UCRL-CONF-216903

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

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

Fotini Katopodes Chow, Branko Kosovic and Stevens T. Chan

2006

86th American Meteorological Society Annual Meeting, Atlanta, GA January 29, 2006 through February 2, 2006.

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

This report has been reproduced
directly from the best available copy.

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (423) 576-8401
http://apollo.osti.gov/bridge/

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161
http://www.ntis.gov/

OR

Lawrence Livermore National Laboratory
Technical Information Department's Digital Library
http://www.llnl.gov/tid/Library.html

J4.4 SOURCE INVERSION FOR CONTAMINANT PLUME DISPERSION IN URBAN ENVIRONMENTS USING BUILDING-RESOLVING SIMULATIONS

Fotini Katopodes Chow[1]*, Branko Kosović[2], and Stevens T. Chan[2]
[1]Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720
[2]Atmospheric, Earth and Energy Department, Lawrence Livermore National Laboratory, Livermore, CA 94551

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 ($\sim$ 100 m) buildings and channeled through urban canyons that are perpendicular to the wind direction (see e.g., IOP 9 in Chan, 2005). The path of wind and scalars in urban environments is difficult to predict even with building-resolving computational fluid dynamics codes, due to the uncertainty in the synoptic wind and boundary conditions and other errors in the models.

Given the difficulties due to the complexity of urban flows, solving an inverse problem becomes quite challening. That is, given measurements of concentration at sensors scattered throughout a city, is it possible to detect the source of the contaminant? The ability to locate a source and determine its characteristics in a complex environment is necessary for emergency response for 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 strictly limited to processes governed by linear equations. In addition, they assume the system is steady-state and that the equations are linear (Enting, 2002). 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 non-linear 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 sparse, poorly-constrained data problems, as well as in the case of high-volume, potentially over-constrained and diverse data streams.

We have developed a more general and powerful inverse methodology based on Bayesian inference coupled with stochastic sampling. Bayesian methods reformulate the inverse problem into a solution based on efficient sampling of an ensemble of predictive simulations, guided by statistical comparisons with observed data. Predicted values from simulations are used to estimate the likelihoods of available measurements; these likelihoods in turn are 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 non-linear 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 sampling for use with a high-resolution building-resolving computational fluid dynamics code, FEM3MP. The inversion procedure is first applied to flow around an isolated building (a cube) and then to flow in Oklahoma City (OKC) using data collected during the Joint URBAN 2003 field experiment (Allwine, 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.

2 Reconstruction procedure

2.1 Bayesian inference and Markov Chain Monte Carlo

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

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)} \propto p(Y|X)p(X) \qquad (1)$$

where X represents possible model configurations or parameters and Y is observed data. For our application, Bayes theorem therefore describes the conditional probability ($p(X|Y)$) of a certain source location and release rate (the model configuration, X) given observed measurements of concentration at sensor locations (Y), also known as the posterior distribution. This conditional probability $p(X|Y)$ is related to $p(Y|X)$, the probability of the data conforming to a given model configuration, and $p(X)$, the possible model configurations before taking into account the measurements. $p(Y|X)$, for fixed Y, is called the likelihood function, while $p(X)$ is the prior distribution. In this application, we initially assume that the source could be located anywhere in the whole domain, so the prior distribution is uniform (though we in effect limit the prior by choosing our domain boundaries). The probability $p(Y)$ distribution is called the *prior predictive distribution* (Gelman et al., 2003) and represents a marginal distribution of Y

$$p(Y) = \int p(X)p(Y|X)dX \qquad (2)$$

Corresponding author address: Department of Civil and Environmental Engineering, University of California, Berkeley, MC 1710, Berkeley, CA 94720-1710, email: chow@ce.berkeley.edu

For a general problem where analytical solutions are not possible, the challenge is in computing the likelihood function. For that purpose we use a stochastic sampling procedure and approximate the posterior distribution ($p(X|Y)$) by the empirical distribution function

$$\pi(X) = \sum_{i=1}^{N} (1/N)\delta(X_i - X) \tag{3}$$

Here, $\delta(X_i - X) = 1$ when $X_i = X$ and 0 otherwise.

