Application Of A Bayesian Inference Method To Reconstruct Short-Range Atmospheric Dispersion Events

Inanc Senocak

Department of Mechanical and Biomedical Engineering
Boise State University
Boise, Idaho, USA, 83725

Abstract. In the event of an accidental or intentional release of chemical or biological (CB) agents into the atmosphere, first responders and decision makers need to rapidly locate and characterize the source of dispersion events using limited information from sensor networks. In this study the stochastic event reconstruction tool (SERT) is applied to a subset of the Fusing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT 07) database. The inference in SERT is based on Bayesian inference with Markov chain Monte Carlo (MCMC) sampling. SERT adopts a probability model that takes into account both positive and zero-reading sensors. In addition to the location and strength of the dispersion event, empirical parameters in the forward model are also estimated to establish a data-driven plume model. Results demonstrate the effectiveness of the Bayesian inference approach to characterize the source of a short range atmospheric release with uncertainty quantification.

Keywords: Atmospheric dispersion, event reconstruction, Gaussian plume modeling
PACS: 92.60.Sz

INTRODUCTION

In their June 2008 report (GAO-08-180) to Congressional requesters, the Government Accountability Office (GAO) has found that "*While the Department of Homeland Security (DHS) and other agencies have taken steps to improve homeland defense, local first responders still do not have tools to accurately identify right away what, when, where, and how much chemical, biological, radiological, or nuclear materials are released in U.S. urban areas, accidentally or by terrorists*" [1]. Given sensor data, there is a need to develop algorithms that identify and quantify the amount of chemical and biological (CB) materials in an accidental or intentional release event.

DHS has deployed the BioWatch program in several major cities in the U.S. to monitor the air for biothreat agents [2]. The number of sensors in urban areas is limited, and a reliable account of the CB dispersion event and its impact on the population cannot be created purely from measurements. However, these measurements can be used to solve an inverse problem and determine the location and strength of a CB agent release event. Once the dispersion event is backtracked in time it can then be projected forward using high-fidelity atmospheric transport and dispersion models to predict the hazard-zone for emergency response and hazard mitigation. The process of using limited and uncertain information from a sensor network, and its fusion into CB dispersion models, is referred to as event reconstruction, source inversion or source term estimation (STE). The inverse problem under consideration is equally significant in defense operations on the battlefield, where estimates on the location, strength and time of CB agent release can support tactical decisions such as areas to avoid, protective gear usage and medical response [3].

Fusing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT 07) was designed to support the development of source term estimation algorithms and evaluate existing ones [4]. Database provides detailed meteorological information and trace gas concentration measurements for short range (~500 meters) dispersion experiments. These experiments were performed for a variety of release types, including single and multiple sources for continuous and puff (instantaneous) releases. In addition, several different sensor layouts were considered to assess the impact of the number of sensors needed in operational use of STE algorithms.

Various inverse methods have been applied to the event reconstruction problem and they can be categorized as either deterministic or probabilistic approaches. Deterministic approaches include the inverted Gaussian plume model [5], variational finite element models for the adjoint problem [6] and genetic algorithms [7]. Algorithms based on the inverted Gaussian plume model are computationally cheap, but the model is not very accurate and these algorithms do not address the uncertainty in the model or in sensor data. Alternatively, the adjoint formulation can be used to analyze the sensitivity of model parameters, but extension to non-linear problems becomes more challenging and demanding in terms of computational resources. Optimal solutions can be found more efficiently with genetic algorithms but uncertainty quantification can be more complicated with these methods.

Probabilistic approaches involve incorporating uncertainty into the event reconstruction problem. Several studies have adopted the Bayesian inference approach for the event reconstruction problem. Johannesson et al. [8] presented dynamic Bayesian models using both the well established Markov chain Monte Carlo (MCMC) method and the sequential Monte Carlo for target tracking and atmospheric dispersion event reconstruction problems. Chow et al. [9] and Neumann et al. [10] extended the method presented in Johannesson et al. to neighborhood scale (building-resolved) atmospheric dispersion events. Chow et al. used computational fluid dynamics (CFD) models, while Neumann et al. used computationally less intensive empirical Gaussian puff models.

