


**Volume 47 Issue 6  
(June 2008)**
[< Previous](#)
[Next >](#)


[Current Issue](#)  
[Available Issues](#)  
[Early Online Releases](#)  
[Author Index](#)

[Share this Article](#)

[Share](#) |

**Journal Information**

Online ISSN: 1558-8432  
 Print ISSN: 1558-8424  
 Frequency: Monthly

[Special Collections](#)

[Staff and Editors](#)

[Instructions to Authors](#)

[Manuscript Submission](#)

[How to Subscribe](#)

[RSS \(What is this?\)](#)

**Previously titled:**

*Journal of Applied Meteorology*, 1988-2005

*Journal of Climate and Applied Meteorology*, 1983-1987

*Journal of Applied Meteorology*, 1962-1982

**Table of Contents**

- [Abstract](#)
- [1. Introduction and background](#)
- [2. Reconstruction procedure](#)
- [3. Isolated building example](#)
- [4. Oklahoma City–Joint Urban 2003 IOP3](#)
- [5. Oklahoma City–Joint Urban 2003 IOP9](#)
- [6. Effect of sensor density](#)
- [7. Discussion and conclusions](#)
- [Acknowledgments](#)
- [REFERENCES](#)

[< Previous Article](#)
**Volume 47, Issue 6 (June 2008)**
[Next Article >](#)
[Add to Favorites](#)
[Email](#)
[Download to Citation Manager](#)
[Track Citations](#)
[Glossary](#)
[Permissions](#)
[Abstract](#)
[PDF](#)
[Figures](#)

Chow, Fotini Katopodes, Branko Kosović, Stevens Chan, 2008: Source Inversion for Contaminant Plume Dispersion in Urban Environments Using Building-Resolving Simulations. *J. Appl. Meteor. Climatol.*, 47, 1553–1572.  
doi: <http://dx.doi.org/10.1175/2007JAMC1733.1>

### Source Inversion for Contaminant Plume Dispersion in Urban Environments Using Building-Resolving Simulations

Fotini Katopodes Chow

*Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, California*

**Branko Kosović and Stevens Chan**

*Atmospheric, Earth and Energy Department, Lawrence Livermore National Laboratory, Livermore, California*

**Abstract**

The ability to determine the source of a contaminant plume in urban environments is crucial for emergency-response applications. Locating the source and determining its strength based on downwind concentration measurements, however, are complicated by the presence of buildings that can divert flow in unexpected directions. High-resolution flow simulations are now possible for predicting plume evolution in complex urban geometries, where contaminant dispersion is affected by the flow around individual buildings. Using Bayesian inference via stochastic sampling algorithms with a high-resolution computational fluid dynamics model, an atmospheric release event can be reconstructed to determine the plume source and release rate based on point measurements of concentration. Event-reconstruction algorithms are applied first for flow around a prototype isolated building (a cube) and then using observations and flow conditions from Oklahoma City, Oklahoma, during the Joint Urban 2003 field campaign. Stochastic sampling methods (Markov chain Monte Carlo) are used to extract likely source parameters, taking into consideration measurement and forward model errors. In all cases the steady-state flow field generated by a 3D Navier–Stokes finite-element code (FEM3MP) is used to drive thousands of forward-dispersion simulations. To enhance computational performance in the inversion procedure, a reusable database of dispersion simulation results is created. It is possible to successfully invert the dispersion problems to determine the source location and release rate to within narrow confidence intervals even with such complex geometries. The stochastic methodology here is general and can be used for time-varying release rates and reactive flow conditions. The results of inversion indicate the probability of a source being found at a particular location with a particular release rate, thus inherently reflecting uncertainty in observed data or the lack of enough data in the shape and size of the probability distribution. A composite plume showing concentrations at the desired confidence level can also be constructed using the realizations from the reconstructed probability distribution. This can be used by emergency responders as a tool to determine the likelihood of concentration at a particular location being above a threshold value.

Keywords: [Urban meteorology](#), [Pollution](#), [Air quality](#), [Dynamics](#)

Received: March 7, 2007; Final Form: August 7, 2007

Corresponding author address: Fotini Katopodes Chow, Department of Civil and Environmental Engineering, University of California, Berkeley, Berkeley, CA 94720-1710. Email: [chow@ce.berkeley.edu](mailto:chow@ce.berkeley.edu)

**1. Introduction and background**

Flow in urban environments is complicated by the presence of buildings, which divert the flow into often unexpected directions. Contaminants released at ground level are easily lofted above tall (~100 m) buildings and channeled through urban canyons that are perpendicular to the wind direction [see, e.g., the ninth intensive observing period (IOP9) in [Chan and Leach \(2007\)](#), hereinafter [CL07](#)]. The path of wind and scalars in urban environments is difficult to predict even with building-resolving computational fluid dynamics codes because of the uncertainty in the synoptic wind and boundary conditions and errors in parameterizations of different physical processes such as turbulence.

Given the difficulties from the complexity of urban flows, solving an inverse problem is quite challenging. That is, given measurements of concentration at sensors scattered throughout a city, is it possible to detect the source and strength of a contaminant release and, if so, can the uncertainty in source characteristics be estimated? The ability to determine

## Related Articles

### Articles Citing this Article

[Google Scholar](#)

### Search for Other Articles

#### By keyword

- Urban meteorology  
 Pollution  
 Air quality  
 Dynamics

#### By author

- Fotini Katopodes Chow  
 Branko Kosović  
 Stevens Chan

#### Search in:

- AMS  
 Google Scholar

[Submit](#)

source location and strength in a complex environment is necessary for emergency response to accidental or intentional releases of contaminants in densely populated urban areas. The goal of this work is to demonstrate a robust statistical inversion procedure that performs well even under the complex flow conditions and uncertainty present in urban environments.

Much work has previously focused on direct-inversion procedures, where an inverse solution is obtained using an adjoint advection–diffusion equation. The exact direct-inversion approaches are usually limited to processes governed by linear equations and also generally assume the system is steady state (Errico 1997; Enting 2002; Keats et al. 2007). With some effort, adjoint procedures can be developed for applications including nonlinear processes but are useful only as long as linearized approximations are valid (Errico 1997). Our approach is general and is geared toward the creation of a library of forward-dispersion model choices for more flexibility. In addition to adjoint models, optimization techniques are also employed to obtain solutions to inverse problems. These techniques often give only a single best answer or assume a Gaussian distribution to account for uncertainties. General dispersion-related inverse problems, however, often include nonlinear processes (e.g., dispersion of chemically reacting substances) or are characterized by non-Gaussian probability distributions (Bennett 2002). Traditional methods also have particular weaknesses for dealing with sparse, poorly constrained data problems as well as high-volume, potentially overconstrained and diverse data streams.

We have developed a more general and powerful inverse methodology based on Bayesian inference coupled with stochastic sampling (Chow et al. 2006). Bayesian methods reformulate the inverse problem into a solution based on a sampling of an ensemble of predictive simulations, guided by statistical comparisons with observed data (see, e.g., Ramirez et al. 2005). Predicted values from simulations are used to estimate the likelihood of available measurements; this likelihood is in turn used to improve the estimates of the unknown input parameters. Bayesian methods impose no restrictions on the types of models or data that can be used. Thus, highly nonlinear systems and disparate types of concentration, meteorological and other data, can be simultaneously incorporated into an analysis.