2.2 Sampling procedure

We use a Markov Chain Monte Carlo procedure with the Metropolis-Hastings algorithm to obtain the posterior distribution of the source term parameters given the concentration measurements at sensor locations. 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 buildings (virtual and real) are also excluded from the prior distribution.

A forward dispersion calculation is performed to provide the initial data for comparison with observed data at sensors. Then the Metropolis-Hastings sampling algorithm is used to advance the Markov chains. 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. A forward calculation is performed for the proposal with these new parameters and results are compared to 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, the proposal is not automatically rejected. Instead, a random (Bernoulli) "coin flip" is used to decide whether or not 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.

Each Markov Chain path is determined using this algorithm at each step, as given in detail in Table 1. Multiple chains are used (typically four) to allow for better statistical sampling of the parameter space. Statistical convergence to the posterior distribution is monitored by computing between-chain variance and within-chain variance. If there are m Markov chains of length n then we can compute between-chain variance B and within-chain vairiances W as:

$$B = \frac{n}{m-1} \sum_{j=1}^{m} (\overline{X}_j - \overline{X})^2 \tag{6}$$

where

$$\overline{X}_j = \frac{1}{n} \sum_{i=1}^{n} X_{ij} \tag{7}$$

and

$$\overline{X} = \frac{1}{m} \sum_{j=1}^{m} \overline{X}_i \tag{8}$$

Table 1: Metropolis-Hastings algorithm used for sampling and advancement of Markov Chains in inversion procedure.

- Given current state X_i draw a new candidate state $\tilde{X}$ from the proposal distribution $T(\tilde{X}, X_i)$.
- Compute acceptance ratio as

$$\rho(\tilde{X}, X_i) = \frac{\pi(\tilde{X})T(X_i|\tilde{X})}{\pi(X_i)T(\tilde{X}|X_i)} \tag{4}$$

- Compute acceptance probability $\alpha(\tilde{X}, X_i)$ as

$$\alpha(\tilde{X}, X_i) = \min\left(\rho(\tilde{X}, X_i), 1\right) \tag{5}$$

- Draw u from uniform distribution $U[0,1]$ and update the state X_{i+1}

$$X_{i+1} = \{ \begin{array}{ll} \tilde{X} & \text{if} \quad u \leq \alpha(\tilde{X}; X_i) \\ X_i & \text{otherwise} \end{array}$$

$$W = \frac{1}{m} \sum_{j=1}^{m} s_i^2 \tag{9}$$

and

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^{n} (X_{ij} - \overline{X}_i)^2 \tag{10}$$

One estimate of variance of X is computed as

$$\text{var}(X) = \frac{n-1}{n} W + \frac{1}{n} B \tag{11}$$

The convergence parameter R, is then computed as a function of two estimates of variance

$$R = \frac{\text{var}(X)}{W} \tag{12}$$

The necessary condition for statistical convergence to the posterior distribution is that R approaches unity.

2.3 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) 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 by considering only steady-state flow conditions. (The methodology remains completely general and can handle unsteady flows.) By assuming that the advection-diffusion problem is linear (e.g., no chemical reactions) we can use the precomputed steady flow field and Green's funcions to carry out

one forward simulation at each of the thousands of locations in our prior distribution using a unit source strength and store the resulting values at the sensor locations in a database. The stored concentrations can be rescaled depending on the proposed source release rate for a particular source location. Thus, during the inversion process, the 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 four Markov chains can be performed in less than five minutes of computational time on two processors.

2.4 Forward model description - FEM3MP

The stochastic inversion procedure relies on a forward model to calculate instances of predicted sensor measurements, Y, for given source term parameters, X. Here we use FEM3MP (Gresho and Chan, 1998; Chan et al., 2001), a three-dimensional, incompressible Navier-Stokes finite-element code able to represent complex geometries and simulate flows in urban environments (Chan and Leach, 2004; Chan, 2005).

For flow around the 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 pre-computed 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 in the region closest to the building (see Fig. 3 later).

