Evaluation of a Fast-Running Urban Dispersion Modeling System Using Joint Urban 2003 Field Data

ERIC A. HENDRICKS,* STEVE R. DIEHL, DONALD A. BURROWS, AND ROBERT KEITH
Advanced Engineering and Sciences, ITT Corporation, Colorado Springs, Colorado

(Manuscript received 7 September 2005, in final form 3 April 2006)

ABSTRACT

An urban dispersion modeling system was evaluated using the Joint Urban 2003 field data. The system consists of a fast-running urban airflow model (RUSTIC, for Realistic Urban Spread and Transport of Intrusive Contaminants) that is coupled with a Lagrangian particle transport and diffusion model (MESO) that uses random-walk tracer diffusion techniques. Surface measurements from fast-response and integrated bag samplers were used to evaluate model performance in predicting near-field (less than 1 km from the source) dispersion in the Oklahoma City, Oklahoma, central business district. Comparisons were made for six different intense operating periods (IOPs) composed of three different release locations and stable nighttime and unstable daytime meteorological conditions. Overall, the models were shown to have an underprediction bias of 47%. A possible influence to this underprediction is that the higher density of sulfur hexafluoride in comparison with air was not taken into account in the simulations. The models were capable of predicting 42% of the sampler data within a factor of 2 and 83% of the data within a factor of 10. When the effects of large-scale atmospheric turbulence were included, the models were shown to be capable of predicting 51% of the data within a factor of 2. The results were further broken down into performance for varying meteorological conditions. For daytime releases, the models performed reasonably well; for nighttime releases the models performed more poorly. Two possible causes of the poorer nighttime comparisons are (a) an inability to model the suppression of vertical turbulence because of the assumption of isotropy in RUSTIC’s \( k-\omega \) turbulence model and (b) difficulty in modeling the light and variable inflow winds. The best comparisons were found for the three continuous daytime releases of IOP-4. It was hypothesized that these good comparisons were a result of steadier inflow conditions combined with the fact that the release site was more exposed and closer to the sodar used for the inflow meteorological conditions.

1. Introduction

Advancing the understanding of dispersion in urban environments is a matter of both practical and scientific importance. The ability to accurately predict the dispersion of a released agent in a city is crucial for saving lives in the event of a terrorist attack or accidental release or, to a lesser extent, for warning people of significant air pollution episodes. However, accurate modeling and simulation of dispersion in these environments has been a great challenge because of their complex radiative, thermodynamic, and aerodynamic characteristics (Oke 1988; Arya 2001). Additionally, the quasi-random behavior of turbulent eddies that arise from mechanical and thermal interactions of the flow with buildings are difficult to predict and also very sensitive to slight perturbations in the inflow. Because of these inherent complexities and for the sake of simply predicting general hazard area definitions, the common approach has been to use models that provide a solution to the advection–diffusion equation using a mean wind. However, these models, while providing a general hazard area definition, do not explicitly model the effects of buildings on the flow and will therefore be less accurate. Another approach involves the use of computational fluid dynamics (CFD) methods to predict gridded solutions for the atmospheric state variables and then to couple these CFD models with particle-based stochastic Lagrangian advection and diffusion codes. This method is more accurate but generally more computationally expensive.

Currently, there are limited field data in urban environments that can be used to evaluate the performance
of these dispersion models. Thus, their utility for hazard prediction is relatively unknown. The Joint Urban 2003 Atmospheric Dispersion Study (JU2003) was completed to address this data void. It was sponsored by the United States Defense Threat Reduction Agency and the Department of Homeland Security and conducted in July 2003 in Oklahoma City, Oklahoma (OKC; Allwine et al. 2004). The objective of the experiment was the collection of meteorological and tracer data resolving atmospheric dispersion at scales of motion including (a) flows around a single city block, (b) flows in and around several blocks in the downtown central business district (CBD), and (c) flows into the suburban Oklahoma City area several kilometers from the CBD. A quality-controlled dataset would be used to evaluate and improve a suite of indoor and outdoor dispersion models, including finescale CFD models and urban parameterizations in numerical weather prediction models. Ten intense operating periods (IOPs) were conducted and approximately three continuous releases and four instantaneous releases of sulfur hexafluoride (SF$_6$) occurred in each IOP. The experiment was a collaboration of numerous universities, national laboratories, and private industry. ITT Corporation’s Advanced Engineering and Sciences (ITT AES) group fielded five portable 1-Hz sulfur hexafluoride detectors and 11 portable 10-Hz sonic anemometers in support of the experiment.

Under the sponsorship of the Defense Advanced Research Projects Agency (DARPA), an urban dispersion modeling system was developed by ITT AES. The models in this system were designed to be both relatively high fidelity and fast running, or more specifically, to fill the gap between the time-consuming full CFD models and very fast, but lower-fidelity mass-consistent flow codes (Fig. 1). The modeling system consists of a fast-running urban wind flow code, Realistic Urban Spread and Transport of Intrusive Contaminants (RUSTIC), that is coupled with a Lagrangian particle advection and diffusion code (MESO). Both codes have selectable fidelity, so that a particular user can choose between a fast, less accurate result and a slower, more accurate result, depending on the particular situation. There exists a great need for these fast-running models, particularly for emergency responders (Brown 2004).