With high-fidelity models, the longer simulation times needed for event re-constructions can limit their applications in emergency response operations. Marzouk et al. [11] reformulated the Bayesian approach to inverse problems by using polynomial chaos (PC) expansions to represent random variables, which then yields a spectral representation of the stochastic forward model. The integrals in the Bayesian formulation are computed via sampling from the PC expansions in a computational fast fashion. In their study, a simple transient diffusion problem was considered. The results have shown that significant gains in computational time can be obtained by adopting the new scheme over direct sampling.

Senocak et al. [12] developed a stochastic event reconstruction tool (SERT) and validated it against both real and synthetic experiments. SERT uses a Bayesian inference algorithm with MCMC sampling for source term estimation. A unique feature of SERT is that it enhances the forward plume model with a data-driven approach whereby empirical turbulence diffusion parameters are estimated as part of the inverse problem in addition to characterizing the dispersion event. The practice leads to substantial improvement over the empirically tuned plume model, Furthermore, Senocak et al. incorporates zero sensors into the probability model in a principled fashion. In most studies zero sensors are either discarded or a negligible small number is assumed arbitrarily. From an operational point of view, treatment of zero sensors can be important because even though trace CB agents below the limit of detection cannot be measured reliably by the sensor, they can be harmful to the population.

In the following, the event reconstruction method of Senocak et al. [12] is described briefly. The method is then applied to the FFT-07 database and results from two trial cases are presented.

DESCRIPTION OF THE EVENT RECONSTRUCTION METHOD

The forward modeling problem can be defined as predicting the response of a system using a physical theory and system parameters. In the inverse modeling problem, an inference is made on the values of system parameters based on observations of the system response [13]. Generally speaking, inverse problems can be formulated as follows:

$$\mathbf{m} \approx F^{-1}(\mathbf{d}), \tag{1}$$

where **d** is a vector of observations, **m** is a vector of forward model parameters, and the operator F is the forward model that governs the system response. Inverse problems are often ill conditioned, because small changes in **d** can lead to large changes in **m**. The present event reconstruction problem requires estimating the model parameters **m** (e.g. release location, emission rate, wind direction etc.) given the observed concentrations **d** from a sensor network.

The Gaussian plume model is used to calculate C_m, the concentration or dosage values at each sensor. For uniform steady wind conditions, this model can be written as follows:

$$C_m(x,y,z) = \frac{Q}{2\pi U \sigma_y \sigma_z} \exp\left(-\frac{y^2}{2\sigma_y^2}\right)\left\{\exp\left(-\frac{(z-H)^2}{2\sigma_z^2}\right) + \exp\left(-\frac{(z+H)^2}{2\sigma_z^2}\right)\right\}. \tag{2}$$

Model parameters include: Q-emission rate or source strength, U-mean wind speed, and H-height of the release. A unique approach that is adopted in SERT is that the standard deviation in the horizontal crosswind and vertical directions, σ_y and σ_z respectively, are estimated as part of the inverse problem. Correlations for standard deviations [14] are enhanced with stochastic parameters as follows:

$$\begin{aligned} \sigma_y &= \varsigma_1 x (1 + 0.0004x)^{-0.5}, \\ \sigma_z &= \varsigma_2 x, \end{aligned} \tag{3}$$