For the Oklahoma City simulations, we use a setup similar to Chan (2005) for the third intensive observation period (IOP3) from Joint URBAN 2003. Again, the flow field is assumed steady, with a logarithmic inflow profile on the southern boundary with magnitude 6.5 m/s at $z = 50$ m and a wind direction of 185° (south). The flow field is pre-computed using FEM3MP. The release rate is constant and simulations are performed until steady-state concentration fields are achieved (after about 10 minutes of simulation time). The atmosphere is assumed to be neutrally stratified since shear production of turbulence is significantly larger then buoyant production. Buildings near the source are explicitly resolved; i.e. the finite element grid lines up with the buildings (see Fig. 9 later). Far from the source, "virtual buildings" are used to reduce the computational cost. In this region, drag is added to the grid cells falling within the building boundaries. Previous work has shown that this approach produces satisfactory dispersion estimates far from the source (Chan, 2005).

3 Isolated building example

We have developed a prototype example of event reconstruction 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 is collected using a forward simulation from the true source location. The data is thus "synthetic" and used in this case only to test the inversion algorithm. Artificial measurement error with a standard log-normal distribution is also added to the synthetic data (in this case with mean $\mu = 0$ and standard deviation $\sigma_{rel} = 0.05$).

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, this 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 due to 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).

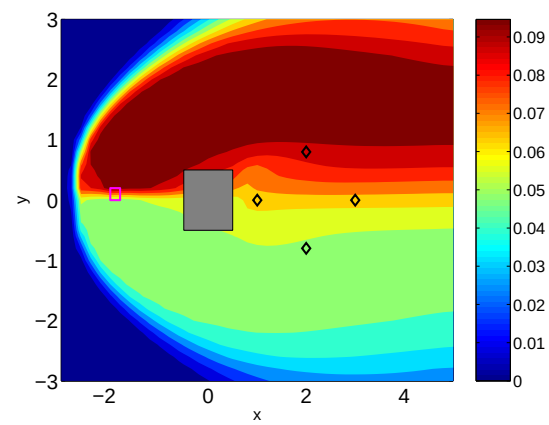

Figure 1: Horizontal concentration contours at the first vertical level generated by forward simulation with FEM3MP for flow around an isolated building (gray shading). Four sensors are placed in the lee of the building (diamonds). The source is indicated by the magenta square.

The domain is discretized using about 19 000 elements (42,32,14). For the purposes of this example, the prior distribution is limited to a region upwind of the sensors to reduce computation time. Forward runs are computed for all possible locations and concentrations values at the sensors are stored in a database for each grid locations. Total computation time for generation of the database was 6 hours using 64 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 described previously. This avoids repeated computations of releases at the same x, y locations by simply scaling the release rate as dictated by the sampling algorithm.

3.1 Source inversion

Figure 2 shows the paths taken 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 the actual concentration plume. The peak of the distribution occurs just upwind of the actual source location. If the error from the measurements is set to zero, 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.

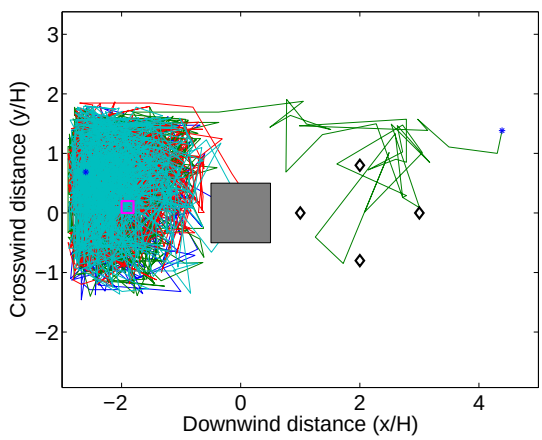

Figure 2: Paths of four Markov chain used for source inversion for flow around an isolated building. Black diamonds show sensor locations.

