

Use of Salt Lake City URBAN 2000 Field Data to Evaluate the Urban Hazard Prediction Assessment Capability (HPAC) Dispersion Model

JOSEPH C. CHANG

George Mason University, Fairfax, Virginia

STEVEN R. HANNA

Harvard School of Public Health, Boston, Massachusetts

ZAFER BOYBEYI AND PASQUALE FRANZESE

George Mason University, Fairfax, Virginia

(Manuscript received 18 February 2004, in final form 9 August 2004)

ABSTRACT

After the terrorist incidents on 11 September 2001, there is a greatly heightened concern about the potential impacts of acts of terrorism involving the atmospheric release of chemical, biological, radiological, and nuclear (CBRN) materials in urban areas. In response to the need for an urban CBRN model, the Urban Hazard Prediction Assessment Capability (Urban HPAC) transport and dispersion model has been developed. Because HPAC is widely used by the Department of Defense community for planning, training, and operational and tactical purposes, it is of great importance that the new model be adequately evaluated with urban datasets to demonstrate its accuracy. This paper describes evaluations of Urban HPAC using the "URBAN 2000" urban tracer and meteorological field experiment data from Salt Lake City, Utah. Four Urban HPAC model configuration options and five plausible meteorological input data options—ranging from data-sparse to data-rich scenarios—were considered in the study, thus leading to a total of 20 possible model combinations. For the maximum concentrations along each sampling arc for each intensive operating period (IOP), the 20 Urban HPAC model combinations gave consistent mean overpredictions of about 50%, with a range over the 20 model combinations from no overprediction to a factor-of-4 overprediction in the mean. The median of the random scatter for the 20 model combinations was about a factor of 3 of the mean, with a range over the 20 model combinations between a factor of about 2 and 9. These performance measures satisfy previously established acceptance criteria for dispersion models.

1. Introduction

The potential impacts of the atmospheric release of chemical, biological, radiological, and nuclear (CBRN) or other hazardous materials are of increasing concern. Hazardous releases can occur due to accidents, such as the release of toxic industrial chemicals in Bhopal, India, in 1984 (e.g., Sharan et al. 1996) and the Chernobyl nuclear power plant disaster in the Ukraine in 1986 (e.g., Puhakka et al. 1990). They can also occur as an unintentional result of military actions, such as the U.S. destruction of rockets with chemical warheads at Khamisiyah, Iraq, after the 1991 Gulf War (Winkenwerder 2002). More recently, terrorist incidents in ur-

ban settings, such as the events on 11 September 2001 in New York City, New York, and Washington, D.C., and military conflicts dramatically raise concerns for the possibility of mass casualties. Therefore, improving the capability to accurately and quickly predict the dispersion of hazardous airborne materials in urban areas is critical for the safety of military and law enforcement personnel, first responders, and urban dwellers, in general.

Several transport and dispersion models are available for application to hazardous gas releases in rural areas. For example, the U.S. Environmental Protection Agency's (EPA's) Industrial Source Complex (ISC3) model (EPA 1995), American Meteorological Society (AMS)–EPA Regulatory Model Improvement Committee (AERMIC) Dispersion Model (AERMOD; Cimorelli et al. 2002), and California Puff (CALPUFF) model (Scire et al. 2000) have the widest use, but are most often applied to routine releases from industrial sources. The Areal Locations of Hazardous Atmo-

Corresponding author address: Dr. Joseph Chang, School of Computational Sciences, MS 5B2, George Mason University, Fairfax, VA 22030-4444.
E-mail: jchang4@scs.gmu.edu

spheres (ALOHA) model (NOAA 2004) and the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLOT) model (Draxler and Hess 1997) are intended for short-term releases from accidents or from natural disasters. The National Atmospheric Release Advisory Center (NARAC) model (Nasstrom et al. 2000) has been developed by the U.S. Department of Energy for use in a wide variety of hazardous chemical releases, and there is a center at Lawrence Livermore National Laboratory that is set up to run the model on a 24 h day⁻¹/7 days week⁻¹ basis.

Fewer models are capable of simulating transport and dispersion in urban areas, where the enhanced roughness and the modified surface energy balance cause lower winds, enhanced turbulence and mixing, and a tendency toward neutral conditions. ISC and AERMOD have urban algorithms that account for these effects. The Urban Dispersion Model (UDM; Hall et al. 2002) was developed in the United Kingdom, based on laboratory and field data, in order to account for enhanced dispersion and wake retention behind groups of buildings.

The Hazard Prediction Assessment Capability (HPAC; DTRA 2001) is a dispersion modeling system that is widely used by the U.S. Department of Defense community for planning, training, and tactical purposes. The dispersion model that is the subject of the current evaluations is known as "Urban HPAC." It combines the standard HPAC developed for rural environments with urban modeling capabilities, such as urban canopy wind and turbulence profiles (Cionco 1972; Macdonald et al. 1998), an urban dispersion model (i.e., UDM), and an urban flow model [i.e., the Urban Wind Field Model (UWM); Lim et al. 2002]. Because the new urban modeling capabilities within HPAC will potentially see many uses, given the heightened awareness of urban terrorism, an independent evaluation of the Urban HPAC model, using the data collected during the Salt Lake City, Utah, "URBAN 2000" urban tracer and meteorological field experiment (Allwine et al. 2002) was conducted. This paper summarizes the findings of the statistical and scientific evaluations for Urban HPAC in general, and for specific model configuration and meteorological input data options. Warner et al. (2004) present additional results based on a different evaluation methodology. The January 2003 version of Urban HPAC was used in this study.

A statistical model evaluation inspects model outputs to see how well they match observations. A scientific evaluation investigates the model algorithms, physics, and assumptions for their consistency, accuracy, and sensitivity (Hanna et al. 1991; Chang et al. 2003a). It also includes comparisons of predictions of model interim outputs, such as cloud width and speed, with observations. The need for an exploratory data analysis is also emphasized, where the basic characteristics and properties of the dataset are studied. Like most sophis-

ticated numerical models, there are a number of model configuration and input data options available for applying Urban HPAC. This paper also examines these options.

2. Brief description of the Urban HPAC model

As stated in the introduction, several urban dispersion models exist. The focus of the current paper is on the HPAC model (DTRA 2001), which has the following three main types of modules: 1) source term, 2) weather, and 3) transport and dispersion. Source-term modules estimate the source emissions characteristics for a variety of source scenarios, including chemical and biological facility damage, chemical and biological weapons, nuclear weapon explosion, missile intercept, etc. Source-term modules were not investigated in this study because the source terms for the URBAN 2000 field data were well known (within about 2%) as a result of careful measurements. Weather or diagnostic wind field modules generate mass-consistent gridded wind fields, based on observations or predictions generated by a mesoscale meteorological model, and knowledge of the underlying topography. HPAC has two weather modules: the Stationary Wind Fit and Turbulence (SWIFT) and the Mass-Consistent Algorithm in Second-Order Closure Integrated Puff model (MC-SCIPUFF). SWIFT is adapted from the MINERVE (méthode d'interpolation et de reconstitution tridimensionnelle d'un champ de vent) diagnostic model (ARIA Technologies 2001), and is the default choice for HPAC. MC-SCIPUFF is invoked in special cases when the input data are inadequate for SWIFT. Neither SWIFT nor MC-SCIPUFF account explicitly for wind speed profiles in urban canopies (i.e., below mean building height). The transport and dispersion module in HPAC is SCIPUFF (Sykes et al. 2000), which is a probabilistic Gaussian puff dispersion model that calculates both the mean and variance of the concentration field. This study emphasized the ensemble mean predictions.

Urban HPAC is based on the basic HPAC described above, in addition to the following three major urban modeling capabilities. First, SCIPUFF was extended for urban applications by accounting for urban canopy wind and turbulence profiles (Cionco 1972; Macdonald et al. 1998). These profiles are applied after the diagnostic wind field model produces its gridded wind fields. Second, UDM, version 2.2 (Hall et al. 2002), was integrated into HPAC as an option to calculate the dispersion in and over an urban area. UDM, originally developed by the Defense Science and Technology Laboratory (DSTL) in the United Kingdom, is based on an ensemble mean Gaussian puff dispersion methodology where surface obstacles within the urban canopy are allowed to modify dispersion patterns. The model treats the urban dispersion problem with three regimes. The "open" regime is for low (<5%) surface

obstacle density (fraction of surface area occupied by obstacles), where a dispersing puff interacts with individual obstacles in succession. The “urban” regime is for high (>5%) surface obstacle density, where the dispersing puff interacts with closely packed obstacles collectively. The “longer range” regime follows the previous two regimes at longer downwind distances where the puff becomes large compared to surface obstacles. UDM treats puff dispersion over an urban area by accounting for the traditional boundary layer turbulence and the aerodynamic disturbance due to the wind flow around surface obstacles. The puff aerodynamic dispersion rates and advection velocities for the urban regime are based on wind tunnel studies, such as Hall et al. (1998) and Macdonald et al. (1998). UDM currently handles terrain effects by importing modified wind field data that are created by separate meteorological models. Third, UWM, version 1.0 (Lim et al. 2002), was also integrated into HPAC as an option to calculate flow patterns as modified by buildings over an urban area. UWM, developed by Titan Corporation, is a computational fluid dynamics (CFD) code that calculates steady-state flow solutions inside the urban boundary layer using a canopy parameterization. Its purpose is to generate urban-scale wind patterns by modifying large-scale wind patterns to reflect the presence of obstacles. UWM spatially averages the Reynolds-averaged equations of motion in order to obtain rapid solutions for operational use. The model, hence, is not designed to resolve finescale winds within the urban roughness sublayer. As a result of the spatial averaging, the momentum and energy equations have extra terms that represent important physical processes due to the presence of obstacles. These physical processes include drag forces and turbulent kinetic energy production, and are parameterized by the urban canopy approach. UWM uses a staggered grid, where vector quantities are defined on cell faces, and scalar quantities are defined on cell centers. The model advances the solution forward in time using the leapfrog scheme, but the diffusion and drag terms are calculated semi-implicitly. UWM relies on other larger-scale gridded meteorological models to provide the initial conditions for its wind field calculations. UWM and UDM share common building databases (e.g., height and geometry) to support flow and dispersion calculations over urban areas. Building databases are usually generated by lidar imaging techniques.

