which is a map of APG and the surrounding region, shows the locations of sound monitors (microphones) and surface mesonet sites. Rawinsondes are launched at the easternmost mesonet site. The APG’s proximity to many population centers means that the test center must balance military testing requirements with citizen concerns, and every effort is made to schedule tests during weather conditions that make noise propagation unlikely to disturb neighboring residents or cause property damage (Clough et al. 2000). Experience has shown that blast noise can propagate very efficiently under the right meteorological conditions, especially across the acoustically “hard,” that is, highly sound reflective, Chesapeake Bay. To schedule tests effectively, sound propagation must be predicted at least 4–5 h in advance, and optimally the day before the test (Clough et al. 2000). Traditionally, this “prediction” has been accomplished using data derived from a rawinsonde, launched some time before the blast is scheduled, as input to a local sound propagation diagnostic model called the Noise Assessment Prediction System (NAPS). Not surprisingly, the NAPS model predictions show large variability in the noise propagation pattern, depending on the details of the atmospheric structure.

An example of the sensitivity of the NAPS-predicted sound propagation pattern to the atmospheric vertical structure is shown in Fig. 2. Four soundings were taken at APG on 22 August 2001 at 1100, 1500, 1700, and 2000 UTC. The vertical profiles obtained from these soundings are shown with the corresponding NAPS-generated sound intensity levels. The propagation patterns are consistent with those expected from examination of the soundings. Propagation is favored by temperature inversions and positive wind shears in the downwind component (see, e.g., Lighthill 1978, section 4.6). The temperature profiles show a persistent inversion at about 2.5 km, trapping sound waves below this level. At earlier times, the wind is mainly from the north with positive wind shears in the 1–2-km layer, accentuating trapping downwind to the south. By 1700 UTC, the low-level winds have shifted to southerly, and the sound propagation pattern expands to the north. Although the low-level winds have shifted toward the northeast by 2000 UTC, the positive wind shear is in the

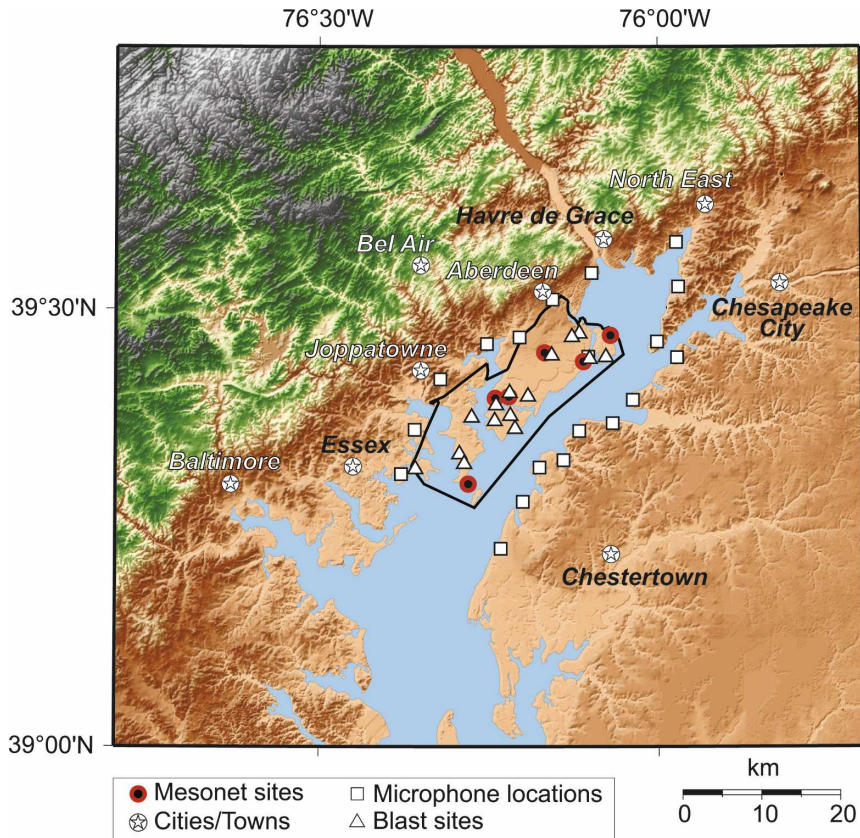


FIG. 1. Topographic map of the area around the APG showing range boundaries (heavy black lines), locations of mesonet stations, test sites, microphone locations, and the surrounding environment including the Chesapeake Bay (in blue) and major cities. The rawinsonde launch site is at the location of the northernmost mesonet site.

westerly component, causing the propagation pattern to shift to the east.

This example demonstrates the variability of the simulated sound propagation pattern over a period of just a few hours. Thus, the use of an early morning sounding may cause the incorrect prediction of sound-focusing effects because of changing mesoscale and synoptic-scale conditions throughout the day. To better predict the sound propagation pattern, model soundings provided by the APG operational RTFDAA model (see Fig. 3) are routinely used as input to NAPS. Thus, forecasts of sound propagation are possible at any time within the RTFDAA forecast window. The feasibility of using mesoscale models to drive sound propagation models has been demonstrated elsewhere (e.g., Hole and Mohr 1999; Heimann and Gross 1999). However, to our knowledge, this is the only *operational* sound propagation forecast system of its type.

#### a. RTFDAA–NAPS statistical performance evaluations

The NAPS model uses empirical relations given by the American National Standards Institute (1983) to

predict peak sound pressure levels (SPLs) as a function of explosive charge weight and distance from the source, to account for spherical spreading and atmospheric absorption. Also, ray-tracing techniques (Thompson 1972, 1974a,b) account for atmosphere-dependent focusing effects. Acoustic intensity is determined by mapping ray tube cross-sectional areas (Krol 1973). The technique is very similar to the Larkhill noise prediction method used in the United Kingdom by the Met Office and other establishments (Kerry et al. 1987; Turton et al. 1988a,b). The cited references estimate root-mean-square errors (RMSEs) in SPLs of about 6 dB, with individual errors of as much as 20 dB. Given the high-frequency variability in atmospheric profiles introduced by turbulence and other effects, this error is deemed acceptable. In fact, it has been suggested that noise complaints should be evaluated statistically because of this variability (e.g., Schomer et al. 1994; Schomer and Luz 1994; Schomer 2001). Given the uncertainties and errors in both the RTFDAA and NAPS models, it is reasonable to evaluate the combined model performance statistically, as well.

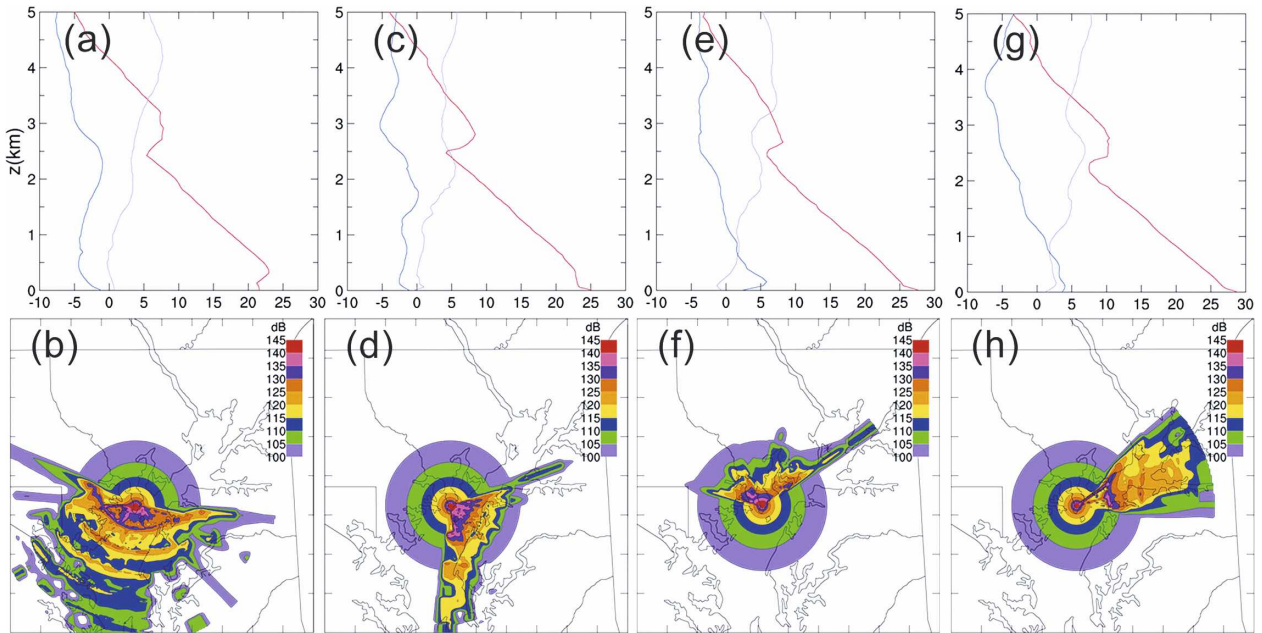


FIG. 2. Example of the diurnal change in boundary layer structure (top) as determined from APG rawinsonde launches and the consequent sound propagation pattern as indicated by NAPS (bottom). The sequence is (left to right) 1100, 1500, 1700, and 2000 UTC 22 Aug 2001. In the sounding plots, temperature ( $^{\circ}\text{C}$ ) is in red,  $u$  ( $\text{m s}^{-1}$ ) is in purple, and  $v$  ( $\text{m s}^{-1}$ ) is in blue. A color table for contours of sound pressure levels (dB) in the lower row is given in the upper-right corner of each panel.

