JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 7, July 2010

Fast, Accurate Defense for Homeland Security: Bringing High-Performance Computing to First Responders

Gopal Patnaik* , Adam Moses† , and Jay P. Boris‡
Naval Research Laboratory, Washington DC 20375

DOI: 10.2514/1.26040

An urban-oriented emergency assessment system, called CT-Analyst® was developed to evaluate airborne contaminant transport threats and to aid in making rapid decisions for complex-geometry environments such as cities where current transport and dispersion methods are slow and inaccurate. Contaminant transport-Analyst was designed for the military prior to 9/11 to incorporate verbal reports, to treat systems with mobile sensors, and to function in realistic situations where the nature, amount, and source location of an airborne contaminant or a chemical, biological, or radiological agent is unknown. Thus contaminant transport-Analyst is well suited to urban defense in the Homeland Security context. Contaminant transport-Analyst gives good accuracy and much greater speed than possible with current alternatives because is it based on entirely new principles and designed to function in information-starved situations, characterizing the first few minutes of a terrorist or accident scenario, where alternate technologies do not. These advantages derive from pre-computed data structures based on three-dimensional large-eddy simulation computational fluid dynamics that includes solar heating and buoyancy, complete building geometry specification, trees, and impressed wind fluctuations. A few detailed urban aerodynamics simulations are pre-computed for each coverage region when contaminant transport-Analyst is installed. These results extend to all wind directions, speeds, likely sources and source locations through a new data structure called Dispersion Nomografs™. Thus a high-performance computing based system can generate Nomografs for cities, military bases, industrial complexes, and other potential danger areas well in advance, removing the need for the emergency first responders and warfighters to wait for supporting analyses. Furthermore, since the full power of high-performance computing is available for the pre-computations, contaminant transport-Analyst provides important, new, real-time, zero-latency functions such as sensor data fusion, backtracking to an unknown source location, and evacuation route planning directly to the first responder. Thus the Department of Homeland Security can avoid the delays and uncertainties of reachback to high-performance computing modeling resources and the associated support and communications infrastructure currently required during airborne contaminant transport emergencies.

Received 22 June 2006; accepted for publication 7 May 2009.

* Center Head, NRL Code 6440, Center for Computational Physics Developments, Laboratory for Computational Physics and Fluid Dynamics, Naval Research Laboratory, Washington, DC 20375, USA, Senior Member. patnaik@lcp.nrl.navy.mil

† Computer Scientist, NRL Code 6440, Center for Computational Physics Developments.

‡ Director and Chief Scientist, NRL Code 6400, Laboratory for Computational Physics and Fluid Dynamics, Fellow.

I. Introduction

THE emergence of increasingly powerful computers stimulated the development of obstacle-resolving microscale flow and transport models based on computational fluid dynamics (CFD). In recent years, these types of models are playing an important role in many applications. They serve as general tools in fluid engineering and wind engineering when complex flow systems have to be designed. Under the general category of urban aerodynamics [1], these models are now commonly applied to predict contaminant transport (CT) in complex structured urban landscapes. Use of such models is made in the licensing of new industrial plants, in safety analysis studies for accidental releases of hazardous materials in the chemical industry, or in the context of crisis management after terrorist attacks in urban environments.

Urban airflow accompanied by CT presents new, extremely challenging modeling requirements [1,2] best met using complex-geometry simulation tools developed by the aerospace industry. Reducing health risks from the accidental or deliberate release of chemical, biological, or radiological (CBR) agents and contaminants from industrial leaks, spills, and fires motivates this work. Configurations with complex geometries and unsteady buoyant flow physics are involved. The wide range of temporal and spatial scales rapidly overwhelms the current modeling capacities. Crucial technical issues that need to be addressed include time-dependent turbulent fluid transport (aerodynamics), environmental boundary condition modeling (meteorology), and the practical post-processing of the simulation results for use by responders in actual emergencies. The advantages of the CFD approach and the large-eddy simulation (LES) representation include the ability to quantify complex-geometry effects, to predict dynamic nonlinear processes faithfully, and to treat turbulent problems reliably in regimes where experiments, and therefore model validations, are impossible or impractical.

CFD solutions to CT can be highly accurate, but are too slow for emergency response purposes. One practical solution to this critical dilemma, described in some detail in this paper, carries out the unsteady CFD simulations in advance and pre-computes compressed databases for specific urban areas incorporating suitably parameterized weather for a full set of wind conditions and distributed test-sources. The relevant information is summarized as *Dispersion Nomograf*™ data sets [3] so that it can be used in a portable system called CT-Analyst® [4], which reproduces the CFD quality results nearly instantly with little loss of fidelity. This paper presents this new methodology bringing the fidelity and accuracy of CFD to the first responder and warfighter fast enough for effective emergency response. As appropriate to a paper in the special issue on Homeland Security, the paper presents an overview of the issues involved in meeting the seemingly contradictory requirements of speed and accuracy. The details of implementation, algorithms employed, etc. are not discussed fully here but are left to references.

A. Standard CFD Simulations

Some “time-accurate” flow simulations that attempt to capture the urban geometry and fluid dynamic details are a direct application of standard (aerodynamic) CFD methodology to the urban scale problem. An example is the finite element CFD simulations of the dispersion of a contaminant in the Atlanta, Georgia metropolitan area [5]. The model includes topology and terrain data and a typical mesh contains approximately 200 million nodes and 55 million tetrahedral elements. These are grand-challenge size calculations and were run on 1024 processors of a CRAY T3E taking up to a whole day to run. Similar approaches are being used by other research groups [6,7]. The chief difficulty with this approach for large urban regions is that the solutions are very computer intensive (days or weeks) and involve severe overhead associated with grid generation for the urban geometry.