Urban HPAC always uses the urban canopy parameterization in SCIPUFF as long as the surface type is specified as urban. In addition, the user also has the option of invoking UDM and UWM independently. When invoked, UDM and UWM only apply to a portion of the Urban HPAC modeling domain. This portion is currently hardwired as $2 \text{ km} \times 2 \text{ km}$ by the model developers. In other words, if the Urban HPAC modeling domain is greater than a $2 \text{ km} \times 2 \text{ km}$ square area, which is usually the case, then UDM and UWM

only apply to the $2 \text{ km} \times 2 \text{ km}$ urban subdomain centered on the source location. Flow and dispersion calculations for the rest of the modeling domain are still handled by the regular SWIFT and SCIPUFF modules. As mentioned previously, UDM and UWM are independent model options in Urban HPAC. The functions of UDM, if not invoked, will be replaced by SCIPUFF; and the functions of UWM, if not invoked, will be replaced by SWIFT.

3. URBAN 2000 field experiment

The URBAN 2000 field experiment was conducted in Salt Lake City (SLC), Utah, during October 2000. Allwine et al. (2002) give a detailed description of the experiment. The following is a brief summary. URBAN 2000 is a comprehensive field tracer study that covered distances from the source ranging from 10 m to 6 km. The source was at ground level in the downtown area. Atmospheric meteorological and tracer experiments were conducted to investigate transport and dispersion around a single downtown building, through the downtown area, and into the greater SLC urban area. The study area was extended beyond the urban scale, making use of the Department of Energy’s Vertical Transport and Mixing (VTMX; Doran et al. 2002) tracer and meteorological study conducted simultaneously on a larger domain in the greater Salt Lake Valley (Fig. 1).

The SLC metropolitan area is located in a large mountain valley about 50 km long and 25 km wide, with mountains to the east (peak $\sim 3400 \text{ m}$), to the west (peak $\sim 3000 \text{ m}$), and to the south (peak $\sim 2000 \text{ m}$); and the Great Salt Lake to the northwest (Fig. 1). The valley is bisected north to south by the Jordan River, which flows from the Utah Lake to the south through a mountain gap to the Great Salt Lake to the north. The average elevation of SLC is approximately 1300–1400 m above sea level, although steep mountain slopes are on the eastern edge of the city.

Sulfur hexafluoride (SF_6) tracer gas was released from a point source or from a 30-m line source near street level in the downtown SLC area during six intensive operating periods (IOPs). There were three SF_6 release trials for each IOP, and all took place at night. A star near the center marks the release point, and black dots mark the SF_6 monitors in Fig. 2, or the “urban domain,” which corresponds to the inner solid square in Fig. 1. The three sampling arcs are visible about 2, 4, and 6 km to the northwest of the release point. Figure 3 shows the $1.3 \text{ km} \times 1.3 \text{ km}$ area known as the URBAN 2000 “downtown domain,” where samplers were located on block intersections and midway along the blocks. These samplers in the downtown domain were used by Hanna et al. (2003), and in the current study, to define four additional arcs at distances from about 0.15 to 0.9 km. Therefore, a total of seven sampling arcs can be defined for URBAN 2000.

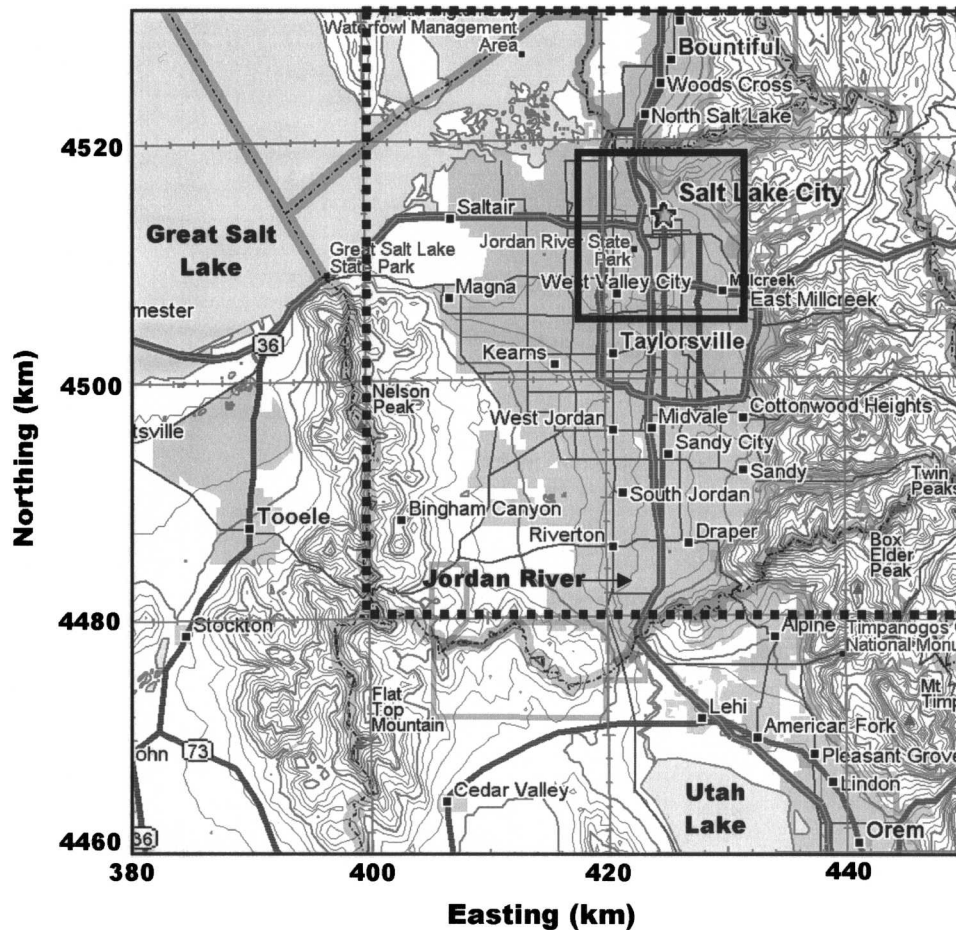


FIG. 1. Map of the Salt Lake valley. The dashed square indicates the Urban HPAC modeling domain (50 km \times 50 km). The small inner square (the urban domain) is further illustrated in Fig. 2, with additional information on terrain, and meteorological and tracer sampling stations. Map originally prepared with the DeLorme Topo USA 4.0 software package.

For each trial, the SF₆ release was maintained at a constant rate for 1 h. The release rate was about 1 g s⁻¹, beginning at 0000, 0200, and 0400 mountain standard time (MST) for all of the IOPs except IOP9. The release rate was 2 g s⁻¹ beginning at 2100, 2300, and 0100 MST for IOP9. The 1-h interval between SF₆ releases was planned to allow sufficient time for a “gap” to appear between the released clouds, so they could be easily distinguished. The SF₆ concentrations analyzed in this paper were reported as 30-min averages over a 6-h period during each night. The 6-h duration was planned to allow sufficient time to capture concentration data from all three releases. However, for some of the light-wind IOPs, the cloud from the third release had not reached the most distant (6 km) arc before the sampling ended.

Meteorological monitors are also shown in Fig. 2. For example, the SLC National Weather Service (NWS) anemometer is at the airport in the northwest corner of the figure. The N01 surface anemometer, N02 sodar, and N03 profiler sites are collocated in a suburban area

known as “Raging Waters” (RGW) about 6 km southwest of the urban area. D11 marks a sodar at the top of a 36-m building. The N02 and D11 sodars provide winds every 5 m, starting 15 m above the sodar. The M02 anemometer is mounted 3 m atop a 121-m building. The surface roughness z_0 is estimated by Hanna et al. (2003) to be about $0.15H_b = 2.25$ m, based on an average building height H_b of 15 m for the SLC downtown area.

4. Urban HPAC model configuration and meteorological input data options

As previously mentioned, a number of enhancements, such as the urban canopy parameterization in SCIPUFF, the UDM, and the UWM, have been implemented in order to provide HPAC with urban modeling capabilities. To fully evaluate model performance, four optional Urban HPAC configurations were considered, as listed in Table 1. Option “UC” refers to the Urban

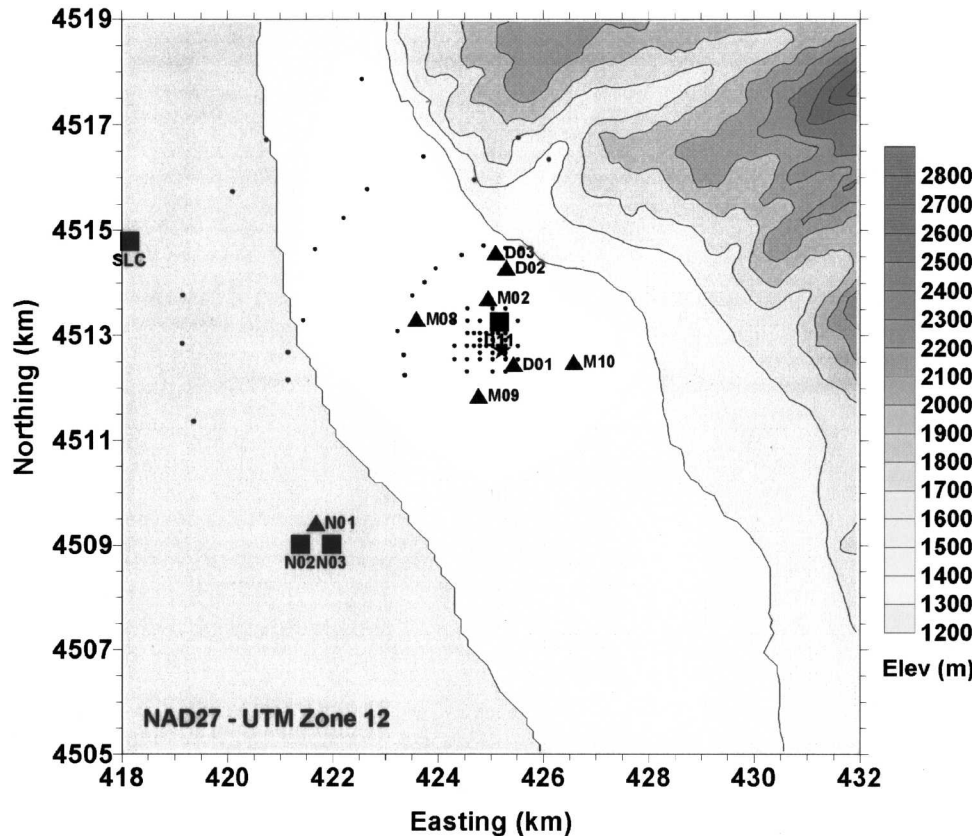


FIG. 2. Map of the Salt Lake City area (urban domain) showing the release location (star), terrain elevations (m), and locations of tracer samplers (small dots) and meteorological measurement sites (triangles indicate surface sites, and squares indicate vertical profile sites, where D11 and N02 are sodar sites, N03 is a profiler site, and SLC is a radiosonde site). Not shown are four sonic anemometers surrounding (within ~ 50 m) the release point.