In this work we have implemented stochastic models based on Markov chain Monte Carlo (MCMC) sampling for use with a high-resolution building-resolving computational fluid dynamics (CFD) code, FEM3MP. The inversion procedure is first applied to flow around an isolated building (a cube) and then to flow in Oklahoma City (OKC), Oklahoma, using data collected from SF<sub>6</sub> tracer gas releases during the Joint Urban 2003 field experiment (Allwine et al. 2004). While we consider steady-state flows in this first demonstration, the approach used is entirely general and is also capable of dealing with unsteady, nonlinear governing equations. Our stochastic approach has previously been applied to other dispersion cases, including canonical test cases, moving sources, and multiple sources, using Lagrangian particle models, Gaussian puff models, and urban puff models (see Johannesson et al. 2004, 2005; Lundquist 2005; Delle Monache et al. 2008). This paper presents the first application of our stochastic approach to urban environments using a building-resolving CFD approach that includes the true geometric and flow complexity inherent in urban areas.

## 2. Reconstruction procedure

### a. Bayesian inference and Markov chain Monte Carlo

The inversion or reconstruction algorithm uses Bayes's theorem combined with an MCMC approach for stochastic sampling of unknown parameters (see, e.g., Gelman et al. 2003). A brief description is given here, and more details can be found in Johannesson et al. (2004, 2005). Bayes's theorem is written as

$$p(M|D) = \frac{p(D|M)p(M)}{p(D)} \propto p(D|M)p(M), \quad (1)$$

where  $M$  represents possible model configurations or parameters and  $D$  is observed data. For our application, Bayes's theorem therefore describes the conditional probability [ $p(M|D)$ ] of certain source parameters (the model configuration  $M$ , including, e.g., source location and release rate) given observed measurements of concentration at sensor locations ( $D$ ). This conditional probability  $p(M|D)$  is also known as the posterior distribution and is related to  $p(D|M)$ , the probability of the data conforming to a given model configuration, and to  $p(M)$ , the possible model configurations before taking into account the measurements. The probability  $p(D|M)$ , for fixed  $D$ , is called the likelihood function, while  $p(M)$  is the prior distribution. In this application, we assume at the outset that the source could be located anywhere in the whole domain, so the prior distribution is uniform over the chosen domain. The probability  $p(D)$  distribution is called the prior predictive distribution (Gelman et al. 2003) and represents a marginal distribution of  $D$ . The probability  $p(D)$  is a normalizing factor and is not needed when computing the posterior distribution. For a general problem in which analytical solutions are not possible, the challenge lies in computing the likelihood function. For that purpose we use a stochastic sampling procedure and approximate the posterior distribution [ $p(M|D)$ ] by the empirical distribution function described below.

### b. Sampling procedure

We use an MCMC procedure with the Metropolis–Hastings algorithm to obtain the posterior distribution of the source term parameters given the concentration measurements at sensor locations (Gelman et al. 2003; Gilks et al. 1996). We thus completely replace the Bayesian formulation with a stochastic sampling procedure to explore the model parameter space and to obtain a probability distribution for the source location and strength. The Markov chains are initialized by taking samples from the prior distribution. To lower the computational cost, we limit the prior distribution to the ground surface (thus ignoring the possibility of elevated sources). All grid cells associated with the footprints of buildings are also excluded from the prior distribution for the Oklahoma City runs.

To begin the iteration process, a forward-dispersion calculation is performed from the initial locations of the Markov chains to provide the initial data for comparison with observed data at the sensors. Each Markov chain path is then determined using the Metropolis–Hastings algorithm at each step as described here (see also Fig. 1 of Delle Monache et al. 2008). First, a sample is taken from a specified Gaussian proposal distribution centered at the current chain location and likewise from a Gaussian proposal distribution for the source strength. The parameters  $x$ ,  $y$ , and  $q$  are each sampled independently and tested against their prior distributions. If the new parameter compares better to its prior than does its previous state, the proposed value is chosen for the next proposed state. If the comparison is worse, the new value is not automatically rejected. Instead, a Bernoulli random variable (a “coin flip”) is used to decide whether to accept the new value. This procedure ensures that each proposed parameter best reflects its prior distribution. Second, a forward

calculation is performed for the proposed state (using the new values of  $x$ ,  $y$ , and  $q$ ) and results are compared with measurements at the concentration sensors. If the comparison is more favorable than the previous chain location, the proposal is accepted and the Markov chain advances to the new location. If the comparison is worse, a coin flip is used to decide whether to accept the new state. This random component is important because it prevents the chain from becoming trapped in a local minimum where comparisons are more favorable than values in the local sampling area but where the chain has not converged on the true source location or release rate.

A log likelihood function is used to quantify the agreement between the model configuration and the data; it is defined as

$$\ln[\mathcal{L}(M)] = -\frac{\sum_i^N (C_i^M - C_i^E)^2}{2\sigma_{\text{rel}}^2}, \quad (2)$$

where  $C_i^M$  are model values at the sensor locations,  $C_i^E$  are the experimentally observed sensor values, and  $\sigma_{\text{rel}}$  is the standard deviation of the combined forward model and measurement errors. The value of  $\sigma_{\text{rel}}$  can be varied depending on the model formulation and observation errors, as described for the applications below. The squared difference in the log likelihood function is summed over the  $N$  sensor locations. In this work, the logarithm of the model and data values is taken before using this formula (zero concentration values are treated according to the sensitivity level of the instrument). This prevents large concentration values from dominating the likelihood calculation when the range of concentrations spreads over several orders of magnitude. The likelihood function is calculated as the forward model for each proposed new state (sample  $x$ ,  $y$ , and  $q$  values) is computed. As described above, the proposed state is accepted if either

$$\ln(\mathcal{L}_{\text{prop}}) \geq \ln(\mathcal{L}) \quad \text{or} \quad \mathcal{L}_{\text{prop}}/\mathcal{L} \geq \text{rnd}(0, 1), \quad (3)$$

where  $\ln(\mathcal{L}_{\text{prop}})$  is the log likelihood of the proposed state,  $\ln(\mathcal{L})$  is the previous likelihood value, and  $\text{rnd}$  denotes a random number generated from a uniform distribution to represent the ‘‘coin flip.’’

The posterior probability distribution in Eq. (1) is computed discretely from the resulting Markov chain paths defined by the algorithm described above and is estimated with

$$p(M|D) = \sum_{i=1}^n (1/n)\delta(M_i - M), \quad (4)$$

which represents the probability of a particular model configuration ( $M$ , including parameters such as source location and strength), giving results that match the observations at sensor locations ( $D$ ). Equation (4) is a sum over the entire Markov chain of length  $n$  of all the sampled values  $M_i$  that fall within a certain ‘‘bin.’’ Thus  $\delta(M_i - M) = 1$  when  $M_i = M$  and 0 otherwise. If a Markov chain spends several iterations at the same location, meaning that multiple proposals were rejected because the given location was more favorable than the proposals, the value of  $p(M|D)$  increases through the summation in Eq. (4), indicating a higher probability for those source parameters.