B. The LES Approach For CT

Direct numerical simulation (DNS) is prohibitively expensive for most practical flows at moderate-to-high Reynolds number, and especially so for urban CT studies. On the other end of the CFD spectrum are the classic aerodynamic methods such as the Reynolds-averages Navier–Stokes (RANS) approach, which simulate the mean flow and approximately model the effects of turbulent scales [8]. These approaches are typically unacceptable for urban CT modeling because they are unable to capture the inherently unsteady plume dynamics driven by the urban geometry. LES constitutes an effective intermediate approach between DNS and the RANS methods [9]. LES is capable of simulating flow features that cannot be handled with RANS such as significant flow unsteadiness and localized vortex shedding, and provides higher accuracy than the industrial methods, but at a lower cost than DNS. LES solutions converge to the solutions of the Navier–Stokes equations as resolution is increased, whereas RANS solutions generally do not. Because the larger-scale unsteady features of the flow govern the unsteady plume dynamics in urban geometries, the LES approximation can capture some key features which the RANS methods and the various Gaussian plume methodologies cannot. Moreover, given its potential for higher computational efficiency, the monotone integrated LES (MILES) approach is well suited for CFD-based plume simulation for urban-scale scenarios, an application where classical LES methods are expensive. See [10,11] for recent reviews.

A practical example of urban-scale MILES is depicted in Fig. 1 which shows contaminant dispersion in Times Square, New York City. The figure demonstrates the typical complex unsteady vertical mixing patterns caused by building vortex and recirculation patterns, and the predicted endangered region associated with this particular release scenario. The large variability of concentration values from minute-to-minute is evident and thus the need for unsteady, time-dependent simulation models.

C. The FAST3D-CT Model

The FAST3D-CT three-dimensional urban aerodynamics model [3,12,13] is based on the scalable, low-dissipation flux-corrected transport (FCT) convection algorithm [14,15]. FCT is a higher-order, monotone, positivity-preserving method for solving generalized continuity equations with source terms. The required monotonicity is achieved by introducing a diffusive flux and later correcting the calculated results with an antidiffusive flux modified by a flux limiter. The version of the convection algorithm implemented in FAST3D-CT is documented in [11,16]. The virtual cell embedding algorithm provides treatment of the complex geometry [17]. Relevant physical processes simulated in FAST3D-CT include complex building vortex shedding, flows in recirculation zones, with algorithms approximating the dynamic subgrid-scale turbulence and stochastic backscatter. The model also incorporates a stratified urban boundary layer with realistic wind fluctuations, solar heating including shadows from buildings and trees, aerodynamic drag and heat losses due to the presence of trees, surface heat variations and turbulent heat transport [11].

Modeling a pollutant as well mixed globally is typically not appropriate in problems where short time spans and large air volumes are involved. It is important to capture the effects of unsteady, buoyant airflow on the evolving pollutant concentration distributions. In typical urban scenarios, both particulate and gaseous contaminants behave similarly insofar as transport and dispersion are concerned, so the contaminant spread can usually be simulated

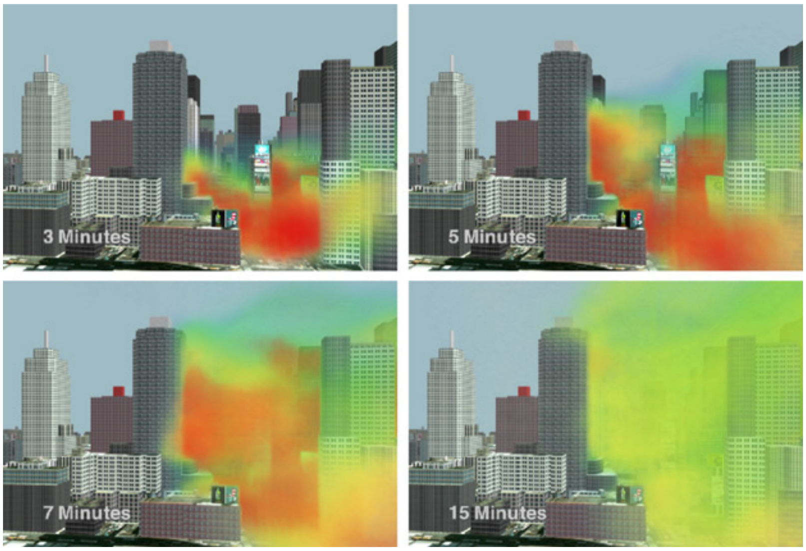


Fig. 1 Contaminant dispersion from an instantaneous release in Times Square, New York City as predicted by the FAST3D-CT model. The frames show concentrations at 3, 5, 7, and 15 min after release. This region was the subject of a 2007 model inter-comparison and blind validation exercise, including FAST3D-CT, based in the MID'05 field trials supported by DHS. To view this figure as a movie, click here.

effectively based on appropriate pollutant tracers with suitable sources and sinks. In other cases, the full details of multigroup particle distributions are required. Additional physics include multigroup droplet and particle distributions with turbulent transport to surfaces as well as gravitational settling, solar chemical degradation, evaporation of airborne droplets, relofting of particles on the ground, and ground evaporation of liquids. Incorporating specific models for these processes in the simulation codes is a challenge but can be accomplished with reasonable sophistication. Details of the physical models in FAST3D-CT are given in [11] and omitted here for brevity. The primary difficulty is the effective calibration and validation of all these physical models since much of the input needed from field measurements of these processes is typically insufficient or even nonexistent.

II. Fast and Accurate CBR Defense

A typical run with the FAST3D-CT model for a complex urban region of 10 square miles area, resolved with 6 m cells, takes about 12 h on a 16-processor SGI shared-memory computer. This is significantly faster per square mile than classical CFD models due to the savings achieved by MILES and other algorithmic improvements. The *critical* dilemma in the CT application is that unsteady urban-scenario flow simulations are currently feasible — but they are still expensive and require a degree of expertise to perform. First responders and emergency managers on site to cope with contaminant release threats have perhaps a minute to make decisions and cannot afford to wait even five or ten minutes while actual simulations and data post-processing are carried out either locally or remotely.

A. CT-Analyst®: An Operational Solution to this Dilemma

An operational solution of this dilemma conducts unsteady CFD simulations in advance and pre-computes compressed databases for specific urban areas incorporating relevant meteorology and a full set of wind conditions and distributed test-sources. The relevant information is summarized as *Dispersion Nomograf*™ data sets [3] so that it can be readily used locally on portable computers with sensors and verbal reports providing current information regarding local presence of contaminants, contaminant concentrations and winds. Thus there is now a methodology making three-dimensional CFD really useful for crisis managers in real time, operational situations. The accuracy of CFD simulations is recovered nearly instantly with little loss of fidelity. The current implementation of this new approach is called CT-Analyst® [4] and is described further below. Near instantaneous CT assessment with high fidelity can reduce the number of people being exposed in urban areas, even for large crowds out in the open, by up to a factor of six once a simple sensor or reporting network is in place. First responders and headquarters staff can use CT-Analyst displays for data fusion to give a minute-by-minute situation assessment.