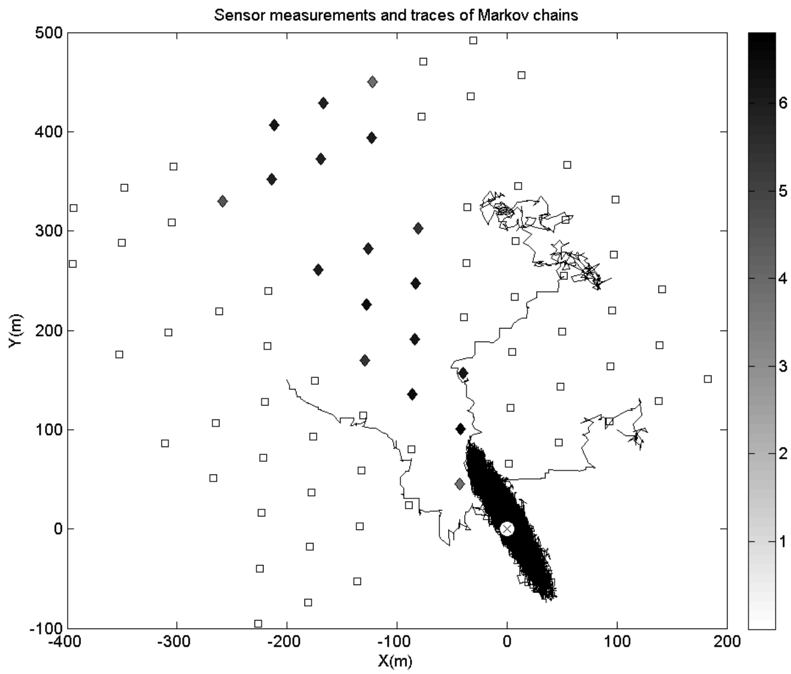

FIGURE 1. Application of SERT model to FFT-07 trial-7 data with continuous dissemination of a tracer gas from a single source. Colored diamond markers indicate concentration measurement by the sensor positive and blank square markers indicate negative (zero) sensor measurements. Traces of three Markov chains are shown. The actual source location is located at (0,0).

where ζ_1, ζ_2 are the stochastic parameters estimated in the inverse solution and x is the distance from the source location along the wind direction. In most air pollution studies these parameters are defined empirically, thus the forward model with standard deviations found in the inversion can be referred to as a data-driven plume model.

In SERT, a Bayesian approach was used to find the inverse. In this approach, a probabilistic inference is made on the forward model parameters **m**, given the observed data **d**. The goal is to find the posterior probability density of the parameters given the data, i.e. to calculate $p(\mathbf{m}\,|\,\mathbf{d})$. This is computationally intensive; therefore, Markov Chain Monte Carlo (MCMC) sampling technique is used to estimate the properties of the posterior probability density by noting the following

$$p(\mathbf{m}\,|\,\mathbf{d}) \propto L(\mathbf{d}\,|\,\mathbf{m})\,p(\mathbf{m}). \tag{4}$$

The observed data **d** enters the Bayesian formulation through the likelihood function $L(\mathbf{d}\,|\,\mathbf{m})$ while $p(\mathbf{m})$ is a prior inference on the model parameters [15].

The likelihood function is formulated by denoting the concentration measured by an ideal sensor i as ξ_i and the concentration observed by an actual sensor by d_i where they are related by:

$$d_i = \begin{cases} 0, & \text{with probability} \exp(-\alpha C_m) \\ \xi_i, & \text{with probability}\; 1-\exp(-\alpha C_m) \end{cases}. \tag{5}$$

It is assumed that the given parameters **m**, the probability ξ_i is observed follows a lognormal distribution with density

$$p(\xi_i \,|\, \mathbf{m}) = \frac{1}{\sqrt{2\pi}\sigma\xi_i} \exp\left(\frac{1}{2\sigma^2} \left(\log \xi_i - \log C_m \right)^2 \right), \tag{6}$$

where σ^2 is the variance representing cumulative measurement and dispersion model errors. Based on these assumptions, the likelihood function for a single datum d_i can then be written as follows:

$$L(d_i \,|\, \mathbf{m}) = \left(1 - \mathrm{sgn}(d_i)\right) \cdot \exp(-\alpha C_m) + \mathrm{sgn}(d_i) \cdot \left(1 - \exp(-\alpha C_m)\right) \cdot \exp\left(\frac{1}{2\sigma^2} \left(\log \xi_i - \log C_m \right)^2 \right). \tag{7}$$

Given sensor data d_i, forward model predictions C_m and prior distributions, the Metropolis algorithm [15,16] was used for MCMC sampling in the SERT algorithm. Further details on SERT can be found in Senocak et al. [12].