Multiple chains are used (typically four) to allow for better statistical sampling of the parameter space and to enable convergence monitoring [thus Eq. (4) is overly simplified]. Statistical convergence to the posterior distribution is monitored by computing between-chain variance and within-chain variance (Gelman et al. 2003). If there are  $m$  Markov chains of length  $n$ , then we can compute between-chain variance  $B$  with

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{M}_j - \bar{M})^2, \quad (5)$$

where

$$\bar{M}_j = \frac{1}{n} \sum_{i=1}^n M_{ij} \quad (6)$$

is the average value along each Markov chain (a sample from a given chain is denoted by  $M_{ij}$ ) and

$$\bar{M} = \frac{1}{m} \sum_{j=1}^m \bar{M}_j \quad (7)$$

is the average of the values from all Markov chains. Between-chain variance represents a measure of difference between chains. The within-chain variance  $W$  is

$$W = \frac{1}{m} \sum_{i=1}^m s_i^2, \quad (8)$$

where

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (M_{ij} - \bar{M}_i)^2. \quad (9)$$

An estimate of the variance of  $M$  is computed as

$$\text{var}(M) = \frac{n-1}{n} W + \frac{1}{n} B. \quad (10)$$

The convergence parameter  $R$  is then computed as

$$R = \frac{\text{var}(M)}{W}. \quad (11)$$

The necessary condition for statistical convergence to the posterior distribution is that  $R$  approaches unity (Gelman et al. 2003). In practice, this is not always a sufficient condition for convergence, as seen below and in other studies (Delle Monache et al. 2008).

#### c. Source strength scaling

Typically, the MCMC sampling requires thousands of iterations (samples) to converge to the posterior distribution, thus requiring thousands of forward-dispersion model calculations. With simple Gaussian puff models (Johannesson et al. 2004) or Lagrangian particle tracking (Delle Monache et al. 2008), it is possible to calculate the forward models on the fly. With a three-dimensional CFD model, the computational cost quickly becomes prohibitive even for the simplest cases. For the current applications, we have simplified the situation for demonstration purposes by considering only steady-state flow conditions. (The chosen methodology remains completely general and can handle unsteady and reactive flows.) Assuming that the advection–diffusion problem is linear (e.g., no chemical reactions) we can use the precomputed steady flow field and Green’s functions to carry out one forward simulation at each of the thousands of locations in our prior distribution using a unit source strength and storing the resulting values at the sensor locations in a database. The source is modeled as a steady flux from one surface grid element. The stored concentrations can be rescaled depending on the proposed source release rate for a particular source location. Thus, during the inversion process, the sampled  $x$  and  $y$  locations are mapped to the corresponding grid element, and dispersion results from each possible source location are obtained from the database and rescaled according to the current sampled value for the source strength. In this way, 20 000 iterations for each of the four Markov chains can be performed in about 10 min of computational time on four 2.4-GHz Xeon processors.

#### d. Forward model description—FEM3MP

The stochastic inversion procedure relies on a forward model to calculate instances of predicted sensor measurements  $D$  for given source term parameters  $M$ . Here we use FEM3MP (Gresho and Chan 1998; Chan and Stevens 2000), a three-dimensional, incompressible Navier–Stokes finite-element code that is able to represent complex geometries and simulate flows in urban environments (Chan and Leach 2004; CL07). Here FEM3MP is used in a Reynolds-averaged Navier–Stokes (RANS) approach.

For the example of flow around an isolated building, the model is driven by a steady logarithmic inflow profile at the upstream (west) boundary. Natural (i.e., zero tangential and normal stress) outflow boundary conditions are applied at the other boundaries. The steady-state flow field is precomputed and is used to drive dispersion from a source with a constant release rate until a steady-state concentration field is obtained. The grid resolution is uniform far from the building, and is doubly fine near the corners of the building (see Fig. 6 later).

For the Oklahoma City simulations, we use setups similar to CL07 for the third and ninth intensive observing periods (IOP3 and IOP9) from Joint Urban 2003. Again, the flow field is assumed steady, with a logarithmic inflow profile on the southern and western boundaries for IOP3 and on the southern boundary for IOP9. The wind speed is set to  $6.5 \text{ m s}^{-1}$  at  $z = 50 \text{ m}$  with a wind direction of  $185^\circ$  (south-southwest) for IOP3 and, similarly,  $7.2 \text{ m s}^{-1}$  and  $180^\circ$  (south) for IOP9. The inflow profiles are based on upwind field observations near the computational domain (CL07). The flow field is precomputed using FEM3MP. The release rate is constant ( $0.005 \text{ kg s}^{-1}$  for IOP3 and  $0.002 \text{ kg s}^{-1}$  for IOP9) and simulations are performed until steady-state concentration fields are achieved (after about 10 min of simulation time). The atmosphere is assumed to be neutrally stratified since shear production of turbulence from buildings is significantly larger than buoyant production (Lundquist and Chan 2007). A standard eddy-viscosity RANS turbulence model was used for IOP3, and a nonlinear model was used for IOP9 (CL07). Buildings near the source are explicitly resolved; that is, velocities and concentrations within the buildings are set equal to zero. Far from the source, “virtual buildings” are used to reduce the computational cost. In this region, a drag force of very large value is added to the momentum equations for grid cells falling within the building boundaries. Previous work has shown that this approach produces satisfactory dispersion estimates far from the source (CL07).

### 3. Isolated building example

We have developed an event reconstruction prototype for a flow around an isolated building (a cube) with a source located upwind from the building (see Fig. 1). Four sensors are placed in a diamond-shaped array in the lee of the building. Data at the sensor locations are collected using a forward simulation from the true source location. The data are thus “synthetic” and used in this case only to test the inversion algorithm. Artificial measurement error with a standard lognormal distribution is also added to the synthetic data (in this case with a mean of  $\mu = 0$  and a standard deviation of  $\sigma_{\text{rel}} = 0.05$ ).

The source release rate was set to 0.1 (nondimensional units). As can be seen from Fig. 1 the actual source is located just above the symmetry line. Because the symmetry line is also the separatrix of this flow, the small deviation of the source location from the line of symmetry results in significant asymmetry in the resulting plume (Fig. 1). This example, while simple in geometry, thus incorporates complexities of its three-dimensional nature that were not accounted for in previous inversion studies. The asymmetry of the plume is generated purely by the presence of the building. More simplistic dispersion models do not explicitly resolve buildings and hence cannot capture such features (Britter and Hanna 2003).

The domain is discretized using about 19 000 elements ( $42 \times 32 \times 14$ ). Forward runs are computed for all possible locations (on  $z = 0$ ), and concentration values at the sensors are stored in a database for each grid location. Total computation time for generation of the database was 6 h using sixty-four 2.4-GHz Xeon processors. The reconstruction or inversion algorithm proceeds as usual, but instead of running a new simulation for each proposed Markov chain step, the results are drawn from the concentration database, as previously described. This avoids repeated computations of releases at the same  $x, y$  locations by simply scaling the release rate as dictated by the sampling algorithm.

#### a. Source inversion

[Figure 2](#) shows the points sampled by the four Markov chains. The chains quickly converge on the source location, sampling more frequently in the northern half of the domain as expected due to the asymmetry of the actual plume. The probability distribution for the source location is given in [Fig. 3](#), which also reflects the asymmetry of flow. The peak of the distribution occurs just upwind of the actual source location. If the error from the measurements is set to zero (i.e.,  $\sigma_{\text{rel}} = 0$ ), the inversion procedure accurately predicts the source location as expected (i.e., the peak of the probability distribution matches the true source; not shown). The probability distribution is constructed using the second half of the MCMC iterations (i.e., 10 000 to 20 000) to allow the Markov chains to “mix” adequately to improve the statistical distribution and to exclude the random initialization from the final statistics. Thus, the so-called burn-in time is 10 000 iterations. The corresponding probability distribution for the source release rate is shown in [Fig. 4](#). The peak of the histogram coincides with the actual release rate of 0.1.

Convergence rates for the  $x$ ,  $y$ , and  $q$  inversions are shown in [Fig. 5](#). All convergence measurements reach a value near 1.1 after about 10 000 iterations, indicating that the sampling procedure was thorough and adequate to generate a meaningful posterior probability distribution. Note that the convergence rate is independent of the spread in the distribution and merely indicates that further sampling will not likely change the results. We are thus able to successfully invert this idealized three-dimensional dispersion problem and determine the source location and release rate to within a tight confidence region.