A number of unique new capabilities are also implemented in the current CT-Analyst 3.3 release including plume tracking, sensor fusion, sensor placement optimization, and geometric “backtrack” to locate undiscovered sources. CT-Analyst can also be used for war games, virtual reality training, site defense planning, and sensor network optimization [18]. This technology requires only limited information, the kind of data and isolated sensor readings that will come in sporadically for situation assessment during the first few minutes of a CBR or contaminant incident. The required model input is minimal by design because a terrorist probably will not tell us the amount and location of a contaminant source or even what the contaminant is. CT-Analyst is designed specifically for use where aerodynamic accuracy (e.g. building deflection and trapping of contaminants) must be combined with very rapid response (seconds rather than minutes). CT-Analyst [3,4] is described further in Sec. II.C below.

B. Nomograf Description

Nomografs™ are compact, pre-computed data structures that capture the aerodynamic and turbulent effects of terrain, buildings, vegetation, and surface types on contaminant plume transport and dispersion. Using Nomografs, improved accuracy and much greater speed are achieved for urban emergency assessment of airborne contaminant plumes. By interpolating into these patented data structures, we can perform plume predictions and related assessments in milliseconds for areas with complex terrain such as cities, military bases, and important facilities.

The Naval Research Laboratory’s (NRL) FAST3D-CT CFD model, as described above, underpins our current implementation of dispersion Nomografs. FAST3D-CT computes the multigigabyte three-dimensional, contaminant flow-path databases from which the high-resolution dispersion Nomografs are extracted. Other models that can provide the same database could also be the source of data to build Nomografs. If enough data were taken in field trials or experiments, equivalent to three-dimensional fields of key variables over the region, Nomografs could then be made from field data [3].

Three physical principles are central to the current dispersion Nomograf representation and its subsequent applications:

1) A CTs dynamically via convection with the local airflow. All relevant diffusion and dispersion arises from resolved (and possible unresolved) fluid motions. Diffusion, per se, plays only a minimal role.
2) Wherever the contaminant goes becomes contaminated and a volume, once contaminated, stays contaminated. This is a conservative approximation favored by first responders to aid in "safe siding" the predictions. It could (perhaps) be relaxed considerably for applications in open regions or over water.
3) Vertical spreading of contaminant is quick on the building scale in an urban environment. Both simulations (see Fig. 1 above) and field trials support this approximation.

The data collected as a basis for Nomograf generation include:

1) The three-dimensional means and standard deviations of the three velocity components of the airflow computed from data collected at six height levels from the ground to above the typical buildings.
2) Six key quantities are also accumulated at two different heights for each of the local sources, that are initialized once the CFD model "spin-up" period is complete and the statistics of the dynamic flow have stabilized.
    a) First arrival time (s) of contaminant density exceeding a threshold value at each location
    b) Time of arrival (s) of the maximum contaminant value at the location
    c) Decay time (s) after the maximum is reached at the location
    d) Local peak contaminant density (gm/m^3) at any time
    e) Integrated contaminant density at the location (dose has units of s gm/m^3)
    f) Local contaminant variability, measured as the integrated total variation (gm/m^3).

Nomografs derive from a weighted average of the contaminant flow behavior on the six levels recorded. The heaviest weighting is at or below the height of the typical buildings, as justified by principle three above. The four steps in generating and using dispersion Nomografs are:

1) An accurate geometry database is compiled from LIDAR, stereo imagery, or shape files. The geometry database used by FAST3D-CT is a two-dimensional (typically one meter resolution) array that returns the heights of terrain, buildings, and trees, and surface composition in the computational domain.
2) Detailed three-dimensional CFD calculations (FAST3D-CT) are repeated for 18 wind directions and the results are captured in an extensive database. These simulations include the appropriate urban boundary layer for the region with realistic turbulent fluctuations imposed at the inflow boundaries. Multiple releases are tracked in each case as described above.
3) The salient features from the CFD database are distilled into Dispersion Nomograf data structures for rapid interactive access. Time integration is thus replaced by interpolations that capture the aerodynamic effects of the full urban geometry through the Nomograf tables.
4) The Nomograf tables are encrypted and input to CT-Analyst, an easy-to-use graphical user interface (GUI) for instantaneous situational analysis. Plume computation, for example, takes less than 50 milliseconds.

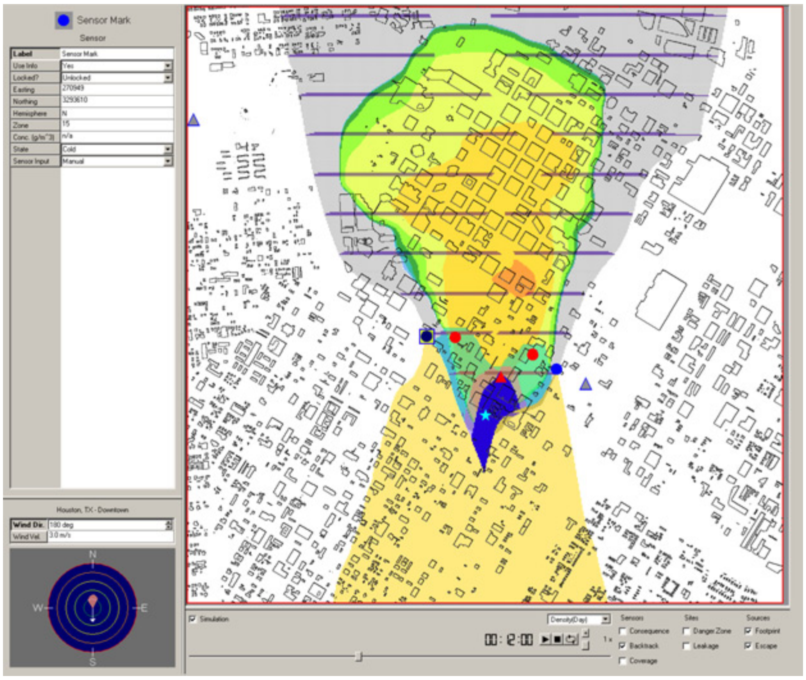
C. The CT-Analyst® Emergency Assessment Tool