In this paper, a statistical evaluation of the performance of MESO–RUSTIC against the Joint Urban 2003 sampler data is presented. In companion papers in this special issue (Burrows et al. 2007; Diehl et al. 2007), the physics and numerics of each model are described in more detail. In section 2, the urban modeling system is described. The data and methods used in the evaluation are discussed in section 3. The model setup is described in section 4. The results of the evaluation are presented in section 5. An examination of the effects of large-scale atmospheric turbulence is discussed in section 6. The conclusions of the study are presented in section 7.

2. Model descriptions

RUSTIC is a fast-running urban airflow model that numerically solves a modified set of the compressible Reynolds-averaged Navier–Stokes (RANS) equations via the finite-volume method. To simplify the computation, the thermodynamic and continuity equations are combined into a pressure tendency equation. In this equation, the speed of sound is reduced to the maximum velocity in the model in order to accelerate the computation (this can be done because acoustic waves do not contribute to the solution). The turbulent fluxes from the RANS equations are parameterized by a gradient transfer process, and a $k$–$\omega$ turbulence model (Wilcox 1998) is used to predict the eddy viscosity coefficient. This turbulence model was chosen because it has been demonstrated to perform better in transitional flows, flows with adverse pressure gradients, and because it is numerically stable. Atmospheric stability effects are incorporated into the model via the pressure tendency equation and a buoyancy production term in the turbulent kinetic energy equation. Buildings are modeled explicitly in RUSTIC, and the model has automatic grid and terrain generators. The exact equations used in RUSTIC are listed below, following standard meteorological conventions for variables (Burrows et al. 2007):
\[
\frac{\partial \mathbf{u}}{\partial t} = -\mathbf{u} \cdot \nabla \mathbf{u} - \frac{1}{\rho} \nabla P - \frac{1}{\rho} (\nabla \cdot \rho K_m \nabla) \mathbf{u},
\]
(1)

\[
\frac{\partial P'}{\partial t} = -\mathbf{u} \cdot \nabla P + w \tilde{g} - \frac{\rho c_p}{\rho \theta} \left[ \nabla \cdot \mathbf{u} - \frac{1}{\rho \theta} (\nabla \cdot \rho \tilde{K}_h \nabla) \theta \right]
- \frac{1}{c_p} \frac{d Q_s}{dt},
\]
(2)

\[
\frac{\partial k}{\partial t} = -u_j \frac{\partial k}{\partial x_j} + \tau_{ij} \frac{\partial u_i}{\partial x_j} - \beta^* k \omega + \frac{\partial}{\partial x_j} \left[ (\nu + \sigma^* \frac{k}{\omega}) \frac{\partial k}{\partial x_j} \right]
- \frac{g}{\theta_0} \tilde{K}_h \frac{\partial \theta}{\partial x_j},
\]
(3)

\[
\frac{\partial \omega}{\partial t} = -u_j \frac{\partial \omega}{\partial x_j} + \frac{\omega}{\kappa} \tau_{ij} \frac{\partial u_i}{\partial x_j} - \beta \omega^2
+ \frac{\partial}{\partial x_j} \left[ (\nu + \frac{k}{\omega}) \frac{\partial \omega}{\partial x_j} \right],
\]
(4)

\[
\tau_{ij} = \frac{k}{\omega} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) - \frac{2}{3} \delta_{ij}, \quad \delta_{ij} = (1, i = j; 0, i \neq j).
\]
(5)

Here, \( \mathbf{u} \) is the three-dimensional velocity vector, \( P' \) is the nonhydrostatic perturbation pressure, \( k \) is the specific turbulent kinetic energy, \( \omega \) is the specific dissipation rate of turbulent kinetic energy, \( \theta \) is the dry potential temperature, \( \nu \) is the molecular viscosity of air, \( \rho \) is the density, \( Q_s \) is the surface heat flux, and \( \tau \) is the Reynolds stress tensor. It is assumed that the heat diffusivity is equal to the eddy viscosity coefficient: \( K_p = K_m = k/\omega \). The closure coefficients used in the \( k-\omega \) model for the simulations presented in this paper are \( \beta = 0.072, \beta^* = 0.09, \sigma = 0.5, \sigma^* = 0.5, \) and \( \alpha = 0.52 \) (Wilcox 1998).

MESO is a mesoscale Lagrangian particle-based transport and diffusion code. For plumes, it uses random-walk techniques (Diehl et al. 1982) for vertical dispersion and Langevin techniques (van Dop et al. 1985) for horizontal dispersion. The random-walk techniques are used because they are numerically fast in comparison with other stochastic techniques. For clouds (or puffs), the tracer dispersion is computed using the random-walk technique with scale-dependent mixing theory. MESO uses state-of-the-art meteorology, and it can handle a suite of chemical and biological agents as well as primary and secondary droplet evaporation. More details on MESO, including some validation results with wind tunnel data, can be found in Diehl et al. (2007).

