Comparisons of Transport and Dispersion Model Predictions of the Joint Urban 2003 Field Experiment

STEVE WARNER, NATHAN PLATT, JEFFRY T. URBAN, AND JAMES F. HEAGY
Institute for Defense Analyses, Alexandria, Virginia

(Manuscript received 11 June 2007, in final form 9 October 2007)

ABSTRACT

For a hazardous material release in a city or densely populated area, effective mitigation requires an understanding of the transport and dispersion of these hazards in the complex urban environment. Improved characterization and understanding of urban transport and dispersion will allow for more robust modeling. The Defense Threat Reduction Agency has developed a Hazard Prediction Assessment Capability (HPAC) that includes features to address hazardous releases within an urban environment. During the summer of 2003, a series of tracer gas releases were carried out in Oklahoma City, Oklahoma, and extensive meteorological and tracer concentration measurements were collected in a field experiment known as Joint Urban 2003 (JU03). This analysis uses the observations of JU03 to evaluate “Urban HPAC.” Twenty sets of simulations, or “predictions,” using four Urban HPAC modes and five meteorological input options, were created and compared using a variety of metrics. Strong consistency was found between the conclusions of this study and those of two previously reported Urban HPAC evaluations. For example, improved predictions were associated with the inclusion of a simple empirically based urban dispersion model within HPAC, whereas improvements associated with the inclusion of a more computationally intensive urban wind module were not found. In this study, two new results are reported. First, there was a substantial difference in the performance of Urban HPAC as a function of release time—day or night—that was not discovered earlier because the previously examined urban field experiments focused on nighttime releases only. Daytime releases tended to be underpredicted and nighttime releases tended to be overpredicted. Second, and with respect to the under- and overpredictions described above, the inclusion of the new “Micro” Stationary Wind Fit and Turbulence (SWIFT) “SPRAY” (MSS) Urban HPAC mode typically led to less underprediction during the day and less overprediction at night than the other Urban HPAC modes. In addition, predictions that included MSS typically resulted in the least scatter between observations and predictions. These improvements warrant further investigation to determine whether this conclusion can be extended to other urban environments.

1. Introduction

The potential effects of the accidental or deliberate release of hazardous materials into the atmosphere continue to be of concern—a concern that is especially acute in more densely populated urban areas. Effective mitigation in urban settings will require an understanding of the transport and dispersion of these hazards in the urban environment. The U.S. Departments of Defense (DoD) and Homeland Security (DHS) require estimates of the effects of hazardous releases within an urban environment on the underlying population to aid planning, emergency response, and recovery efforts. Especially desired are estimates of where and when relatively low level human-effects thresholds are exceeded (i.e., hazard regions). Models that predict the transport and dispersion of hazardous materials in urban settings are needed to develop doctrine, plan counterproliferation operations, determine and characterize the source of the hazardous release based on limited observations, develop casualty estimates and emergency response plans, conduct forensic analyses, and, perhaps, help to guide near-real-time emergency response activities.

Improved characterization and understanding of urban transport and dispersion will allow for more robust modeling. The buildings (which are sometimes large...
and are often closely packed) located in urban regions serve as significant roughness elements that can disperse toxic materials in ways that are not completely understood. Other features of the urban environment, including traffic-induced turbulence, heat-island formation, flows associated with the deep street canyons of some large cities, the relative lack of moisture, and differential heating on building faces, also can have varied effects on the transport and dispersion of hazardous materials.

Modeling of urban transport and dispersion has been a subject of study since the late 1970s (Britter and Hanna 2003). There are many atmospheric transport and dispersion models available for application to open-field (rural area) hazardous gas releases, but only a few models have been developed specifically to simulate transport and dispersion in urban environments. Until recently, urban transport and dispersion models could be divided into two main categories: 1) low-fidelity models (e.g., urban canopy models) that account for the large-scale effects of urban terrain, such as drag from buildings and boundary layer perturbations, and 2) high-fidelity models that include detailed representations of buildings, streets, and other urban features and, in the case of computational fluid dynamics (CFD) models, integrate some form of the Navier–Stokes equations. Between the lower-fidelity and higher-fidelity models described above lies a class of turbulence and wind profile parameterizations derived from these urban details. Examples include the particle-based “MESO”/Realistic Urban Spread and Transport of Intrusive Contaminants (RUSTIC) model (Diehl et al. 1982, 2007; Hendricks et al. 2004, 2007; Burrows et al. 2004, 2007) and the Quick Urban Industrial Complex (QUIC) models—“QUIC-URB” (Pardyjak and Brown 2001) and “QUIC-PLUME” (Williams et al. 2004; Kaplan and Dinar 1996)—and the Urban Hazard Prediction Assessment Capability (HPAC) model (Defense Threat Reduction Agency 2001). In this paper, we evaluate on the urbanized version of HPAC (Urban HPAC), which has been the subject of validation efforts within our group for several years.

A few recent field experiments have included the release of environmentally safe, inert, tracer gases in urban environments. For example, tracer gases were released in Salt Lake City, Utah, in 2000 in a field experiment referred to as Urban 2000 (Allwine et al. 2002) and were released during the Mock Urban Setting Test (MUST) at Dugway Proving Ground, Utah (Biltoft 2002). An important use of the data collected during these field experiments is to provide support for the evaluation of transport and dispersion models. Data collected during Urban 2000 have been used to aid assessments of the validity of HPAC (Warner et al. 2004a; Chang et al. 2005). Likewise, data collected during the MUST field experiment have been used to evaluate Urban HPAC simulations (Warner et al. 2006).

a. Brief description of the Joint Urban 2003 (JU03) field experiment

Under the joint sponsorship of the DoD [through the Defense Threat Reduction Agency (DTRA)] and DHS, a series of tracer gas releases were carried out in Oklahoma City, Oklahoma, starting on 28 June and ending on 31 July 2003. This field experiment included 10 intensive operating periods (IOPs), in which the tracer gas sulfur hexafluoride (SF$_6$) was released in downtown Oklahoma City (Allwine et al. 2004). As part of JU03, 29 thirty-minute continuous SF$_6$ releases were accomplished during the 10 IOPs, with 2 h of sampler monitoring following the start of each release. JU03 represents a major urban study and included components that examined the following five subjects:

1) the urban atmospheric boundary layer (Berg et al. 2004; De Wekker et al. 2004),
2) flows within a street canyon to include traffic-induced turbulence (Brown et al. 2004a;b; Pol et al. 2004; Ramamurthy et al. 2004; Pol and Brown 2006),
3) flows within, and downwind (to 4 km) of, the tall-building core (Leach et al. 2004),
4) surface energy balance in an urban area (Grimmond et al. 2004; Zajic et al. 2004), and
5) dispersion of tracer into, out of, and within buildings (Allwine and Flaherty 2006).