To solve the critical speed vs accuracy dilemma and to meet real operational requirements, NRL developed an integrated CBR emergency assessment tool that is much faster than current "common use" models with accuracy comparable to three-dimensional, physics-based flow simulations for scenarios involving complex and urban landscapes. The focus is on situation assessment through sensor fusion of qualitative and incomplete data. A terrorist probably will not tell us the amount and location of an agent source or even what the agent is. Therefore we should not expect this information early enough for action in a crisis unless we somehow can generate what we need from the hints that will be available. The only existing software tool with these capabilities is called CT-Analyst (contaminant transport analyst) and is both zero-latency (meaning nearly zero computing delay) *and* high fidelity. CT-Analyst is entirely visual, i.e., "point-and-click," in application. Beta-test versions, implemented in modest laptop and workstation versions, treating all of the buildings and structures in a multiple-square-mile area of downtown, have been delivered to the cities of Chicago, New York, Houston, Washington DC, and to other officials in the Department of

Defense and Department of Homeland Security. A corresponding capability has been delivered to civil emergency-management authorities in the District of Columbia. The Missile Defense Agency has incorporated CT-Analyst into its post-engagement ground effects model [19] and a commercial implementation for law enforcement has been developed by Defense Group Incorporated [20].

Each point in a domain of interest, if considered as a source location, has a downwind region called the footprint that can become contaminated by an airborne agent reaching that source point. Any selected location (considered as a site of interest) also has an upwind region (the danger zone) within which contaminant would have to be released to reach and contaminate that site. These two classes of regions are completely complementary, being effectively each other's inverse. The interlocked source footprints and site danger zones have boundaries that change continuously as the source location or target site location is moved continuously within in the domain. All assessments in CT-Analyst are "computed" by manipulating these two distinct regions for sensor report locations, for selected site locations, and for source locations. The dispersion Nomograf representation is designed to make these manipulations very fast while requiring only a minimum amount of tabulated data for each wind direction. The CT-Analyst implementation integrates these capabilities in a graphically oriented framework to treat airborne scenarios requiring higher spatial and temporal resolution than current operational tools. The focus in this new urban capability is on the first few minutes to an hour after a CBR release and the first few miles from the source. Beyond these limits, spatially varying prevailing winds and weather data plays a progressively greater role and the user also has much more time to respond.

As indicated above, the dispersion Nomograf representation and processing algorithms also allow some new features. Multiple sensor fusion for instantaneous situation assessment is an automatic consequence of the Nomograf representation. The methodology can accept qualitative and anecdotal input and does not require knowledge of a source location or even a source type or amount. A backtrack to unknown source locations is performed graphically with zero delay by overlap operations on the upwind danger zones of the "hot" and "cold" sensor reports.



Label: Sensor Mark
Use Info: Yes
Locked?: Unlocked
Easting: 270949
Northing: 3293610
Hemisphere: N
Zone: 15
Conc. (g/m^3): n/a
State: Cold
Sensor Input: Manual

Wind Dir.: 180 deg
Wind Vel.: 3.0 m/s

☑ Simulation
Density(Day)
00 : 2:00
☐ Consequence
☑ Backtrack
☐ Coverage
☐ Danger Zone
☐ Leakage
☑ Footprint
☑ Escape

Fig. 2 CT-Analyst display for downtown Houston showing contaminant concentration contours (yellow, green, and blue), contamination footprint (gray), and evacuation routes (magenta/purple lines). Observations near the edges of the evolving plume (represented as red and blue circles) allow backtracking to the source location (blue star in the purple backtrack diamond). To view this figure as a movie, click here.

Figure 2 above shows a typical CT-Analyst display for an urban area, in this case a section of downtown Houston. The contaminant concentration display (orange–yellow–green–blue contours) above fills the same area as the corresponding FAST3D-CT plume shown in Fig. 3 below and is overlaid on the contamination footprint (gray region above and salmon region in Fig. 3). Star-shaped nodes are sources, triangular and circular nodes are sensor reports, and square nodes indicate specific sites. When a source node is active it is colored light blue, as shown above. Footprints, plume envelopes, contaminant concentration plots, and escape routes can be displayed for sources by activating the buttons shown on the lower portion of the CT-Analyst screen. Triangular sensor report nodes inside an active plume envelope are "hot" (red) while those still uncontaminated are "cold" (blue). Downwind consequence regions (for active "hot" reports) and upwind backtrack estimates (for all active "hot" and "cold" reports) can be displayed for the active sensor nodes, indicated by filled triangles. Contamination zones from down wind leakage and upwind danger zones can be plotted for all square site nodes (bright green when they are active). The diagonal purple lines are the recommended evacuation (escape) routes.

To compute displays such as danger zones, plume envelopes, and backtracks to unknown source locations, knowing the actual concentration of the airborne agent is not necessary. Indeed, until the total amount of the contaminant is known, plotting the actual concentration distribution is not even possible. Therefore, CT-Analyst provides a relative concentration until the mass of the agent from a specific source can be determined. Fortunately, this relative concentration and its time history are all that is needed to select civil defense options that minimize the inhaled dose of contaminant. The normalization used for Fig. 2 was chosen to correspond to the integrated mass of the source used in the FAST3D-CT simulation shown in Fig. 3 below. This normalization also accounts for the contaminant that leaves the grid through an analytic extension of the Nomograf tables. The contour levels in Fig. 2 are similar