SCIPUFF baseline case prior to the addition of UDM or UWM, that is, SCIPUFF with its urban canopy (UC) parameterization. Option “DM” refers to the use of UDM for the dispersion calculations over the $2 \text{ km} \times 2 \text{ km}$ urban subdomain, where UWM is not invoked. Option “WM” refers to the use of UWM for the flow calculations over the $2 \text{ km} \times 2 \text{ km}$ urban subdomain, where UDM is not invoked. Option “DW” means that both UDM and UWM are invoked for the dispersion and flow calculations, respectively, over the $2 \text{ km} \times 2 \text{ km}$ urban subdomain.

All Urban HPAC model runs for this study involved a $50 \text{ km} \times 50 \text{ km}$ modeling domain (the dashed square in Fig. 1), enclosing the SLC downtown area. The vertical domain extends to 2.5 km above ground. HPAC uses adaptive nested grids to efficiently account for contributions from puffs to surface dosage and deposition integrals. The local grid resolution is determined in such a way to adequately resolve those puffs whose contributions to surface integrals are significant. The terrain, land use, and building morphology data that came with the Urban HPAC software package were

used. In general, runs were made using default model options whenever possible.

In addition to the four model configuration options discussed above, five plausible meteorological input options, ranging from data-sparse to data-rich scenarios, were also considered (Table 2). The justifications for each of the five meteorological input data options are further provided below. Option “SLC” uses only the SLC airport wind data. Airport data are routinely used in many dispersion calculations, because they are usually the only available quality-controlled data. However, there is a potential problem with airport observations because they may be unrepresentative of the conditions near a release, because the airport may be very far from the release location and the time of radiosonde launches may not coincide with the time of the release. As seen in Fig. 2, the SLC airport is located about 10 km from the downtown area.

Option “LDS” uses only the wind data from the top (121 m) of the tallest building—the Latter-Day Saints (LDS) building. This option explores the feasibility of installing a single wind sensor at the top of a represen-

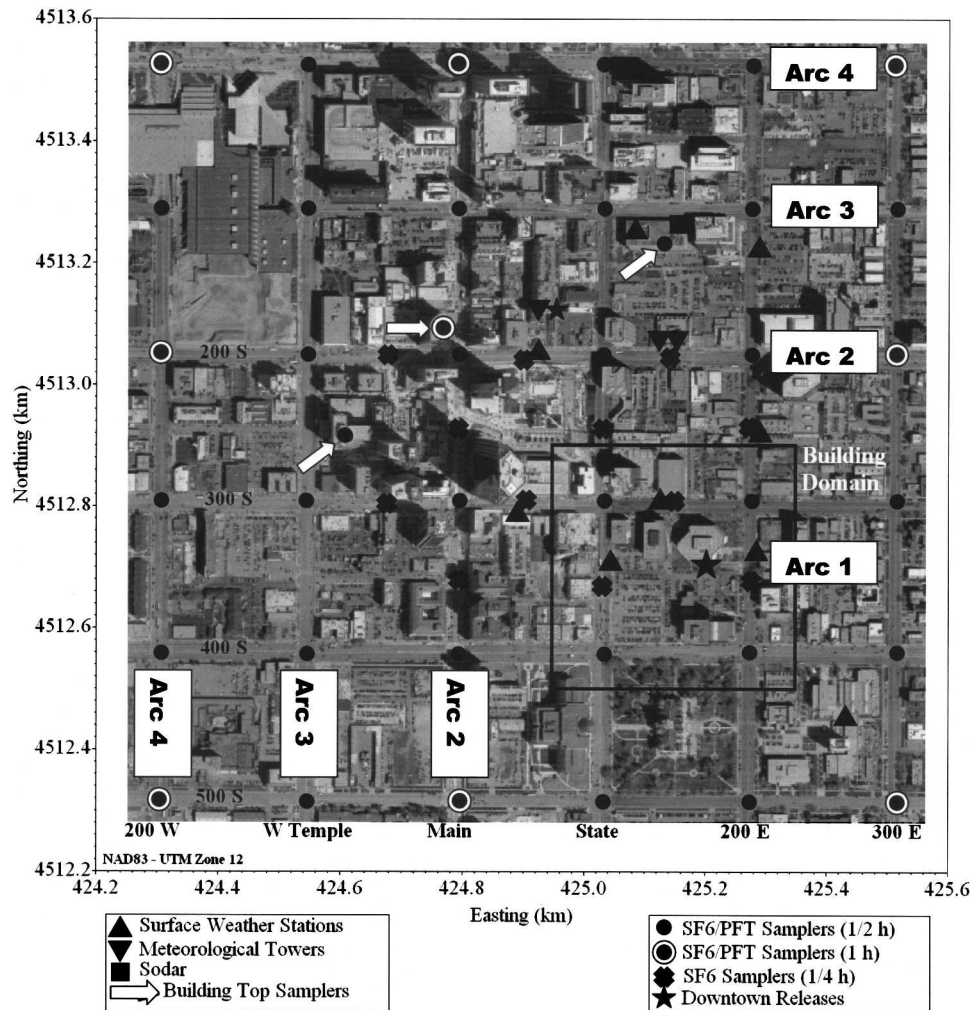


FIG. 3. Map of the downtown Salt Lake City area (downtown domain), with locations of tracer samplers (solid dots), meteorological instruments (triangles and squares), and source (star). The hexagonal building just north of the source is the Heber–Wells building. The four hypothetical arcs are also labeled. Figure courtesy of Allwine et al. (2002).

tative downtown building in each major city in order to provide the necessary weather data for urban dispersion modeling. Option “RGW” uses the wind data only from the upwind site. This option investigates the use-

fulness of the data from a single upwind profile (Raging Waters) for urban dispersion modeling, and the ability of Urban HPAC to adjust the upwind flow pattern that is approaching an urban area. Option “ALL” uses all of

TABLE 1. Urban HPAC model configuration options considered in this study.

Model configuration option	Modules			Remarks
	UWM	UDM	HPAC/ SCIPUFF*	
UC	Not used	Not used	Used	Baseline urban capability available to current operational HPAC users; probabilistic output available
DW	Used	Used**	Used	Expected normal operational urban mode for future versions of HPAC
DM	Not used	Used	Used	Might be used if response time is critical
WM	Used	Not used	Used	Probabilistic output available

* HPAC/SCIPUFF with surface type entered as “urban” to invoke empirical parameterizations of urban canopy wind and turbulence profiles.

** UDM used with the terrain channeling effects disabled to avoid double counting urban effects on wind fields.

TABLE 2. Urban HPAC meteorological input data options considered in this study. See Fig. 2 for the locations of various stations.

Meteorological input data option	Remarks
SLC	Hourly surface and twice-daily upper-air data from Salt Lake City airport (station SLC).
LDS	Measurements from a single meteorological observation station (station M02) at the top of the LDS administration building (121-m height) in the downtown area. The LDS building is one of the highest buildings in SLC.
RGW	Data from a single boundary layer meteorological profile at the Raging Waters site (station RGW), located about 5 km upwind (southwest) of the urban area. The site includes a surface anemometer (station N01, 10 m above ground), a sodar (station N02, collecting data from 40 m up to 300 m above ground), and a profiler (station N03, collecting data from 170 m up to 3700 m above ground).
ALL	All available meteorological observations, except for sonic anemometer measurements in street canyons, collected during URBAN 2000. This option includes stations SLC, M02, N01, N02, and N03 mentioned above, plus surface stations D01, D03, M08, M09, and M10, and sodar station D11 in the downtown area.
OMG	Data from the simulations by the OMEGA meteorological model. The horizontal grid interval for the OMEGA runs was $\sim 2\text{--}3$ km.

the available URBAN 2000 wind observations in the Salt Lake City area shown in Fig. 2. This option aims to evaluate model performance in a data-rich observational environment. Note that the sonic anemometer data measured at street level (1.5-m height) were not included in this option. Option “OMG” makes use of the Operational Multiscale Environment Model with Grid Adaptivity (OMEGA) prognostic meteorological model (Bacon et al. 2000). This option is being explored because, in an operational environment, observed data are often sparse or missing over the area of interest, and a plausible source of necessary high-resolution meteorological inputs would be from a prognostic meteorological model (Anthes et al. 1989; Boybeyi and Bacon 1996). OMEGA was run at an $\sim 2\text{--}3$ km grid interval to provide meteorological inputs for Urban HPAC. Chang et al. (2003b) describe how OMEGA was configured in detail.

Therefore, a total of 20 combinations of Urban HPAC model configuration options and meteorological input options were considered in this study. These 20 combinations represent a range of possible ways of applying Urban HPAC. Table 3 summarizes the keywords used later in this paper to indicate various model and weather options, and their combinations. Most of the model performance results are expressed as a median and a range over the 20 model combinations listed in Table 3.

Intuitively, it is expected that the most “sophisticated” combination, that is, the DW model configura-

tion option (i.e., UDM and UWM) coupled with the ALL meteorological input data option (i.e., all available meteorological observations), would lead to the best model performance. However, as indicated later, this is not the case. This can be due to a variety of reasons, such as different assumptions in processing and interpreting meteorological data, possible incompatibility among modules, and shortcomings in module algorithms. Because urban modeling is still in its infancy and much subject matter experience is still lacking, the technical manuals of SCIPUFF (Sykes et al. 2000), UDM (Hall et al. 2002), and UWM (Lim et al. 2002) do not provide guidance on the optimal configuration and meteorological input data options for an urban scenario. It is hoped that the results of this study will provide some preliminary guidance.