To calculate the dispersion, RUSTIC is first run to approximate convergence in order to obtain a steady-state solution to the modified set of the RANS equations. Then, MESO uses the RUSTIC steady-state wind field and eddy viscosities to predict the transport and diffusion of a released agent. These eddy viscosities change significantly with atmospheric stability, strongly affecting the dispersion. As an example of this, simulated surface concentrations are shown in Fig. 2 for stable (surface upwind heat flux of \(-30 \text{ W m}^{-2}\)) and unstable (+250 \text{ W m}^{-2}) conditions in the Oklahoma City CBD for a surface release. In the stable case, the mixing is suppressed and surface concentrations are higher. In the unstable case, the vertical and horizontal mixing is enhanced, causing the plume to spread much more.
RUSTIC and MESO can also be run in a transient mode to more accurately simulate turbulent boundary layer flows. In this scenario, RUSTIC is first run to an approximate steady state. Then, the input winds can be varied in time sinusoidally or directly specified by the user. MESO can be utilized in the same transient manner, taking meteorological files representing the different conditions from RUSTIC at various times during the simulation. The focus of this paper will be on the steady-state runs because a more complex state of the atmosphere cannot be specified for Joint Urban 2003 because of limited upwind spatial and temporal measurements. However, an excursion into some simulations that were done modeling real input wind variability is presented in section 6.

3. Data and methods

a. Data used

The sampler data used to evaluate the models may be found in Table 1, which includes data from both integrated bag and continuous fast response samplers. The data were provided by the following Joint Urban 2003 participants: the National Oceanic and Atmospheric Administration/Air Resources Laboratory/Field Research Division in Idaho Falls, Idaho (NOAA/ARL/FRD); Lawrence Livermore National Laboratory in Livermore, California (LLNL); Volpe National Transportation Systems Center in Cambridge, Massachusetts (VOLPE); ITT AES in Colorado Springs, Colorado; and Washington State University (WSU) in Pullman, Washington. Most of the data are from surface samplers (approximately 1–2 m above ground level), with the exception of the vertical profiler from WSU.

b. Releases modeled

Six 30-min-long continuous releases were modeled. These releases are listed in Table 2. They include all three releases locations (Botanical Gardens, Westin Hotel, and Park Street) and stable nighttime and daytime releases. The configuration of all samplers for each release configuration is shown in Fig. 3. These are all surface samplers. A comparison was also made to the WSU vertical profiler, and this is discussed in section 5b. The vertical profiler was located approximately 1.2 km north of the Botanical Gardens (not shown in the maps in Fig. 3).

c. Method of comparison

Comparisons are made at the discrete locations of the samplers. The quantity used in the comparison is the release mean concentration, defined as the time-integrated concentration during the release period divided by the release period:

$$C = \frac{1}{T} \int_{t=0}^{t=T} C(t) \, dt \approx \frac{1}{T} \sum_{i=0}^{N} C(t_i) \Delta t_i.$$  (6)

For the Joint Urban 2003 continuous releases, \(T = 30\) min. In the discretization, time-integrated concentrations from fast-response samplers are generally calculated with \(\Delta t \approx 0.5–1\) s and time-integrated concentrations from integrated bag samplers are generally calculated with \(\Delta t \approx 5–15\) min (depending on the particular sampler). Because the modeled time-integrated concentration is output on the RUSTIC Cartesian grid, that is, \(C = f(x, y, z, t)\), it is necessary to interpolate this value in space to the sampler location. This is done by the inverse-distance weighting method. In this method, the model-predicted concentrations at the discrete sam-

| IOP | Date     | Time (UTC)   | Release location | Release type/rate (continuous or puff/g s⁻¹) | \(H_u\) upwind/\(H_p\) CBD (W m⁻²) | WS/WD at \(z = 20\) m (m/s²/°) | \(z_0/u*|/d\) (m/s²/m) |
|-----|----------|--------------|------------------|------------------------------------------|---------------------------------|-----------------------------|----------------------|
| I0P-2| 2 Jul 2003| 2000–2030 (daytime) | Westin Hotel | Continuous/4.99 | 166/60 | 2.18/182 | 6.8/1.02/4.0 |
| I0P-4| 9 Jul 2003| 1600–1630 (daytime) | BG | Continuous/3.13 | 153/68 | 2.89/198 | 3.6/0.90/7.2 |
| I0P-4| 9 Jul 2003| 1800–1830 (daytime) | BG | Continuous/2.99 | 240/75 | 3.05/187 | 2.2/0.92/10.7 |
| I0P-4| 9 Jul 2003| 2000–2030 (daytime) | BG | Continuous/2.77 | 162/79 | 2.20/195 | 3.4/1.02/9.9 |
| I0P-7| 20 Jul 2003| 0400–0430 (nighttime) | BG | Continuous/2.99 | -22/6 | 1.03/185 | 6.7/0.21/10.0 |
| I0P-9| 27 Jul 2003| 0800–0830 (nighttime) | Park Ave. | Continuous/2.09 | -18/11 | 1.13/185 | 5.0/0.54/9.1 |
pler locations are weighted inversely by the distance from those locations. This method was chosen because it provides interpolated values that retain the effects of the local area, while not smoothing excessively. Mean concentrations hereinafter are expressed in parts per trillion by volume (pptv). The conversion factor used was 1 pptv $\text{SF}_6 = 5.803 \times 10^{-12}$ kg m$^{-3}$ $\text{SF}_6$, and this includes variations in altitude and mean temperature of Oklahoma City from normal temperature and pressure (NTP; 1 atm and 20°C).