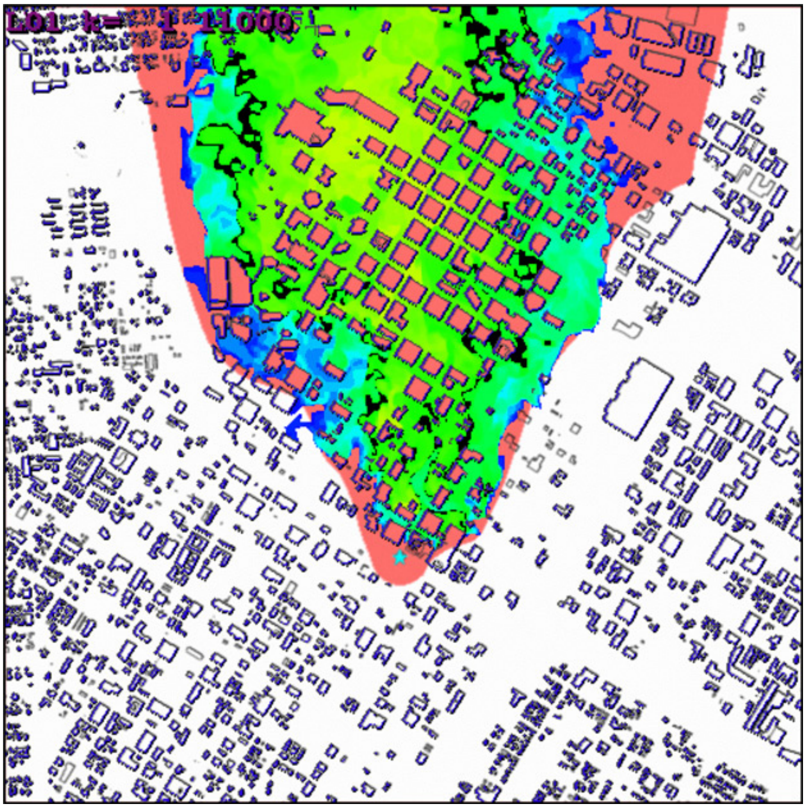


Fig. 3 FAST3D-CT simulation for downtown Houston. The contaminant cloud is shown 30 min after release superimposed on the CT-Analyst footprint (i.e., the overall contamination region). The average wind is from 180 deg at 3 m/s. To view this figure as a movie, click here.

but not identical to those in Fig. 3 throughout the range of concentrations since the CT-Analyst representation must be generic.

Visually comparing the CT-Analyst and FAST3D-CT solutions in Fig. 3 shows how well the compression process used to generate the Dispersion Nomografs captures the urban geometry-induced deviations from a smooth plume shape. For example, the geometry-induced spreading of the concentration contours shown on the left of the plume corresponds quite closely in both figures. A quantitative Figure of Merit measuring the congruence of the CT-Analyst rendering with the underlying database is described in [1, Fig. 7; 17]. For the case above, the Cumulative Figure of Merit is about 80%, meaning that CT-Analyst over predicts (on the conservative side as shown) the actual plume contamination area by at most a factor of 20% and under predicts, which is does not occur in the figure, by at most 10%. Under prediction is more dangerous and this is weighted more heavily. These comparisons represent a kind of validation of CT-Analyst but more extensive formal validation programs for both FAST3D-CT and CT-Analyst are continuing. Comparisons between field trial data, FAST3D-CT simulations, and CT-Analyst predictions are compared for a region of downtown Los Angeles in [1,21] and for Oklahoma City (Joint Urban 2003) in [22–24].

The contamination footprints plotted by CT-Analyst are chosen to provide plausible worst cases, that is, they are designed to “safe-side” the resulting situation assessments. The plume envelopes, which expand in time to fill the footprint, share this conservatism in the “predictions.” This means that the edges of the plume envelope and the footprint are smoothed to maintain continuity in such a way that the predicted contamination areas should always be slightly larger than observed in the field. This is an interpretation designed for first responders. In practice this means that any particular realization, e.g., Fig. 3, may only fill a part of the plume envelope depending on the structure of the wind gusts for that particular run. CT-Analyst attempts to indicate all regions that may be dangerously contaminated with a minimal degree of uncertainty. This is different from an ensemble average because the edges of the plume envelope are quite sharp and this is reflected in the concentration plots provided. This also is a feature of the individual high-resolution realizations during the first few minutes of any scenario.

III. High-Performance Computing Implementation of Data Generation for Nomografs

The Nunn–Lugar–Domenici domestic preparedness program [25] initially identified one hundred and twenty cities as likely terrorist targets. This number has since been increased to over 150. Military installations and other potential targets further increase this number. For each of these cities and bases, Nomograf data has to be generated for each of eighteen wind angles and up to four environmental conditions. Each combination of wind angle and environmental condition requires a separate CFD run to obtain the required data for the corresponding Nomograf. Clearly, this results in a large number of *independent* runs that will be required. In addition, the city could also be divided up into a number of distinct tiles. When treated separately, these tiles lead to even more independent CFD runs than can be executed in parallel.

In performing production runs to develop the necessary Dispersion Nomograf data sets, a number of levels of parallelism can be exploited to optimize the processing and permit the data sets to be developed in a rational sequence, working from the center of the city outward. Because of these multiple levels of parallelism, the version of the model (and/or the parallel architecture) used for each case only needs to be moderately scalable. A great deal of effort has gone into optimizing FAST3D-CT for performing multiple runs simultaneously using OpenMP with the main computational kernel, LCPFCT [16] placed in parallel loops so that the overhead of parallelization is minimal. However, measures and metrics such as gigaflops, parallel-speed-up, etc. are largely academic. The key, and arguably the only metric of importance, is the overall time to solution, and thus we have focused our efforts on reducing the time required to provide a robust and accurate system for protecting our cities and military installations.

It takes less than half a day to perform a full FAST3D-CT computation at six-meter resolution for a given wind direction and environmental condition on 16 nodes of an SGI Altix or comparable multicore cluster for a modest region of a city (e.g., ten square miles). To perform CFD computations for the compressed Nomografs used by CT-Analyst, a separate run for each of 18 different wind directions thus takes about 150 node days. This is the basic unit of computation used to construct Table 1 above. The table also assumes that four different data sets would be generated, day and night for two seasons of the year. Day and night and stable and unstable atmospheres are the conditions that are being considered. These variations are reflected in the input wind fluctuation profiles and the input temperature profile. The table also assumes that four cities could be processed simultaneously when in production and that lower resolution Nomografs for 300 square miles of suburbs take only four times as much computing at

Table 1 Estimated computing required for Nomograf table preparation (November 2008 Technology)

	First usable Nomograf data set	Operational all year Nomograf data sets	Make high-resolution Nomograf tables	Nomograf validation (four sets)
10 Square-mile-downtown	150 nodes: 1 day	600 nodes: 1 day	1 node: 2 h	1 node: 8 h
(6 m) 150 cities	150 nodes: 1/2 yr	600 nodes: 1/2 yr	3 nodes: 100 h	3 nodes: 400 h
30 Square-mile full city	150 nodes: 1/2 week	600 nodes: 1/2 week	1 node: 6 h	1 node: 8 h
(6 m) 150 cities	150 nodes: 1.5 yr	600 nodes: 1.5 yr	3 nodes: 300 h	3 nodes: 400 h
300 square-mile suburbs	600 nodes: 1/2 week	600 nodes: 2 weeks	1 node: 24 h	1 node: 8 h
(12 m) 150 cities	600 nodes: 1.5 yr	600 nodes: 6 yr	12 nodes: 300 h	3 nodes: 400 h

12-meter resolution as 30 square miles does at six-meter resolution. For each complete data set, some modest amount of processing is required to actually generate the Nomograf tables from the contaminant source computations. There also has to be some validation and testing for each data set. In production this would take about 2 h for each of the four operational data sets. Estimates of these additional requirements are included in the last two columns and note that these last two categories of work are manpower intensive.