#### b. Composite plume

In addition to probabilistic predictions of the source location, emergency responders need predictions of concentrations over the entire plume area. A “most likely plume” could easily be constructed by performing a forward simulation from the peak of the probability distribution for the source location. This, however, would be one realization and would not contain the probabilistic information inherent in the reconstruction procedure.

We therefore construct a probabilistic, composite plume from the plume realizations corresponding to all the samples from the posterior probability distribution of source term parameters. The composite plume is obtained by first creating histograms of concentration values at each spatial location in the domain using results from all iterations beyond the burn-in time. This step is followed by determining the concentration value at each location for which a certain prespecified probability is exceeded. Contours of the 90% confidence interval are shown in [Fig. 6](#). For values above the threshold (chosen to be 0.03), the plot shows 90% confidence that the concentration at a given location is higher than the contoured value. For values below the threshold, the contours indicate 90% confidence that the concentration is less than the contoured value.

The shape of this composite plume is quite different from that of the actual plume in [Fig. 1](#). Notice, for example, that the high concentration contour lines in [Fig. 1](#) extend upstream of the source in the upper half of the plot. In the composite plume in [Fig. 6](#) the highest valued contour line extends only downstream of the source. The composite plume represents a probabilistic estimate of concentrations and could aid in emergency-response decisions for evacuation or sheltering in place, depending on a chosen confidence interval and whether an area lies above or below a threshold value for toxicity.

### 4. Oklahoma City–Joint Urban 2003 IOP3

The OKC domain for IOP3 includes the central business district, with a maximum building height of 120 m and an average building height of 30 m. [Figure 7](#) shows the complexity of the wind flow in the downtown area during IOP3 generated using FEM3MP with steady inflow boundary conditions on the southern and western edges of the domain. Comparisons of dispersion results are made to 30-min averages of concentration measured at fifteen sensors within this domain. The domain is discretized using about 580 000 elements (132, 146, 30) covering a region of approximately  $x = [-260, 346]$ ,  $y = [-68, 590]$ . The prior distribution is limited to a somewhat smaller domain ( $x = [-150, 130]$ ,  $y = [80, 410]$ ) to reduce computation time. The source strength was allowed to vary from 0.000 01 to 1.0  $\text{kg s}^{-1}$  with a mean of 0.5 and standard deviation of  $q_s = 0.5$ . Standard deviations for the location sampling were set to  $x_s = y_s = 100$  m with means near the center of the sample domain at  $x = 0$  m and  $y = 80$  m. Standard deviations for source location and strength were determined by the problem domain size and refined with a trial and error procedure to ensure that the Markov chains had access to realistic ranges with minimal occurrences of “stuck” chains. Stuck chains can occur when the standard deviations chosen for the next iteration lead to a large number of rejected samples such that the chain remains in a given position for many iterations.

In addition, the cell spacing was effectively doubled by only considering sources in every other grid cell in a checkerboard pattern. Total computation time for 2560 forward runs (from each possible source location in the concentration database) was over 12 h using 1024 2.4-GHz Xeon processors (equivalent to 17 days on 32 processors). Each forward run of FEM3MP simultaneously calculated 20 different source locations, requiring 128 different launches of the model. Each instance of the model used 32 processors. The required computational time for all the forward runs is currently too large to allow for immediate use of our inversion algorithm with CFD for real-time emergency-response scenarios. A database of forward runs, however, could be created ahead of time for a small urban area based on local climatology for a full range of possible source locations. Especially with the continued rapid growth of computational resources, it is easy to envision the creation of such CFD databases for source inversion in urban areas. Initially databases could be created for the purpose of critical facility protection applications, and over time they could be extended to also include wider areas of densely populated urban centers. After generating a database, the inversion process itself requires less than 10 min of computation time on four processors, a cost that is already affordable today.

#### a. Source inversion

[Figure 8](#) shows the location of the buildings and 15 sensors in the downtown OKC area, together with four Markov chain paths. The chains quickly converge from four random initial locations to the general vicinity of the actual source location, where they spend the remainder of the time sampling the parameter space and refining the probability distribution. Using the Markov chain paths, we construct the probability distribution for the source location, as shown in [Fig. 9](#). Recall that a checkerboard pattern, reflected in the shaded squares in the figure, was used for selecting possible sources. The peak of the distribution is located approximately 70 m south of the actual source location. Reasons for this will be discussed below. The accompanying release rate histogram is given in [Fig. 10](#). The peak of the distributions falls near 0.001  $\text{kg s}^{-1}$ , while the actual source strength was 0.005  $\text{kg s}^{-1}$ .

[Figure 11](#) shows convergence rates for  $x$ ,  $y$ , and  $q$  during the 20 000 iterations of the inversion procedure for OKC IOP3. The values for  $x$ ,  $y$ , and  $q$  converge after 10 000 iterations and only change slightly after that. The value for  $y$  is sometimes more difficult to pinpoint in the inversion process. Here,  $y$  is the streamwise direction, where a change in the distance to the source can sometimes be accommodated by a corresponding change in release rate. That is, a weaker source closer to the sensor can sometimes produce similar results to a stronger source farther away. Therefore, a value of  $R = 2$  for the  $y$  location of the source is considered acceptable.

A closer look at the individual plumes predicted by different source locations gives insight into the location of the peak of the  $x$ ,  $y$  probability distribution. [Figure 12](#) shows the plume predicted by FEM3MP for a source at the actual source location for IOP3 with the actual release rate. Contours of concentrations predicted by FEM3MP are shown together with small squares at the sensor locations colored according to the 30-min averaged observed concentrations during IOP3. [Figure 13](#) shows the plume from the inverted source location, that is, the peak of the  $x$ ,  $y$  probability distribution for the source location. While the plumes predicted by the code seem reasonable, there are clearly discrepancies between the predicted concentrations and the observations for both simulated plumes. These can be seen more clearly in a comparison of observed and modeled values at the 15 sensor concentrations. The inverted source location was determined by the stochastic inversion algorithm, which minimizes the absolute error between modeled and observed values. Indeed, the sum of the absolute errors ([Fig. 14](#)) at the sensor locations is smaller using the inverted source location ( $\sim 986$  ppb total) than the true source location ( $\sim 2733$  ppb total). A discussion of model errors is given below.

#### *b. Composite plume*

We again construct a probabilistic, composite plume that represents the probability of concentration at a specific location being higher or lower than a certain value. Contours of the 90% confidence interval are shown in [Fig. 15](#) with the threshold chosen at 10 ppb. Again we note that the shape of this composite plume is quite different from any individual realization or plume prediction, such as those shown in [Figs. 12, 13](#). The white region indicates a lack of information and the inability to specify a 90% confidence interval at those locations (this region is dependent on the choice of the threshold value). The dark blue region envelopes the composite plume, indicating regions where there is 90% confidence that the concentrations are less than 0.01 ppb.

#### *c. Treatment of model errors*

The inversion procedure clearly relies heavily on the accuracy of the sensor measurements as well as the accuracy of the forward model used for dispersion simulations. While the FEM3MP code has been evaluated and tested for many urban flows, there are several possible sources of error. To obtain a good probabilistic distribution for the source location and strength, all sources of error must in theory be included a priori. Since model errors are sometimes difficult to control or isolate, individual errors are treated as described below.