d. Performance measures

A set of statistical performance measures (Hanna 1989, 1993; Chang and Hanna 2004) is used to assess the prediction skill of the models. These measures are the (a) fractional bias (FB), (b) geometric mean bias (MG),
(c) normalized mean square error (NMSE), (d) geometric variance (VG), (e) correlation coefficient $R$, (f) fraction of predictions within a factor of 2 (FAC2), and (g) fraction of predictions within a factor of 10 (FAC10). These measures have been used in the statistical evaluation of other dispersion models (e.g., Hanna et al. 2004), and they are defined below:

$$\text{FB} = \frac{\overline{C_o} - \overline{C_p}}{0.5(\overline{C_o} + \overline{C_p})},$$

(7)

$$\text{MG} = \exp(\ln \overline{C_o} - \ln \overline{C_p}),$$

(8)

$$\text{NMSE} = \frac{(\overline{C_o} - \overline{C_p})^2}{\overline{C_p}},$$

(9)

$$\text{VG} = \exp[(\ln \overline{C_o} - \ln \overline{C_p})^2],$$

(10)

$$R = \frac{(\overline{C_o} - \overline{C_p})(\overline{C_p} - \overline{C_o})}{\sigma_{\overline{C_p}} \sigma_{\overline{C_o}}},$$

(11)

$$\text{FAC2} = \text{Fraction of data that satisfies } 0.5 \leq \frac{\overline{C_p}}{\overline{C_o}} \leq 2.0, \text{ and}$$

(12)

$$\text{FAC10} = \text{Fraction of data that satisfies } 0.1 \leq \frac{\overline{C_o}}{\overline{C_p}} \leq 10.0.$$  

(13)

Here, the overbar denotes the mean over the dataset, $C_o$ denotes the observed concentration and $C_p$ denotes the predicted concentration, and $\sigma$ denotes the standard deviation over the dataset. A background concentration of 3 ppt was added to all predictions, and observations less than 3 ppt were set to 3 ppt as has been done in past work (e.g., Allwine et al. 2002; Warner et al. 2004.)

4. Model setup

For the simulations, a Cartesian RUSTIC grid was used with $x$, $y$, and $z$ dimensions of 1.4 km, 1.4 km, and 200 m. The total number of grid points was 1 191 300 (190 $\times$ 190 $\times$ 33), and the smallest spacing was 5 m $\times$ 5 m $\times$ 5 m. The dimension of 5 m was chosen to retain good resolution in street canyons (5–6 grid points across). The mesh was nonuniform; the finest resolution occurred in the CBD where the plume was released while coarser resolution (10 m) existed near the lateral domain boundaries. A section of the model grid showing the resolution around buildings is shown in Fig. 4.

The initial inflow velocity profiles were provided by the Botanical Gardens (BG) minisodar. The BG minisodar was used for the vertical profiles of the upwind velocity because it was in an ideal location for the representative upwind condition, just south of the CBD. The minisodar was developed in-house at Argonne National Laboratory (ANL). Some details of the instrument are as follows (Coulter and Martin 1986): (a) frequency, 4.5 kHz; (b) minimum and maximum measurement altitudes, 5 and 200 m (however, the first two altitudes are often ignored because of the ground clutter effect); and (c) range gate spacing, 5 m. For the RUSTIC initialization, the minimum measurement altitude used was generally $z = 15$ m AGL; however, in some cases it was $z = 20$ m.

Although the logarithmic velocity profile is strictly only valid in neutral boundary layers, it was found to be a good approximation to the observed inflow profiles in the daytime releases (Fig. 5). Thus, RUSTIC was initialized with a logarithmic curve fit to the observed wind speed profile that yielded the necessary constants of surface roughness $z_0$, friction velocity $u_*$, and the zero-plane displacement $d$:

$$u(z) = \frac{u_*}{k} \ln \left( \frac{z - d}{z_0} \right).$$

(14)

In Eq. (14), $z$ is restricted to be greater than $d + z_0$ so that the natural logarithm remains positive. Below the displacement height, the velocity was set to 10% of the velocity at the displacement height.

It is known that least squares logarithmic fitting of data to obtain $d$ and $z_0$ is a difficult and sensitive task.
However, the values obtained seemed reasonable for flow just upwind of a dense urban CBD and therefore they were used. Although we have not performed a sensitivity test by varying these parameters, it is likely that the model results will be more sensitive to variances in $z_0$ than variances in $d$. This is because $z_0$ is used in calculating the shear at the surface and affects the turbulent kinetic energy (TKE) budget equation.

For the nighttime releases, the logarithmic fit was found to be a poor approximation near the surface (where very light winds were found) but better at upper levels (above approximately 60 m). Because of this, RUSTIC was initialized with the exact sodar profile (Fig. 5, bottom two panels). The wind direction was held constant at the approximate mean value for these simulations (Fig. 6). The initial meteorological parameters used in the RUSTIC initialization for each release are summarized in Table 2. In addition, a summary of the observed meteorological conditions for each release is shown in Table 3.