5. Exploratory analysis of URBAN 2000 data

This section shows some results of exploratory analysis of the URBAN 2000 data in order to summarize the field data and to identify straightforward physical facts that may aid understanding of the statistical model evaluation results. More details are given in Hanna et al. (2003) and Chang et al. (2003b).

a. Wind observations

Table 4 lists the vector-averaged wind observations from 12 locations (see Fig. 2) below and above the

TABLE 3. Keywords for the four Urban HPAC model configuration options and the five meteorological input options, and the resulting 20 combinations.

Meteorological input options	Urban HPAC model configuration options			
	UC	DM	WM	DW
SLC	UC-SLC	DM-SLC	WM-SLC	DW-SLC
LDS	UC-LDS	DM-LDS	WM-LDS	DW-LDS
RGW	UC-RGW	DM-RGW	WM-RGW	DW-RGW
ALL	UC-ALL	DM-ALL	WM-ALL	DW-ALL
OMG	UC-OMG	DM-OMG	WM-OMG	DW-OMG

TABLE 4. Observed wind speeds (m s^{-1}) from fixed URBAN 2000 anemometer sites in Salt Lake City, averaged over each IOP and all IOPs for each anemometer, and averaged over anemometers D01, D03, M02, M08, M09, and M10 for each IOP. See Fig. 2 for anemometer locations.

Wind instrument height (m AGL)	D01	D03	G01	G02	G04	G06	M02	M08	M09	M10	N01	SLC	Avg u over D and M anemometers
IOP2	0.29	0.60	0.15	0.30	0.12	0.21	1.06	1.28	0.46	0.52	1.15	2.57	0.70
IOP4	0.44	1.26	0.30	0.50	0.21	0.25	1.43	1.67	0.59	0.77	1.60	3.57	1.03
IOP5	0.42	1.65	0.37	0.21	0.27	0.33	0.63	1.27	0.53	0.44	1.12	2.71	0.82
IOP7	0.42	1.33	0.32	0.39	0.27	0.35	1.32	1.77	0.78	0.77	1.22	3.43	1.07
IOP9	1.33	2.43	0.73	1.29	0.66	0.49	5.45	3.64	1.91	1.11	4.47	7.38	2.64
IOP10	0.75	1.56	0.51	0.67	0.57	0.41	3.40	2.75	1.27	0.73	3.37	6.00	1.72
Avg	0.63	1.51	0.38	0.52	0.33	0.33	2.29	2.10	0.95	0.73	2.18	4.31	1.37

urban canopy layer during each IOP of URBAN 2000. The four “G” instruments in Table 4 were sonic anemometers and were mounted at a height of 1.5 m in the area around the large hexagonal building near the source in Fig. 3. The average wind speed listed in Table 4 is based on the two “D” anemometers and the four “M” anemometers, and is used in the general analysis. Table 4 shows that observed wind speeds were very light (about $0.2\text{--}0.5 \text{ m s}^{-1}$) at street level (1.5-m height) and were about $1\text{--}2 \text{ m s}^{-1}$ at a height of 50 m for IOPs 2, 4, 5, and 7. Wind speeds were higher (about 1 m s^{-1} at street level and $4\text{--}5 \text{ m s}^{-1}$ at 50 m) for IOPs 9 and 10. The NWS anemometer at the SLC airport, located in flat, open terrain, has wind speeds consistently about 3 times as large as in the urban area. Monitor N01, at the Raging Waters suburban site upwind of the city, has wind speeds in between those at the SLC airport and those in the urban area.

Two of the meteorological input options in this study involve observations from the downtown area (see Table 2). As mentioned earlier, the LDS option was wind monitor M02 mounted 3 m atop the 121-m-tall LDS administration building. The ALL option included eight wind anemometers (D01, D03, M02, M08, M09, M10, N01, and SLC) and four upper-air stations (D11, N02, N03, and SLC). As an illustration of the time and space variability of the winds in this urban area, Fig. 4 shows time series of 15-min-averaged wind vectors for the near-surface anemometers for IOP7, a low-wind case (see Table 4). Relatively large variability in speed and direction is evident, with an apparent 1.5-to 2-h periodicity or “pumping” of the winds during this nighttime period. This periodicity is roughly in phase at all anemometer sites. Possible reasons for the 2-h period are drainage flow fluctuations and “sloshing” of the stable air mass in the valley. It is found later that there was not an improvement in HPAC predictions when all the meteorological stations are input (the ALL weather input option), and this may be due to the variability seen in Fig. 4.

b. Concentration observations

The maximum observed 30-min-averaged concentration C_{max} was identified on each of the seven sampling

arcs. In some cases, there were problems because concentrations on an arc were all near or below the instrument threshold (say, $< 45 \text{ ppt}$), or there was perhaps only a single high observation, or the plume was obviously on the edge of the sensor network on that arc. The C_{max} values for these problem cases were not used in subsequent analysis. As a further check, the concentration observations for each IOP were analyzed for continuity in space and time, as seen in Fig. 5 for IOP4 for each of the seven arc distances. For each of the seven arcs, time series of 30-min-averaged arc-maximum concentrations normalized by the emission rate (C_{max}/Q) are plotted. The three source releases (from 0000 to 0100, 0200 to 0300, and 0400 to 0500 MST) can be distinguished on the figure, and a delay

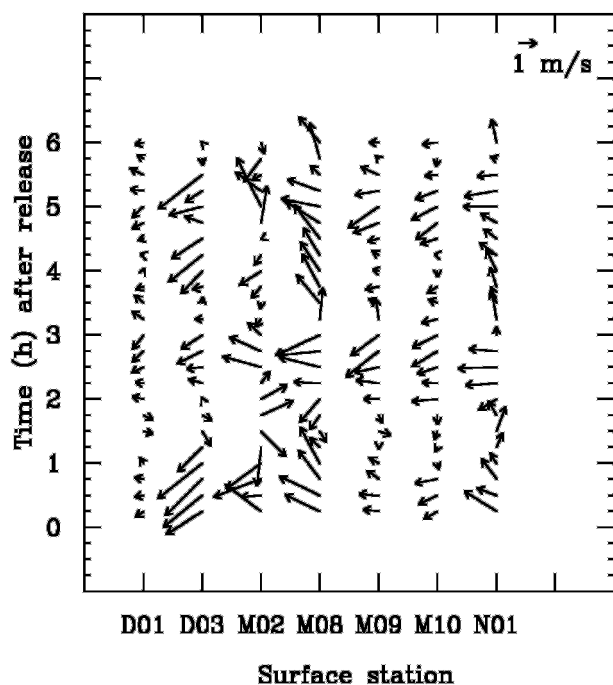


FIG. 4. Fifteen-minute vector-averaged surface wind fields for anemometers D01, D03, M02, M08, M09, M10, and N01 for IOP7, which began at 0000 MST 18 Oct 2000. Time refers to period ending.

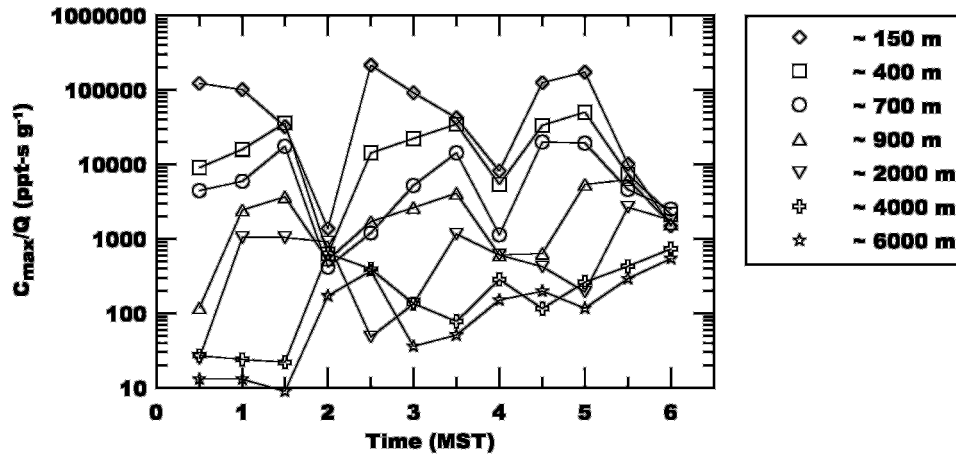


FIG. 5. Time series of observed 30-min-averaged arc-maximum concentrations normalized by the emission rate (C_{\max}/Q) for IOP4 during URBAN 2000, for the seven monitoring arcs, at downwind distances in meters given in the legend; SF_6 releases occurred between 0000 and 0100, 0200 and 0300, and 0400 and 0500 MST 9 Oct 2000.

can be seen in the arrival of the cloud at distant arcs. For example, the delay of about 1.5 h in the arrival of the peak at the 6-km arc is consistent with the 1 m s^{-1} observed average wind speed (see Table 4). The figure also shows that the time scale is about 30–60 min for a decrease of concentration by a factor of 10 at the closest ($\sim 150 \text{ m}$) monitoring arc.

The observed 1-h-averaged C_{\max}/Q values are listed in Table 5 for each of the seven arcs in each of the 18 trials. The average wind speed within the urban canopy is listed for each IOP and trial. The observed C_{\max}/Q on each arc averaged over the 18 trials is given in the bottom row. Hanna et al. (2003) show that the mean C_{\max}/Q values follow an approximate $x^{-1.5}$ power law, in agreement with observations at other field studies.