This study examines the 29 continuous JU03 SF$_6$ releases and describes comparisons of the tracer gas sampler observations with Urban HPAC simulations, or “predictions.” Several sets of predictions based on different configurations of Urban HPAC and varying meteorological inputs are examined. The preparation of these predictions—for example, the protocols required to create comparable predictions—is also described. In addition, comparisons between sets of Urban HPAC predictions are examined and performance among configurations is assessed.

As shown in Fig. 1, four release locations, all in downtown Oklahoma City, were used for the 29 continuous SF$_6$ releases. The duration of each of these releases was 30 min, with sampler monitoring for 2 h after
the start of the release. The altitude for each of the releases was about 1.9 m above ground level (AGL). Release rates varied from 1.9 to 5.0 g s$^{-1}$, indicating that total released SF$_6$ masses varied from 3.42 to 9.00 kg per release.

This study focused on comparisons of the spatial distribution of predicted and observed concentrations near the surface, both in the central business district (CBD; i.e., the city center near the release points) and on outer arcs at 1, 2, and 4 km. We chose to use data collected by the programmable integrating gas samplers (PIGS) and enhanced PIGS, referred to as SuperPIGS, which, for the vast majority, provided 30-min average concentration observations. These samplers provided relatively complete, dense, long- and short-range, ground-based coverage. In addition, all of the PIGS and SuperPIGS were analyzed at the same laboratory and were processed following the same protocols, thus mitigating concerns about cross correlation between different sampler types or analysis procedures. This study did not consider measured and modeled concentrations for averaging times shorter than 30 min and therefore cannot be directly applied to assessments of human-effects models that might require detailed information on shorter-term concentration exposures. Figure 2 shows the typical locations of the SF$_6$ ground-based samplers that were used. These “ground based” samplers were located at approximately 3 m AGL and were typically mounted on light poles. Fifty-five samplers were located in the CBD, and 23 were located on an arc located about 1 km from the release points. In a similar way, 21 samplers were located on a 2-km arc and 21 were located on the 4-km arc. Additional description of the releases and samplers can be found in Warner et al. (2007).
b. Brief description of Urban HPAC configurations examined in this study

The first goal of this study was to create predictions of JU03 using the HPAC software. A comprehensive set of HPAC predictions of the JU03 field experiment has not previously been reported. DTRA’s HPAC is composed of a suite of software modules that can generate source terms for hazardous releases, retrieve and prepare meteorological information for use in a prediction, model the transport and dispersion of the hazardous release over time, and plot and report the results of these calculations. HPAC has been applied to various national defense problems, including military studies and operational planning. For transport and dispersion of hazardous material, HPAC uses the Second-Order Closure Integrated Puff (SCIPUFF) model and an associated mean wind field model (Sykes et al. 1996). Within HPAC, two weather modules can be used to prepare mass-consistent gridded wind fields: Stationary Wind Fit and Turbulence (SWIFT) and a mass-consistent SCIPUFF (MC-SCIPUFF) algorithm. Note that neither SWIFT nor MC-SCIPUFF accounts explicitly for wind speed profiles below the mean building height.

For this study, four types of Urban HPAC predictions are discussed: 1) HPAC (version 4.04 Service Pack 3—the HPAC version available at the time of the study) with surface type entered as “urban” [referred to as baseline or UC (for urban canopy)], 2) HPAC with the Urban Dispersion Model (UDM) toggled on (denoted DM), 3) HPAC with the Urban Windfield Module (UWM) and UDM toggled on (denoted DW), and 4) the newest Urban HPAC component, “Micro” SWIFT “SPRAY” (MSS) (denoted MS).

For fast-running urban applications of HPAC—in particular, when a detailed urban database is not avail-
able, the vertical wind and turbulence profiles can be modified to account for urban effects. The basic exponential profile shape used is described in Cionco (1972) and has been shown to fit a variety of canopy types, including plant canopies as well as discrete objects. This mode of operation (UC) is considered as a baseline for comparison and is further described in Warner et al. (2007).

The UDM component of Urban HPAC was created by the U.K. Defence Science and Technology Laboratory (Hall et al. 2002). The UDM is designed to compute the transport and dispersion of an instantaneous discharge [a “puff” or train(s) of puffs] of pollutant released over a surface containing a mixture of open and urban areas. UDM considers variations in dispersion rates as a function of surface changes and direct interaction of the pollutant cloud with surface obstacles. UDM is based on ensemble-mean Gaussian puff dispersion methods but allows surface obstacles to modify the dispersion patterns. UDM’s empirical parameters are set based on extensive wind-tunnel experiments.

The UWM model, resident within Urban HPAC, predicts steady-state winds (speed and direction) inside the urban boundary layer using a canopy parameterization (Lim et al. 2003; Sykes and Henn 1989). UWM-generated average winds can then be used by Urban HPAC (with or without UDM toggled on) to drive material transport and dispersion. UWM is designed to provide a computationally fast solution, within a CFD framework, by considering spatially averaged obstacle effects.

MSS is a new component of Urban HPAC (introduced in 2007). It is meant to provide a fast computation of the wind field within the urban environment while accounting for a detailed representation of the buildings [e.g., as generated by geographic information systems; Moussafir et al. (2004)], and it is based on the work of Röckle (1990) and Kaplan and Dinar (1996). Given information about the local buildings (locations, shapes, and sizes), Micro SWIFT creates a modified wind field by creating zones in which the flow is modified according to the buildings’ locations, and the flow is adjusted to satisfy the continuity equation and impermeability on the ground and on building walls. Micro SWIFT also derives a diagnostic turbulence—diffusive coefficients and the turbulent kinetic energy dissipation rate—by considering the distance to the nearest obstacle as a mixing length.

Micro SPRAy is a Lagrangian particle dispersion model, derived from SPRAy (Tinarelli et al. 1994), that can account for obstacles. Dispersion is simulated by following the movement of a large number of particles, each representing a portion of the original released mass. The motions of the particles are obtained by considering two components: a mean component, which follows the local winds, and a stochastic component (Thomson 1987). Within HPAC, MSS is used to compute transport and dispersion of hazardous material over the first several hundred meters from the release point (at least, as we employed MSS). Beyond this distance, concentration information in the form of Gaussian puffs is “handed off” to SCIPUFF for the rest of the longer-range calculation. We used a horizontal grid spacing of 3 m over a square domain with 1-km sides, 31 vertical grid points with a 2-m spacing for the first 11 grid points (near the surface) and stretched out geometrically for the next 20 grid points to a maximum domain height of about 700 m, and 100 000 particles for the Micro SPRAy simulation.

2. Study method

a. Meteorological input options used for Urban HPAC mode comparisons

A large variety of meteorological measurements was collected during JU03. The main goal of this paper is to describe the comparative results of JU03 predictions created with varying Urban HPAC configurations—UC, DM, DW, and MS. Several meteorological (“MET”) input options were examined, in part, to understand how relative Urban HPAC mode performance varies for different reasonable MET options. This section provides a brief description of the five MET input options that were selected. The shorthand notations for the five MET options are BAS, GCT, PNA, ACA, and PO7 (described below). Additional details associated with these MET options are provided in Warner et al. (2007).