To evaluate the RTFDFA–NAPS combined model performance, a number of cases were chosen for which good microphone data were available. The cases selected include blasts with TNT-equivalent charge weights of from 3 to 1200 lb (1.4–544.3 kg), discharged at seven different locations over different times of the day/year for a total of 16 separate blasts and 85 microphone readings. At the APG, there are many tests each day, often taking place simultaneously at different test locations. Therefore, it is sometimes difficult to correlate microphone readings to a particular test. Also, wind noise may activate the microphones, so most microphones are set to a fairly high recording threshold. For these reasons, only microphone data having a peak SPL of greater than 110 dB are used in this evaluation. The cases selected for comparisons were all associated with special tests, therefore ensuring that a rawinsonde launch was fairly close in time to that of the event, and that the microphone data could be correlated with the event.

For each blast case, NAPS was executed using the atmospheric profiles from both the rawinsonde and the RTFDFA model to obtain the area distribution of the SPL. Two NAPS runs were conducted using profiles applicable at the location of the rawinsonde sounding: one based on the rawinsonde sounding itself and one using a profile from the model analysis of the current conditions (i.e., no forecast). NAPS runs were also

conducted using model forecasts, with different lead times, of profiles at the time and location of the blast. The microphone peak-SPL data are compared with the NAPS output in Tables 1 and 2. In Table 1, the RTFDFA soundings are extracted from the operational D3 grid having a 3.3-km grid increment (see Fig. 3), and in Table 2 the RTFDFA soundings are obtained from a version of the system with a D4 grid having a 1.1-km grid increment.<sup>1</sup> In terms of RMSEs, the NAPS output that is based on RTFDFA D3 forecasts for all lead times is about as skillful as that based on the use of the model analysis and the rawinsonde sounding taken near the blast time (Table 1). So, with the use of RTFDFA forecasts, NAPS output for times up to 12 h in the future is as accurate as the estimation of current sound levels using rawinsonde data as input. When higher-resolution atmospheric input model is used from D4 (Table 2), the RTFDFA–NAPS system produces errors that are somewhat smaller than when data from the coarser-resolution D3 grid are used, depending on the lead time. This is probably due to the ability of the finer D4 (1.1-km resolution) grid to resolve sea-breeze and other local small-scale circulations

<sup>1</sup> The APG RTFDFA system had a domain 4 in the nest at the time of this study. Since then, this grid has been temporarily eliminated to enable extension of the length of the forecast. Domain 4 will be reinstated at the time of the next hardware upgrade.

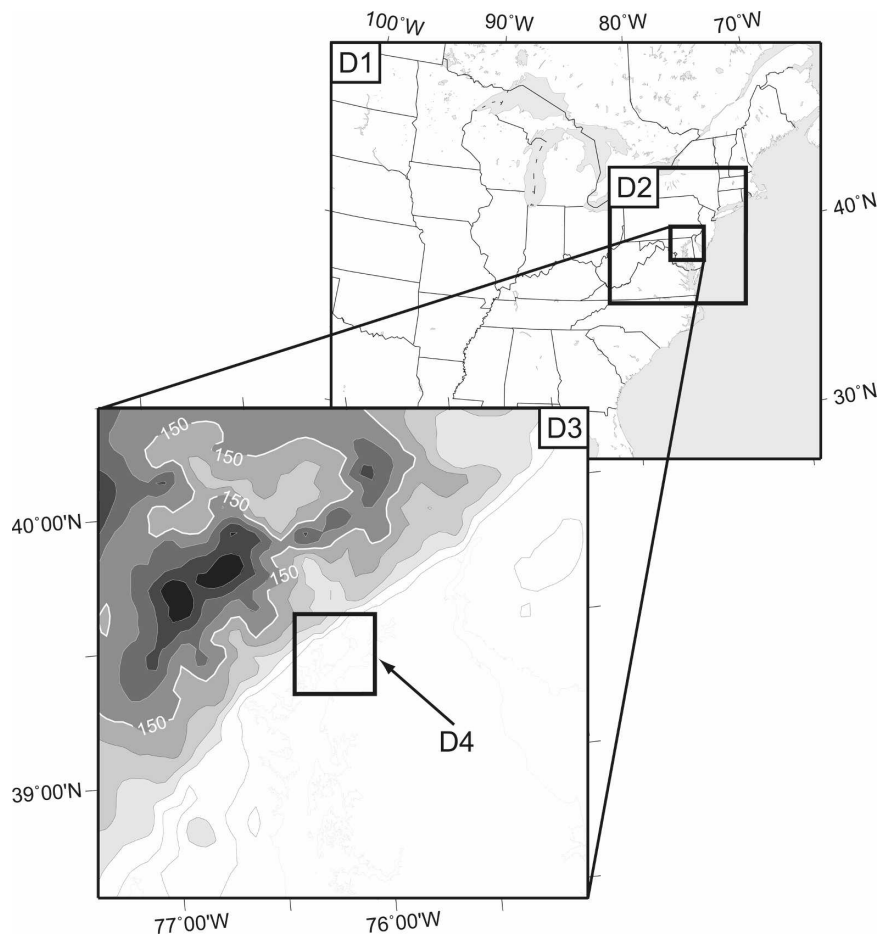


FIG. 3. RTFDFA grid configuration for the APG. (top) Grid areas for domains (D) 1, 2, and 3, and (bottom) a topographic map with domains 3 and 4. Contours of terrain height (shaded at 30-m intervals) are also shown in the bottom panel. The horizontal resolutions of each domain are 30, 10, 3.3, and 1.1 km, respectively. Note that D4 is temporarily not part of the operational APG system, but at the time of this study it was being used.

in the complex land–sea boundaries of the Chesapeake Bay. This speculation is supported by local forecaster experience. Note that there is no strong dependence of RMSE on forecast lead time, and in fact given there are

only 85 microphone readings available, the expected standard deviation of the scatter based on chi-square statistics is about 1 dB, or about as much as the differences between forecast lead times. The operational D3

TABLE 1. Summary of NAPS SPL RMSE and directional performance statistics for APG, using rawinsonde soundings (location DC1) and RTFDFA profiles as input to NAPS in comparison with 85 microphone readings from 16 cases. For RTFDFA profiles, the D3 output is used from model forecast runs for various forecast lead time windows ( $f_x - f_y$ ). That is, at the time of each blast, there are four 3-hourly RTFDFA forecasts available, and the table shows the statistics for each in terms of the lead-time window. The use of the RTFDFA analysis is evaluated by interpolation to the APG sounding site (DC1) and also to the blast site (i.e., the location of the blast). All RTFDFA–NAPS analyses and forecasts are based on interpolation of RTFDFA profiles to the rawinsonde location or blast site. The sample size for each forecast lead time is given in the first row.

	APG DC1 rawinsonde	$f_0$ DC1	$f_0$ blast site	$f_{1.5-4.5}$ blast site	$f_{4.5-7.5}$ blast site	$f_{7.5-10.5}$ blast site	$f_{10.5-13.5}$ blast site	Avg
RMSE (dB)	12.0	11.9	12.4	12.0	12.0	12.5	12.4	12.2
% in 90° wedge	49.8	53.3	49.4	57.1	54.7	46.2	45.8	50.9
% in 135° wedge	66.2	61.5	68.8	68.0	67.5	63.2	62.0	65.3
% in 180° wedge	84.2	80.3	82.7	83.7	83.6	80.9	76.1	81.6

TABLE 2. As in Table 1 except that statistics are based on an experimental RTFDFA D4 with a 1.1-km grid increment.

	APG DC1 radiosonde	$f_0$ DC1	$f_0$ blast site	$f_{1.5-4.5}$ blast site	$f_{4.5-7.5}$ blast site	$f_{7.5-10.5}$ blast site	$f_{10.5-13.5}$ blast site	Avg
RMSE (dB)	12.0	10.6	10.6	12.4	12.2	11.6	11.4	11.3
% in 90° wedge	49.8	63.3	54.8	57.3	60.1	54.3	55.1	56.4
% in 135° wedge	66.2	73.2	67.7	70.3	75.2	71.8	65.7	70.0
% in 180° wedge	84.2	84.6	86.7	84.3	83.0	83.9	80.9	84.0

0–12-h forecasts produce SPL RMSEs that are about equivalent to that, while the D4 0–12-h forecasts are about 10% smaller than those resulting from the use of the rawinsonde data, again depending on the lead time. Thus, for the cases examined here, by this performance metric, driving NAPS with as much as a 12-h forecast prior to the blast time is as reliable as driving NAPS with a sounding taken near the blast time.