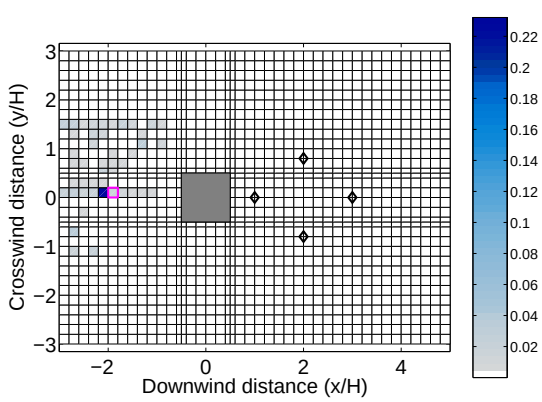

Figure 3: Probability distribution of source location for flow around an isolated building.

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.

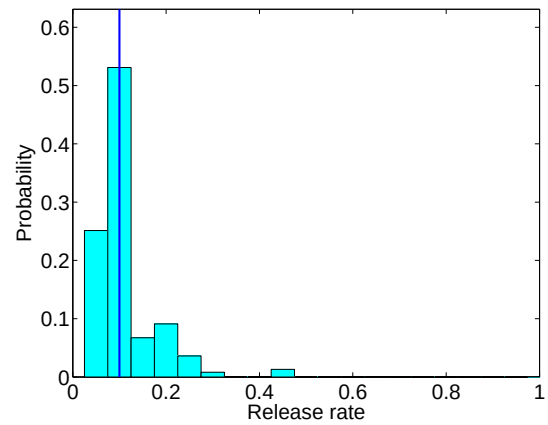

Figure 4: Histogram of source strengths for flow around an isolated building. Vertical blue line shows actual release rate.

Convergence rates for the x, y and q inversions are shown in Fig. 5. All convergence measures 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 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.

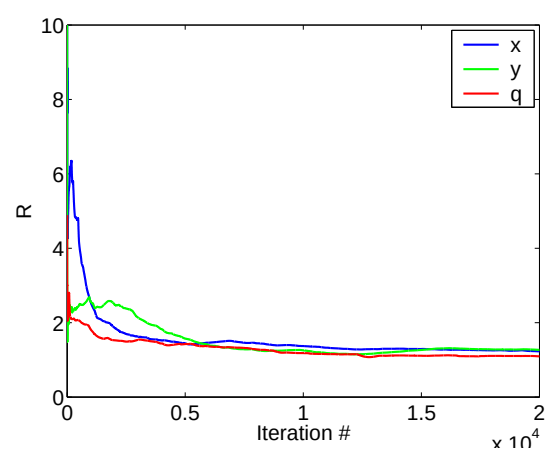

Figure 5: Convergence rates for horizontal position (x, y) and source strength q for flow around an isolated building.

3.2 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. This step is followed by determining the concentration value at each location for which a certain pre-specified 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.

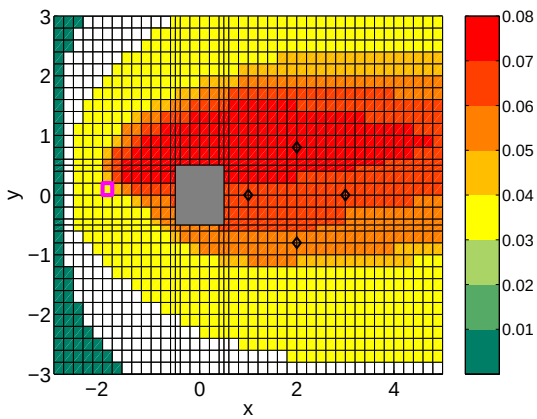

Figure 6: Composite plume showing 90% confidence intervals for concentration levels for flow around an isolated building. The threshold is set at 0.03. For concentrations above the threshold, there is 90% confidence that the concentration is higher than the contoured value. For values below the threshold, there is 90% confidence that the concentration is less than the contoured value. White regions indicate that a 90% confidence interval cannot be established.