The BAS MET option was designed to correspond to a baseline situation in which the meteorological information is consistent with what could have been retrieved from the DTRA meteorological server at some point (~2 h or more) after the release. That is, this information corresponds to assimilated observations during the release as opposed to forecast information. Meteorological sources within 30 km of Oklahoma City were considered for BAS. Surface wind velocity observations from four stations between 12 and 28 km from the downtown area and upper-air wind velocity observations from a station 28 km southeast of Oklahoma City were used as shown in Fig. 3. The diagnostic wind field model SWIFT was invoked to create gridded wind fields for the BAS MET option.

The GCT MET option was designed to correspond to
a baseline situation in which a gridded numerical weather assimilation was used as input to HPAC. The GCT MET option corresponds to a Global Climatological Analysis Tool (GCAT) prediction of wind velocity profiles (i.e., wind velocity as a function of height AGL) at many grid locations. These files can be thought of as surrogates for gridded numerical weather assimilations that could be available several hours after an event. GCAT was developed by the National Center for Atmospheric Research (NCAR) and the National Ground Intelligence Center. GCAT combines the fifth-generation Pennsylvania State University–NCAR (regional atmospheric) Mesoscale Model (MM5) with the four-dimensional data assimilation technique to produce finescale climatological analyses anywhere in the world (Vandenberghe et al. 2006). This approach incorporates observations that are available at irregular times and locations. For GCT, GCAT was run with 36
vertical levels, with the first level at 20 m AGL, and a horizontal grid spacing of 3.3 km. The observations from all four BAS stations were included as input to create the GCT MET option. SWIFT was invoked to create gridded wind fields for the GCT MET option.

The PNA MET option corresponds to using both the sodar and vertical profiler observations that were available at a meteorological site located 1.6 km upwind from the release points and operated by the Pacific Northwest National Laboratory. Figure 4 shows the location of the PNA site. The sodar and profiler provide wind speeds and directions (two dimensional) as a function of altitude at a single geographic location—that is, a “vertical profile.” Previous studies of Urban HPAC suggested that a single measured upwind vertical profile could provide a satisfactory input for the creation of urban predictions (Warner et al. 2004a). Thus, the PNA MET option allows further testing of this hypothesis.

For the PNA MET option, MC-SCIPUFF was invoked instead of SWIFT to create gridded wind fields because SWIFT intermittently aborted for some of the 29 releases (Warner et al. 2007). SWIFT software issues related to this problem continue to be investigated by the developer.

The ACA MET option corresponds to using both the sodar and profiler observations that were available at a meteorological site located 4 km downwind from the release points and operated by the Argonne National Laboratory (see Fig. 4). This MET option was considered to be particularly useful for comparison and contrast with the PNA upwind site option. For the ACA MET option, MC-SCIPUFF was invoked to create gridded wind fields, for the same reason that it was used for the PNA option.

The PO7 MET option corresponds to a set of observations from a single location 40 m AGL on the roof of
the Oklahoma City Post Office building, just upwind of the downtown urban environment (Fig. 4). A portable weather information and display system was located on a tower 25 m above the five-story building roof and collected (two dimensional) wind speed and direction data every 10 s. These observations were vector averaged over 15 min and were time tagged with the midpoint time of the 15-min interval (Warner et al. 2007). This option corresponds to using a single close-in observation as input for the Urban HPAC predictions. During a previous study of urban atmospheric transport and dispersion (Urban 2000; Warner et al. 2004a), it was found that using a single downtown building-top measurement resulted in relatively degraded predictions in terms of fits to the observations when compared with the other MET options. Therefore, the PO7 MET option allows for the reconsideration of this concept, that is, using a set of observations from a single building top as input for hazardous material transport and dispersion predictions, albeit this time an upwind building as opposed to a downtown building. SWIFT was invoked to create gridded wind fields for the PO7 MET option. Additional description of the MET options used for this study can be found in Warner et al. (2007). A summary of the 20 sets of JU03 Urban HPAC predictions that were compared is listed in Table 1.

### b. Summary of metrics used for this comparative study

The methods used for the comparisons in this study have been previously described in Warner et al. (2004a). Various statistical metrics to assess bias, scatter, and correlation were examined as well as a user-oriented measure of effectiveness (MOE; Warner et al. 2004b) that allowed for assessments of the ability of the model to predict the “hazardous” region (i.e., the region above a concentration threshold of interest). The metrics discussed in this paper are fractional bias (FB), normalized absolute difference (NAD), normalized mean square error (NMSE), fraction of predictions within a factor of 2 (FAC2), and the MOE. These metrics are defined as

\[
FB = \frac{(C_p - C_o)}{0.5(C_o + C_p)},
\]

(1)

\[
NAD = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{C_p(i) - C_o(i)}{C_p(i) + C_o(i)} \right|
\]

(2)

\[
NMSE = \frac{(C_p - C_o)^2}{C_o C_p}
\]

(3)

\[
FAC2 = \text{fraction of data for which } \frac{1}{2} \leq \frac{C_p}{C_o} \leq 2
\]

and

\[
MOE = (x, y) = \frac{A_{OV}}{A_{OB}} - \frac{A_{OV}}{A_{PR}} = \left(1 - \frac{A_{FN}}{A_{OB}}, 1 - \frac{A_{FP}}{A_{PR}} \right).
\]

(4)

where \(C\) is the observation/prediction of interest (e.g., concentration), \(C_p\) corresponds to model predictions, \(C_o\) corresponds to observations, a bar above the quantity (e.g., \(C\)) denotes the average, \(n\) is the number of data points used in the comparisons, \(C_o(i)\) refers to the \(i\)th observed concentration, \(C_p(i)\) refers to the \(i\)th predicted concentration, \(A_{FN}\) is the region of false negative, \(A_{FP}\) is the region of false positive, \(A_{OV}\) is the region of overlap, \(A_{PR} = A_{FP} + A_{OV}\) is the region of the prediction, and \(A_{OB} = A_{FN} + A_{OV}\) is the region of the observation.

FB measures the systematic bias in a model in terms of absolute differences: \(FB > 0\) indicates overprediction, and \(FB < 0\) indicates underprediction. NMSE and NAD measure the scatter associated with the predictions relative to observations. Perfect agreement with a set of observations would result in \(FB, NMSE, \) and \(NAD = 0.0\) and \(FAC2 = 1.0\).