The directional error associated with the combined RTFDFA–NAPS forecasts was evaluated separately by computing the percentage of the triggered microphones that are contained in an angular wedge of specified angular width, that is, centered on the blast site and oriented in the direction of the maximum acoustic power calculated by NAPS. The average acoustic power ( $\int p^2/r^2 dA$ , where  $p$  is acoustic pressure,  $r$  is radial distance from the blast, and  $dA$  is an area increment) within a wedge was calculated for angular wedges with widths of 90°, 135°, and 180°. Each wedge was rotated through 360° at 5° increments to define the direction of maximum power (the wedge with the maximum average acoustic power). Because the minimum microphone distance from blast sources was about 6 km, the acoustic power was computed only in those portions of the wedge for which the radial distance from the source was greater than 6 km. Using this technique, it was found that, averaged over all of the blasts considered, about 50% of the triggered microphones were located within the 90° wedge when NAPS used the rawinsonde profile as input (see Table 1). This coverage increased to ~66% when the wedge angular width was increased to 135°, and to ~84% when the wedge was a semicircle. Even though this skill level has been historically sufficient for the APG test directors, the lack of lead time when using the sounding has been an issue. Tables 1 and 2 provide the same performance statistics for the RTFDFA–NAPS forecasts, and as can be seen, the forecasts provide about the same performance as NAPS driven by the rawinsonde input. Thus, at least for the cases examined here, in terms of both RMSEs and directional error metrics, the combined RTFDFA–NAPS model noise sound level forecasts out to 12 h have equivalent or less error than the no-

lead-time estimates available from the NAPS predictions using APG rawinsonde data.

### b. Spatial ensembles

As mentioned in the introduction, computing RTFDFA ensembles is not practical in the current operational setting because of the computational burden associated with providing tens of RTFDFA runs to derive probabilities of exceedance of specified SPLs. Therefore, as an alternative, a spatial (or “poor person’s”) ensemble of model soundings is constructed, with each model sounding derived from a different grid point within a certain radius of the blast site (as shown schematically in Fig. 4) to drive NAPS. The radius is defined by the user and is typically 10–40 km, depending in part on the location of the blast, and is meant to capture the spatial variability of the environment across the Chesapeake Bay. The ensemble of RTFDFA–NAPS predictions can then be used to derive mean propagation patterns as well as information about uncertainties in the mean pattern by providing a standard deviation and probabilities of exceedance.

Figure 5 shows examples from a spatial ensemble of RTFDFA–NAPS forecasts. Each panel shows a different NAPS prediction based on the particular model sounding grid point indicated in the upper-left corner of the panel. Although there are obvious differences in the details, most of the patterns show sound focusing toward the southeast. This feature is reflected in the summary statistics and probabilities of exceedance in Fig. 6, which are based on all 108 model soundings used in this exercise. The SPL average (arithmetic and geometric), median, standard deviation, and probabilities of exceedance are all highest to the southeast. The arithmetic and geometric means are almost identical in this case, and are similar in appearance to the median, although the means show a slightly more intense area of focusing to the southeast than does the median.

These RTFDFA–NAPS spatial ensemble forecasts are currently used daily at the APG to help staff meteorologists advise range officials about whether to delay or cancel tests that would adversely impact surrounding neighborhoods. Usually the RTFDFA–

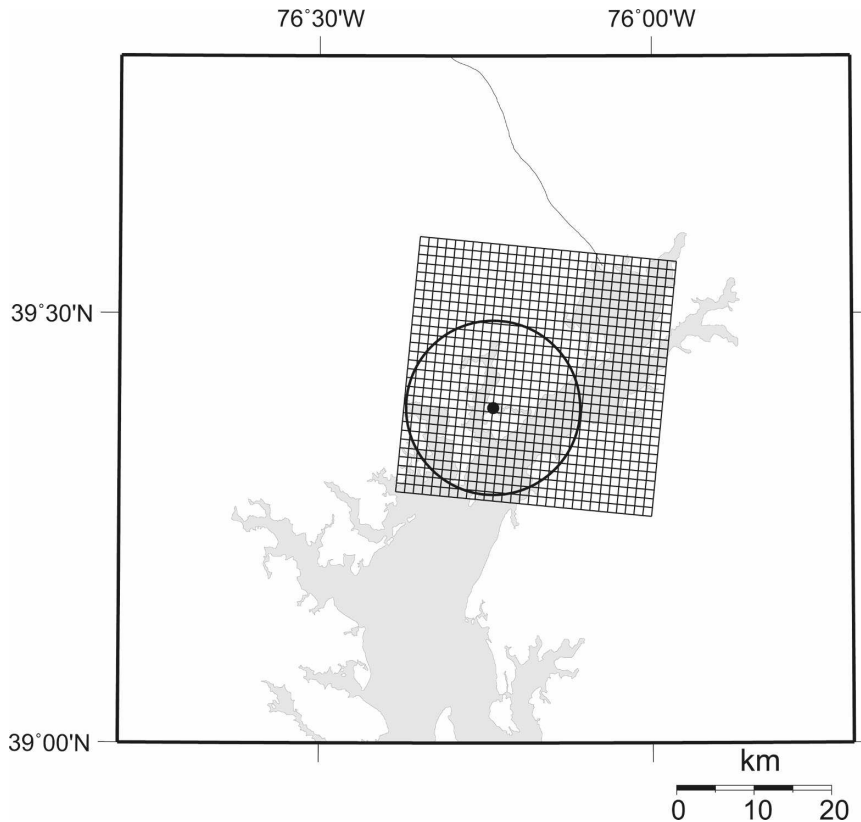


FIG. 4. Schematic of spatial ensemble selection criterion. All model soundings at D4 grid points within the radius specified (10 km in this example) are used to drive the NAPS sound propagation model.

NAPS 3–6-h forecasts are examined in the morning for tests that are to take place later in the day. However, some tests that involve larger blasts may require a 24-h forecast to schedule. The ensemble average, standard deviation, and the 120- and 130-dB probabilities of exceedance are routinely reviewed for decision making.

### 3. Missile trajectory modeling

Testing guided and unguided missiles and rockets is one of the major activities at the WSMR, where the RTFDDA system is used to help forecast wind conditions and possibly turbulence that may cause a missile to stray outside the range boundaries. The WSMR RTFDDA grid configuration (D1–D3) is shown in Fig. 7, along with the range boundaries in the D3 domain. Because of the range's overall south–north orientation, missiles are usually fired northward from a launch location near the southern end of the range. The RTFDDA-forecast winds are used as input to various missile trajectory models to provide the range directors some assurance that the projectile will remain within the range boundaries after being launched. The trajec-

tory models are typically either three degrees of freedom (3DOF, three translational modes) or six degrees of freedom (6DOF, three translational plus three rotational modes) rigid-body dynamics models. However, for WSMR operational use, the 6DOF model is actually executed in a 5DOF mode where vehicle roll is neglected. The 3DOF and 6DOF equations are discussed in standard dynamics and aerodynamics texts, such as Etkin (1972).

The rockets launched at WSMR often reach apogees of nearly 300 km and, therefore, respond to winds from the surface to the thermosphere. However, the missile lateral acceleration due to winds is proportional to  $\rho_a |\mathbf{V}|/B$ , where  $B$  is a ballistic coefficient that depends on the vehicle mass, drag coefficient, and surface area, and the atmospheric density  $\rho_a$  decreases with height more rapidly than the magnitude of the wind vector,  $|\mathbf{V}|$ , increases with height. Thus, the trajectory model is most sensitive to winds in the troposphere and lower stratosphere where  $\rho_a$  is largest. The accuracy of the RTFDDA-produced winds in these regions was evaluated for several time periods in 2004 for which WSMR rawinsonde flights were available for comparison.