There are several reasons for the mismatch in predicted and observed concentrations. First of all, most of the observed concentrations are averaged values from a 30-min release, whereas model predictions are steady-state results. Additionally, there are uncertainties in the lateral boundary conditions prescribed in the simulation. Steady inflow has been specified for the inflow boundary, whereas in reality the wind at the domain boundary has fluctuations in space and time. A slight change in mean wind direction can greatly affect dispersion results. [Chan and Leach \(2004\)](#) demonstrated that time-varying inflow boundary conditions significantly changed the concentration plume in simulations of dispersion in Salt Lake City. In addition, to save computation time, the domain size used for the IOP3 simulations is smaller than for those performed by [CL07](#) for OKC, which perhaps increases the influence of the boundaries. We also use a simplified linear eddy-viscosity turbulence model for computational cost reasons, whereas [CL07](#) used a nonlinear eddy-viscosity model that gives better agreement with the data but at a much higher computational cost. The nonlinear eddy-viscosity model often better represents dispersion in regions of building-induced turbulence, hence giving better agreement with observed concentrations as in [CL07](#). This eddy-viscosity model is used for IOP9 below.

Another potential source of error is in the specification of the source term in the simulation. While the tracer gas was released from a point source in the experiment, the model distributes the source over a grid cell, where the vertical injection velocity and concentration are specified at the boundary to match the release rate from the experiment. This yields a nearly steady concentration flux over the grid cell but with numerical oscillations (see region near the source in [Fig. 12](#)) in the neighboring cells because of the strong concentration gradients and insufficient grid resolution in the source area.

It is difficult to quantify the individual contributions of the multiple sources of error in FEM3MP. Model errors are therefore incorporated into the inversion process in a simple, lump-sum fashion by adjusting  $\sigma_{rel}$  of the standard lognormal distribution, the relative error allowed in the comparison between different realizations of the simulation and the observed values. For the OKC simulations,  $\sigma_{rel}$  was set to the relatively high value of 0.5. Our experience with real and synthetic data indicates that the model error dominates over measurement error and that the model error is represented well with  $\sigma_{rel}$  in the range between 0.1 and 0.5 depending on the model used. Lower values of  $\sigma_{rel}$  correspond to a tighter fit of the modeled and observed values and lead to poor reconstruction of the source probability distribution in this application because of the large model errors. Higher values lead to a broader probability distribution, which indicates more flexibility in the fit between the modeled and observed concentration values. When synthetic data are used, model error vanishes and  $\sigma_{rel}$  then effectively represents assumed measurement error. Recall that the value of  $\sigma_{rel}$  was set to 0.05 in the case of flow around the isolated building above; with  $\sigma_{rel} = 0.0$ , the model was able to perfectly reconstruct the synthetic data for that idealized case.

### **5. Oklahoma City–Joint Urban 2003 IOP9**

As a further example of the building-resolving inversion procedure, we have also applied the source inversion to observations obtained during IOP9. The IOP9 simulations use the full-size domain as well as the more sophisticated three-equation nonlinear eddy-viscosity closure of [CL07](#) to eliminate the modeling compromises made in the IOP3 case. The experiment conditions (in particular the true source location), however, lead to a much more challenging situation for the inversion procedure and demonstrate a case in which the procedure is much more sensitive to the data and the forward model errors.

The IOP9 domain covers approximately  $x = [-498, 530]$ ,  $y = [-430, 2580]$  using a grid of  $201 \times 303 \times 45$  (approximately 2.75 million grid points). Grid spacing in the horizontal is as fine as 1–2 m in the vicinity of resolved buildings and 1 m near the surface in the vertical direction. Again, for computational reasons, the prior distribution is restricted to a slightly smaller domain  $\{x = [-180, 200], y = [310, 525]\}$  and a checkerboard pattern is used to limit the total number of forward runs to 3360. The forward simulations required about 100 h of wall clock time using 1024 processors. Standard deviations for the location sampling were again set to  $x_\sigma = y_\sigma = 100$  m with means near the center of the sample domain at  $x = 0$  m and  $y = 400$  m. The source strength was allowed to vary from 0.00001 to  $1 \text{ kg s}^{-1}$ . The inversion procedure for IOP9 required approximately 10 min of computation time on four processors. The inversion time depends on the choice of  $x_\sigma$  and  $y_\sigma$  and the proximity to buildings; the density of buildings in the IOP9 source region slows down the sampling procedure since samples that fall within buildings are not allowed. Thus, a chain located in a narrow gap between buildings is limited in its choices for the next iteration; this requirement that samples not be located within buildings does not count toward a sample rejection or acceptance and only slightly slows the algorithm.

#### a. Source inversion

The resulting Markov chain paths and  $x, y$  probability distribution are shown in [Figs. 16, 17](#), respectively. The IOP9 experiment collected data from only eight sensors (as compared with 15 in IOP3), as shown in [Fig. 16](#); model results are compared with 15-min averages of concentration at the sensor locations from 15 to 30 min after the release. The availability of fewer sensors, combined with the location of the source between two buildings, creates difficulties for the inversion procedure. [Figure 17](#) indicates three probability peaks, clustered between different sets of buildings. The four Markov chains converge to locations far south of the true source location, though time series of the  $x, y$ , and  $q$  values do not clearly identify a single final source choice but continue to jump within the three peak regions of the probability distribution (not shown), contrary to the convergence rate plots shown in [Fig. 18](#) that indicate a trend of convergence (values of  $R$  less than 2; [Delle Monache et al. 2008](#)). Probability distributions of the  $x, y$ , and  $q$  values are shown independently in [Fig. 19](#). The  $y$  distribution shows three distinct peaks corresponding to the gaps between the buildings. The  $q$  distribution indicates good agreement in source strength; the inversion is able to pinpoint the source strength to within a narrow range from the original distribution of 0 to  $1 \text{ kg s}^{-1}$ : the peak indicates values between 0.001 and  $0.004 \text{ kg s}^{-1}$ , which compares reasonably well to the true source strength for IOP9 of  $0.002 \text{ kg s}^{-1}$ .

The peak of the full probability distribution (from [Fig. 17](#)) near  $x = 5$  m,  $y = 385$  m and the diffuse peak in the region of  $x = 5$  m,  $y = 320$  m give smaller errors (as indicated by the reconstruction) relative to simulation results from the actual source location at  $x = 30$  m,  $y = 435$  m. When restricting the  $x, y$  probability distribution to source strength values  $q$  within 50% of the true source strength ( $0.001\text{--}0.003 \text{ kg s}^{-1}$ ), the resulting conditional probability distribution shows a peak (near  $y = 440$  m) within 50 m west of the actual source location (see [Fig. 20](#)).

#### b. Composite plume

[Figure 21](#) shows the resulting composite plume using the IOP9 inversion data. The shape of the plume is entirely different from any single realization. The broadness of the composite plume shapes reflects the uncertainty in the inversion procedure, a natural property of our stochastic inversion procedure. The composite plume is unable to indicate concentration levels with any specificity in this case; it merely delineates regions where it is 90% likely that the concentration will be greater than 10 ppb (green) and 90% likely that it will be less than 0.1 ppb (blue).

#### c. Discussion of model errors

The complexity added by the presence of the source location between two buildings perpendicular to the flow direction appears to challenge the inversion procedure more than in the case of IOP3, but this is largely due to errors in the model predictions in comparison with the sensor observations because of reasons mentioned above. A test of this hypothesis was performed using synthetic sensor data. With the source placed at the true location and using the true source strength, the forward model results were collected at the eight sensors to create a synthetic observation dataset. Inversion results using the synthetic data are shown in [Fig. 22](#). All four Markov chains used in the inversion correctly identified the true source under these circumstances, using the same inversion parameters as those for the standard IOP9 run. Several values of  $\sigma_{rel}$ ,  $x_\sigma$ , and  $y_\sigma$  were tested to determine sensitivity to these choices. One chain occasionally converged to a location far from the source (depending on  $x_\sigma$  and  $y_\sigma$ ; not shown), indicating that the building geometry also adds complexity to the flow so that multiple source locations are possible even when model error is largely removed by using synthetic data. The choice of inversion parameters is also important in determining the rate and accuracy of convergence, as discussed further below.