The MOE described above has two dimensions and includes directional effects; that is, the prediction of the location of a hazard and not just the shape and size of the plume is critical to obtaining a high MOE “score.” The perfect MOE score is \((1, 1)\) and indicates complete overlap of the predictions and the observations. From Eq. (5) it can be seen that MOE values along the “diagonal” of the two-dimensional MOE space indicate equal sizes (e.g., areas or amounts of material) of the prediction and the observation (i.e., \(A_{PR} = A_{OB}\)), even if the locations differ. The quantities \(A_{FN}\), \(A_{FP}\), and \(A_{OV}\) can be computed directly from the predictions and field trial observations paired in space and time. These
Table 2. Urban HPAC modes, for five MET input options, that led to improved predictive performance of JU03 releases based on measures of predicted/observed scatter (c-MOE, NAD, and NMSE).*

<table>
<thead>
<tr>
<th>Condition</th>
<th>BAS (SWIFT)</th>
<th>GCT (SWIFT)</th>
<th>PO7 (SWIFT)</th>
<th>PNA (MC-SCIPUFF)</th>
<th>ACA (MC-SCIPUFF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Day CBD</td>
<td>DW/DM</td>
<td>Mixed</td>
<td>Mixed</td>
<td>(UC, DM, DW)/MS</td>
<td>DW/MS</td>
</tr>
<tr>
<td>Day arcs (CBD)</td>
<td>(MS, DW)/(UC)</td>
<td>Mixed</td>
<td>Mixed</td>
<td>(MS, DM, DW)/(UC)</td>
<td>No differences</td>
</tr>
<tr>
<td>Night CBD</td>
<td>(MS, DM, DW)/(UC)</td>
<td>Mixed</td>
<td>Mixed</td>
<td>(MS, DM, DW)/(UC)</td>
<td>No differences</td>
</tr>
<tr>
<td>Night arcs</td>
<td>(MS, DM, DW)/(UC)</td>
<td>Mixed</td>
<td>Mixed</td>
<td>(MS, DM, DW)/(UC)</td>
<td>MS/(UC, DM)</td>
</tr>
</tbody>
</table>

* The XX/YY nomenclature denotes that model(s) XX had statistically significant relative improvement over model(s) YY. For this table, the designation means that, for at least two of the three scatter-related metrics (c-MOE, NAD, and NMSE), XX showed a statistically significant improvement relative to YY. The word mixed means that there was not a consistent finding of one model or models over others.

quantities can be computed for comparisons of average concentration and for exceeding threshold values. Additional discussions associated with the computation of average concentration–based and threshold-based MOE values can be found in Warner et al. (2004a,b). Nonparametric hypothesis test methods for detecting statistically significant differences between sets of predictions for a given metric and procedures for estimating confidence intervals are described in Warner et al. (2004a,b, 2006). In addition, and as important as any metric, comparative plots of model predictions and observations, including bar charts and contour plots, were created and scrutinized for all releases (Warner et al. 2007).

For each comparison, comparisons of 30-min average concentration were examined in the CBD and separately for all of the arc-based samplers together (1-, 2-, and 4-km arcs). It was recognized early in the analysis that model performance varied greatly as a function of the release time (i.e., day or night). Therefore, analyses were done separately for the day and night releases; hence, four situations were examined: daytime releases with samplers in the CBD, daytime releases with samplers on the arcs, nighttime releases with samplers in the CBD, and nighttime releases with samplers on the arcs. There were 17 daytime and 12 nighttime continuous 30-min releases.

For this study, the above metrics were computed and examined and the inherent limitations of each were recognized. It is expected that the measures, taken together, can provide good insight into the relative performance of a given model. For example, NMSE assesses scatter between observations and predictions but considers normalized squared differences as opposed to the absolute differences, reflected in NAD. These differences in scatter metrics can result in somewhat different sensitivities (Warner et al. 2004a). In the results section of this paper, we consider three metrics—NAD, NMSE, and the concentration-based MOE (c-MOE)—as measuring scatter between predictions and observations, and scatter is considered to be an important measure of a model’s predictive performance. When applied to observations and predictions paired in space and time, these scatter-based metrics get at the question of how well the model predicted the location and timing (at least for the 30-min averages here) of the observations. We demand that at least two of the three scatter-related metrics reveal statistically significant improvements, with a significance level ($p$ value) of 0.05, before we declare that one mode outperforms another for a particular situation. Note that, for the three scatter-related metrics discussed above, we examined 360 hypothesis tests ($360 = 3 \times 6 \times 6$ pairwise comparisons among the 4 modes $\times 4$ combinations of day/night and CBD/arc conditions $\times 5$ MET options), and if, for example, one demanded a significance level of 0.05 before highlighting a difference as an improvement then one would expect about 18 false significance findings. However, by demanding that two of three differences result in statistical significance, these false-positive findings are greatly reduced, perhaps to as low as about three.

3. Results and discussion

a. Urban HPAC mode comparisons

Table 2 shows, for each of the five MET input options that were considered, the Urban HPAC modes that resulted in the least scatter—that is, the best fit to the observations. In addition, Table 2 identifies the Urban HPAC modes that resulted in relative (and statistically significant) improvement for the five MET options and the four conditions (day and night, CBD and arcs) that were examined. For example, the “(MS, DM,
entry in the cell at the intersection of the BAS column and the night CBD row in Table 2 indicates that, for two of the three scatter-related metrics, MS, DM, and DW outperformed UC. Several specific findings are discussed below.

1) DAY VERSUS NIGHT RELEASES

There was a substantial difference in the performance of Urban HPAC as a function of day and night. Figure 5 shows comparative NAD results for the five MET input options and four Urban HPAC modes for both day and night and for the CBD samplers. For the SWIFT-associated MET options (BAS, GCT, and PO7), Urban HPAC predictions resulted in substantially more scatter at night than during the day, with the exception of MS. For the MC-SCIPUFF-associated MET input options (PNA and ACA), the scatter results were much more similar for the day and night Urban HPAC predictions, with perhaps some evidence of improved performance during the day. Examinations of arc-based results showed similar behavior to that described above for the CBD. A full discussion of arc-based specific results as a function of Urban HPAC mode and MET option can be found in Warner et al. 2007.

For all five MET options, daytime releases tended to be underpredicted and nighttime releases tended to be overpredicted in the CBD and on the arcs. Figure 6 compares day and night FB values for the CBD samplers and for the 20 sets of predictions. Examinations of arc-based results showed behavior that was similar to that described above for the CBD.

2) MSS MODEL PERFORMANCE BEHAVIOR DIFFERS FROM OTHER URBAN HPAC MODES

With respect to the under- and overpredictions described above, the MS mode typically led to less under-
prediction during the day (in fact, slight overprediction) and less overprediction at night than did the other Urban HPAC modes. There were some minor exceptions in which the DM and DW modes were similar to MS. The MS mode typically resulted in the least biased predictions of the 30-min average concentrations at the surface samplers (CBD and arcs).

Tables 3 and 4 list the NAD and NMSE values, respectively, from least scatter (best) to most scatter (worst). MS-based predictions resulted in the least scatter in seven of the eight categories (day/night, CBD/arcs, and NAD/NMSE). The prediction for PO7_DW is the sole exception to the above, having the best NAD value for the day–CBD condition.