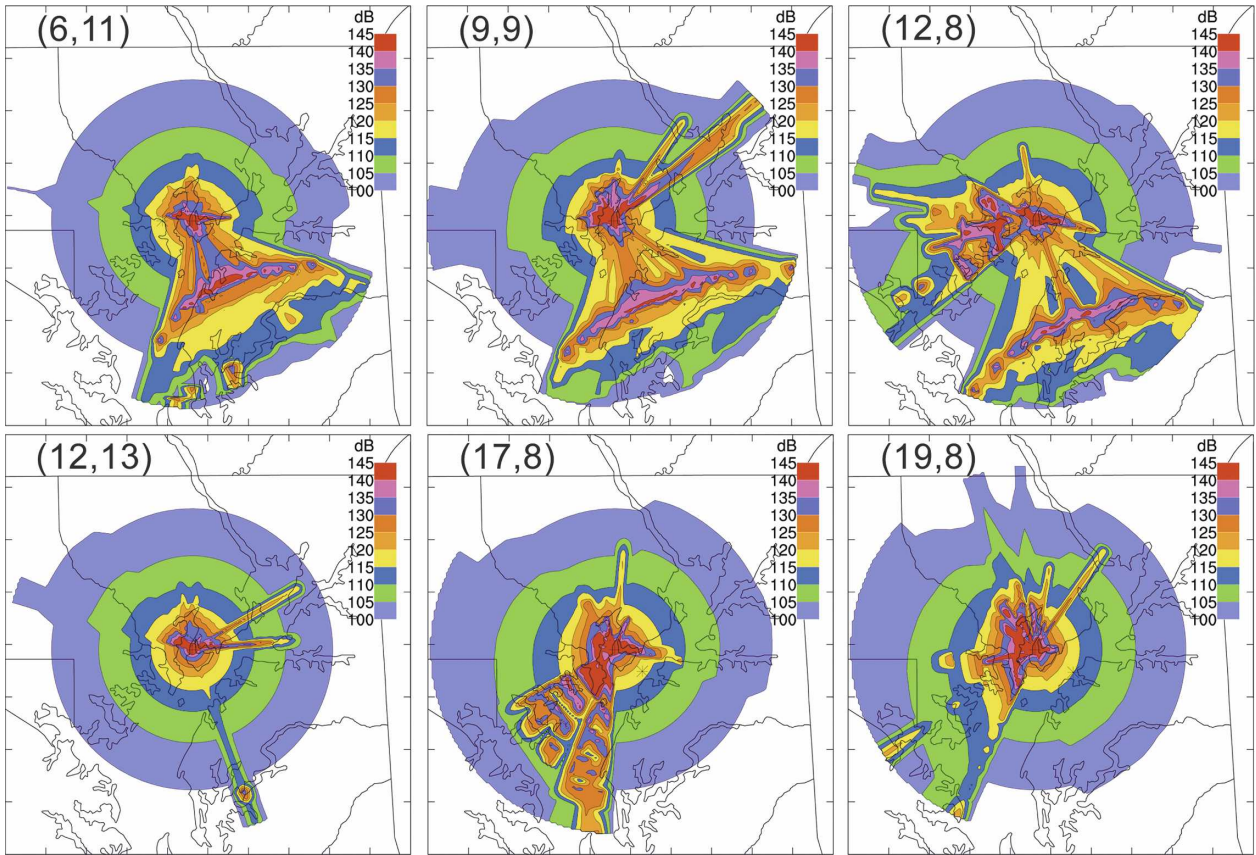


FIG. 5. Samples of contours of SPL (dB) derived from NAPS based on RTFDFA 18-h forecast model soundings at the D4 gridpoint locations indicated in the upper-left corner. The SPL contour color table is given in the upper-right corner of each panel.

Table 3 summarizes the results of that comparison in terms of aggregate 1–12- and 13–24-h forecast wind speed bias and RMSE, wind direction bias and RMSE, and RMS vector wind error (RMSVE), with

$$\text{bias} = \frac{1}{N} \sum (q_p - q_o),$$

$$\text{RMSE} = \left[ \frac{1}{N} \sum (q_p - q_o)^2 \right]^{1/2}, \text{ and}$$

$$\text{RMSVE} = \left\{ \frac{1}{N} \sum [(u_p - u_o)^2 + (v_p - v_o)^2] \right\}^{1/2},$$

where  $q$  denotes wind speed or direction,  $u$  and  $v$  are the east–west and north–south wind components, respectively, the subscripts  $p$  and  $o$  indicate predicted and observed values, and the summation is over all  $N$  observations. From the table, wind speed errors are generally modest and tend to increase with increasing altitude, whereas the wind direction errors tend to decrease with increasing altitude. The magnitude of these errors is comparable to those of the National Centers for Environmental Prediction’s 20-km grid increment

Rapid Update Cycle model (see Benjamin et al. 2004, their Fig. 9), which show the same trend of increasing RMSVE with increasing altitude up to about 250 hPa. It should be mentioned that these errors may be larger than those from coarse-resolution NWP models, since as discussed in Part II (Liu et al. 2008b) traditional model error statistics tend to be larger with higher resolution (e.g., Mass et al. 2002; Davis et al. 2006). In any event, since, as mentioned above, the missile trajectory is most sensitive to lower-level winds, these errors are considered acceptable for input to trajectory models.

However, above the middle stratosphere RTFDFA-derived winds are not available because the model top is at 50 hPa, or about 20-km elevation above sea level. Thus, it was necessary to merge the RTFDFA upper-level winds with climatological winds derived from previous WSMR rocketsonde launches. The climatological winds have been tabulated into a range reference atmosphere (RRA), which provides the monthly mean and standard deviation of winds and temperature over the range from the surface to about 137 km. The RRA climatological winds were merged with the RTFDFA

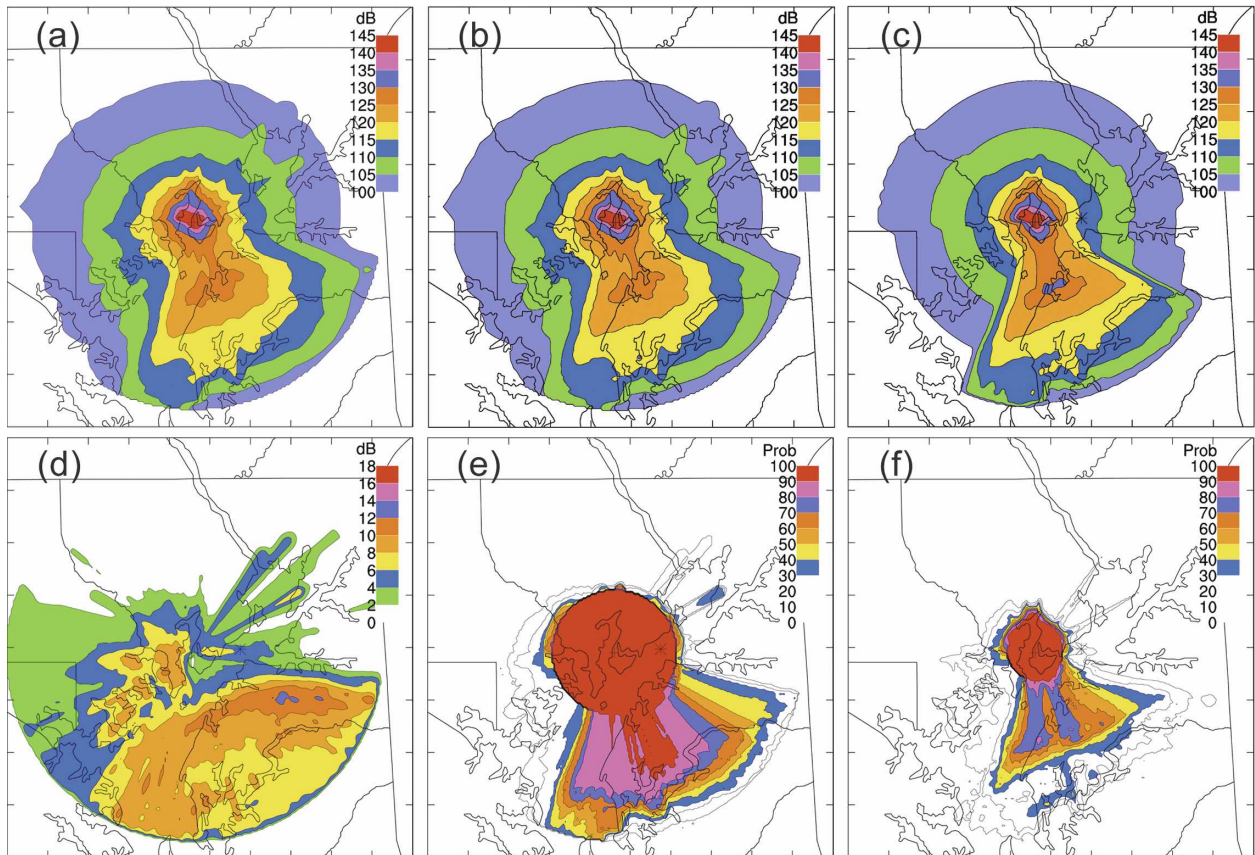


FIG. 6. Contours of (a) arithmetic mean, (b) geometric mean, (c) median, and (d) standard deviation of the SPL derived from NAPS, based on RTFD DA 18-h forecast model soundings at the D4 grid points within 10-km radius from the blast location, shown in Fig. 4. Corresponding probabilities (%) of exceeding (e) 110 and (f) 120 dB are also shown. All figures are based on a total of 108 model soundings surrounding the blast site. A color table for the contours is given in the upper-right corner of each panel.

forecast winds, by shifting the RRA mean profile to be consistent with the RTFD DA winds near the model top.

To account for uncertainties in the forecast winds, the operational procedure adopts the same methodology used for the sound propagation forecasts at APG; that is, a spatial ensemble of model soundings is used (contained in a radius of typically 50–100 km within the center of the range), where the RRA profile is fitted to each ensemble member. An example of the merger of the RRA winds and RTFD DA 24-h forecast winds is provided in Fig. 8. The mean and plus/minus one standard deviation ( $\pm 1 \sigma$ ) values from the RRA profile are indicated by the dashed lines, and the level identified as “M” is where the forecast and RRA profiles are merged. The gray area shows the ensemble of 24-h forecast wind profiles up to about 20 km, valid at about the time of a rocket launch, with the RRA profile used above that level. In this particular case, the forecast wind profiles agree fairly well with the rawinsonde profile up to the model top, where the  $u$  component is

about  $+1 \sigma$  and the  $v$  component is about  $-1 \sigma$  from the average according to the RRA. But above about 20 km, rawinsonde profiles deviate back toward the mean, showing that the extrapolation does not always produce the best possible fit. However, without increasing the model top, this pragmatic procedure seems to provide reasonable results.