6. Brief description of statistical model performance measures

Standard statistical model performance measures (e.g., Hanna et al. 1991, 1993; Chang and Hanna 2004; Chang et al. 2003a) have been used in this study to appraise model performance. In particular, the following five statistical performance measures were calculated, which include the fractional bias (FB), the geometric mean bias (MG), the normalized mean square error (NMSE), the geometric variance (VG), and the fraction of predictions within a factor of 2 of observations (FAC2):

$$\text{FB} = \frac{(\overline{C_o} - \overline{C_p})}{0.5(\overline{C_o} + \overline{C_p})}, \tag{1}$$

$$\text{MG} = \exp(\overline{\ln C_o} - \overline{\ln C_p}), \tag{2}$$

$$\text{NMSE} = \frac{(\overline{C_o - C_p})^2}{\overline{C_o} \overline{C_p}}, \tag{3}$$

$$\text{VG} = \exp[\overline{(\ln C_o - \ln C_p)^2}], \text{ and} \tag{4}$$

$$\text{FAC2} = \text{fraction of data that satisfy } 0.5 \leq \frac{C_p}{C_o} \leq 2.0, \tag{5}$$

where C_p is model predictions, C_o is observations, and an overbar ($\overline{}$) is the average over the dataset. Depending on the need, additional performance measures related to false positives and false negatives can also be defined (e.g., Warner et al. 2001, 2004). The following three model concentration outputs were considered in this study: 1) the overall maximum 30-min-averaged concentration anywhere in the domain over all IOPs, that is, no pairing at all; 2) the maximum 30-min-averaged concentration along a sampling arc for each IOP, that is, paired only by distance and by IOP; and 3) the concentration at each sampler location and for each averaging (30 min) period, that is, fully paired in space and time.

The fractional bias is a normalized mean bias that estimates the systematic overprediction or underprediction of a model with respect to measurements. A model might give $\text{FB} = 0.0$ (i.e., no mean bias) even though predictions are completely out of phase with observations. Also, FB can be overly influenced by a single large over- or underprediction at high concentrations. The geometric mean bias MG is similar to FB in that it measures systematic bias, but on a logarithmic scale. For a dataset with a wide range in values, this logarithmic form of relative bias is probably more appropriate because low and high values are weighted equally. However, logarithmic measures are found to be sensitive to extremely low values. NMSE and VG are measures of scatter and reflect both systematic and unsystematic (random) errors, where NMSE is on a linear

TABLE 5. Observed hourly average maximum concentration normalized by the emission rate (C_{\max}/Q ; ppt g^{-1}) for the seven monitoring arcs and the 18 trials in URBAN 2000. The symbol “N/a” means that the data did not meet acceptance criteria. The third column also lists the average wind speed within the urban canopy (i.e., the average over anemometers D01, D03, M02, M08, M09, and M10). The bottom row of the table contains the observed C_{\max}/Q for each monitoring arc averaged over all IOPs and trials.

IOP	Trial	u (m s^{-1})	Sampling arc and distance (m) from source						
			Arc 1	Arc 2	Arc 3	Arc 4	Arc 5	Arc 6	Arc 7
			~150	~400	~700	~900	~2000	~4000	~6000
2	1	0.81	58 294	14 606	2606	657	717	361	86
2	2	0.61	77 284	19 028	523	N/a	N/a	N/a	N/a
2	3	0.5	67 174	15 927	1332	255	N/a	N/a	N/a
4	1	1.13	111 248	22 037	6495	2972	1040	516	272
4	2	0.94	153 413	28 385	5413	2330	806	172	174
4	3	0.76	105 156	34 257	11 119	3872	2073	464	418
5	1	0.64	27 450	14 183	4073	2532	710	218	547
5	2	0.91	45 761	14 771	3541	2495	750	248	242
5	3	1.06	73 798	21 706	3688	2349	1484	275	229
7	1	1.01	95 413	34 440	5945	3284	1116	363	N/a
7	2	1.04	36 807	4734	5303	1681	1890	453	475
7	3	1.21	38 092	13 908	7670	6679	1927	574	339
9	1	2.69	23 725	9101	8239	875	N/a	116	90
9	2	2.47	44 716	10 422	5963	N/a	N/a	194	185
9	3	3.23	21 156	6844	2018	1387	483	275	N/a
10	1	1.51	28 991	6055	1945	743	372	218	N/a
10	2	2.16	28 147	5853	1806	1486	633	343	N/a
10	3	2.31	13 376	4165	774	327	272	106	N/a
Avg over all IOPs and trials		1.39	58 330	15 578	4367	2128	1020	306	279

scale and VG is on a logarithmic scale. FAC2 is probably the most robust of the five measures, because high and low outliers do not overly influence it. The bootstrap resampling procedure (Efron 1987; Efron and Tibshirani 1993) can be used to estimate the confidence limits of various performance measures in order to test such hypotheses as 1) whether the FB for a model is significantly different from zero, and 2) whether the FBs for two models are significantly different.

For the five performance measures defined above, FB is bounded between -2 and 2 , and FAC2 is bounded between 0 and 1 . On the other hand, the remaining three measures are unbounded, with $\text{NMSE} \geq 0$, $\text{MG} > 0$, and $\text{VG} \geq 1$.

Values of performance measures can also be expressed in more understandable terms. For example, an FB of -0.67 represents a relative overprediction bias of a factor of 2, and an FB of 0.67 represents a relative underprediction bias of a factor of 2. A geometric mean bias of 0.5 and 2.0 represent a factor of 2 in the mean over- and underprediction, respectively. A relative or normalized mean square error (NMSE) of 1 and 4 roughly represent a ratio of the scatter to the mean of 1.0 and 2.0 , respectively. A geometric variance VG of 1.6 and 6.8 represent a scatter that is roughly a factor of 2 and 4 of the mean, respectively.

A key question of the study is whether the Urban HPAC performance meets criteria for model acceptance. Chang and Hanna (2004) have estimated the criteria for an “acceptable” model, based on the evaluations by them and by others of many models with many field databases. For example, a “good” or acceptable

model would have a relative mean bias (FB) with a magnitude ranging from -0.67 to $+0.67$ (plus and minus a factor of 2), a relative mean square scatter (NMSE) of less than about 4 (corresponding to scatter equal to 2 times the mean), and the fraction of predictions within a factor of 2 (FAC2) more than about 0.5. These suggestions are for maximum concentrations at given arc distances, and the criteria should be relaxed somewhat if the data are paired in time and space. Furthermore, these suggestions primarily apply to research-grade tracer studies, such as the Salt Lake City URBAN 2000 dataset being used here.

7. Statistical evaluation of concentration predictions by Urban HPAC

The primary objective of the statistical evaluation exercise is to evaluate the performance of Urban HPAC against the URBAN 2000 field data. Because 20 possible ways of running Urban HPAC were considered (see section 4), that is, combining four model configuration options with five meteorological input data options, the results are given as lists of the ranges of model performance measures over the 20 combinations. The model combinations that yield more satisfactory results for URBAN 2000 are identified.

It is important to emphasize that the evaluation results here are based on a single research-grade field experiment, URBAN 2000, with a limited range in the prevailing wind direction (i.e., from the southeast). The experiments were also limited to nighttime in the early

fall. Because model performance is likely to change once different field datasets are used, more robust conclusions on the Urban HPAC performance can only be obtained after considering a wider array of conditions in Salt Lake City and expanding the evaluations to additional field datasets.

a. Determination of sets of observations and predictions to be evaluated

In any statistical exercise for evaluating model performance, there are many ways to define the model outputs to be used. In particular, as discussed in section 6, this paper mainly focuses on the following three options for data pairing:

- single overall maximum 30-min concentration anywhere in the domain for all IOPs, that is, unpaired in space and time;
- maximum 30-min concentration on each sampling arc for each IOP, that is, paired by arc distance and by IOP;
- 30-min concentration at each sampler location for each time period, that is, fully paired in space and time. {For this option, we further considered only those cases when both observed and predicted SF₆ concentrations are greater than 45 ppt—the value [i.e., the method limit of quantitation (MLOQ)] above which the experimenters have confidence that the observed value is accurate within about 30% (R. Carter 2003, personal communication).}

It is evident that the number of data pairs involved for evaluation increases from the first to the last analysis option, and that the data acceptance criterion mainly impact the last analysis option. The SF₆ concentrations from 66 whole-air samplers (Fig. 2) were included in the evaluation. These 66 samplers were further grouped according to arc distances in order to determine the spatial dependence of model performance. As described above, there are seven arcs in total, with arc distances of ~150, 400, 700, 900, 2000, 4000, and 6000 m from the source. The first four arcs, which are considered to be in the downtown domain (Fig. 3), were grouped by Hanna et al. (2003) in their analysis of the basic physical characteristics of the URBAN 2000 data. The last three arcs are well defined, as seen in Fig. 2.

b. Results of statistical analysis

1) SINGLE OVERALL MAXIMUM CONCENTRATION ANYWHERE IN THE DOMAIN AND FOR ALL IOPs (I.E., UNPAIRED IN SPACE AND TIME)

The simplest comparison is of the single, unpaired in time and space maximum observed 30-min-averaged concentration anywhere in the domain for all IOPs. In other words, the single maximum value is selected from observations and predictions for each of the 20 model combinations. This maximum concentration is of im-

TABLE 6. Median and range of the overall maximum predicted 30-min SF₆ concentrations anywhere in the domain over all IOPs for each Urban HPAC model configuration option, where each model option is further coupled with five weather input options. Thus, the range refers to over five weather input options. The observed overall maximum is 173 430 ppt.

Urban HPAC model configuration option	Median (ppt)	Range (ppt)
UC	450 350	60 270–969 170
DM	142 800	93 600–230 730
WM	412 420	136 370–505 910
DW	216 750	146 800–287 030

portance because it defines the maximum expected health impact to an individual. These maximum values always occurred on the closest ($x \sim 150$ m) monitoring arc. The observed maximum concentration was 173 430 ppt. For each of the four model configuration options, there are five weather input options. Table 6 lists the median and the range of the five predictions (due to different weather inputs) for each of the four model options. It is seen that the DM and DW options produced predictions of the overall maximum that were always within a factor of 2 of the observed maximum. The UC and WM options tended to overpredict by a factor of 2–3, because these options do not satisfactorily account for the enhanced dispersion due to buildings and obstacles. This trend will be seen to continue with the evaluations based on other options of data pairing.