## 6. Effect of sensor density

A common question for urban planners to consider is the placement of chemical-detecting sensors in regions of high interest: for example, near dense or high-occupancy buildings in urban areas. Sensor network design is easily evaluated using our stochastic algorithm, which can be used to indicate the importance of a sensor to source inversion in a particular region ([Lundquist 2005](#)). A related question exists with regard to the number of sensors required for accurate source inversion. The appropriate number generally depends on the complexity of building geometries and ambient wind conditions.

We have evaluated the latter question using IOP3 as our test case. Of the available 15 sensors, inversion procedures were carried out using eight, four, two, and then just one of these sensors. The corresponding probability distributions are shown in [Fig. 23](#). As expected, the probability distribution broadens significantly as the number of sensors is reduced, reflecting the increased uncertainty because of the fewer data points involved. Nevertheless, results with two sensors are still able to identify the general region of the source, thus indicating that even as few as two sensors may be useful in an urban environment, provided they are deployed at the appropriate locations. Interestingly, the peak probability in the case of two sensors ([Fig. 23c](#)) is closer to the actual source than with four sensors ([Fig. 23b](#)), indicating the sensitivity to the particular chosen sensor locations in addition to the number of sensors. With one sensor, the probability distribution becomes much more sensitive to model and observation errors. [Figure 23d](#) shows that if a sensor is used with a zero reading, it simply outlines the region where the source cannot be located. In this case, however, this outline is incorrect since the model is not able to reproduce the zero reading even when the source is in the correct location. Choosing another sensor ([Fig. 23e](#)) with a nonzero reading produces better results.

## 7. Discussion and conclusions

Our stochastic methodology for source inversion is based on a Bayesian inference combined with a Markov chain Monte Carlo sampling procedure. The stochastic approach used in this work is computationally intensive, but the method is completely general and can be used for time-varying release rates and flow conditions, nonlinear problems, and problems characterized by non-Gaussian distributions. The results of the inversion, specifically the shape and size of the posterior probability distribution, indicate the probability of a source being found at a particular location with a particular release rate, thereby inherently reflecting uncertainty in observed data or the data's insufficiency with respect to quality, or spatial or temporal resolution.

We have demonstrated successful inversion of a prototype problem with flow around an isolated building. Application to the complex conditions present during IOP3 and IOP9 of the Joint Urban 2003 experiment in Oklahoma City also proved successful. Despite the many sources of error present in the comparison of model predictions with observed data during the inversion procedure, the peak of the probability distribution for the source location was within 70 m of the true source location for IOP3, and the actual source location was contained within the top percentiles of the probability distribution. For IOP9, model errors and other uncertainties limited the ability of the inversion procedure to exactly pinpoint the true source, though the source was contained within the broader distribution. A composite plume showing concentrations at the 90% confidence level was created for all three cases using plume predictions from the realizations given by the reconstructed probability distribution. This composite plume contains probabilistic information from the iterative inversion procedure and can be used by emergency responders as a tool to determine the likelihood of concentration at a particular location being above or below a threshold value. The effect of sensor density was also evaluated for IOP3 and was found to give expected increases in the spread of the source probability distribution with a decrease in the number of available sensors.

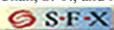
Uncertainties in the inversion procedure increase with the complexity of the domain, paralleling the errors in the forward model. Because the probability distributions are able to reflect the uncertainty in source location, the source inversion procedure demonstrated here indicates high potential to be a useful tool for emergency responders regardless of model limitations. The building-resolving capability introduced here will enable source locations to be pinpointed with high resolution. The limiting factor to real-time response situations is currently the large computation time required. It is, however, conceivable that the building-resolving CFD model could be coupled with a simpler forward model [e.g., the Lagrangian Operational Dispersion Integrator (LODI), the Lagrangian particle model used by [Delle Monache et al. \(2008\)](#); see also [Nasstrom et al. \(2000\)](#) and [Ermak and Nasstrom \(2000\)](#)] to reduce the prior distribution to a reasonable size. Thus, LODI can be used over a large urban region with our stochastic inversion algorithm and inversion with FEM3MP could follow to pinpoint the source within the subregion identified by the LODI inversion. It is also conceivable that databases could be generated in advance for specific urban areas for plume predictions from CFD simulations, similar to the databases generated for IOP3 and IOP9 in this work. Forward runs could be performed for a variety of prevailing wind directions. In an emergency situation, the inversion procedure could be performed in a very reasonable time frame, on the order of 10 min.

Efforts to reduce forward model errors are under way in parallel research programs; forward model capabilities do not inhibit the stochastic algorithm in any way, though the practical applications of course depend on both. Further experience with inversion procedures in urban areas will lead to a better grasp of the range of the inversion parameter space most suitable for dense urban areas with complex building geometries. Future work will also include investigation of unsteady releases, unsteady flow conditions, reactive chemistry, and elevated sources. Meteorological uncertainty will also be incorporated to allow for errors induced by lack of sufficient information at the lateral boundaries, such as errors in the specified mean wind direction.

### Acknowledgments

Thanks are extended to Roger Aines, Luca Delle Monache, Kathy Dyer, William Hanley, and John Nitao for their contributions to this work. Computations were performed on Linux clusters at the Lawrence Livermore National Laboratory (LLNL) computing center. This work was performed under the auspices of the U.S. Department of Energy by the University of California, LLNL under Contract W-7405-Eng-48. The project (04-ERD-037) was funded by the Laboratory Directed Research and Development Program at LLNL.

### REFERENCES

- Allwine, K. J., M. J. Leach, L. W. Stockham, J. S. Shinn, R. P. Hosker, J. F. Bowers, and J. C. Pace, 2004: Overview of Joint Urban 2003—An atmospheric dispersion study in Oklahoma City. Preprints, *Fourth Symp. on Planning, Nowcasting, and Forecasting in the Urban Zone*, Seattle, WA, Amer. Meteor. Soc., J7.1. 
- Bennett, A. F., 2002: *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, 234 pp. 
- Britter, R. E., and S. R. Hanna, 2003: Flow and dispersion in urban areas. *Annu. Rev. Fluid Mech.*, **35**, 469–496. [\[CrossRef\]](#) 
- Chan, S. T., and D. E. Stevens, 2000: An evaluation of two advanced turbulence models for simulating flow and dispersion around buildings. *Proc. Millennium NATO/CCMS Int. Tech. Meeting on Air Pollution Modeling and its Application*, Boulder, CO, NATO/CCMS, 355–362. 
- Chan, S. T., and M. J. Leach, 2004: Large eddy simulation of an URBAN 2000 experiment with various time-dependent forcing. Preprints, *Fifth Symp. on the Urban Environment*, Vancouver, BC, Canada, Amer. Meteor. Soc., 13.3. 
- Chan, S. T., and M. J. Leach, 2007: A validation of FEM3MP with Joint Urban 2003 data. *J. Appl. Meteor. Climatol.*, **46**, 2127–2146. [\[Abstract\]](#) 

Chow, F. K., B. Kosović, and S. T. Chan, 2006: Source inversion for dispersion in urban environments using building-resolving simulations and Bayesian inference with stochastic sampling. Preprints, *Sixth Symp. on the Urban Environment*, Atlanta, GA, Amer. Meteor. Soc., J4.4.