The spatial ensemble used to construct Fig. 8 was composed of 304 individual profiles within 100 km of the center of the WSMR. Each model sounding was used to drive a 5DOF version of the rocket trajectory model known as General Electric Missiles and Satellites Systems (GEMASS). Figure 9 compares the trajectories computed using the GEMASS model for one mission with 1) the 304 RTFD DA 24-h forecast wind profiles, 2) the rawinsonde-measured winds at the time of launch, and 3) zero winds. The actual rocket impact point is also shown. In this case, the agreement between the rawinsonde-derived trajectory and the RTFD DA 24-h forecast-derived trajectories is very good, even though, as already indicated, the upper-level winds are

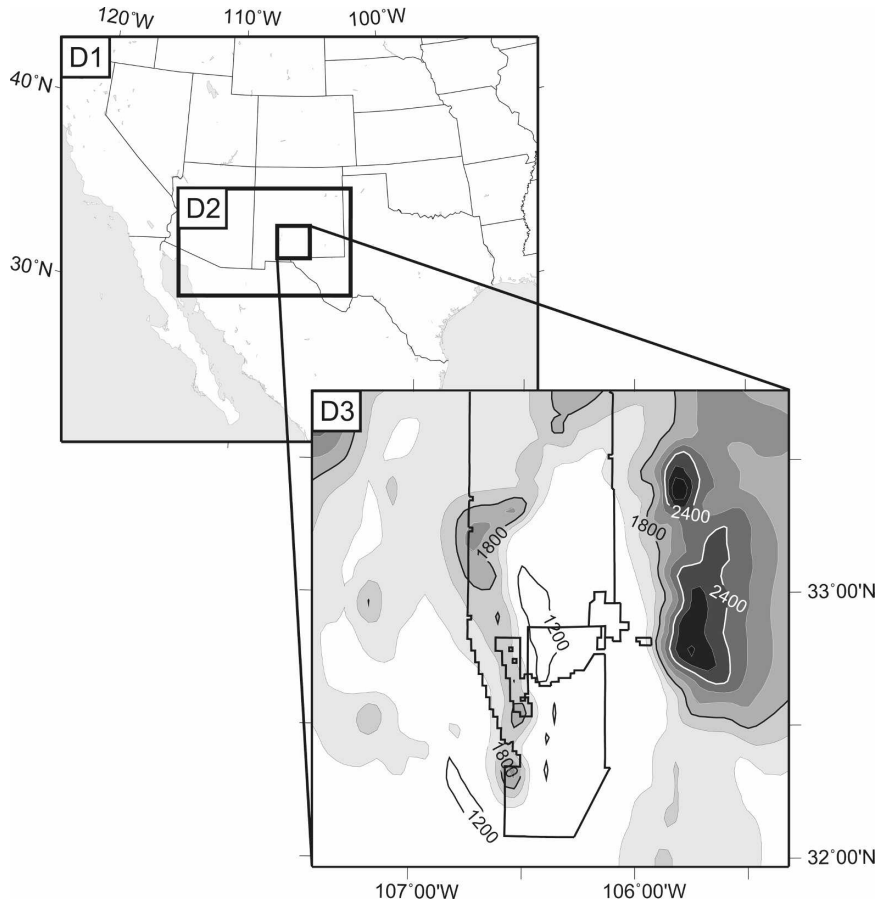


FIG. 7. RTFDADA grid configuration for the WSMR. (top) Grid areas for domains (D) 1, 2, and 3, and (bottom) a topographic map and the WSMR range boundaries in D3. The horizontal resolutions for D1–D3 are 30, 10, and 3.3 km, respectively.

somewhat misrepresented in the merged RTFDADA–RRA sounding. Table 4 lists the impact point errors for four different missions (including the one shown in Figs. 8 and 9) for which impact data were available. Of course, four cases are not enough to provide statistical significance, but based on this limited sample, it can be seen that the impact errors associated with winds from the 24-h RTFDADA forecasts to drive the missile trajectory model are about the same as those when using winds from a rawinsonde at the time of missile launch. Because of these results and observations of other cases by range personnel, RTFDADA-forecast-driven trajectory models are now routinely used by range meteorologists for test planning.

#### 4. Transport and diffusion modeling

The DPG, and occasionally other ATEC ranges, use atmospheric transport and diffusion models to predict the concentrations of smokes, obscurants, and simulants for chemical or biological (CB) agents, as the

clouds or plumes move across the range. The transport and diffusion model normally used for this purpose is the Second-Order Closure Integrated Puff (SCIPUFF) model, a Lagrangian puff dispersion model that represents a plume by a series of Gaussian puffs and that uses a second-order parameterization for the turbulent dispersion (Sykes and Gabruk 1997). The SCIPUFF model differs from other operational dispersion models in its capability to predict both the mean concentration and the concentration variance. The meteorological input includes gridded fields of spatially and temporally varying wind, temperature, surface heat flux, and planetary boundary layer depth from mesoscale-model forecasts. Hence, the spatial ensemble approach used with the sound propagation model or the missile trajectory model cannot be applied to SCIPUFF since SCIPUFF uses output from the entire three-dimensional RTFDADA grid. However, in principle, ensembles could still be generated from separate RTFDADA runs with different initializations, boundary layer parameterizations, and surface treatments to cre-

TABLE 3. RTFDFA model wind speed and direction performance statistics as a function of altitude (pressure; hPa) based on comparisons with rawinsonde data at WSMR for five different time periods during 2004 for which several rawinsonde launches per day were available. The data are separated into combined statistics for 1–12- and 13–24-h forecasts. The bias and RMSE for wind speed and direction, and RMS vector error, are provided. The total number of observations used at each level to compute the statistics is given in the last column.

$p$ level (hPa)	Speed ( $\text{m s}^{-1}$ )		Direction ( $^{\circ}$ )		RMSVE ( $\text{m s}^{-1}$ )	$N$
	Bias	RMSE	Bias	RMSE		
1–12-h forecast wind statistics						
850	1.4	3.6	–18.5	69.4	3.9	170
700	1.8	4.4	–0.5	36.8	4.9	153
500	1.1	5.0	9.2	29.7	5.0	106
400	0.6	4.1	12.1	40.1	4.8	94
300	0.8	4.3	8.2	19.9	4.7	55
250	2.8	4.8	11.2	16.0	5.7	42
200	5.4	6.9	17.7	24.6	7.5	23
150	3.8	7.3	11.4	19.2	7.4	24
100	6.7	8.0	4.4	23.6	7.7	43
13–24-h forecast wind statistics						
850	1.9	3.7	–18.6	90.6	4.6	179
700	2.7	4.8	–16.8	61.3	7.2	166
500	–0.3	4.8	11.1	43.2	6.1	110
400	0.0	5.2	12.6	51.2	6.9	101
300	–0.5	5.6	4.6	18.0	6.0	57
250	2.8	5.3	18.7	23.6	7.1	46
200	3.0	4.5	29.1	36.7	6.7	22
150	4.1	7.9	14.7	25.3	8.8	25
100	7.0	8.6	2.6	24.1	8.5	38

ate the ensemble members, as was done in a special case study by Warner et al. (2002). In this approach, each RTFDFA member is used to drive a SCIPUFF run, and the ensemble of plumes is used to quantify the uncertainty in the surface concentration or dosage (time-integrated concentration) prediction. However, in practice this approach is currently too computationally intensive to be used operationally. Another approach would be to use ensembles of transport and diffusion models as was done in the McKeen et al. (2007) study, but this would require many models (at least 10) that we do not have at our disposal in the operational setting. At the moment then, SCIPUFF is the only transport and diffusion model used operationally at the ranges, and thus only examples of its use are provided here.

The SCIPUFF model has been evaluated as a component of the Hazard Prediction and Analysis Capability (HPAC) suite of models using data from field dispersion studies with controlled releases (e.g., Sykes and Gabruk 1997; Chang et al. 2003). However, most of these SCIPUFF model evaluation studies used observed meteorological data as input, rather than gridded output from a mesoscale model forecast. A first

step in the evaluation of the RTFDFA–SCIPUFF combined model is to compare RTFDFA boundary layer winds with observations. Figure 10 provides two examples of comparisons of D4 RTFDFA analyses and short-term (3–6 h) wind forecasts with data from a 915-MHz wind profiler at DPG. The comparison is for 850 hPa, or roughly 200–350 m AGL, for two time periods for which profiler data were fairly continuous. Comparisons are shown for both the final analysis and 3–6-h forecasts. Two results are evident for these cases: 1) the RTFDFA analyses follow the observations fairly well for both time periods, indicating that the RTFDFA observation-nudging process is able to capture the boundary layer wind evolution satisfactorily, and 2) the model short-term forecasts generally capture the wind episodes seen by the profiler. Other levels (not shown) show similar results. The combined RMSVEs over these two time periods (containing about 200 separate observations) are 2.8 and 4.3  $\text{m s}^{-1}$  for the final analysis and the 3–6-h forecasts, respectively. Some of this error is due to phase differences. This is somewhat higher than a statistical analysis of RTFDFA near-surface performance at DPG shown in Rife et al. (2004), but the period shown here was a little more variable with summertime convection. The magnitude of these errors is comparable to that of other coarser-scale NWP models. This result is related in part to the difficulty mentioned above that conventional performance statistics do not tend to show much, if any, improvement, by using higher resolution (e.g., Mass et al. 2002; Davis et al. 2006). However, as shown in the Rife et al. (2004) study, higher-resolution models such as RTFDFA do in fact provide a better representation of the spatial variability of winds due to terrain and other localized effects.