Although it typically took less than 5 min to generate a UC- or DM-mode prediction of a single release (tracking the plume for 2 h), MS predictions, at the resolutions used here (Warner et al. 2007), took, on average, about 60 min per release. We also created a lower-resolution set of MS predictions using a horizontal grid spacing of 5 m over a 0.8 km × 0.8 km domain, 21 vertical grid points (vs 31 for the higher resolution), and 50 000 particles (Warner et al. 2007) for which each prediction took about 30 min, and the results were substantially similar to those reported here for the higher-resolution runs.

3) RELATIVE URBAN HPAC MODE PERFORMANCE FOR NIGHTTIME RELEASES: MS, DM, AND DW REPRESENTED IMPROVEMENTS OVER UC

An additional important result is that for the nighttime releases the MS, DM, and DW modes offer improvement over the UC mode for the three MET input options that invoked SWIFT. This result is true for the samplers in the CBD and along the arcs. This result can be considered especially important because SWIFT corresponds to the recommended and default mode of operating Urban HPAC. In addition, these MET options—in particular, BAS and GCT—appear to correspond to reasonably realistic and potential operational applications of Urban HPAC. We also found that adding UWM to UDM to create the DW

Fig. 6. As in Fig. 5, but for FB.
mode did not lead to substantial or consistent significant improvements relative to using UDM alone (i.e., DM). This result is entirely consistent with past studies of the Urban 2000 (Warner et al. 2004a) and MUST (Warner et al. 2006) field trials. Note also that the DW predictions (as we ran them) took, on average, approximately 80 min longer per release than the corresponding DM prediction. These results, and past findings, call into question the value of including UWM.

For the nighttime releases and the MC-SCIPUFF-associated MET options (PNA and ACA) results were mixed, with no Urban HPAC mode consistently offering improvement, although the MS mode did so for the ACA MET option when considering the arc-based samplers relative to UC and DM.

### Table 3. NAD values for 20 sets of Urban HPAC predictions of JU03 ordered from least to most scatter. Values are shown for day and night and for CBD and arcs.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Day mode</th>
<th>CBD MET</th>
<th>Metric NAD</th>
<th>Night mode</th>
<th>CBD MET</th>
<th>Metric NAD</th>
<th>Day mode</th>
<th>Arcs MET</th>
<th>Metric NAD</th>
<th>Night mode</th>
<th>Arcs MET</th>
<th>Metric NAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>DW PO7</td>
<td>0.42</td>
<td>MS PO7</td>
<td>0.49</td>
<td>MS PNA</td>
<td>0.31</td>
<td>MS ACA</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>UC PO7</td>
<td>0.44</td>
<td>MS GCT</td>
<td>0.57</td>
<td>DW PNA</td>
<td>0.32</td>
<td>DW ACA</td>
<td>0.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>UC BAS</td>
<td>0.48</td>
<td>DM PNA</td>
<td>0.58</td>
<td>DW PO7</td>
<td>0.34</td>
<td>UC ACA</td>
<td>0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>DW BAS</td>
<td>0.49</td>
<td>UC ACA</td>
<td>0.59</td>
<td>UC PNA</td>
<td>0.34</td>
<td>DM ACA</td>
<td>0.41</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>DM PO7</td>
<td>0.49</td>
<td>MS BAS</td>
<td>0.60</td>
<td>UC PO7</td>
<td>0.35</td>
<td>DM PNA</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>MS PO7</td>
<td>0.49</td>
<td>MS ACA</td>
<td>0.60</td>
<td>DM PO7</td>
<td>0.35</td>
<td>DW PNA</td>
<td>0.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>UC GCT</td>
<td>0.53</td>
<td>DM ACA</td>
<td>0.61</td>
<td>MS PO7</td>
<td>0.36</td>
<td>MS PNA</td>
<td>0.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MS BAS</td>
<td>0.54</td>
<td>DW ACA</td>
<td>0.61</td>
<td>DM PNA</td>
<td>0.36</td>
<td>MS PO7</td>
<td>0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>DW GCT</td>
<td>0.54</td>
<td>DM GCT</td>
<td>0.62</td>
<td>MS BAS</td>
<td>0.37</td>
<td>UC PNA</td>
<td>0.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>DW PNA</td>
<td>0.55</td>
<td>DW PNA</td>
<td>0.62</td>
<td>DW ACA</td>
<td>0.39</td>
<td>DM GCT</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>UC PNA</td>
<td>0.55</td>
<td>MS PNA</td>
<td>0.63</td>
<td>MS ACA</td>
<td>0.40</td>
<td>MS GCT</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>DW ACA</td>
<td>0.56</td>
<td>UC PNA</td>
<td>0.64</td>
<td>DM ACA</td>
<td>0.41</td>
<td>DW GCT</td>
<td>0.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>MS GCT</td>
<td>0.57</td>
<td>DM PO7</td>
<td>0.67</td>
<td>UC ACA</td>
<td>0.41</td>
<td>DM BAS</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>DM BAS</td>
<td>0.57</td>
<td>DM MAS</td>
<td>0.68</td>
<td>DW BAS</td>
<td>0.41</td>
<td>DM PO7</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>DM GCT</td>
<td>0.57</td>
<td>DW GCT</td>
<td>0.70</td>
<td>UC BAS</td>
<td>0.44</td>
<td>MS BAS</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>UC ACA</td>
<td>0.59</td>
<td>DW PO7</td>
<td>0.73</td>
<td>UC GCT</td>
<td>0.44</td>
<td>DW PO7</td>
<td>0.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>DM ACA</td>
<td>0.60</td>
<td>DW BAS</td>
<td>0.78</td>
<td>DW GCT</td>
<td>0.45</td>
<td>DW BAS</td>
<td>0.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>DM PNA</td>
<td>0.62</td>
<td>UC GCT</td>
<td>0.86</td>
<td>DM BAS</td>
<td>0.46</td>
<td>UC GCT</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>MS PNA</td>
<td>0.66</td>
<td>UC PO7</td>
<td>0.86</td>
<td>DM GCT</td>
<td>0.47</td>
<td>UC PO7</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>MS ACA</td>
<td>0.69</td>
<td>UC BAS</td>
<td>0.91</td>
<td>MS GCT</td>
<td>0.57</td>
<td>UC BAS</td>
<td>0.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4. As in Table 3, but for NMSE.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Day mode</th>
<th>CBD MET</th>
<th>Metric NAD</th>
<th>Night mode</th>
<th>CBD MET</th>
<th>Metric NAD</th>
<th>Day mode</th>
<th>Arcs MET</th>
<th>Metric NAD</th>
<th>Night mode</th>
<th>Arcs MET</th>
<th>Metric NAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MS PO7</td>
<td>16.91</td>
<td>MS PO7</td>
<td>27.68</td>
<td>MS PNA</td>
<td>7.69</td>
<td>MS ACA</td>
<td>9.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>DW BAS</td>
<td>17.68</td>
<td>MS BAS</td>
<td>28.58</td>
<td>MS PO7</td>
<td>9.34</td>
<td>MS PNA</td>
<td>11.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>MS GCT</td>
<td>17.94</td>
<td>MS GCT</td>
<td>33.38</td>
<td>DW PNA</td>
<td>9.40</td>
<td>DW ACA</td>
<td>12.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>DW PO7</td>
<td>20.08</td>
<td>UC PNA</td>
<td>60.41</td>
<td>MS BAS</td>
<td>10.29</td>
<td>DM ACA</td>
<td>12.47</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>DW GCT</td>
<td>20.61</td>
<td>MS PNA</td>
<td>63.36</td>
<td>DM PO7</td>
<td>10.47</td>
<td>DM ACA</td>
<td>13.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>DM GCT</td>
<td>21.08</td>
<td>UC ACA</td>
<td>71.10</td>
<td>DW PO7</td>
<td>11.27</td>
<td>DW PNA</td>
<td>14.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>DM PO7</td>
<td>22.70</td>
<td>MS ACA</td>
<td>77.88</td>
<td>UC PNA</td>
<td>11.54</td>
<td>UC PNA</td>
<td>14.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>MS BAS</td>
<td>24.37</td>
<td>MS PNA</td>
<td>78.59</td>
<td>UC PO7</td>
<td>11.76</td>
<td>UC ACA</td>
<td>14.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>UC BAS</td>
<td>27.22</td>
<td>DW ACA</td>
<td>90.38</td>
<td>DM PNA</td>
<td>11.80</td>
<td>MS PO7</td>
<td>14.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>UC GCT</td>
<td>29.07</td>
<td>DW PNA</td>
<td>91.96</td>
<td>DM GCT</td>
<td>11.85</td>
<td>MS BAS</td>
<td>18.32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>UC PO7</td>
<td>29.33</td>
<td>DM ACA</td>
<td>96.66</td>
<td>DW GCT</td>
<td>12.11</td>
<td>DM BAS</td>
<td>24.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>DM BAS</td>
<td>31.48</td>
<td>DM GCT</td>
<td>133.79</td>
<td>DM ACA</td>
<td>12.88</td>
<td>DM GCT</td>
<td>26.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>DM PNA</td>
<td>33.17</td>
<td>DM PO7</td>
<td>143.00</td>
<td>DW BAS</td>
<td>13.07</td>
<td>MS GCT</td>
<td>33.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>DW ACA</td>
<td>33.97</td>
<td>DW GCT</td>
<td>150.00</td>
<td>DM BAS</td>
<td>13.42</td>
<td>DW GCT</td>
<td>34.99</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>DM ACA</td>
<td>35.84</td>
<td>DM PO7</td>
<td>159.87</td>
<td>UC GCT</td>
<td>13.45</td>
<td>DM PO7</td>
<td>35.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>DM PNA</td>
<td>38.81</td>
<td>DM BAS</td>
<td>185.90</td>
<td>DW ACA</td>
<td>15.30</td>
<td>DW PO7</td>
<td>49.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>UC PNA</td>
<td>45.22</td>
<td>DW BAS</td>
<td>218.92</td>
<td>MS ACA</td>
<td>15.57</td>
<td>DW BAS</td>
<td>50.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>UC ACA</td>
<td>50.18</td>
<td>UC BAS</td>
<td>346.48</td>
<td>UC ACA</td>
<td>17.04</td>
<td>UC GCT</td>
<td>120.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>MS ACA</td>
<td>100.40</td>
<td>UC GCT</td>
<td>488.84</td>
<td>MS GCT</td>
<td>17.94</td>
<td>UC PO7</td>
<td>143.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>MS PNA</td>
<td>119.28</td>
<td>UC PO7</td>
<td>506.26</td>
<td>UC BAS</td>
<td>19.18</td>
<td>UC BAS</td>
<td>162.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4) Relative Urban HPAC Mode Performance for Daytime Releases was Mixed and Inconsistent