2) MAXIMUM CONCENTRATION ON EACH SAMPLING ARC FOR EACH IOP (I.E., PAIRED BY ARC DISTANCE AND IOP)

In this case, the maximum observed and predicted 30-min concentration along each of the seven arcs was selected for each IOP. Note that each IOP lasted 6 h and had three 1-h releases, but we have picked the maximum 30-min concentration on each arc for the total 6-h period. Therefore, the total number of data pairs involved is 42 (7 arcs \times 6 IOPs). Figure 6 provides a concise summary of the performance measures FB, NMSE, MG, and VG [see Eqs. (1)–(4) for definitions], together with their confidence intervals based on the bootstrap resampling, for the 20 Urban HPAC combinations. Considering all of the 20 model combinations, the fractional bias FB has a median of -0.37 (i.e., $\sim 50\%$ mean overprediction), and a range from -0.02 to -1.17 (i.e., between a 2% and a factor of 4 overprediction) ignoring the UC–OMG combination. (It is the only model combination that has a positive FB, or a mean underprediction.) The median geometric mean MG is 0.49, and suggests a median overprediction bias of about a factor of 2. The range in MG is 0.24–0.75, corresponding to a mean overprediction of 33% to a factor of 4. Excluding the outlier, the median normalized mean square error NMSE is 1.7, implying a median random scatter of roughly a factor of 3 of the mean. The

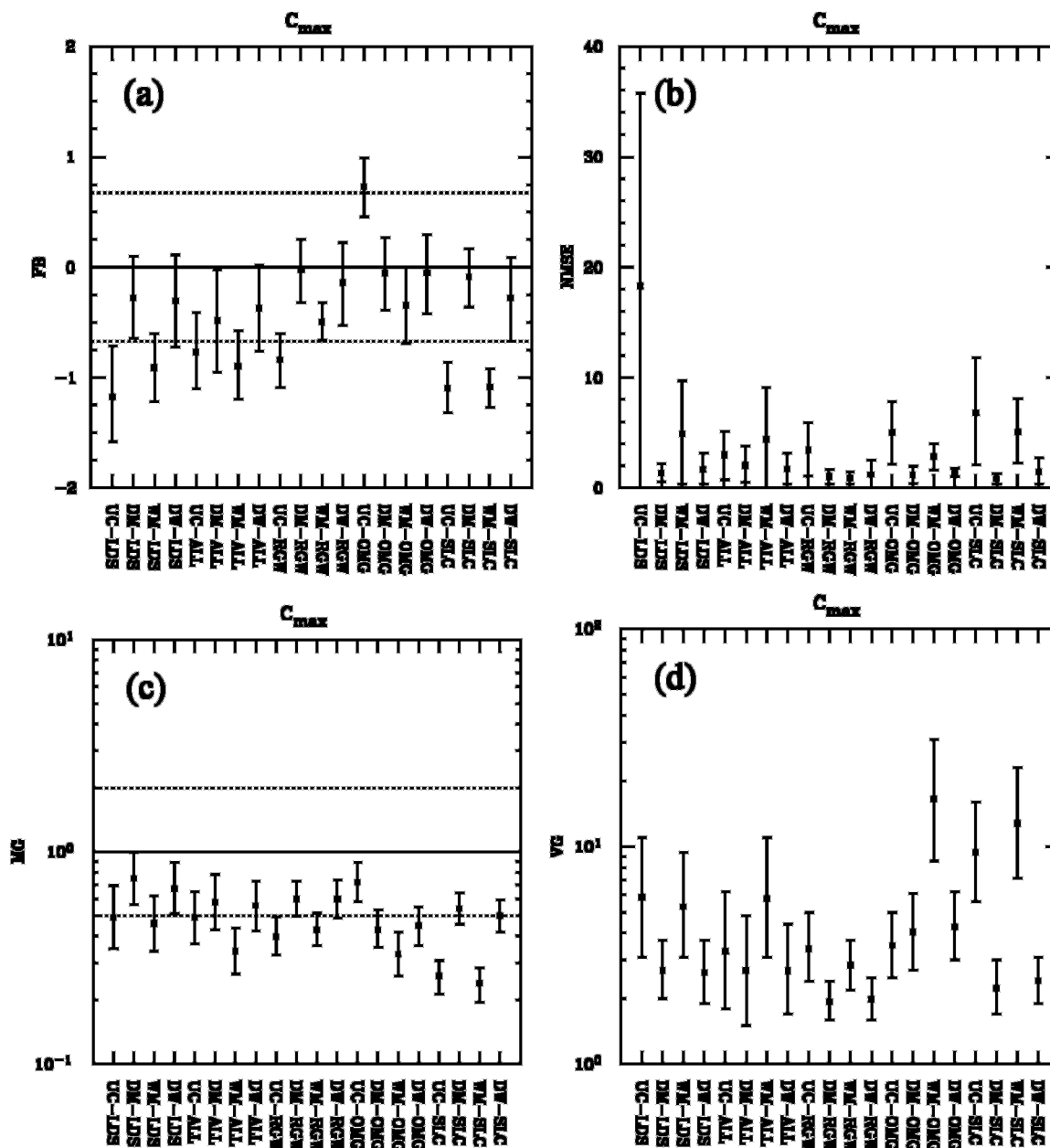


FIG. 6. Statistical measures (a) FB, (b) NMSE, (c) MG, and (d) VG, together with their confidence intervals, based on the arc-maximum 30-min SF_6 concentrations (C_{\max}) over each IOP, for each combination of the model configuration and meteorological input data options (20 combinations in total). See Eqs. (1)–(4) for definitions. Negative FB and $\text{MG} < 1$ indicate overprediction. Positive FB and $\text{MG} > 1$ indicate underprediction.

range in NMSE is 0.8–6.8, implying a random scatter of about a factor of 2–9 of the mean. For all the 20 model combinations, the values of the geometric variance VG are between 1.9 and 16.5, that is, between a factor of 2 and 5 random scatter, with a median scatter of a factor of 3 of the mean. These median biases and scatters are within the model acceptance criteria suggested by Chang and Hanna (2004) that are listed earlier.

For this dataset, the median FAC2 for the 20 model

combinations (not shown on the figure) is $\sim 40\%$ with a range between 5% and 58%. This median FAC2 is also comparable to the range of acceptable model performance.

For all the 20 model combinations considered in this study, the DM–RGW model combination (i.e., UDM coupled with the upwind Raging Waters wind input) has slightly better performance, followed closely by the DM–SLC (i.e., UDM coupled with the Salt Lake City

airport observations) and the DW-RGW (i.e., UDM and UWM coupled with the upwind Raging Waters wind input) model combinations. Without the urban upgrades (i.e., the UDM and/or UWM) and onsite, research-grade meteorology, the operational mode of HPAC is equivalent to the UC-SLC model combination (i.e., basic urban canopy algorithm coupled with the airport weather data). This combination tends to yield larger (a factor of 2–4) mean overpredictions.

It can also be seen from Fig. 6 that overall the DM model configuration option yields slightly better performance (i.e., less overprediction) than do the WM and UC model options. The RGW weather option generally leads to better model performance. The OMG weather option with data from the OMEGA mesoscale meteorological model also appears to lead to better agreement, except for when combined with the urban canopy parameterization (combination UC-OMG). The LDS weather option, with data from downtown Salt Lake City, and the ALL weather option, with data from many stations (Table 2), including downtown, do not improve model performance, which is an unexpected result. A possible reason is that downtown observations exhibited relatively large fluctuations in wind direction and speed (e.g., see Fig. 4), which caused the predicted plume to be spread out too much. In fact, the σ_y evaluations reported in the next section confirm the tendency to overpredict σ_y when the LDS or ALL weather inputs were used.

The model performance during lower-wind IOPs (2, 4, 5, and 7) was slightly better than during higher-wind IOPs (9 and 10). There does not appear to be a consistent trend with arc distance for the model performance based on the arc maximum.

3) CONCENTRATION AT EACH SAMPLER LOCATION FOR EACH TIME PERIOD (I.E., FULLY PAIRED IN SPACE AND TIME)

For this performance evaluation option, observed and predicted 30-min concentrations from each sampler location and each time period are paired for comparisons. Therefore, the total maximum number of data pairs would be 4752 (66 samplers \times 12 30-min time periods per IOP \times 6 IOPs). Of course, the actual number of data pairs used in this evaluation is less because of bad or missing observed data, and because of the 45 ppt concentration threshold used for accepting data. The 45 ppt value is the so-called MLOQ for the whole-air samplers (R. Carter 2003, personal communication). About 47% of observed data are greater than 45 ppt. When the observed value at a sampler location for a time period is bad or missing or below the 45 ppt threshold, the corresponding predicted value is not considered either. Figure 7 summarizes the performance measures FB, NMSE, MG, and VG [see Eqs. (1)–(4) for definitions], together with their confidence intervals, for the 20 Urban HPAC combinations based on

the paired-in-space-and-time comparisons. As in Fig. 6, which contains the results for the maximum 30-min concentration on each arc for each IOP, Fig. 7 again shows that, except for the UC-OMG combination, all model combinations have negative fractional bias FB, that is, a mean overprediction. These 19 model combinations have a median FB suggesting about a 50% mean overprediction, with a range between a 25% and a factor of 4 mean overprediction. Although UC-OMG has an FB that is closest to zero (0.205), it underpredicts the arc maximum by roughly a factor of 2 (Fig. 6), suggesting that the model underpredicts the observed arc-maximum values, but overpredicts lower observed values. All of the 20 model combinations in Fig. 7 have a geometric mean MG less than 1 (i.e., an overprediction tendency), which is consistent with the FB results. The values of MG have a median corresponding to roughly a 60% overprediction, and have a range corresponding to a 33% to a factor of 2.5 overprediction.

The NMSE values plotted in the figure suggest that, except for the outliers UC-LDS and UC-ALL, the remaining 18 model combinations yield a factor of 9–25 random scatter, which is larger than the results based on the arc maximum. The values of VG for the 20 model combinations correspond to a factor of 6–12 random scatter, which again is larger than the results based on the arc maximum. The larger scatter is expected for the more restrictive pairing in space and time. The values of FAC2 for all the model combinations are between 20% and 30% with little variations among different model combinations. This is mainly due to the fact that many of the cases where predictions are within a factor of 2 of observations result from low predicted and observed values when more data pairs are included in the paired-in-space-and-time comparisons.

Figure 7 also suggests that the RGW weather input option generally yields better performance than other weather options, and that the DM model configuration option generally yields better performance than other model configuration options. Also, DM-RGW is among the better-performing models.