To quantify the differences between the logarithmic curve fit and the observations for the daytime releases,

![Fig. 5. Vertical profiles of wind speed observed at the Botanical Gardens minisodar (black circles) and the logarithmic curve fit used to initialize RUSTIC (dashed line).](image)
root-mean-square error (RMSE) values were calculated. Overall, RUSTIC matched the sodar data well. For the daytime releases with a logarithmic approximation, the RMSE value was 0.37 m s\(^{-1}\). The following are the RMSE values recorded for each release: (a) IOP-2 at 2000–2030 UTC, RMSE = 0.38 m s\(^{-1}\); (b) IOP-4 at 1600–1630, RMSE = 0.38 m s\(^{-1}\); (c) IOP-4 at 1800–1830 UTC, RMSE = 0.31 m s\(^{-1}\); and (d) IOP-4 at 2000–2030 UTC, RMSE = 0.40 m s\(^{-1}\).

To estimate the level of atmospheric stability properly, upwind and downtown sonic anemometers were used to calculate the surface-sensible heat flux, that is, 

\[
H_s = \rho c_p (w'q')_b, \quad \text{where the brackets denote temporal averaging.}
\]

The perturbation quantities were calculated using Reynolds averaging, where the slowly varying mean fields were defined with running averages of 10 min. For RUSTIC, two bulk area-averaged heat fluxes were used: 1) upwind suburban and 2) grid. The upwind heat flux was calculated from the Indiana University’s (IU) Tyler Media tower location located 5 km upwind of the city in a suburban area. The grid fluxes were calculated from the average of 33 surface sonic anemometers throughout the central business district (operated by Dugway Proving Ground, ITT AES, and the...
University of Utah). These heat fluxes are also listed for each release in Table 2. The net upwind surface heat flux was larger upwind than in the grid primarily because of the reduction of incoming shortwave radiation by shadowing from tall buildings. The heat fluxes were used to obtain the inflow vertical profiles of turbulent kinetic energy (via a buoyancy production term) and potential temperature. They were also used in the computation of TKE and perturbation pressure tendency. For Meso, 100 000 tracers were used to simulate the SF$_6$ transport and dispersion.

### 5. Comparison results

**a. Surface samplers**

Scatterplots depicting the comparison for each of the six releases can be found in Fig. 7. In these plots, each point represents a sampler in the CBD (shown in the maps in Fig. 2). Some amount of prediction skill can be seen for each release from these scatterplots. From these plots, it appears that the IOP-4 at 2000–2030 UTC release had the best correlation. Good correlations were also found for the other daytime releases; however, the IOP-4 at 1600–1630 UTC and the IOP-4 at 1800–1830 UTC releases exhibited more scatter at the lower concentrations. There was significantly more scatter and a less favorable correlation for the two nighttime releases: IOP-7 at 0400–0430 UTC and IOP-9 at 0800–0830 UTC. Since winds and eddy viscosities computed by RUSTIC are passed to Meso to predict these concentration levels, it is useful to examine the comparison of RUSTIC-computed turbulence kinetic energy (TKE) and velocity with measurements as well. In general, RUSTIC did a fair job at predicting TKE in the daytime releases, but a poorer job for the nighttime releases. Comparisons with CBD sodar velocity profiles were generally very good for most releases (Burrows et al. 2007).

The statistical performance measures (based upon the data shown in Fig. 7) for this study are shown in Table 4. There were 13 sets of data examined. The first six sets were each of the releases, and the final seven sets were combinations of the first six made up of varying meteorological and release conditions. Overall, the FB of $-0.39$ indicates that the models generally overpredicted concentrations by 33% [rearrangement of Eq. (7)]. However, the MG indicated a 47% underprediction bias [rearrangement of Eq. (8)]. These measures clearly contradict one another. This occurs because FB is overly influenced by high concentrations, while MG may be overly influenced by low concentrations. According to Chang and Hanna (2004), MG and VG may be more suitable for dispersion modeling because of concentrations spanning many orders of magnitude. Because of this fact, we feel it is more likely that the models have an underprediction bias. A possible factor for the underprediction bias is that the molecular weight of the denser SF$_6$ (as compared with air) was not taken into account in Meso. (Meso is currently not able to model gaseous releases of substances with variable densities).

The arithmetic and geometric variance measures (NMSE and VG, respectively) were very large. However, these measures can be very sensitive to a few outliers (Chang and Hanna 2004). For all of the releases, in fact, the systematic component of these measures was very small: $\text{NMSE}_{sys} = 0.16$ and $\text{VG}_{sys} = 1.83$. This indicates that random error accounted for a large portion of these total values. The correlation coefficient $R$ was $0.06$ ($0.0 < R < 0.16$ with 95% confidence), indicating no significant correlation between the model predictions and measurements. Overall, 42% of the model predictions were within a factor of 2 of the measurements, and 83% were within a factor of 10.

Examining individual datasets in Table 4, it is apparent that the models did a fairly good job of prediction for IOP-4 releases (datasets 2, 3, 4, and 12), but generally did a significantly poorer job for the other releases (datasets 1, 5, 6, 9, and 11). This is particularly evident in the $R$ values: $R = 0.89$ ($0.86 < R < 0.92$ with 95% confidence) for the IOP-4 releases and $R = 0.04$ ($0.0 < R < 0.19$ with 95% confidence) for the others. Two of these other releases were nighttime releases (IOP-7 and IOP-9) and the other was a Westin Hotel release loca-