For the daytime releases, no consistent trend was found. For example, for the BAS-associated predictions on the arcs, the MS and DW modes offered improvement (e.g., less scatter) over the UC mode, but, for the PNA-associated predictions in the CBD, the UC, DM, and DW modes resulted in improved scatter relative to the MS mode. However, in the latter PNA-based case, the observed improvements in scatter for the UC, DM, and DW predictions came at the cost of a large under-prediction relative to MS.

b. Concentration-based versus threshold-based MOE values

Predictions of exceeding a relatively low concentration threshold \([5 \times \text{background} \text{ and } 50 \times \text{background}]\), with the \(\text{SF}_6\) background assumed to be 5 ppt (parts per trillion); Warner et al. 2007\] were substantially more accurate (as measured by the MOE) than predictions of absolute 30-min average concentrations. Figure 7 shows an example comparison of the concentration-based and threshold-based MOE values for the BAS_MS predictions. MOE point estimates (of the average) for the 17 daytime and 12 nighttime releases are shown, and substantial movement of the MOE values toward the per-

![Figure 7. Comparisons of c-MOE and threshold-based MOE (25-ppt MOE and 250-ppt MOE) values for BAS_MS predictions of the daytime (red) and nighttime (blue) releases of JU03 within the CBD.](image-url)
fect value of (1, 1) is observed. These results, and similar ones found for the other Urban HPAC mode–MET option combinations, indicate substantial improvements when the thresholds are examined. This result is consistent with past studies of the Urban 2000 (Warner et al. 2004a) and MUST (Warner et al. 2006) field trials. An important implication of the above finding is that using Urban HPAC to predict the extent (in time and space) to which a relatively low threshold is (or might be) exceeded is likely to lead to a more accurate representation of a hazardous release than using Urban HPAC to predict the actual concentrations in time and space (e.g., perhaps needed for a detailed and complete assessment of expected casualties, given a human-effects model that requires concentration–time histories as a function of location). The basic finding is that, even in urban environments, the prediction of low-level hazard areas can be accomplished with relatively simple dispersion models (e.g., UDM). Although this might not be particularly surprising, it must be noted that, for many of the potential users of Urban HPAC and many of their specific operational applications, low-level hazard areas are of critical interest. This conclusion gets at the heart of the question of what level of performance is “good enough” (i.e., acceptable false-positive and false-negative fractions) for certain applications.

c. Brief comparison of MET options

As has been previously discussed, day releases were generally underpredicted and night releases were generally overpredicted both within the CBD and on the arcs. The MS mode represented an exception to the above—in particular for the CBD samplers. A substantial difference between the two MC-SCIPUFF-associated MET options (PNA and ACA) and the SWIFT-associated MET options (BAS, GCT, and PO7) was revealed. For the UC, DM, and DW modes, the MC-SCIPUFF-based MET options resulted in a smaller amount of material being predicted at the surface samplers (e.g., less overprediction at night and more underprediction during the day) relative to the other MET options. In addition, examination of Tables 3 and 4 shows that, for the nighttime releases on the arcs, the best 7 and 8 (of 20) NAD and NMSE values, respectively, were associated with the MC-SCIPUFF-based predictions.