Delle Monache, L., and Coauthors, 2008: Bayesian inference and Markov chain Monte Carlo sampling to reconstruct a contaminant source at continental scale. *J. Appl. Meteor. Climatol.*, in press.

Enting, I. G., 2002: *Inverse Problems in Atmospheric Constituent Transport*. Cambridge University Press, 392 pp.

Ermak, D. L., and J. S. Nasstrom, 2000: A Lagrangian stochastic diffusion method for inhomogeneous turbulence. *Atmos. Environ.*, **34**, 1059–1068. [\[CrossRef\]](#)

Errico, R. M., 1997: What is an adjoint model? *Bull. Amer. Meteor. Soc.*, **78**, 2577–2591. [\[Abstract\]](#)

Gelman, A., J. Carlin, H. Stern, and D. Rubin, 2003: *Bayesian Data Analysis*. Chapman and Hall/CRC, 668 pp.

Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, 1996: *Markov Chain Monte Carlo in Practice*. Chapman and Hall/CRC, 486 pp.



Gresho, P., and S. Chan, 1998: Projection 2 goes turbulent and fully implicit. *Int. J. Comput. Fluid Dyn.*, **9**, 249–272. [\[CrossRef\]](#)

Johannesson, G., B. Hanley, and J. Nitao, 2004: Dynamic Bayesian models via Monte Carlo—An introduction with examples. LLNL Tech. Rep. UCRL-TR-207173, Lawrence Livermore National Laboratory, 53 pp.

Johannesson, G., and Coauthors, 2005: Sequential Monte Carlo-based framework for dynamic data-driven event reconstruction for atmospheric release. *Proc. 2005 Joint Statistical Meetings*, Minneapolis, MN, American Statistical Association, 73–80.

Keats, A., E. Yee, and F-S. Lien, 2007: Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.*, **41**, 465–479. [\[CrossRef\]](#)

Lundquist, J. K., and S. T. Chan, 2007: Consequences of urban stability conditions for computational fluid dynamics simulations of urban dispersion. *J. Appl. Meteor. Climatol.*, **46**, 1080–1097. [\[Abstract\]](#)

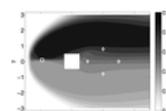
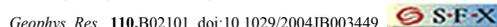
Lundquist, J. K., B. Kosovic, and R. Belles, 2005: Synthetic event reconstruction experiments for defining sensor network characteristics. LLNL Tech. Rep. UCRL-TR-217762, Lawrence Livermore National Laboratory, 32 pp.

Nasstrom, J. S., G. Sugiyama, J. M. Leone Jr., and D. L. Ermak, 2000: A real-time atmospheric dispersion modeling system. Preprints. *11th Joint Conf. on the Applications of Air Pollution Meteorology*, Long Beach, CA, Amer. Meteor. Soc., 84–89.

Neuman, S., L. Glascoe, B. Kosović, K. Dyer, W. Hanley, and J. Nitao, 2006: Event reconstruction for atmospheric releases employing urban puff model UDM with stochastic inversion methodology. Preprints, *Sixth Symp. on the Urban Environment*, Atlanta, GA, Amer. Meteor. Soc., J4.5.



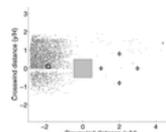
Ramirez, A. L., and Coauthors, 2005: Stochastic inversion of electrical resistivity changes using a Markov chain Monte Carlo approach. *J. Geophys. Res.*, **110**.B02101, doi:10.1029/2004JB003449.



[View larger version \(28K\)](#)

**FIG. 1.**

Horizontal concentration contours at the first vertical level generated by forward simulation with FEM3MP for flow around an isolated building (white box). Four sensors are placed in the lee of the building (white diamonds). The source location is indicated by the white square.



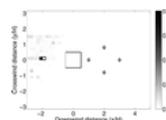
[View larger version \(28K\)](#)

**FIG. 2.**

Gray dots show locations sampled by the four Markov chains used for source inversion for flow around an isolated building (gray square at origin). Black diamonds indicate the four sensor locations. Black stars show the random starting points of the Markov chains (two are collocated at the origin). Small black square shows true source location.

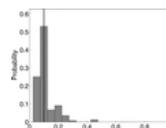
**FIG. 3.**

Probability distribution of source location for flow around an isolated building. Black diamonds indicate the four sensor locations. Small black square outline shows true source



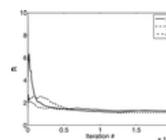
[View larger version \(19K\)](#)

location adjacent to peak of probability distribution. Each small shaded square corresponds to the probability of the source being located within that grid element (see grid lines in [Fig. 6](#)).



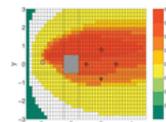
[View larger version \(14K\)](#)

**FIG. 4.** Histogram of source strengths for flow around an isolated building. Solid vertical line shows actual release-rate magnitude.



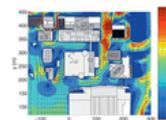
[View larger version \(15K\)](#)

**FIG. 5.** Convergence rates for horizontal position  $(x, y)$  and source strength  $q$  for flow around an isolated building.



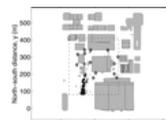
[View larger version \(114K\)](#)

**FIG. 6.** Composite plume showing 90% confidence intervals for concentration levels for flow around an isolated building (gray box). The threshold is set at 0.03. For concentrations above the threshold, there is 90% confidence that the concentration is higher than the contoured value. For values below the threshold, there is 90% confidence that the concentration is less than the contoured value. White region indicates that a 90% confidence interval cannot be established. Black diamonds indicate the four sensor locations. Small magenta square shows true source location.



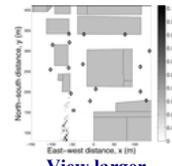
[View larger version \(117K\)](#)

**FIG. 7.** Surface wind vectors (every third point shown in each direction) and contours of velocity magnitude ( $\text{m s}^{-1}$ ) predicted by FEM3MP for flow in the central business district of Oklahoma City during IOP3 of the Joint Urban 2003 field experiment. Buildings are indicated with various shades of gray.



[View larger version \(38K\)](#)

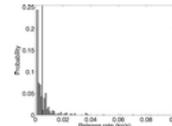
**FIG. 8.** Black dots show locations sampled by the four Markov chains used for source inversion for flow in Oklahoma City during IOP3. Black diamonds indicate sensor locations. Black stars show the random starting points of the Markov chains. Small black square shows true source location. Buildings that are treated explicitly are outlined in black in addition to shading; others are treated as virtual buildings. Dashed line shows zoomed-in region near the source used for later figures.



[View larger version \(38K\)](#)

**FIG. 9.**

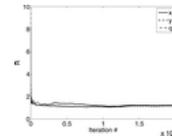
Probability distribution of source location for flow in OKC during IOP3. Only subdomain indicated by dashed line in Fig. 8 is shown. Actual source location is shown by black square outline. Buildings are shaded in gray. Each small shaded square corresponds to the probability of the source being located within that grid element. A checkerboard pattern was used in selecting possible source locations and this is reflected in the distribution here.



[View larger version \(15K\)](#)

**FIG. 10.**

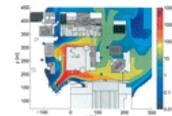
Histogram of source strengths for flow in OKC during IOP3. Solid vertical line shows magnitude of actual release rate.