As can be seen, to compute Nomografs for 150 cities out into the suburbs (300 square miles) will require an estimated six years with 600 nodes working more or less continuously. 600 nodes is not a daunting number and could easily be tripled in light of the seriousness of the problem, reducing the preparation time to two years. If we take into account another doubling of computer capability over the next 24 months, this time reduces further to well less than two years. Therefore we would recommend that Nomografs for all the downtown (10 or 30 square mile areas) be prepared before the larger areas are tackled. This would allow the downtown areas to be covered in less than two years. In order to generate all the data to cover 150 cities will require many independent large, but not huge, shared-memory parallel jobs. In this application the scalability of individual jobs is not so critical, and an effective architecture would be a cluster of shared-memory nodes, each with 16–32 processors. This problem is limited currently by processor speed and memory bandwidth, especially between processors on each node. In this case, other than for file management and subsequent Nomograf generation, communication within the cluster is negligible. Such computers are readily available today.

Special cases will arise that need to be considered for detailed “forensic” analysis. These studies may involve higher resolution and include much more detailed and complex physics, especially if the agent has complicated physics or chemistry. Agent fate, chemical reactions, and deposition may also have to be considered. These cases will be far fewer in number but will be required due to special circumstances and will usually be time-critical. The individual runs for these cases may have to be completed as expeditiously as possible. Here, massive scalability becomes much more important. For the shared-memory-based FAST3D-CT code, the ideal computer would have flat memory access from all processors. For detailed high-resolution runs we may be required to re-develop an up-to-date distributed memory version of FAST3D-CT. This is complicated by the fact that optimal performance in the sub-models in FAST3D-CT requires a number of different data structures. Detailed simulations of buoyant and neutral gas contaminants, multigroup droplet problems, and multigroup particle sources for biological and radiological (“dirty bomb”) scenarios for example cause severe load-balancing problems for realistic problems in which much of the grid may have no contaminants at all. To provide a high degree of fidelity in solar deposition, a ray-trace algorithm was implemented so buildings and trees cast realistic shadows. To keep this cost to a few percent of the overall running time, this piece of complex geometry physics was knowingly implemented in a manner only conducive to a shared-memory implementation. Furthermore, the largest problem is what numbers to use for all this physics. For optimization on a distributed parallel architecture this translates into a great deal of memory being required to replicate all the tables and pre-computed intermediate computations in a number of different processors. This, in turn, again suggests the simplicity and economy of a shared-memory implementation.

Current CT models implemented operationally for the DoD, NOAA, and DHS employ a different high-performance computing (HPC) approach also based on parallel processors or nonportable workstation clusters. This approach is generically known as “reachback” because it reaches back to a centralized software system that runs on HPC resources with an attendant infrastructure. A defining example is given in references [26–28]. These

CT models are sophisticated and require highly skilled personnel, who are available continuously around the clock, to deal with emergencies. The modelers receive notification from the emergency responders *in situ*, process what information there is about the release, and collect meteorological data relevant to the release location and time. They then execute relatively efficient three-dimensional models on the remote-site supercomputer(s), generally taking 5–10 min. The reachback experts then generate automatic displays that depict the size and location of the plume, and perhaps affected population, health risks, and proposed emergency responses, and deliver these results back to the first responders and emergency managers via network, fax, or e-mail.

In support of this approach, CT experts are processing the technical information with the best predictive models they have that run in 5–10 min. They have access to a wide variety of meteorological resources and can be expected to bring a depth of experience to the emergency that the responders on site are unlikely to possess. This approach also introduces a possible single point of failure as all emergency response modeling is expected to funnel through the center. In multiple simultaneous emergencies, this facility might be overwhelmed leading to unacceptable response times. To avoid this, sufficient reserve capacity must be maintained, at considerable ongoing expensive. Obviously reliable communications must be ensured which may be difficult to ensure in a large-scale crisis. The main difficulties with reachback, however, are that the response times are simply too long in a crisis and the overall costs higher than one would like for comprehensive planning.

The CT-Analyst system takes a different approach in which all the detailed CFD computations are performed ahead of time with the requisite expertise present and with time to carefully test the results. CT-Analyst merely looks up and interpolates results using the dispersion Nomograf tables. Even though HPC assets have to be used to generate the Nomograf databases for the cities and military bases, extremely large resources do not need to be concentrated in one location. Smaller HPC facilities already present throughout the country can be used and the process can even be distributed and/or privatized. As CT-Analyst is run locally using Nomografs computed with the local urban geometry, every jurisdiction can have its own state-of-the-art CFD-based plume modeling system. Since the costs of this system are essentially all incurred once during the Nomograf generation phase, CT-Analyst can be used freely on a daily basis for "routine" emergencies such as fires and hazardous spills. With CT-Analyst, first responders have the tools they need right at their fingertips, in the police car or fire truck. The response is local and instant, with no requirements for reach-back facilities, personnel, or communications.

IV. Conclusions

In the previous section we described the high-performance computing and personnel resources needed to develop the compact Nomograf representations for the urban regions where CT-Analyst application could be necessary. This is the single largest cost and therefore the main deterrent to wider use of CT-Analyst. On the hardware side, the extensive computing workload is exactly what the big computers have been designed to do. On the human side, it is far better to prepare well ahead of time rather than to have our crisis managers wait for results while hundreds or even thousands of people perish. Even if the Nomografs are never used in a terrorist incident, they also apply to everyday incidents like fires and industrial leaks and spills. Furthermore, the accumulated time saved by all the people involved in planning, sensor optimization, and training, people who now wait at least 5–10 min per scenario computed, will save far more in actual dollars than the Nomograf pre-computation costs. Nevertheless, reducing the computer requirements and the time delay to implement CT-Analyst for a new area is very important and thus the subject of continuing research and development. The technology has progressed to the point where a turnkey Nomograf generation system could be implemented for relatively untrained personnel (not PhDs) to use on smaller HPC systems that are already present at a number of sites.