---

**Table 3. Observed weather conditions for each release (from KPWA: Wiley Post Airport, Oklahoma City, OK).**

<table>
<thead>
<tr>
<th>IOP</th>
<th>Date</th>
<th>Time (UTC)</th>
<th>Conditions</th>
<th>$T/T_d$ (°C)</th>
<th>Prevailing wind at 10 m AGL:</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOP-4</td>
<td>9 Jul 2003</td>
<td>1600–1630 (daytime)</td>
<td>Partly cloudy</td>
<td>30/21</td>
<td>SSW, 6.7</td>
</tr>
<tr>
<td>IOP-4</td>
<td>9 Jul 2003</td>
<td>1800–1830 (daytime)</td>
<td>Scattered clouds</td>
<td>32/21</td>
<td>S, 7.2</td>
</tr>
<tr>
<td>IOP-4</td>
<td>9 Jul 2003</td>
<td>2000–2030 (daytime)</td>
<td>Scattered clouds</td>
<td>34/21</td>
<td>S, 10.3</td>
</tr>
<tr>
<td>IOP-7</td>
<td>20 Jul 2003</td>
<td>0400–0430 (nighttime)</td>
<td>Scattered clouds</td>
<td>32/16</td>
<td>S, 6.2</td>
</tr>
<tr>
<td>IOP-9</td>
<td>27 Jul 2003</td>
<td>0800–0830 (nighttime)</td>
<td>Clear</td>
<td>26/17</td>
<td>S, 5.1</td>
</tr>
</tbody>
</table>

---

**Table 4.** There were 13 sets of data examined. The first six sets were each of the releases, and the final seven sets were combinations of the first six made up of varying meteorological and release conditions. Overall, the FB of $-0.39$ indicates that the models generally overpredicted concentrations by 33% [rearrangement of Eq. (7)]. However, the MG indicated a 47% underprediction bias [rearrangement of Eq. (8)]. These measures clearly contradict one another. This occurs because FB is overly influenced by high concentrations, while MG may be overly influenced by low concentrations. According to Chang and Hanna (2004), MG and VG may be more suitable for dispersion modeling because of concentrations spanning many orders of magnitude. Because of this fact, we feel it is more likely that the models have an underprediction bias. A possible factor for the underprediction bias is that the molecular weight of the denser SF$_6$ (as compared with air) was not taken into account in Meso. (Meso is currently not able to model gaseous releases of substances with variable densities).

The arithmetic and geometric variance measures (NMSE and VG, respectively) were very large. However, these measures can be very sensitive to a few outliers (Chang and Hanna 2004). For all of the releases, in fact, the systematic component of these measures was very small: $\text{NMSE}_{sys} = 0.16$ and $\text{VG}_{sys} = 1.83$. This indicates that random error accounted for a large portion of these total values. The correlation coefficient $R$ was $0.06$ ($0.0 < R < 0.16$ with 95% confidence), indicating no significant correlation between the model predictions and measurements. Overall, 42% of the model predictions were within a factor of 2 of the measurements, and 83% were within a factor of 10.

Examining individual datasets in Table 4, it is apparent that the models did a fairly good job of prediction for IOP-4 releases (datasets 2, 3, 4, and 12), but generally did a significantly poorer job for the other releases (datasets 1, 5, 6, 9, and 11). This is particularly evident in the $R$ values: $R = 0.89$ ($0.86 < R < 0.92$ with 95% confidence) for the IOP-4 releases and $R = 0.04$ ($0.0 < R < 0.19$ with 95% confidence) for the others. Two of these other releases were nighttime releases (IOP-7 and IOP-9) and the other was a Westin Hotel release loca-
tion. The models did better for the Botanical Gardens releases than the Park Avenue or Westin Hotel releases (datasets 8 and 9) and performed better for daytime releases than nighttime releases (datasets 10 and 11). The inclusion of large-scale atmospheric turbulence effects (discussed further in section 6) improved most performance measures (MG, VG, R, FAC2, and FAC10) but worsened a few (FB and NMSE) (dataset 13). The most significant improvement was the increase in FAC2 from 0.42 to 0.51.

Examining the results, the following main conclusions may be drawn: (a) the models were significantly better at prediction in daytime than nighttime, (b) the models performed much better for IOP-4 releases than any others, and (c) the models performed better for Botanical Gardens releases than the others. While it is a difficult task to fully understand what the reasons are for the variances in model performance, we can point to some possible causes. A possible reason for the poorer comparison at night is the $k−\omega$ model’s assumption of isotropy; the stable nighttime conditions may suppress the vertical turbulence component more than the horizontal components [the TKE comparisons were also poorer at night; Burrows et al. (2007)]. The inflow velocity profiles were also very different between day and night (as expected by theory); the daytime releases were nearly logarithmic, while the nighttime releases were more linear (Fig. 5). Although it cannot be verified directly because of the 15-min-averaging times for the minisodar, it is possible that very light inflow winds at night (e.g., approximately 1 m s$^{-1}$ at $z = 30$ m) were also more variable. This variability would not be cap-

**Fig. 7.** Scatterplots depicting the comparison of the modeled and the measured release-averaged concentration at the surface samplers shown in Fig. 2.
tured by a steady-state RUSTIC run. The IOP-4 releases were chosen initially because the wind direction was less variable in time than some other daytime releases (e.g., IOP-2). This is illustrated clearly in Fig. 8. Here, two 15-min averages of the BG minisodar wind direction are shown for the IOP-2 at 2000–2030 release and the IOP-4 at 1800–1830 UTC release. For IOP-2, note that the second 15-min-average wind direction was significantly different than the first 15-min average; for IOP-4, the two 15-min averages are similar. If this variability signal was present to this extent in the 15-min averages, one would expect that if higher-frequency measurements were taken even more variability would be evident. This may have been a factor in the poorer daytime comparison for IOP-2. A likely reason that the comparison with the Botanical Gardens releases was more favorable than the Park Avenue and Westin Hotel releases was that the Botanical Gardens site was closest to the BG minisodar, used for all of the initializations. Also, the Park Avenue and Westin Hotel sites were sheltered in the CBD near tall buildings, while the Botanical Gardens site was more exposed near the upwind edge of the city.