Two examples of combined RTFDFA–SCIPUFF performance are shown in Figs. 11 and 12 based on the Dipole Pride 26 (DP26) experiment at the Yucca Flat, Nevada Test Site, Nevada, during November 1996. The DP26 experiment is described in detail in Chang et al. (2003). Briefly, sulfur hexafluoride ( $\text{SF}_6$ ) was released as instantaneous puffs at several locations over several days, and concentrations were measured by three lines of whole-air samplers (30 per line, with a 15-min sampling interval) downwind from the release site. Meteorological information was available from eight surface stations, two pilot balloon locations, and one rawinsonde station (see Fig. 1 in Chang et al. 2003). The DP26 experiments were chosen for use in this study because of the large amount of available tracer and meteorological data. Also, the fairly irregular terrain near the test site makes the meteorological and plume modeling more challenging.

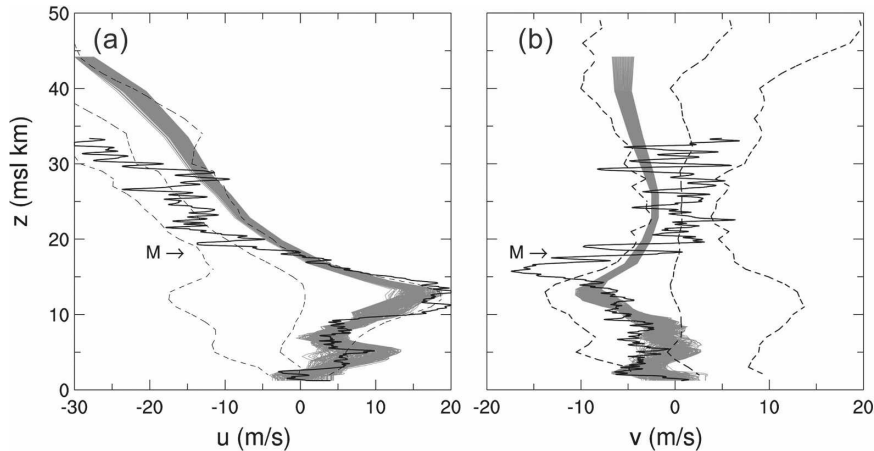


FIG. 8. Example of RTFDAA 24-h forecast (valid at 1600 UTC) spatial ensemble profiles (gray) of (a) the east–west wind component  $u$  and (b) the north–south wind component  $v$ , derived from 304 model grid points within 100 km of the midpoint of the test range. The light dashed lines indicate the WSMR RRA values (with the middle line indicating the mean and the lines to the left and right of the mean indicating the negative and positive one standard deviation values, respectively). The launch-time-coincident rawinsonde (heavy line) is also shown. This particular case is for a rocketsonde launch at 1622 UTC 7 Jul 2005. The letter M indicates the level where the RTFDAA winds were merged with the RRA winds.

Figure 11 compares maps of the SCIPUFF-produced surface dosages with the bag sampler dosages for trials 4 and 11 (3.5 h after the release), with SCIPUFF driven by 1) the meteorological observations and 2) the RTFDAA final analysis (three-dimensional winds and temperature, plus boundary layer height and surface heat flux, updated hourly) at 3.5 h after the release. The

observations used to drive SCIPUFF were quality controlled, and as part of that process rawinsonde data in the first 500 m above ground level were not used. This strategy is consistent with that described in the Chang et al. study. As can be seen, the overall structure of the SCIPUFF-predicted plumes is very different when using RTFDAA output when compared with using actual

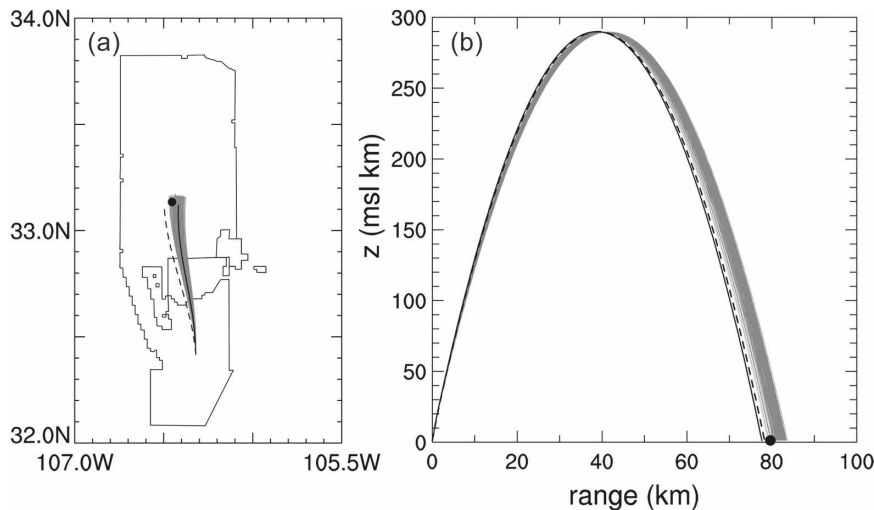


FIG. 9. (a) Horizontal and (b) vertical projections of rocket trajectories obtained using the GEMASS 5DOF model. The gray lines are the trajectories derived from the 304 spatial ensemble members of the RTFDAA 24-h forecast winds plotted in Fig. 8. The dark solid line is the trajectory computed from the rawinsonde-measured winds (Fig. 8), and the dashed line is the trajectory computed assuming zero winds. The black dot indicates the actual rocket impact location.

TABLE 4. Cross-track (“X”) and down-track (“Y”) mean absolute error (MAE) and RMSE metrics based on differences between the GEMASS-computed trajectory impact point and the actual recorded impact point. The columns labeled “obs” are based on rawinsonde data taken near the time of the missile launch, and the columns labeled “fcst” are from the 24-h RTFDFA forecast spatial ensembles (303 members), to drive GEMASS.

Metric (km)	Cases									
	Mission 1		Mission 2		Mission 3		Mission 4		All missions	
	Obs	Fcst	Obs	Fcst	Obs	Fcst	Obs	Fcst	Obs	Fcst
XMAE	3.33	3.12	9.41	11.67	0.85	1.84	1.17	2.40	3.69	4.75
YMAE	1.78	2.63	7.25	2.49	8.49	6.12	1.00	2.06	4.63	3.33
XRMSE		1.87		2.10		1.14		1.94	3.97	4.40
YRMSE		0.81		1.46		0.88		1.76	3.78	2.08

observations, but because of the limited area of the bag samplers it is difficult to say which structure is more correct.

Figure 12 shows the same results in a different format within the area covered by the bag samplers. In particular, it compares the sampler dosage measurements along each line with the SCIPUFF predictions along the same lines, using the two sources of meteorological inputs. Trial-11 SCIPUFF dosages produced from the RTFDFA analysis provide better agreement with the observed dosages, probably because of the space- and time-varying data available from the model. However, the benefit of RTFDFA is not as clear for trial 4. A more quantitative comparison is provided in Table 5, which shows the same performance metrics as used in the Chang et al. (2003) paper. These metrics include the fractional bias (FB), geometric mean bias (MG), nor-

malized mean square error (NMSE), geometric variance (VG), and the fraction of predictions within a factor of 2 of the observations (FAC2). They are defined as follows:

$$FB = \frac{(\overline{D_o} - \overline{D_p})}{0.5(\overline{D_o} + \overline{D_p})},$$

$$MG = \exp(\overline{\ln D_o} - \overline{\ln D_p}),$$

$$NMSE = \frac{(\overline{D_o - D_p})^2}{\overline{D_o D_p}},$$

$$VG = \exp[(\overline{\ln D_o} - \overline{\ln D_p})^2], \text{ and}$$

FAC2 = fraction of data for which  $0.5 \leq D_p/D_o \leq 2.0$ , where  $D_p$  and  $D_o$  are, respectively, the SCIPUFF dosage predictions and the observations of the maxi-

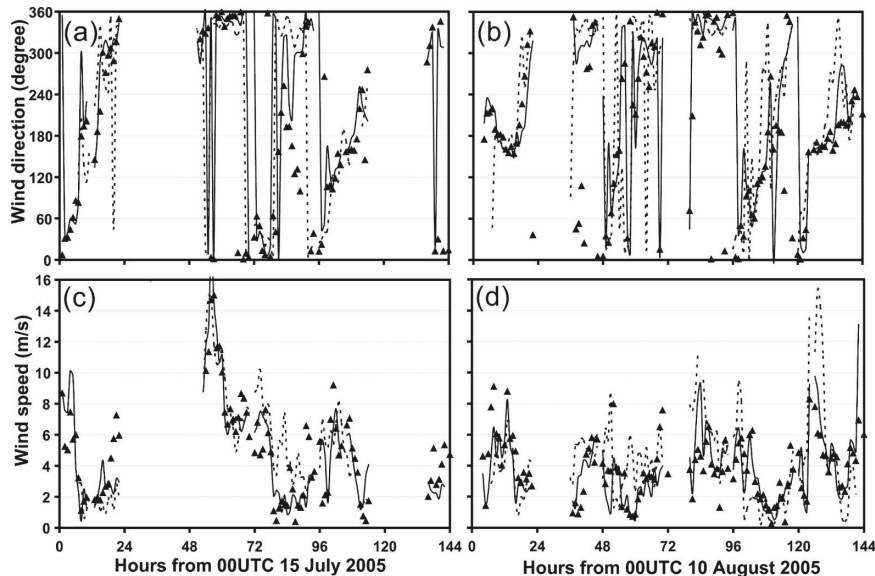


FIG. 10. Comparison of (a), (b) the hourly evolution of wind direction and (c), (d) wind speed at 850 hPa from the RTFDFA analyses (solid lines), 3–6-h forecasts (dashed lines), and DPG profiler measurements (black triangles) for two periods: 15–21 Jul 2005 (left) and 10–16 Aug 2005 (right). Time periods for which profiler data were not available are not shown.