In this general approach, every agency, major military command, or state could have a state-of-the-art CFD-based plume modeling system while producing the CT-Analyst representations for regions under their control. This would give them the freedom to compute detailed scenario realizations for planning and training, to conduct field trials with simulants to validate their own capabilities, and to expand coverage to outlying regions on a schedule of their choosing. Today we are aiming at the ability to produce a set of high-fidelity Nomografs in three days for a 30 square-mile region once the geometry is available. Other models are also being considered for rapid generation of lower-fidelity Nomografs for even larger regions. Once several of these advanced capabilities are available, the problem of generating coverage for the main cities, bases, and facilities could be reduced to two years or less.

Validation always has and always will be an important consideration. This gives an additional reason to distribute the CT-Analyst capability and the ability to generate Nomografs. Only by performing their own tests on local coverage regions will cities, police, fire departments, and hazmat units gain a sense of confidence in the new tools, understand their limitations, and have an easy familiarity with their use. Appreciable validation has been performed showing that CT-Analyst represents a significant step up over other available models, but there is room for much more validation. The Missile Defense Agency has recently funded Independent Verification and Validation of CT-Analyst by outside agents. Two distinct DTRA projects are in their final stages supporting efforts to compare the detailed FAST3D-CT model with field trial data and extensive wind-tunnel experiments for Oklahoma City and to independently verify and validate the latest release of CT-Analyst.

This work has also led to further development of CT-Analyst. Extensive data have been collected from FAST3D-CT runs relating to temporal variability and multirealization variability and these measures of variability have been validated with detailed wind tunnel data [11,21–24]. New probability of contamination and agent consequence displays have been developed in prototype and are being integrated into CT-Analyst 4.0 to relate this information to the user so they will have a context for understanding the uncertainty inherent in such predictions and to approximate the human consequences assuming the amount, location, and type of agent is known. CT-Analyst developers also have demonstrated algorithms, modules, and software for inclusion in future releases to provide: 1) zero-latency line-of-sight (obstructed view) displays for a number of applications, 2) zero-latency line-of-sight integrations in cities for estimating and interpreting "standoff" sensor performance, 3) zero-latency "keepout zone" analyses around buildings for unmanned aerodynamic vehicles using the flow data collected for Nomograf generation, 4) zero-latency explosion damage estimates for urban geometry, and 5) zero-latency analyses of airborne plume threats against building heating, ventilation and air-conditioning settings to give immediate assessments for individual buildings of when and if sheltering-in-place is a safer alternative than evacuation.

Physically realistic urban aerodynamics simulations are now possible but still require some compromises due to time, computer, and manpower resource limitations. The necessary trade-offs result in sometimes using simpler models, algorithms, and geometry representations than we would wish. It is understood that the building- and larger-scale fluid dynamics effects that are captured today [1,11,21] by FAST3D-CT govern the turbulent dispersion and thus the quality of computed answers will get better with time because the MILES and FCT algorithms are convergent as computational resolution is improved. Inherent uncertainties in simulation inputs and model parameters beyond the environmental conditions also lead to errors that need to be further quantified by comparison with high quality reference data but we believe that the largest source of errors remaining is knowing what the wind input initial conditions and boundary conditions are when comparing with field trial data or analyzing an emergency situation. In spite of inherent uncertainties and model trade-offs, it is possible to achieve a reasonable degree of predictability and to understand the uncertainty and variability [21].

The NRL FAST3D-CT model can be used to simulate sensor and system response to postulated threats, to evaluate and optimize new systems, and to conduct sensitivity studies for relevant processes and parameters. Moreover, the simulations constitute a virtual test range for microscale and nanoscale atmospheric fluid dynamics and aerosol physics, to interpret and support field experiments, and to evaluate, calibrate, and support simpler models. Moreover, using this HPC-based LES model as a detailed scenario generator, we have invented a process to make the three-dimensional CFD really useful in real time for crisis managers in operational situations. As a bottom line, the increased speed and accuracy of using dispersion Nomografs in CT-Analyst can reduce the number of people being exposed in urban CT scenarios, even for large crowds in the open, by 85 to 95% once an effective sensor or reporting network is in place. First responders and headquarters staff can use this tool for data fusion to give a minute-by-minute situation assessment. It can also be used for war games, virtual reality training, site defense planning, and sensor network optimization. The technology requires only limited information, the kind of partial data and isolated sensor readings that will come in sporadically for situation assessment during the first few minutes of a CBR release scenario. Incomplete model input situations must be treated efficiently because terrorists will probably not announce ahead of time the amount and location of a contaminant source or even what the agent is.

By performing all the major computing ahead of time on HPC computers well suited to the application, the results of a number of complete, high-resolution three-dimnesional simulations can be recalled for operational use with no sensible delay for integration of even simple models. In this way we have solved the usual dilemma of more computer time being required to obtain better answers. The best available answers are now being presented instantly with full urban geometry in a readily comprehended format.

Acknowledgments

The authors wish to thank Bob Doyle, Keith Obenschain, Julie Pullen, Lisa Williams, and Theodore Young, Jr., the NRL contaminant transport team who made this technology practical and easy to use. Further thanks are due Carey Cox, Bo Cybyk, Jack Fulton, John Iselin, Sandy Landsberg, Charles Lind, and Rob Scott for their technical discussions and contributions to various components of this effort. Aspects of the work discussed here were supported by ONR through NRL, and in the past by the DoD High Performance Computing Modernization Office, NSWC Crane, DARPA, DTRA, MDA, TSWG, and the University of Houston.

References

[1] Boris, J. P., "Dust in the Wind: Challenges for Urban Aerodynamics," AIAA paper 2005-5953, AIAA, 2005.