[View larger version \(13K\)](#)

**FIG. 11.**

Convergence rates for horizontal position ( $x, y$ ) and source strength ( $q$ ) for flow in OKC during IOP3.



[View larger version \(92K\)](#)

**FIG. 12.**

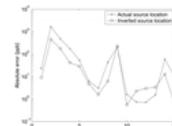
Concentration plume predicted by FEM3MP with actual source location (small black square) and release rate for OKC IOP3 in comparison with averaged concentration measurements (small squares colored by concentration value).



[View larger version \(86K\)](#)

**FIG. 13.**

As in Fig. 12, but source location and strength are from peak of reconstructed probability distribution.

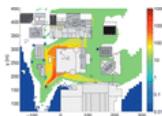


[View larger version \(22K\)](#)

**FIG. 14.**

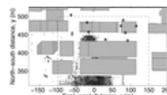
Absolute error of FEM3MP predictions in comparison with observed concentrations at the 15 sensor locations for actual and inverted source locations.

**FIG. 15.**



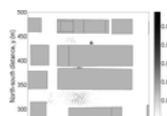
[View larger version \(76K\)](#)

Composite plume showing 90% confidence intervals for concentration levels for flow in OKC during IOP3. Observed concentrations are also shown as small colored squares. The threshold is set at 10 ppb.



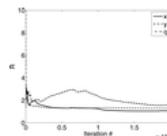
[View larger version \(43K\)](#)

**FIG. 16.**  
As in [Fig. 8](#), but for IOP9.



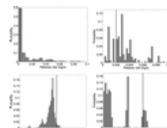
[View larger version \(31K\)](#)

**FIG. 17.**  
As in [Fig. 9](#), but for IOP9.



[View larger version \(15K\)](#)

**FIG. 18.**  
Convergence rates for horizontal position  $(x, y)$  and source strength  $q$  for flow in OKC during IOP9.



[View larger version \(30K\)](#)

**FIG. 19.**  
Histogram of source strengths and  $x, y$  positions for flow in OKC during IOP9. Vertical solid lines denote release rates and location coordinates of actual source.



[View larger version \(31K\)](#)

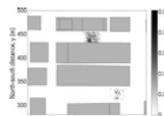
**FIG. 20.**  
As in [Fig. 17](#), but for conditional probability distribution of source location for flow in OKC during IOP9 with source strength values within 50% of true value.



**FIG. 21.**  
Composite plume showing 90% confidence intervals for concentration levels for flow in OKC during IOP9. Observed concentrations are also shown as small colored squares. The threshold is set at 10 ppb.



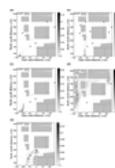
[View larger version \(78K\)](#)



[View larger version \(29K\)](#)

**FIG. 22.**

Probability distribution of source location for flow in OKC during IOP9 using synthetic sensor data generated from a forward simulation at the actual source location. Otherwise, as in [Fig. 17](#).



[View larger version \(49K\)](#)

**FIG. 23.**

Probability distribution of source location for flow in OKC during IOP3 using (a) eight, (b) four, and (c) two sensors, and (d), (e) two different instances with one sensor chosen from the original 15. Sensors chosen are shown with black diamonds. The true source is indicated by the black square. Otherwise, as in [Fig. 9](#).

#### Cited by

Denglong Ma, Simin Wang, Zaoxiao Zhang. (2014) Hybrid algorithm of minimum relative entropy-particle swarm optimization with adjustment parameters for gas source term identification in atmosphere. *Atmospheric Environment*. Online publication date: 1-May-2014.

[CrossRef](#)

Denglong Ma, Jianqiang Deng, Zaoxiao Zhang. (2013) Comparison and improvements of optimization methods for gas emission source identification. *Atmospheric Environment* **81**, 188-198. Online publication date: 1-Dec-2013.

[CrossRef](#)

Derek Wade, Inanc Senocak. (2013) Stochastic reconstruction of multiple source atmospheric contaminant dispersion events. *Atmospheric Environment* **74**, 45-51. Online publication date: 1-Aug-2013.

[CrossRef](#)

Nan Wu, Rui Yang, Hui Zhang, Lifeng Qiao. (2013) Decentralized Inverse Model for Estimating Building Fire Source Location and Intensity. *Journal of Thermophysics and Heat Transfer* **27**:3, 563-575. Online publication date: 1-Jul-2013.

[CrossRef](#)

Andrew J. Annunzio, George S. Young, Sue Ellen Haupt. (2012) Utilizing state estimation to determine the source location for a contaminant. *Atmospheric Environment* **46**, 580-589. Online publication date: 1-Jan-2012.

[CrossRef](#)

Eugene Yee. (2012) Inverse Dispersion for an Unknown Number of Sources: Model Selection and Uncertainty Analysis. *ISRN Applied Mathematics* **2012**, 1-20. Online publication date: 1-Jan-2012.

[CrossRef](#)

B. Addepalli, K. Sikorski, E.R. Pardyjak, M.S. Zhdanov. (2011) Source characterization of atmospheric releases using stochastic search and regularized gradient optimization. *Inverse Problems in Science and Engineering* **19**:8, 1097-1124. Online publication date: 1-Dec-2011.

[CrossRef](#)

Ivan V. Kovalets, Spyros Andronopoulos, Alexander G. Venetsanos, John G. Bartzis. (2011) Identification of strength and location of stationary point source of atmospheric pollutant in urban conditions using computational fluid dynamics model. *Mathematics and Computers in Simulation* **82**:2, 244-257. Online publication date: 1-Oct-2011.

[CrossRef](#)

Xiaoping Zheng, Zengqiang Chen. (2011) Inverse calculation approaches for source determination in hazardous chemical releases. *Journal of Loss Prevention in the Process Industries* **24**:4, 293-301.

Online publication date: 1-Jul-2011.

[CrossRef](#)

Shaodong Guo, Rui Yang, Hui Zhang, Xin Zhang. (2010) New Inverse Model for Detecting Fire-Source Location and Intensity. *Journal of Thermophysics and Heat Transfer* **24**:4, 745-755.

Online publication date: 1-Oct-2010.

[CrossRef](#)

Luca Delle Monache, Jeffrey Weil, Matthew Simpson, Marty Leach. (2009) A new urban boundary layer and dispersion parameterization for an emergency response modeling system: Tests with the Joint Urban 2003 data set. *Atmospheric Environment* **43**:36, 5807-5821.

Online publication date: 1-Nov-2009.

[CrossRef](#)

Luca Delle Monache, Julie K. Lundquist, Branko Kosović, Gardar Johannesson, Kathleen M. Dyer, Roger D. Aines, Fotini K. Chow, Rich D. Belles, William G. Hanley, Shawn C. Larsen, Gwen A. Loosmore, John J. Nitao, Gayle A. Sugiyama, Philip J. Vogt. (2008) Bayesian Inference and Markov Chain Monte Carlo Sampling to Reconstruct a Contaminant Source on a Continental Scale. *Journal of Applied Meteorology and Climatology* **47**:10, 2600-2613.

Online publication date: 1-Oct-2008.

[Abstract](#) . [Full Text](#) . [PDF \(918 KB\)](#)



© 2014 American Meteorological Society [Privacy Policy and Disclaimer](#)

Headquarters: 45 Beacon Street Boston, MA 02108-3693

DC Office: 1120 G Street, NW, Suite 800 Washington DC, 20005-3826

[amsinfo@ametsoc.org](mailto:amsinfo@ametsoc.org) Phone: 617-227-2425 Fax: 617-742-8718

[Allen Press, Inc.](#) assists in the online publication of AMS journals

Technology Partner - Atypon Systems, Inc.