The shape of this composite plume is quite different from that of the actual plume (Fig. 1). 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 IOP 3

The OKC domain 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 generated using FEM3MP with constant inflow boundary conditions on the southern edge 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). The prior distribution is limited to a slightly smaller domain ($x = [-150, 130], y = [80, 410]$) to reduce computation time. 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 hours using 1024 2.4 GHz Xeon processors (equivalent to 17 days on 32 processors). Each forward run of FEM3MP simulataneously calculated 20 different source locations, requiring 128 different launches of the model. Each instance of the model used 32 processors. After generation of the database, the inversion process itself requires only five minutes of computation time on two processors.

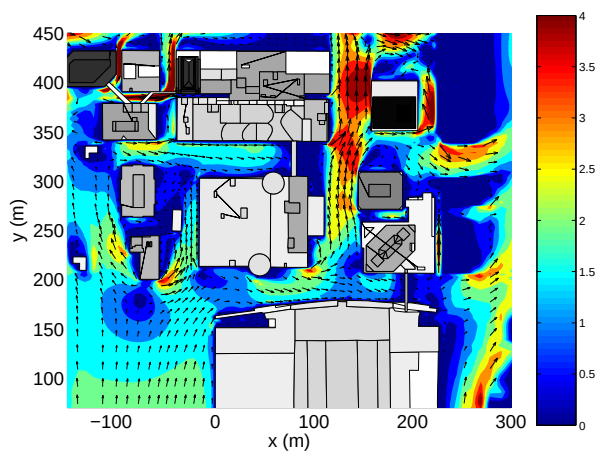

Figure 7: Surface wind vectors (every third point shown in each direction) and contours of velocity magnitude predicted by FEM3MP for flow in the central business district of Oklahoma City during IOP 3 of the Joint URBAN 2003 field experiment. Buildings are indicated with various shades of gray.

4.1 Source inversion

Figure 8 shows the location of buildings and 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 their 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. The peak of the distribution is located approximately 70 meters 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 between 0.003 and 0.004 kg/kg, while the actual source strength was 0.005 kg/kg.

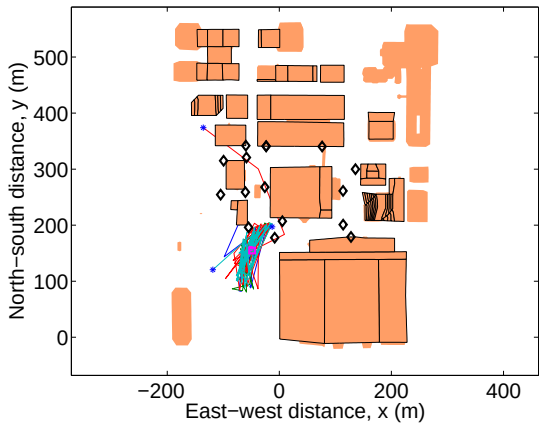

Figure 8: Paths of four Markov chain used for source inversion for flow in Oklahoma City during IOP 3. Black diamonds show sensor locations.

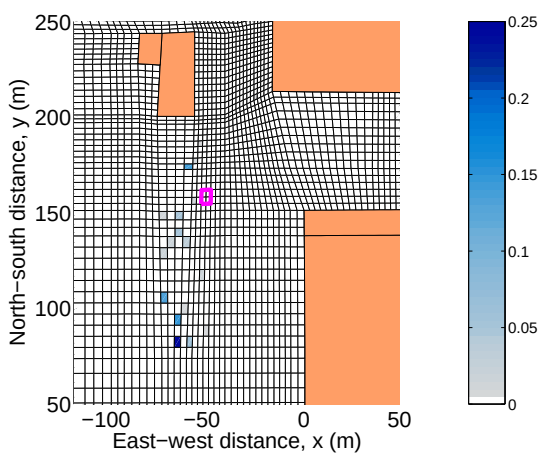

Figure 9: Probability distribution of source location for flow in OKC during IOP 3.

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 and q converge after 10 000 iterations and only change slightly after that. The value for y is more difficult to pinpoint in the inversion process. Here y is the stream-wise 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 the sensor can sometimes produce similar results to a stronger source further away.