[2] Britter, R. E., and Hanna, S. R., "Flow and Dispersion in Urban Areas," *Annual Review of Fluid Mechanics*, Vol. 35, 2003, pp. 469–496.
doi: 10.1146/annurev.fluid.35.101101.161147

[3] Boris, J. P., "The Threat of Chemical and Biological Terrorism: Preparing a Response," *Computating in Science and Engineering*, Vol. 4, 2002, pp. 22–32, see also "Dust in the Wind: Challenges for Urban Aerodynamics," AIAA paper 2005-5393, AIAA, 2005.
doi: 10.1109/5992.988644

[4] Boris, J. P., Obenschain, K., Patnaik, G., and Young, T. R., "CT-ANALYST™, Fast and Accurate CBR Emergency Assessment," *Proceedings of the 2nd International Conference on Battle Management*, DTRA, Williamsburg, VA, 2002.

[5] Aliabadi, S., and Watts, M., "Contaminant Propagation in Battlespace Environments and Urban Areas," *AHPCRC Bulletin*, Vol. 12, No. 4, 2002, http://www.ahpcrc.org/publications/archives/v12n4/Story3/ [retrieved 11 May 2009].

[6] Chan, S., "FEM3C — An Improved Three-Dimensional Heavy-Gas Dispersion Model: User's Manual," Lawrence Livermore National Lab., Rept. UCRL-MA-116567, rev. 1, 1994.

[7] Camelli, F., and Löhner, R., "Assessing Maximum Possible Damage for Release Events," *Seventh Annual George Mason University Transport and Dispersion Modeling Workshop*, GMU, Fairfax, VA, 2003.

[8] Hendricks, E., Burrows, D. A., Diehl, S., and Keith, R., "Dispersion in the Downtown Oklahoma City Domain: Comparisons Between the Joint Urban 2003 Data and the RUSTIC/MESO Models," *5th AMS Symposium on the Urban Environment*, AMS, Boston, MA, 2004.

[9] Sagaut, P., *Large Eddy Simulation for Incompressible Flows*, 2nd ed., Springer-Verlag, Berlin, 2004.

[10] Grinstein, F. F., and Fureby, C., "From Canonical to Complex Flows: Recent Progress on Monotonically Integrated LES," *Computing Science and Engineering*, Vol. 6, 2004, pp. 37–49.

[11] Patnaik, G., Boris, J. P., Grinstein, F. F., and Iselin, J., "Large Scale Urban Simulations with FCT," *High-Resolution Schemes for Convection-Dominated Flows: 30 Years of FCT*, edited by D. Kuzmin, R. Löhner, and S. Turek, Springer-Verlag, Berlin, 2005, pp. 105–130.

[12] Cybyk, B. Z., Boris, J. P., Young, T. R., Lind, C. A., and Landsberg, A. M., "A Detailed Contaminant Transport Model for Facility Hazard Assessment in Urban Areas," AIAA paper 99-3441, AIAA, Washington, DC, 1999.

[13] Cybyk, B. Z., Boris, J. P., Young, T. R., Emery, M. H., and Cheatham, S. A., "Simulation of Fluid Dynamics Around Complex Urban Geometries," AIAA paper 2001-0803, AIAA, Washington, DC, 2001.

[14] Boris, J. P., and Book, D. L., "Flux-Corrected Transport I, SHASTA, A Fluid Transport Algorithm that Works," *J. Computational Physics*, Vol. 11, 1973, pp. 8–69.

[15] Boris, J. P., and Book, D. L., "Solution of the Continuity Equation by the Method of Flux-Corrected Transport," *Methods in Computational Physics*, Vol. 16, 1976, pp. 85–129.

[16] Boris, J. P., Landsberg, A. M., Oran, E. S., and Gardner, J. H., "LCPFCT — A Flux-Corrected Transport Algorithm for Solving Generalized Continuity Equations," U.S. Naval Research Lab. Memorandum Rept. NRL/MR/6410-93-7192, 1993.

[17] Young, T., Boris, J., Hooper, S., Lind, C., Obenschain, K., and Patnaik, G., "Emergency Assessment with Sensors and Buildings: Advances in CT-Analyst Technology," *Science and Technology for Chem-Bio Information Systems*, DTRA, Arlington, VA, 2004.

[18] Obenschain, K., Boris, J. P., and Patnaik, G., "Using CT-ANALYST to Optimize Sensor Placement", SPIE paper 5416-02, 2004.

[19] Lind, C. A., Rose, H. E., and Moore, W. K., "Application of Post-Engagement M&S Capabilities to Homeland Security," *2003 Advanced Simulation Technologies Conference*, SCS, San Diego, CA, 2003.

[20] CoBRA Decision-Support Toolset for Crisis Response and Emergency Response Management to All Hazard Incidents, http://www.defensegroupinc.com/cobra/ [retrieved 11 May 2009].

[21] Boris, J. P., Cheatham, S., Patnaik, G. and Young, T., "Variability, Validation, and CFD: There is No Truth," *8th Annual George Mason University Conference on Transport and Dispersion Modeling*, Fairfax, VA, 2004.

[22] Lee, M-Y., Young, Jr., T., and Patnaik, G., "Statistical Study of Puff Releases from FAST3D-CT Simulations and JU2003 IOP8," *12th Annual George Mason University Conference on Atmospheric Transport and Dispersion Modeling*, Fairfax, VA, 8–12 July 2008.

[23] Leitl, B., Harms, F., Herbst, I., Schatzmann, M., Patnaik, G., and Boris, J., "Development of Model- and Application-Specific Validation Data Sets for LES-Based Micro-Scale Flow and Dispersion Models," *12th Annual George Mason University Conference on Atmospheric Transport and Dispersion Modeling*, Fairfax, VA, 8–12 July 2008.

[24] Harms, F., Leitl, B., Schatzmann, M., and Patnaik, G., "Validating LES-based Flow and Dispersion Models — A Challenge on its Own," *Proceedings of the 12th Annual George Mason University Conference on Atmospheric Transport and Dispersion Modeling*, Fairfax, VA, 8–12 July 2008.

[25] Davis, R., "Observations on the Nunn–Lugar–Domenici Domestic Preparedness Program," United States General Accounting Office Rept. GAO/T-NSIAD-99-16, 1999.

[26] National Atmospheric Release Advisory Center, http://narac.llnl.gov/ [retrived 11 May 2009].

[27] NARAC Emergency Response System: Overview, https://narac.llnl.gov/ers_overview.php [retrieved 15 December 2008].

[28] NARAC Emergency Response System: Real-Time Operational Models, https://narac.llnl.gov/modeling.php [retrieved 11 May 2009].

Gerard Parr
Associate Editor