8. Scientific evaluation of Urban HPAC algorithms

The scientific evaluation of Urban HPAC in this study consisted of a peer review of the technical documentation. The documents for SCIPUFF (Sykes et al. 2000), UDM (Hall et al. 2002), and UWM (Lim et al. 2002) were reviewed at a level similar to that for a scientific journal. It was found that these models represent the “state of the science” in that up-to-date scientific algorithms are included and their technical explanations are adequate in their documentation (Chang et al. 2003b). The scientific evaluation also included assessments of model components, such as cloud speed, vertical and lateral cloud spread, and wake retention

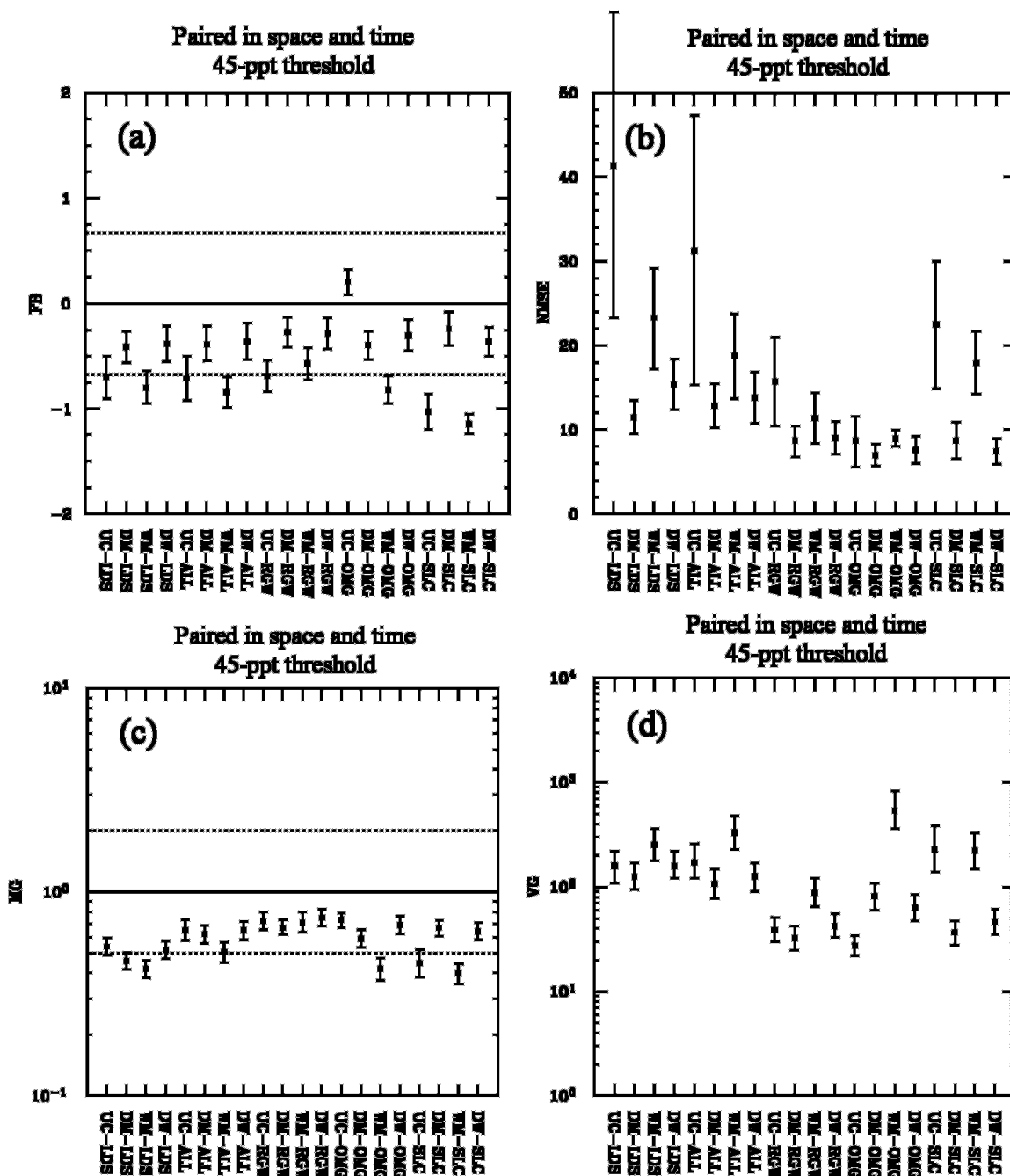


FIG. 7. Statistical measures (a) FB, (b) NMSE, (c) MG, and (d) VG, together with their confidence intervals, based on 30-min SF_6 concentrations paired in space and time with a lower data acceptance threshold of 45 ppt, for each combination of the model configuration and meteorological input data options (20 combinations in total). See Eqs. (1)–(4) for definitions. Negative FB and $\text{MG} < 1$ indicate overprediction. Positive FB and $\text{MG} > 1$ indicate underprediction.

time. These variables are sometimes called “intermediate variables” because they are not directly output by the model, but have to be inferred from concentration predictions. Alternatively, the model code could be modified to output these variables. This was not done in the current evaluation because the model source code was not directly available.

The lateral plume width σ_y was estimated from the concentration predictions on the 6–10 samplers on a given arc distance (see Fig. 2). The maximum concentration C_{\max} was identified, and then the lateral distance to the estimated point of $C_{\max}/10$ on either side of the plume was estimated. The known distance to $C_{\max}/10$ in a Gaussian distribution is approximately equal to

$2\sigma_y$. The vertical plume depth σ_z could not be estimated in such a way because of insufficient data coverage in the vertical direction. The analysis of vertical depth was based on the ratio of concentration at the building top to concentration at ground level. The cloud speed was estimated by studying time series of 30-min-averaged C_{\max} on each monitoring arc, and estimating the delay between the time of the observed cloud peak at distant arcs and close arcs. The wake retention time was estimated by determining the rate of decrease of concentration in the above-mentioned time series at the closest ($x \sim 150$ m) arc.

For σ_y , the values of FAC2 for all 20 Urban HPAC model combinations were 0.5 and higher, meaning that over 50% of the time the predicted σ_y is within a factor of 2 of the observed σ_y . The only exception is the UC–SLC option (i.e., the urban canopy parameterization coupled with the meteorological data from the SLC airport), whose FAC2 was 0.333. In the UC–SLC runs, σ_y was overpredicted on average by a factor of 2.

A limited evaluation of the vertical dispersion was carried out. There were only three sites where concentrations were measured on building tops (heights of 36, 56, and 64 m) and at ground level not too far from the buildings. However, there was a ~ 100 m horizontal displacement between the building-top samplers and the nearby ground-level samplers. The median ratio of the observed concentration at the building top to that at ground level was about 0.5. This implies that the observed σ_z was about 30–50 m at a distance of about 200 m downwind of the source. This is consistent with the standard McElroy–Pooler urban σ_z dispersion coefficients derived from tracer observations in St. Louis, Missouri, in the 1960s (McElroy and Pooler 1968; Briggs 1973). Most model combinations provided rough agreement with the observations. For example, the DM–RGW option (i.e., UDM coupled with the meteorological data at the Raging Waters site) estimated the ratio of the predicted concentration at the building top to that at ground level to be around 0.6.

Comparisons were made of observed tracer cloud speeds with observed wind speeds for each of the IOPs. Cloud speeds were estimated by determining the delay in arrival time of the cloud at various downwind monitoring arcs. It was found that there was good agreement (i.e., within a factor of 2 most of the time and with little mean bias) between the mean tracer cloud speeds and the mean wind speeds as defined by the average of the observations from six anemometers (the D and M anemometers in Fig. 2 and Table 4) in the downtown urban canopy. Typical cloud speeds were about 1 m s^{-1} for the slow-wind IOPs 2, 4, 5, and 7, and were about 2 m s^{-1} for the moderate-wind IOPs 9 and 10. Preliminary comparisons show that predicted cloud speeds derived from model outputs agree with observed cloud speeds within about 30%.

In addition, the time series of observed concentrations on the closest monitoring arc ($x \sim 150$ m) were

studied to determine the typical time scale associated with the decrease in concentration after the release ceased in each trial. Typically the e -folding time scale is about 30–60 min. This is a factor of 30 or more larger than the building wake retention time scale assumed in UDM. However, the difference is probably due to the relatively low wind speeds in URBAN 2000, which caused the along-wind turbulence component σ_u to be approximately equal to the mean wind speed u . When σ_u/u is of order 1 or greater, the upwind dispersion further contributes to the slow decrease in concentration after the release ceases.

9. Conclusions and recommendations

Twenty combinations of the Urban HPAC model configuration options and the meteorological input options have been evaluated with the URBAN 2000 tracer data from Salt Lake City. These 20 model combinations lead to fairly consistent mean overpredictions, with a median of about 50% overprediction, and a range from no overprediction to about a factor-of-4 overprediction for the 30-min arc-maximum concentrations for each IOP. This performance satisfies the model acceptance criteria as suggested by Chang and Hanna (2004), based on the results for many models and many field experiments. In other words, Urban HPAC performs satisfactorily for the URBAN 2000 field data.

In general, the statistical results are quantitatively similar for data paired in space and time versus data paired only by arc distance and IOP. As expected, model performance is reduced for the paired-in-space-and-time comparisons. Nevertheless, the basic conclusions, such as mean overpredictions and the DM model option (i.e., the Urban Dispersion Model) showing improved performance over the UC option (i.e., the urban canopy parameterization), remain the same.

Considering the model performance for the overall domain-wide maximum concentration for all of the IOPs, the DM model option (i.e., the Urban Dispersion Model) and the DW model option (i.e., the Urban Dispersion Model combined with the Urban Wind Field Model) provide agreement within a factor of 2 for all weather input options. The UC model option (i.e., the urban canopy parameterization) and the WM model option (i.e., the Urban Wind Field Model) lead to overpredictions of the overall maximum by a factor of about 2–3, because these options do not adequately account for the enhanced dispersion due to buildings.

Moreover, for the paired-by-arc-and-IOP and paired-in-space-and-time comparisons, the newer DM and DW model options suggest slightly better agreement (i.e., less overprediction bias) than the older UC model option. It can be concluded that the improved science in urban modeling appears to improve performance, although the improvement is not statistically significant at the 95% confidence level.