When examining the comparison and statistical measures, it is instructive to recall that MESO–RUSTIC were not designed to be of the highest fidelity possible (Fig. 1). In keeping the models fast running, certain physical processes present in real urban environments were not represented. For example, uneven surface heating (including building walls) was not modeled, vegetation was parameterized by an enhanced area of surface roughness, and turbulence induced by traffic was not modeled. Future work will be focused on adding relevant physics to more accurately model urban areas without making great sacrifices in speed.

b. Altitude samplers

Comparisons were also made between the WSU crane SF₆ vertical profiler with bag samplers placed at altitudes of $z = 10, 17, 24, 34, 48, 62, \text{ and } 75 \text{ m AGL}$. The profiler was located approximately 1.2 km north of the Botanical Gardens release site, downwind of the

![Fig. 8. The 15-min averages of wind direction from the BG minisodar for the (left) IOP-2 at 2000–2030 UTC and (right) IOP-4 at 2000–2030 UTC releases.](image-url)
tall buildings of the CBD. Because of the location of the WSU profiler, a new RUSTIC grid (3 km × 3 km in size, with a minimum cell size of 8 m) was used so that the location of the profiler would be inside the RUSTIC grid. Vertical profiles of the average concentration seen here in comparison with the models are shown for the IOP-7 at 0400–0430 UTC nighttime release in Fig. 9. In general, the comparison was favorable; however, better agreement was seen at upper levels than at lower levels. A possible reason for this is that the decreased resolution (8 m instead of 5 m) may have been sufficient for resolving the upper-level flow, but not sufficient for resolving the flow at low levels amid the buildings. The observations indicated the plume is well-mixed vertically; MESO–RUSTIC on the other hand was not able to predict the ground-level concentrations very well. This is not too surprising because the nighttime comparison with surface CBD samplers was poor as well. More releases need to be examined to determine the significance of this result.

6. Inclusion of large-scale atmospheric turbulence

The inclusion of variations in the upwind velocity caused by large convective motions in the boundary layer (hereinafter large-scale atmospheric turbulence, LST) was examined. The inclusion of this effect is important because the spectral TKE decomposition indicates that most of the power is from eddies of larger time scales, with the peak power occurring at a frequency of approximately 0.006 Hz (measured from the daytime releases of IOP-4). Running the models to steady state is therefore an oversimplification of the real atmosphere (Venkatram et al. 2004) and it may have significant effects on the dispersion, particularly at samplers far from the plume centerline. The stochastic nature of LST also yields large variances in the number of times a given sampler will be hit (Fig. 10), and it is therefore a measurement uncertainty when comparing 30-min-average concentrations. The net effect of meandering upwind velocity is to spread the plume more horizontally and make the plume more uniform (Yuan and Venkatram 2005). Other dispersion models have modeled LST by this empirical characteristic; for example, the American Meteorological Society–Environmental Protection Agency Regulatory Model (AERMOD) defines the horizontal concentration distribution as a function of both Gaussian and uniform components (Environmental Protection Agency 2005). As an example of LST in Joint Urban 2003, three plots are shown in Fig. 11 from a 10-Hz sonic anemometer mounted on Indiana University’s Tyler Media tower (z = 19 m AGL, 5 km upwind of the CBD) for IOP-4 from 1800 to 1900 UTC: (a) correlation of wind speed with wind direction, (b) probability density function of the wind speed, and (c) probability density function of the wind direction.

Some simulations were performed while incorporating LST for the IOP-4 releases using the Indiana University data. Although the Tyler Media tower was located 5 km upwind of the urban CBD, it was assumed that the velocity variance there was approximately equal to the velocity variance at the BG minisodar. A postprocessing method was developed where simulated concentrations from steady-state simulations at multiple different wind directions were weighted by the
amount of time spent at those wind directions during the 30-min release. A similar method was not used initially for wind speed because it is known to be less significant on the overall dispersion.

The different wind directions used spanned 1.7σ (approximately 25°) from the mean of the probability density functions of wind direction. The probability density functions were shown to be approximately Gaussian, with 1σ accounting for 68% of the wind direction (WD) points. In the discretization, 5° bins were defined about this 1.7σ range, and the amount of time spent at in a bin was determined. For example, for the IOP-4 at 1800–1830 release, the mean wind direction was 187°. Therefore, the weighting spanned the 162°–202° range, and the fractional amounts of time spent in the bins 162°–167°, 167°–172°, 172°–177°, 177°–182°, 182°–187°, 187°–192°, 192°–197°, and 197°–202° were assessed. These fractions were then normalized with the assumption that 100% of the time was spent in the ±1.7σ range. To be more accurate, a ±2-σ or 3-σ range could be used without this normalization; however, this method was used to reduce the number of runs (since most of the WD points were within 1.7σ of the mean). The determination of the weighted concentration was then, $C = A_1C_{164.5} + A_2C_{169.5} + A_3C_{174.5} + \cdots + A_8C_{199.5}$. Here, the $A$ coefficients represent the fractional amount of time spent in the bins and $C$ represents the MESO–RUSTIC-predicted concentration for the wind direction at the middle of the bin.