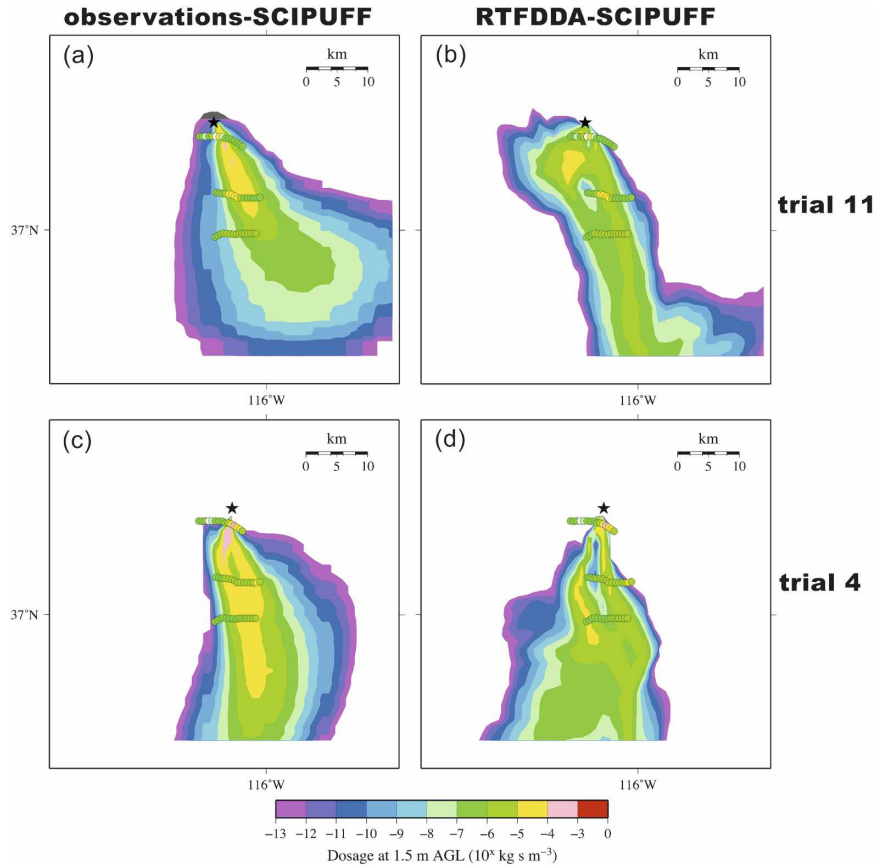


FIG. 11. Total dosage contour maps at 1.5 m AGL, 3.5 h after release, as calculated by SCIPUFF (top) for trial 11 on 15 Nov and (bottom) for trial 4 on 9 Nov 1996 during the DP26 experiment at the Yucca Flat, Nevada Test Site, Nevada. The release point is indicated by the star in the plots. Shown are the predictions using SCIPUFF driven by (left) available meteorological observations and (right) RTFDDA final analyses. For comparison, the dosage data collected along the three lines of bag samplers are shown with the same color scale. For presentation, only every other bag sample value in each line is plotted. The dosage contours and sampler readings follow the color scale given at the bottom of the figure.

imum dosage along each sampling line. The overbar represents the average over all samples taken (three lines  $\times$  30 bags per line). For both cases, and for almost all performance metrics computed, the RTFDDA-SCIPUFF modeling system provides better agreement with the dosage observations than does SCIPUFF using observed meteorology. The implication is that RTFDDA-SCIPUFF forecasts may provide adequate information about the spread of intentional or inadvertent releases that can be used by first and second responders and other agencies in evacuation and decontamination planning.

## 5. Summary and conclusions

The feasibility of using high-resolution meteorological forecast models to drive secondary-applications models at the ATEC test ranges was investigated using

three different applications models. Not enough observational data were available to make rigorous quantitative statistical estimates of the accuracy of some of the secondary applications driven by RTFDDA, but for the case studies presented there was comparable accuracy for the application models that were driven by current observations and by 12–24-h RTFDDA forecasts. Further statistical evaluations are required to quantify the performance for different applications and different forecast lead times. The results presented here describe an initial capability; incremental improvements in performance are to be expected with time.

The forecast lead times (0–12 h) provided by the RTFDDA application model are sufficient for most test planning purposes. The RTFDDA sound propagation forecasts at the APG are usually examined in the early morning for late morning and afternoon tests (i.e., lead

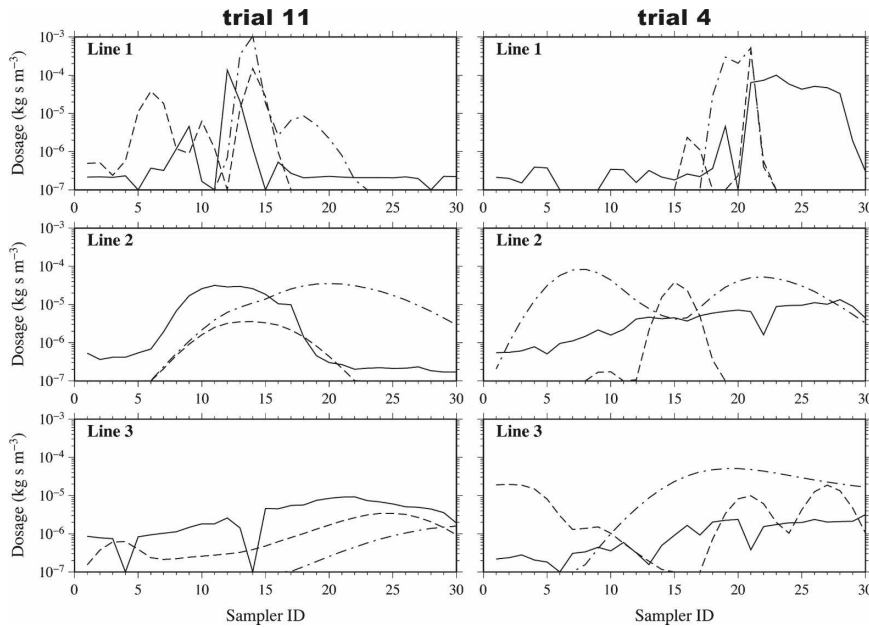


FIG. 12. Total dosage (time-integrated concentration) at each of the 30 sampler sites in each of the three sampled lines at the time corresponding to Fig. 11; i.e., 3.5 h after release for (left) trial 11 and (right) trial 4. The measured dosages are indicated by the solid lines, the dosages calculated by SCIPUFF driven by available meteorological observations by the dash-dot lines, and the dosages predicted by SCIPUFF driven with RTFDFA final analyses by the dashed lines.

times of 3–6 h). The missile launches at WSMR are very expensive, so the RTFDFA-driven missile trajectory forecasts are examined from several days out up to an hour or two before the launch. Indications of rapidly changing conditions in the model output are cause for concern. Costs of conducting transport and diffusion campaigns at the DPG can also be very expensive, depending on the equipment allocated and the number of personnel involved. However, the required lead times for a short-range plume transport forecast are typically 6–8 h and the required lead times for a long-range forecast are typically 12–24 h. Regardless of the specific application, even a 1–2-h lead time can translate into considerable time and money savings by minimizing the

need for last minute delays due to unexpected weather changes.

Uncertainties associated with the forecast model can be accounted for, in part, by the use of spatial ensembles, in which individual model soundings over the forecast domain are used to drive separate executions of the application model or models. There are several reasons for using the spatial ensemble approach.

- Computing meteorological ensembles is not practical in the current operational setting because of the computational burden associated with providing tens of RTFDFA runs to derive probabilities.
- For most secondary applications in use at the ranges,

TABLE 5. Comparisons of model performance metrics of dosage-related parameters at 3.5 h after release using SCIPUFF driven by available quality controlled observations and using SCIPUFF driven by RTFDFA for DP26 trials 4 and 11. The “perfect” value for each metric is given in the second column.

Metric	Perfect	Trial 4: 9 Nov 1996		Trial 11: 14 Nov 1996		Both trials	
		Obs–SCIPUFF	RTFDFA–SCIPUFF	Obs–SCIPUFF	RTFDFA–SCIPUFF	Obs–SCIPUFF	RTFDFA–SCIPUFF
FB	0.0	–1.3902	–1.2399	–1.4403	0.1281	–1.4214	–0.7581
MG	1.0	0.1253	0.2355	0.8779	2.7972	0.3316	0.8116
NMSE	0.0	7.0974	6.0116	13.0406	0.1045	11.9715	3.6575
VG	1.0	95.787	8.9700	11.3766	6.7695	33.0111	7.7925
FAC2	1.0	0.0000	0.0000	0.3333	0.3333	0.1667	0.1667



the regions of interest span several grid points, and the meteorological conditions at all grid points of interest need to be somehow included in assessing uncertainties of the combined RTFDDA–application model forecasts.