We note that the current recommended default meteorological preprocessor for HPAC is SWIFT, and thus we create and compare Urban HPAC predictions that use SWIFT whenever possible. However, we further investigated these differences by creating additional sets of JU03 Urban HPAC predictions. For each of the SWIFT-based options (BAS, GCT, and PO7), a corresponding set (designated BASM, GCTM, and PO7M, respectively) was created using MC-SCIPUFF as the meteorological preprocessor. For this particular analysis, we are interested in detecting model performance differences that can be attributed to the meteorological preprocessor (SWIFT or MC-SCIPUFF). We examined all four dispersion modes (UC, DM, DW, and MS), but we focus these discussions on UC to allow for the clearest evaluation of differences between SWIFT- and MC-SCIPUFF-generated predictions. We do this because DW (through UWM) and MS (through Micro SWIFT) utilize their own additional wind models that are initiated by gridded SWIFT or MC-SCIPUFF outputs to predict the urban winds. Furthermore, DM (through UDM) uses empirical fits to slow down and widen the predicted plumes. Figure 8 compares FB (left panel) and NAD (right panel) for UC-based daytime and nighttime predictions. Significant differences between the SWIFT-based and MC-SCIPUFF-based predictions can be seen. First, the SWIFT-based FB values always show more material at the surface samplers—that is, somewhat less underprediction during the day and substantially more overprediction at night—relative to the corresponding MC-SCIPUFF-based FB values. With respect to scatter, differences in day/night behavior are observed between the predictions based on the differing meteorological preprocessors. For PO7M, BASM, and GCTM, there is little difference between model performance during the day and at night. However, for the SWIFT-based predictions, the daytime results show substantially lower (improved) NAD values relative to the corresponding nighttime predictions. This finding that the meteorological preprocessor can have a significant impact on Urban HPAC prediction performance is in apparent disagreement with Cox et al. 2005, whose study found that the wind fields created by SWIFT and MC-SCIPUFF were in essential agreement. However, there are substantial differences between Cox et al. 2005 and our current effort. First, we use the current version of Urban HPAC that integrates SWIFT and MC-SCIPUFF. Next, we compare predicted concentrations output by Urban HPAC, not meteorological preprocessor wind field outputs—in essence, for this study we treat MC-SCIPUFF and SWIFT like “black boxes.” We have reported these findings to the appropriate model developers. Considerations of fundamental differences between SWIFT and MC-SCIPUFF, how each model component is implemented within Urban HPAC, and how such differences may account for the above findings are ongoing research topics.

Unauthenticated | Downloaded 08/23/22 12:40 AM UTC
d. Comparison of JU03 (Oklahoma City) and previous Urban HPAC studies: Urban 2000 (Salt Lake City) and MUST

The releases associated with the Urban 2000 and MUST field experiments were conducted at night or in the very early morning while atmospheric conditions were still relatively stable. Because there were substantial differences in Urban HPAC model behavior and performance as a function of day and night during JU03, the most appropriate JU03 results for comparison with Urban 2000 and MUST are those associated with the nighttime releases. In a similar way, previous studies included the UC, DM, and DW modes but not the MS mode because it was not available at those times. Therefore, we do not consider the more recent MS JU03 predictions here. In addition, the Urban 2000 and MUST predictions included the use of SWIFT in all cases. Thus, only SWIFT-based JU03 predictions were considered for these comparisons.

In general, Urban HPAC modes led to overpredictions of the surface sampler concentrations. For JU03, the median nighttime FB values (not considering the MS mode) for the CBD and the arcs were 0.95 and 0.79 (or overprediction factors of about 2.8 and 2.3), respectively. For Urban 2000, the comparable FB values (for the CBD and on the arcs, respectively) were 0.48 (overprediction factor of about 1.6) and 0.70 (overprediction factor of about 2.1). For MUST, overprediction factors, across the entire 200-m square of samplers, varied by Urban HPAC mode from about 1.0 to 3.0.

We also found for all three field experiments that adding UWM to UDM to create the DW mode did not lead to substantial or consistent significant improvements relative to using UDM alone (i.e., DM). In addition, for all three field experiments, predictions of exceeding a relatively low concentration threshold were more accurate (as measured by the MOE) than predictions of absolute average concentrations.

With respect to NAD, the JU03 Urban HPAC predictions generally resulted in less scatter than the comparable Urban 2000 predictions. For example, for Urban 2000, the downtown and arc-based NAD values were always greater than 0.60 for any of the 20 sets of predictions that were examined and when considering 30-min average concentrations. A few sets of Urban HPAC JU03 nighttime predictions (after excluding MS for these comparisons) resulted in NAD values less than 0.60, including SWIFT-based values on the arcs at night of 0.53 and 0.58 for the GCT-based DM and DW predictions, respectively. For MUST, NAD values are not easily comparable because the time resolutions that were examined—10, 60, and 300 s and about 15 min—are substantially different from those examined during JU03 and Urban 2000. Nonetheless, for the ≈15-min time resolution, NAD values for the 20 sets of Urban HPAC MUST predictions were never less than 0.45 and were typically above 0.55. The relative improvement in model fit (less scatter as measured by NAD) for the JU03 predictions relative to those of Urban 2000 could be partially explained by improved Urban HPAC models available since the time of the Urban 2000 study and/or by improved MET inputs used in this JU03 study (i.e., MET inputs that better represent the actual winds that transport the plume). We further discuss this “better MET input” hypothesis using the FAC2 metric below.

The FAC2 metric considers the ratios between the predictions and the observations at each point in space.

![Figure 8](https://example.com/fig8.png)

**Fig. 8.** Comparisons of (left) FB and (right) NAD values for Urban HPAC UC surface predictions of the daytime (red) and nighttime (blue) releases of JU03 using the six MET input options (labeled along the x axes of each chart as PO7M, BASM, GCTM, PO7, BAS, and GCT). The smaller colored points correspond to the FB or NAD values for each of the individual releases (17 day and 12 night). The larger colored diamonds correspond to the average FB or NAD values (day = red and night = blue), with the large black diamond representing the overall average for all 29 releases.
and time—here, for 30-min average concentrations. This metric is equally sensitive to the smaller and larger concentrations, whereas metrics such as NAD, NMSE, and the c-MOE can be dominated by the larger concentrations. As such, FAC2 can be particularly sensitive to mismatches in plume transport direction (e.g., several samplers with relatively small observed concentrations missed on one side of the plume because of a 10°–20° transport error could easily lead to predictions and observations differing by more than a factor of 2). This metric can also be very sensitive to the underlying sampling distribution and the data processing protocols. For different field experiments, we followed consistent protocols to allow for fair comparisons between experiments. This required recomputation of the Urban 2000 values (Warner et al. 2007). Table 5 compares the range of FAC2 values for the comparable JU03 and Urban 2000 predictions. Table 5 indicates a substantial improvement associated with the JU03 predictions. We speculate that this is due to improved MET inputs, in terms of plume transport wind directions and/or speeds, available to our study team for the creation of Urban HPAC predictions during JU03 relative to Urban 2000.