Scatterplots showing the comparison with the IOP-4 releases before inclusion and after inclusion of LST are shown in Fig. 12. The main benefit of the inclusion of LST was the improvement of the comparison at samplers far from the release site (and plume centerline) (Fig. 12, circled area). This brought the FAC2 statistic...
up from 0.42 to 0.51 (Table 4, dataset 13). For another perspective, the same data with the inclusion of LST are shown against the angle from the plume centerline in Fig. 13.

This postprocessing method is not as accurate as direct simulation of the variable inflow winds. However, that level of accuracy was unobtainable because there were no high-frequency upwind measurements in the vicinity of the BG minisodar that could be used to drive the models. More sensitivity tests need to be conducted besides varying the inflow velocity. Quantitative results on the solution sensitivity to the following important input parameters would be justified: heat fluxes, grid resolution, and number of tracers.

7. Conclusions

The performance of a fast-running urban dispersion modeling system was evaluated with Joint Urban 2003 sampler data. The modeling system consists of an urban airflow model (RUSTIC) and a Lagrangian particle-based dispersion model using stochastic tracer techniques (MESO). The focus was on near-field dispersion, in this case less than 1 km from the source, in the central business district of Oklahoma City. Comparisons were made between both surface and altitude samplers. Statistical measures were used to evaluate the model performance. The evaluations were made for six continuous releases, composed of both stable and unstable atmospheric conditions, as well as three different release configurations.

Overall, the models were shown to have an underprediction bias of 47% when using the geometric mean bias statistic. A possible influence to the underprediction is that the higher density of SF₆ in comparison with air was not taken into account in the simulations. The models were capable of predicting 42% of the sampler data within a factor of 2 and 83% of the data within a factor of 10. The results were further broken down into performance for varying meteorological conditions. For daytime releases, the models performed reasonably well; for nighttime releases, the models performed more poorly. Two possible causes of the poorer nighttime comparisons were (a) an inability to model the suppression of vertical turbulence because of the assumption of isotropy in RUSTIC’s $k-\omega$ turbulence model and (b) difficulty in modeling the light and variables inflow winds. The best comparisons were found for the three continuous daytime releases of IOP-4. It was hypothesized that these good comparisons were a result of steadier inflow conditions combined with the

![Fig. 12. Comparison of concentrations from (left) all IOP-4 releases for the steady-state runs and (right) using the WD weighting method to account for large-scale turbulence effects. The points are organized by distance from the source.](image1)

![Fig. 13. Plume centerline plot for the IOP-4 releases using the weighting WD method. The centerline is defined by the mean WD, and the ordinate is the modeled concentrations divided by the observed. The perfect correlation is the 1:1 line.](image2)
fact that the release site was more exposed and closer to the sodar used for the inflow meteorological conditions. The inclusion of large-scale atmospheric turbulence was investigated. A postprocessing method was used where concentrations from simulations with different wind directions were weighted according to the amount of time that was spent at a given direction. This improved the results significantly for samplers that were far from the plume centerline and was evident in the improvement of the FAC2 statistic from 42% to 51%.

The meteorological inputs are the most important factor for determining the urban flow and turbulence fields, which drive the dispersion of SF6 with MESO. However, the measurements were not sufficient (particularly temporal) to specify an upwind boundary condition more complex than an average profile. It is likely that the results presented here would be improved if a more complex atmospheric state could be specified to drive the models. In future urban field campaigns, we therefore recommend that a line of 10-Hz sonic anemometers mounted on towers at least 20 m above ground level (or profilers with high-frequency measuring capacity) be placed across the entire upwind boundary of the central business district (less than 500 m in) so that the most accurate representation of the upwind condition can be ascertained to drive these higher-fidelity models. This will remove a source of measurement uncertainty in the evaluation, leading to more definitive guidance for choosing between the more time-consuming CFD and fast-running Gaussian models.

The results presented herein are based from modeling six continuous releases. There were approximately 30 continuous releases conducted in Joint Urban 2003, as well as a number of instantaneous releases that have not been examined yet. Future work will be focused on comparing the models with these new data to obtain a better overall picture of the model performance.

Acknowledgments. This research was sponsored by the Defense Advanced Research Projects Agency. We thank Mr. Roger Gibbs of DARPA for his support of this research and development project. We thank Drs. Nathan Platt and Steve Warner of the Institute for Defense Analyses (IDA) for helpful discussions. Field data provided by the following Joint Urban 2003 participants were invaluable for this research effort: NOAA/ARL/FRD, VOLPE, LLNL, LBNNL, WSU, UU, DPG, ANL, and PNNL. We thank the following ITT personnel for helping to field the experiment: John Betz, John LeSage, Charles Tobin, Les Mann, Greg Boyle, Leldon Troyer, and Vernon Smith. This manuscript was improved by the helpful comments of three anonymous reviewers.

REFERENCES