- The approach accounts for RTFDDA forecast errors in the timing and location of synoptic and mesoscale features (e.g., sea-breeze fronts) that would affect the secondary-applications results.
- The approach allows the results to be presented in terms of a probability of exceeding a given threshold, which is easy for test directors to understand and apply to everyday testing needs.

Certainly, using meteorological ensembles is more appealing and work is currently under way using high-performance clusters to compute 50–200 ensembles (Liu et al. 2007). Each of these ensemble members could be used to drive a secondary application. However, until this approach is feasible operationally, initial comparisons of this approach with the spatial ensemble approach have shown good overall agreement in the ensemble spread and therefore in the probability of exceedance metrics of the sound propagation model, bolstering our confidence in the veracity of the spatial ensemble approach.

Although three specific application models were presented here, in principle any secondary-applications model that requires meteorological input data could be used in this way. Examples include

- atmospheric refractivity and ducting models (e.g., Burk et al. 2003),
- dust transport models (e.g., Barnum et al. 2004),
- upper-level turbulence prediction models (e.g., Sharman et al. 2006),
- aircraft icing diagnoses (e.g., Bernstein et al. 2005),
- aircraft exhaust condensation trail (contrail) prediction (e.g., Appleman 1953; Schumann 1996), and
- ceiling and visibility models (e.g., Stoelinga and Warner 1999).

Note that model performance of a combined NWP model secondary application may be another way of evaluating mesogamma-scale forecast models in general, and may suggest possible modifications to model physics or configurations that are optimal compromises between range forecasting requirements and application programs requirements.

*Acknowledgments.* This research was funded by the U.S. Army Test and Evaluation Command through an interagency agreement with the National Science Foun-

ation. Carol Park provided editorial assistance. Careful readings of an earlier version of the manuscript by three anonymous reviewers and their helpful suggestions for improvements are much appreciated.

#### REFERENCES

- American National Standards Institute, 1983: Airblast characteristics for single point explosions in air with a guide to evaluation of atmospheric propagation and effects. ANSI S2.20-1983, Acoustic Society of America, 32 pp.
- Appleman, H. S., 1953: The formation of exhaust condensation trails by jet aircraft. *Bull. Amer. Meteor. Soc.*, **34**, 14–20.
- Barnum, B. H., and Coauthors, 2004: Forecasting dust storms using the CARMA-dust model and MM5 weather data. *Environ. Model. Softw.*, **19**, 129–140.
- Benjamin, S. G., and Coauthors, 2004: An hourly assimilation–forecast cycle: The RUC. *Mon. Wea. Rev.*, **132**, 495–518.
- Bernstein, B. C., F. McDonough, M. K. Politovich, B. G. Brown, T. P. Ratvasky, D. R. Miller, C. A. Wolff, and G. Cuning, 2005: Current icing potential: Algorithm description and comparison with aircraft observations. *J. Appl. Meteor.*, **44**, 969–986.
- Burk, S. D., T. Haack, L. T. Rogers, and L. J. Warner, 2003: Island wake dynamics and wake influence on the evaporation duct and radar propagation. *J. Appl. Meteor.*, **42**, 349–367.
- Chang, J. C., P. Franzese, K. Chayantrakom, and S. R. Hanna, 2003: Evaluations of CALPUFF, HPAC, and VLSTRACK with two mesoscale field datasets. *J. Appl. Meteor.*, **42**, 453–466.
- Clough, C., J. K. Leurs, and E. J. Hall, 2000: Development of an acoustic ray-trace model, high-resolution boundary-layer measurements, and meso- $\Gamma$ -scale forecasts driven by off-range, blast-noise management requirements. Preprints, *Ninth Conf. on Aviation, Range, and Aerospace Meteorology*, Orlando, FL, Amer. Meteor. Soc., 415–420.
- Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784.
- Etkin, B., 1972: *Dynamics of Atmospheric Flight*. John Wiley and Sons, 579 pp.
- Frehlich, R. G., 2006: Adaptive data assimilation to include the spatial variations in observation error. *Quart. J. Roy. Meteor. Soc.*, **132**, 1225–1257.
- , and R. Sharman, 2003: Improving the small scale turbulence structure for fluid dynamics computations. *Proc. 41st AIAA Aerospace Sciences Meeting and Exhibit*, Reno, NV, AIAA Paper 2003-0195.
- Grell, G. A., J. Dudhia, and D. R. Stauffer, 1995: A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR/TN-398, NCAR, 122 pp.
- Heimann, D., and G. Gross, 1999: Coupled simulation of meteorological parameters and sound levels in a narrow valley. *Appl. Acoust.*, **56**, 73–100.
- Hole, L. R., and H. M. Mohr, 1999: Modeling of sound propagation in the atmospheric boundary layer: Application of the MIUU mesoscale model. *J. Geophys. Res.*, **104**, 11 891–11 901.
- Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 342 pp.
- Kerry, G., D. J. Saunders, and A. G. Sills, 1987: The use of me-

- teological profiles to predict the peak sound-pressure level at distance from small explosions. *J. Acoust. Soc. Amer.*, **81**, 888–896.
- Krol, H. R., 1973: Intensity calculations along a single ray. *J. Acoust. Soc. Amer.*, **53**, 864–868.
- Lighthill, J., 1978: *Waves in Fluids*. Cambridge University Press, 504 pp.
- Liu, Y., M. Xu, J. Hacker, T. Warner, and S. Swerdlin, 2007: A WRF and MM5-based 4-D Mesoscale Ensemble Data Analysis and Prediction System (E-RTFDAA) developed for ATEC operational applications. Preprints, *18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 7B.7.
- , and Coauthors, 2008a: The operational mesogamma-scale analysis and forecast system of the U.S. Army Test and Evaluation Command. Part I: Overview of the modeling system, the forecast products, and how the products are used. *J. Appl. Meteor. Climatol.*, **47**, 1077–1092.
- , and Coauthors, 2008b: The operational mesogamma-scale analysis and forecast system of the U.S. Army Test and Evaluation Command. Part II: Interrange comparison of the accuracy of model analyses and forecasts. *J. Appl. Meteor. Climatol.*, **47**, 1093–1104.
- Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430.
- McKeen, S., and Coauthors, 2007: Evaluation of several PM<sub>2.5</sub> forecast models using data collected during the ICARTT/NEAQS 2004 field study. *J. Geophys. Res.*, **112**, D10S20, doi:10.1029/2006JD007608.
- Rife, D. L., C. A. Davis, and Y. Liu, 2004: Predictability of low-level winds by mesoscale meteorological models. *Mon. Wea. Rev.*, **132**, 2553–2569.
- Schomer, P. D., 2001: A statistical description of blast sound propagation. *Noise Control Eng. J.*, **49**, 79–87.
- , and G. A. Luz, 1994: A revised statistical analysis of blast sound propagation. *Noise Control Eng. J.*, **42**, 95–100.
- , L. R. Wagner, L. J. Benson, E. Buchta, K.-W. Hirsch, and D. Krahé, 1994: Human and community response to military sounds: Results from field-laboratory tests of small-arms, tracked vehicle, and blast sounds. *Noise Control Eng. J.*, **42**, 71–84.
- Schumann, U., 1996: On conditions for contrail formation from aircraft exhausts. *Meteor. Z.*, **5**, 4–23.
- Sharman, R., C. Tebaldi, G. Wiener, and J. Wolff, 2006: An integrated approach to mid- and upper-level turbulence forecasting. *Wea. Forecasting*, **21**, 268–287.
- Stauffer, D. R., and N. L. Seaman, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor.*, **33**, 416–434.
- Stoelinga, M. T., and T. T. Warner, 1999: Nonhydrostatic, meso-beta-scale model simulations of cloud ceiling and visibility for an East Coast winter precipitation event. *J. Appl. Meteor.*, **38**, 385–404.
- Sykes, R. I., and R. S. Gabruk, 1997: A second-order closure model for the effect of averaging time on turbulent plume dispersion. *J. Appl. Meteor.*, **36**, 1038–1045.
- Thompson, R. J., 1972: Ray theory for an inhomogeneous moving medium. *J. Acoust. Soc. Amer.*, **51**, 1675–1682.
- , 1974a: Ray-acoustic intensity in a moving medium I. *J. Acoust. Soc. Amer.*, **55**, 729–732.
- , 1974b: Ray-acoustic intensity in a moving medium II. *J. Acoust. Soc. Amer.*, **55**, 733–737.
- Turton, J. D., D. A. Bennetts, and D. J. W. Nazer, 1988a: The Larkhill noise assessment model. Part I: Theory and formulation. *Meteor. Mag.*, **117**, 145–154.
- , —, and —, 1988b: The Larkhill noise assessment model. Part II: Assessment and use. *Meteor. Mag.*, **117**, 169–179.
- Warner, T. T., R.-S. Sheu, J. F. Bowers, R. I. Sykes, G. C. Dodd, and D. S. Henn, 2002: Ensemble simulations with coupled atmospheric dynamic and dispersion models: Illustrating uncertainties in dosage simulations. *J. Appl. Meteor.*, **41**, 488–504.