Among the five meteorological input options, RGW (Raging Waters; an upwind suburban site) produces better results than SLC (Salt Lake City airport), OMG (the OMEGA mesoscale meteorological model outputs), and LDS and ALL, which include downtown observations. It is suggested that the downtown sites were characterized by much variability and, therefore, produced too broad a plume. The average 2–3-km OMEGA horizontal grid resolution probably did not accurately capture the local-scale flow patterns associated with the complex terrain features in the Salt Lake City valley. The outputs from more recent preliminary OMEGA runs with a finer horizontal grid resolution (average ~1 km) appear to lead to better performance for Urban HPAC, although a formal performance evaluation has not been conducted.

The above findings are consistent with those reported by Warner et al. (2004), who use a different evaluation methodology, and primarily focus on paired-in-space comparisons.

Evaluations have also been carried out for model predictions of the lateral distance scale of the concentration distribution (σ_y). It is found that the Urban HPAC model combinations predict σ_y that is within a factor of 2 of that observed about 50% of the time. Model predictions of the vertical distance scale of the concentration distribution (σ_z) appear to agree approximately with the limited observations at three downtown building tops. A more detailed analysis of σ_z was not possible because of a ~100 m horizontal displacement between the building-top samplers and the nearby ground-level samplers. The limited σ_z observations are also consistent with the standard McElroy and Pooler urban σ_z dispersion coefficients. It is found that the mean observed tracer cloud speed agrees well (within a factor of 2 with little mean bias) with the mean observed urban canopy wind speed. Additional analysis is necessary to investigate whether Urban HPAC can satisfactorily predict the tracer cloud speed. Moreover, the observed data at the closest sampling arc suggested a typical plume retention time scale of 30–60 min. There is also a need to compare these observations with the Urban HPAC model predictions in a future study.

It is emphasized that the Urban HPAC evaluation results presented here are based on a single field experiment (URBAN 2000), where synoptic flows were predominantly from the southeast, and where all tracer releases took place during the nighttime. Therefore, model performance should be interpreted within this limited context. A more robust conclusion can only be reached after the model has been evaluated with the data from a number of other field experiments.

Acknowledgments. The Defense Threat Reduction Agency (DTRA) supported this work. The authors thank the DTRA Urban Program Manager, John Pace, for his support, and his successor, Richard Fry, for his

assistance in many contributions to this effort. Drs. Steve Warner and Nathan Platt of the Institute for Defense Analyses conducted some of the Urban HPAC simulations and contributed to many helpful discussions.

REFERENCES

- Allwine, K. J., J. H. Shinn, G. E. Streit, K. L. Clawson, and M. Brown, 2002: Overview of URBAN 2000: A multiscale field study of dispersion through an urban environment. *Bull. Amer. Meteor. Soc.*, **83**, 521–536.
- Anthes, R. A., Y.-H. Kuo, E.-Y. Hsie, S. Low-Nam, and T. W. Bettge, 1989: Estimation of skill and uncertainty in regional numerical models. *Quart. J. Roy. Meteor. Soc.*, **115**, 763–806.
- ARIA Technologies, 2001: General design manual, MINERVE Wind Field Model, version 7.0. ARIA Technologies, 72 pp.
- Bacon, D., and Coauthors, 2000: A dynamically adapting weather and dispersion model: The Operational Multiscale Environment Model with Grid Adaptivity (OMEGA). *Mon. Wea. Rev.*, **128**, 2044–2076.
- Boybeyi, Z., and D. P. Bacon, 1996: The accurate representation of meteorology in mesoscale dispersion models. *Environmental Modeling III*, P. Zannetti, Ed., Computational Mechanics Publications, 109–143.
- Briggs, G. A., 1973: Diffusion estimation for small emissions. Atmospheric Turbulence and Diffusion Laboratory, National Oceanic and Atmospheric Administration ATDL Contribution File 79.
- Chang, J. C., and S. R. Hanna, 2004: Air quality model performance evaluation. *Meteor. Atmos. Phys.*, **87**, 167–196.
- , P. Franzese, K. Chayantrakom, and S. R. Hanna, 2003a: Evaluations of CALPUFF, HPAC, and VLSTRACK with two mesoscale field datasets. *J. Appl. Meteor.*, **42**, 453–466.
- , S. R. Hanna, Z. Boybeyi, P. Franzese, S. Warner, and N. Platt, 2003b: Independent evaluation of Urban HPAC with the Urban 2000 field data. Report prepared for the Defense Threat Reduction Agency by George Mason University and the Institute for Defense Analyses, 40 pp.
- Cimorelli, A. J., and Coauthors, 2002: AERMOD: Description of model formulation (version 02222). U.S. Environmental Protection Agency, OAQPS, EPA-454/R-02-002d, 85 pp.
- Cionco, R. M., 1972: A wind-profile index for canopy flow. *Bound.-Layer Meteor.*, **3**, 255–263.
- Doran, J. C., J. D. Fast, and J. Horel, 2002: The VTMX 2000 Campaign. *Bull. Amer. Meteor. Soc.*, **83**, 537–551.
- Draxler, R. R., and G. D. Hess, 1997: Description of the HYSPLIT-4 modeling system. NOAA Tech. Memo. ERL ARL-224, 24 pp.
- DTRA, 2001: The HPAC user's guide, version 4.0.3. Prepared for Defense Threat Reduction Agency, Contract DSWA01-98-C-0110, by Science Applications International Corporation, Rep. HPAC-UGUIDE-02-U-RAC0, 602 pp.
- Efron, B., 1987: Better bootstrap confidence intervals. *J. Amer. Stat. Assoc.*, **82**, 171–185.
- , and R. J. Tibshirani, 1993: *An Introduction to Bootstrap Statistics and Applied Probability Monogr.*, No. 57, Chapman & Hall, 436 pp.
- EPA, 1995: Description of model algorithms. Vol. II, User's guide for the Industrial Source Complex (ISC3) dispersion models, Environmental Protection Agency, OAQPS, EPA-454/B-95-003b, 128 pp.
- Hall, D. J., R. Macdonald, S. Walker, and A. M. Spanton, 1998: Measurements of dispersion within simulated urban arrays—A small scale wind tunnel study. Building Research Establishment, Ltd., Rep. CR 244/98, 70 pp.
- , A. M. Spanton, I. H. Griffiths, M. Hargrave, and S. Walker, 2002: The Urban Dispersion Model (UDM): Version 2.2. Defence Science and Technology Laboratory Tech. Doc. DSTL/TR04774, 106 pp.

- Hanna, S. R., D. G. Strimaitis, and J. C. Chang, 1991: Evaluation of commonly-used hazardous gas dispersion models. Vol. II, Hazard Response Modeling Uncertainty (A Quantitative Method), Rep. A119/A120 prepared by Earth Tech, Inc., for Engineering and Services Laboratory, Air Force Engineering and Services Center, and for the American Petroleum Institute, 334 pp.
- , J. C. Chang, and D. G. Strimaitis, 1993: Hazardous gas model evaluation with field observations. *Atmos. Environ.*, **27A**, 2265–2285.
- , R. Britter, and P. Franzese, 2003: A baseline urban dispersion model evaluated with Salt Lake City and Los Angeles tracer data. *Atmos. Environ.*, **37**, 5069–5082.
- Lim, D. W., D. S. Henn, and R. I. Sykes, 2002: UWM version 1.0 technical documentation. Titan Research and Technology Division Tech. Doc., 37 pp.
- Macdonald, R. W., D. J. Hall, S. Walker, and A. M. Spanton, 1998: Wind tunnel measurements of wind speed within simulated urban arrays. Building Research Establishment Rep. CR 243/98, 65 pp.
- McElroy, J. L., and F. Pooler, 1968: St. Louis dispersion study. U.S. Public Health Service, National Air Pollution Control Administration Rep. AP-53, 51 pp.
- Nasstrom, J. S., G. Sugiyama, D. Ermak, and J. M. Leone Jr., 2000: A real-time atmospheric dispersion modeling system. *Proc. 11th Joint Conf. on the Application of Air Pollution Meteorology with the Air and Waste Management Association*, Long Beach, CA, Amer. Meteor. Soc., CD-ROM, 5.1.
- NOAA, 2004: ALOHA (Areal Locations of Hazardous Atmospheres) user's manual. NOAA Hazardous Materials Response Division, 224 pp.
- Puhakka, T., K. Jylhä, P. Saarikivi, J. Koistinen, and J. Koivukoski, 1990: Meteorological factors influencing the radioactive deposition in Finland after the Chernobyl accident. *J. Appl. Meteor.*, **29**, 813–829.
- Scire, J. S., D. G. Strimaitis, and R. J. Yamartino, 2000: A user's guide for the CALPUFF dispersion model (version 5.0). Earth Tech, Inc., 521 pp. [Available online at <http://www.src.com/>]
- Sharan, M., S. G. Gopalakrishnan, R. T. McNider, and M. P. Singh, 1996: Bhopal gas leak: A numerical investigation of the prevailing meteorological conditions. *J. Appl. Meteor.*, **35**, 1637–1657.
- Sykes, R. I., and Coauthors, 2000: PC-SCIPUFF version 1.3 technical documentation. Titan-ARAP Rep. 725, 259 pp.
- Warner, S., N. Platt, and J. F. Heagy, 2001: Application of user-oriented measure of effectiveness to HPAC probabilistic predictions of Prairie Grass field trials. Institute for Defense Analyses Paper P-3554, 275 pp. [Available from Steve Warner at swarner@ida.org, or Steve Warner, Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311-1882.]
- , —, and —, 2004: Comparison of transport and dispersion model predictions of the URBAN 2000 field experiment. *J. Appl. Meteor.*, **43**, 829–846.
- Winkenwerder, W., Jr., 2002: Case narrative: U.S. demolition operations at Khamisiyah. Special Assistant to the Under Secretary of Defense (Personnel and Readiness) for Gulf War Illnesses, Medical Readiness, and Military Deployments, U.S. Department of Defense (DOD) Final Rep. 2001137-0000055, 242 pp. [Available online at http://www.gulflink.osd.mil/khamisiyah_iii/]