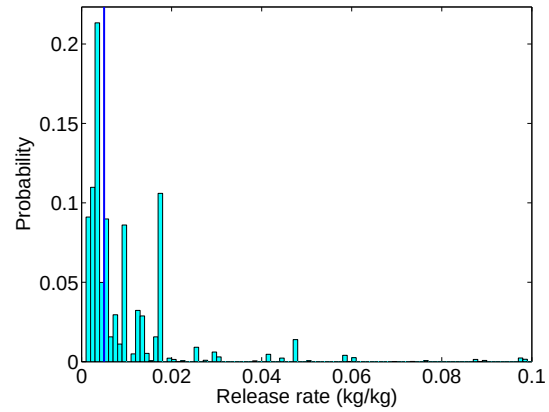

Figure 10: Histogram of source strengths for flow in OKC during IOP 3. Vertical blue line shows actual release rate.

Therefore, a value of $R = 2$ for the y location of the source can be considered acceptable.

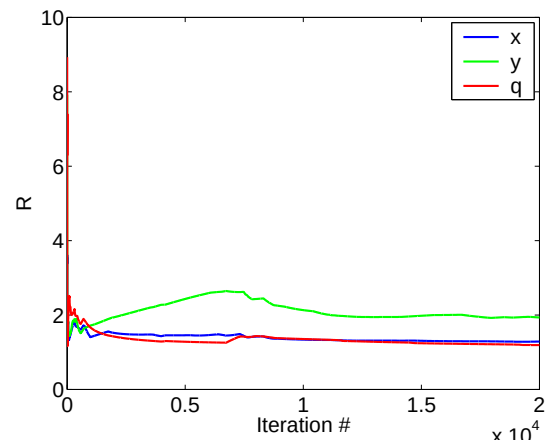

Figure 11: Convergence rates for horizontal position (x, y) and source strength q for flow in OKC during IOP 3.

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 observed concentrations during IOP3. Figure 13 shows the plume from the inverted source location, i.e. the peak of the x, y probability distribution for the source location. While the general plumes predicted by the code seem reasonable, there are clearly discrepancies between the predicted concentrations and observations for both simulated plumes. These can be seen more clearly in a one-to-one comparison of observed and modeled values at the 15 sensor concen-

trations, as seen in Fig. 14. The inverted source location was determined by minimizing the absolute error between modeled and observed values. The sum of the absolute errors (Fig. 15) at the sensor locations is smaller using the inverted source location ($\sim$ 1090 ppb total) than the true source location ($\sim$ 2860 ppb total).

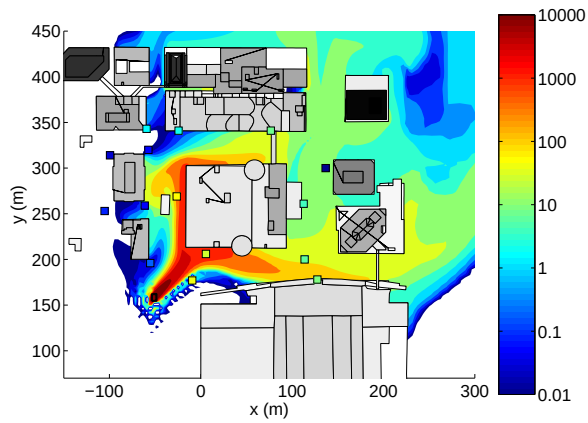

Figure 12: Concentration plume predicted by FEM3MP with observations (small squares colored by concentration value) for actual source location (small black square) and release rate for OKC IOP 3.

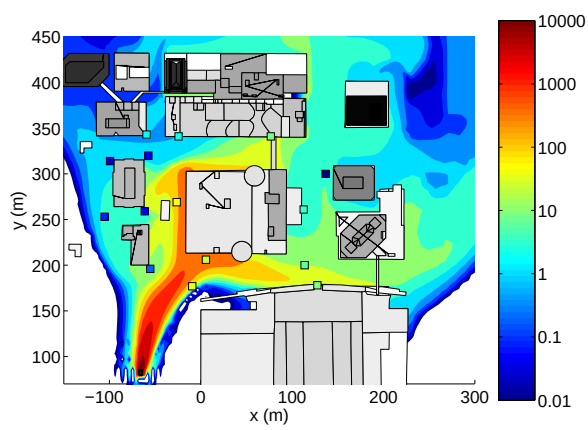

Figure 13: Concentration plume predicted by FEM3MP with source location at peak of reconstructed probability distribution (small black square) compared to observations for OKC IOP 3.