It was postulated during the previous analysis of the Urban 2000 field experiment that the terrain associated with Salt Lake City represented a challenging and important feature that could influence the wind fields substantially. For example, a mesoscale numerical forecast that was used as MET input for the Urban 2000 study [Operational Multiscale Environment Model with Grid Adaptivity (OMEGA)] led to some of the best Urban 2000 predictions yet was recognized as missing the plume direction on the arcs (Warner et al. 2004a). This apparent inconsistency can be explained by noting that the majority of the Urban 2000 samplers were located in the close-in CBD, where these OMEGA-based predictions did very well, and not on the longer-range (2, 4, and 6 km) arcs.

We also examined a set of JU03 Urban HPAC predictions that used minisodar wind observations from the botanical gardens located in downtown Oklahoma City near the release sites (Fig. 4). Our analysis of the predictions that resulted from using this botanical-gardens (BGS) MET option suggested that the wind directions were not well matched to the actual directions relative to the other MET options that were examined. This MET option was rejected for the final comparative analyses because, based on our examination of scatter metrics, MOE values, and—of importance—concentration contours, we found that the predicted and observed plume directions often differed (Warner et al. 2007). The FAC2 values for the BGS-based Urban HPAC predictions of JU03 are shown in the last column of Table 5.

The nighttime FAC2 values associated with the flawed BGS predictions are very similar to those previously reported for Urban 2000. The suggestion here is that if improved MET input that better represented the actual plume transport winds were available and were used for the JU03 predictions then improvements in FAC2 of the magnitude observed here (approximately a factor of 2 in FAC2) between Urban 2000 and JU03 might be expected. Future efforts are planned to evaluate the latest version of Urban HPAC (including MSS) using the Urban 2000 field experiment. Such efforts may shed light on the relative differences discussed above.

### 4. Conclusions

The observations of the JU03 field experiment were used to evaluate HPAC, including its four component urban modes: baseline urban canopy (UC), UDM (DM), UDM + UWM (DW), and MSS (MS). Twenty sets of predictions of the 29 continuous releases, using five different MET options and four Urban HPAC modes, were created, and an examination of metrics associated with bias, scatter, correlation, and a user-oriented two-dimensional MOE was used to identify differences and similarities between modes. Several new findings were reported. First, there were noticeable differences observed in the transport and dispersion of SF6 between the nighttime and daytime releases. Figure 9 compares contours of the first-hour average concentration for a typical continuous daytime release (the second release of the fourth IOP) with a typical nighttime release (the second release of the ninth IOP). In these plots, the release rate is scaled to 2 g s⁻¹ to enable consistent comparison (i.e., equal release mass comparisons because release times were identical). As can be seen from Fig. 9, SF6 tended to drift farther...
downrange (at the surface) at night than during the day. Also, the contour of the observed surface average concentration indicated that the daytime plumes are slightly wider than the nighttime plumes. These differences in nighttime versus daytime behavior are akin to stability-category differences between nighttime and daytime transport and dispersion for open-area releases. In a similar way, HPAC-based prediction performance shows daytime versus nighttime differences.

For all five MET options, daytime releases tended to be underpredicted (30-min average concentrations at the surface samplers in the CBD and on the arcs) and nighttime releases tended to be overpredicted. Based solely on the 30-min predicted average concentrations, UDM (DM), UDM + UWM (DW), and baseline (UC) Urban HPAC predictions significantly exaggerated these differences relative to the observations (Fig. 6); that is, for the nighttime releases there is a significant overprediction and for the daytime releases there is an underprediction. With respect to the under- and overpredictions described above, the MS mode typically led to less underprediction during the day and less overprediction at night than the other Urban HPAC modes (Fig. 6). The MS mode typically resulted in the least biased predictions of the 30-min average concentrations at the surface samplers (CBD and arcs). The MS-based predictions resulted in the least scatter in seven of the eight categories (day/night, CBD/arcs, and NAD/NMSE) that were examined. The improvement associated with the inclusion of MSS warrants further investigation to determine whether this conclusion can be extended to other urban environments.

For the more-difficult-to-predict nighttime releases, the MS, DM, and DW modes offered improvement, in terms of scatter and the user-oriented MOE, over the UC mode for the three MET input options that invoked SWIFT. We also found that adding UWM to UDM to create the DW mode did not lead to substantial or consistent significant improvements relative to using UDM alone (i.e., DM). This result is entirely consistent with past studies of the Urban 2000 (Warner et al. 2004a) and MUST (Warner et al. 2006) field trials. For the

![Figure 9: Average surface concentration contours for the first hour of a representative (left) daytime (IOP 4, release 2) and (right) nighttime (IOP 9, release 2) release. Contour levels shown are 100, 250, 500, and 1000 ppt, release rates are normalized to 2 g s$^{-1}$, and a Delaunay-based interpolation scheme is used to create the contours (Warner et al. 2005). (top) Results for all surface samplers out to the 4-km arcs, and (bottom) a zoom view of the CBD surface samplers.]
nighttime releases and the MC-SCIPUFF-based MET options (PNA and ACA), results were mixed. For the daytime releases, no consistent trends among Urban HPAC modes were found.

The improvement in HPAC predictive performance for the daytime releases relative to the nighttime releases and the observation, in general, of nighttime overpredictions relative to the daytime releases are both consistent with past evaluations of HPAC using the Prairie Grass field experiment. In the case of the previous Prairie Grass evaluations (Warner et al. 2004b), the results were improved predictive performance for the unstable atmospheric condition relative to the stable condition and the observation of HPAC overpredictions for the stable releases relative to the unstable releases.

Comparisons with previous analysis that used observations from the Urban 2000 (Salt Lake City) field experiment—in particular, estimates of FAC2—suggested that improved MET options that better represented the actual winds that transported the plume were available for the creation of the JU03 predictions. One hypothesis is that the terrain associated with Salt Lake City (Urban 2000) represented a particularly challenging feature relative to the relatively flat terrain in Oklahoma City (JU03) and that this feature may not have been fully accounted for in the MET input options used for our previous simulations. Future research is planned to address this hypothesis.

Acknowledgments. This effort was supported by the Defense Threat Reduction Agency and the Central Research Program of the Institute for Defense Analyses. The authors thank Richard J. Babarsky (National Ground Intelligence Center, Charlottesville, Virginia) for providing the GCAT predictions and acknowledge R. Ian Sykes and Doug S. Henn of Titan Corporation (Princeton, New Jersey), Cathy Dougherty and Thomas B. Harris of Science Applications International Corporation (King of Prussia, Pennsylvania), Jacques Mousafir of ARIA Technologies (Boulogne-Billancourt, France), and John R. Hanan and Richard N. Fry of DTRA/Joint Science and Technology Office (Fort Belvoir, Virginia) for valuable discussions throughout this study. No official endorsement by the U.S. Department of Defense is intended or should be inferred.

REFERENCES


De Wekker, S. F. J., L. K. Berg, K. J. Allwine, J. C. Doran, and