4.2 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 validated for many urban flows, there are several possible sources of error. To

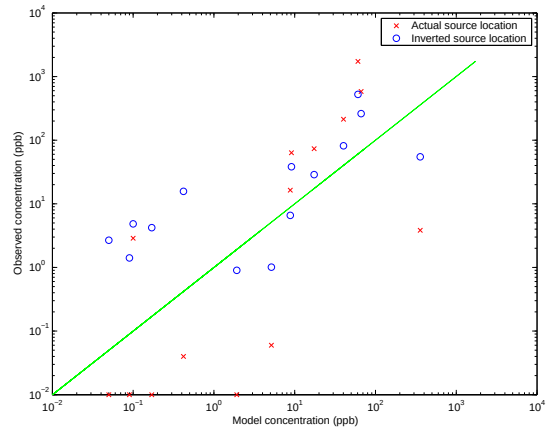

Figure 14: Scatter plot of FEM3MP predictions vs. observed concentrations at the 15 sensor locations for actual and inverted source locations.

obtain a good probabilistic distribution for the source location and strength, all sources of error must be included a priori.

There are several reasons for the mismatch in predicted and observed concentrations. First of all, 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 also 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, the domain size used for these simulations is smaller than that of Chan (2005) (to save computation time) which perhaps increases the influence of the boundaries. We also use a simplified linear eddy-viscosity turbulence model, whereas Chan (2005) uses a non-linear eddy-viscosity model which gives better agreement with the data but at a much higher computational cost. Results from Chan (2005) are reproduced in Fig. 16. The non-linear eddy-viscosity model better represents dispersion in regions of building-induced turbulence, hence giving better agreement with observed concentrations.

Another 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 steady concentration flux over the grid cell but creates numerical oscillations (see region near the source in Fig. 12) in the neighboring cells due to the strong concentration gradients and inherent limitations of the numerical scheme.

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 lump sum fashion by adjusting σ_{rel} of the standard log-normal distribution, the relative error allowed in the comparison between different realizations of the simulation

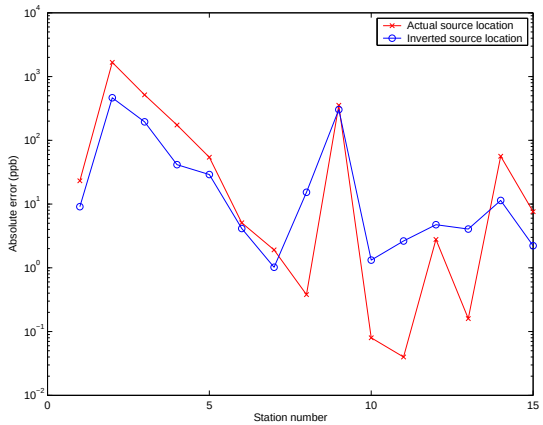

Figure 15: Absolute error of FEM3MP predictions compared to observed concentrations at the 15 sensor locations for actual and inverted source locations.

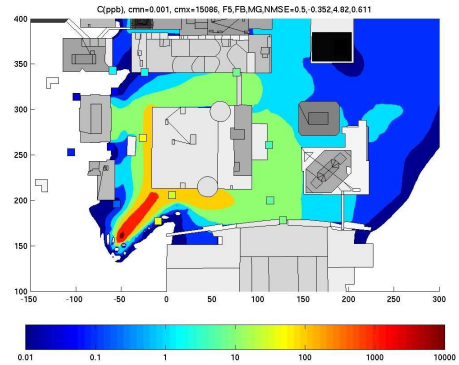

Figure 16: Concentration plume from Chan (2005) predicted by FEM3MP with larger grid and more advanced turbulence model compared to observations (small squares colored by concentration value) for actual source location and release rate for OKC IOP 3.

and the observed values. For the OKC simulations, σ_{rel} was set to the relatively high value of 0.5.

4.3 Composite plume

We again construct a probabilistic, composite plume, representing the probability of concentration at a specific location being higher or lower than a certain value. Due to memory limitations in the post-processing step, only the final third of the iterations are used to construct the composite plume (as opposed to the second half of iterations used for the flow around the isolated building). Contours of the 90% confidence interval are shown in Fig. 17 with the threshold chosen at 100 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 and 13.

5 Discussion and conclusions

Our stochastic methodology for source inversion is based on 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. 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, 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 IOP 3 of the Joint URBAN 2003 experiment in Oklahoma City also proved successful. Despite the many sources of error present in the CFD model (FEM3MP) used during the inversion procedure, the peak of the probability distribution for the source location was within 70 m of the true source location, and the actual source location was contained within the probability distribution. A composite plume showing concentrations at the 90% confidence level was created 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.

Future work will include investigation of source inversion using a smaller subset of the 15 sensors available during the OKC IOP 3 experiment. Unsteady releases and unsteady flow conditions will also be considered. Finally, meteorological uncertainty will 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.

6 Acknowledgments

Thanks are extended to Kathy Dyer, John Nitao, and Bill Hanley for their contributions to this work. Computations were performed on linux clusters at LLNL's computing center. This work was funded by LDRD project number 04-ERD-037 and was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

References

Allwine, K. J., 2004: Joint Urban 2003 field study and urban mesonets. *Fourth Symposium on Planning, Nowcasting,*

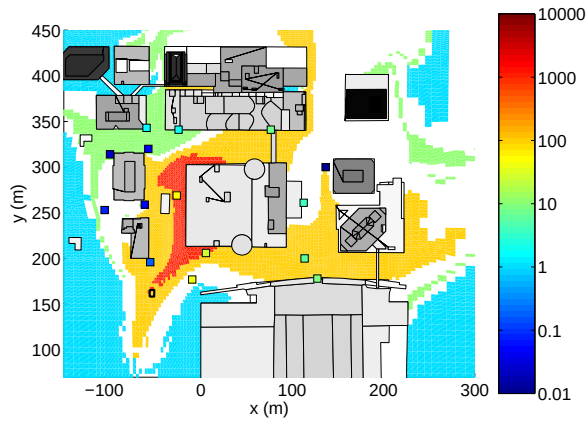

Figure 17: Composite plume showing 90% confidence intervals for concentration levels for flow in OKC during IOP 3. Observed concentrations are also shown as small colored squares. The threshold is set at 100 ppb.

and Forecasting in the Urban Zone, American Meteorological Society.

Bennett, A., 2002: *Inverse Modeling of the Ocean and Atmosphere*. Cambridge Univ. Press, 234 pp.

Britter, R. and S. Hanna, 2003: Flow and Dispersion in Urban Areas. *Ann. Rev. Fluid Mech.*, **35**, 469–496.

Chan, S., D. Stevens, and W. Smith, 2001: Validation of two CFD urban dispersion models using high resolution wind tunnel data. *Third International Symposium on Environmental Hydraulics*, 107.

Chan, S. T., 2005: A validation of FEM3MP with Joint Urban 2003 data. *J. Appl. Meteor.*, **submitted for publication**.

Chan, S. T. and M. J. Leach, 2004: Large eddy simulation of an Urban 2000 experiment with various time-dependent forcing. Paper 13.3. *Fifth Symposium on the Urban Environment, American Meteorological Society*.

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

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

Gresho, P. and S. Chan, 1998: Projection 2 goes turbulent and fully implicit. *International Journal of Computational Fluid Dynamics*, **9**, 249–272.

Johannesson, G., W. Hanley, and J. Nitao, 2004: Dynamic Bayesian Models via Monte Carlo - An Introduction with Examples. Technical Report UCRL-TR-207173, Lawrence Livermore National Laboratory, Livermore, CA.