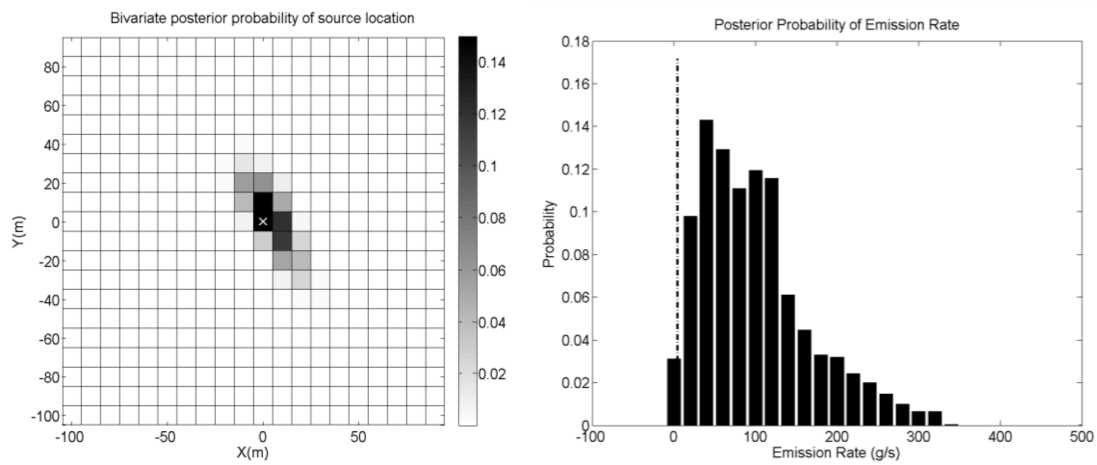

FIGURE 2. Application of SERT model to FFT-07 trial-7 data with continuous dissemination of a tracer gas from a single source. Left: Probability map for source location. The actual source is located at (0,0). Right: Marginal probability distribution of emission rate. The actual emission rate is represented by the vertical dashed line.

RESULTS

Figure 1 presents the reconstruction results for Trial-7 of the FFT-07 database. Sensor locations are indicated by markers. The sensors were deployed as a square matrix. All of the zero sensors were retained in the computations, but only those sensors that were deemed unreliable (e.g. -999 values for measurements) by the FFT-07 organizers were excluded from the matrix. Traces of three Markov chain relative to all the sensors used in the calculations are shown in Figure 1. After a short burn-in time, the Markov chains sample from the vicinity of the true source location. All Markov chains starting from different locations are able to sample from the true source location indicate that the event reconstruction algorithm does not depend on initialization of the chain. It should be noted that the Metropolis algorithm is very efficient in avoiding the local peaks. Additionally, the burn-in time for MCMC chains are typically short.

In Figure 2, the plot on the left shows the probability map for the source location. The true source location is at (0,0) and the calculated probability region for the source is shaded around it. The highest probability region coincides with the true source location meaning that the actual source location is reconstructed with proximity of less than 10 m. Such probabilistic results are very useful to first-responders, because in emergency situations first responders are interested in zones instead of coordinates of a point.

In Figure 2, the plot on the right presents the calculated marginal probability distribution of the emission rate, where the true emission rate is indicated with a dashed line. The peak of the calculated distribution is reasonably close to the true value, however, this is difficult to estimate with the Gaussian plume model, because the emission rate (Q) is jointly estimated with wind speed (U) and these two quantities appear as a ratio in the model.

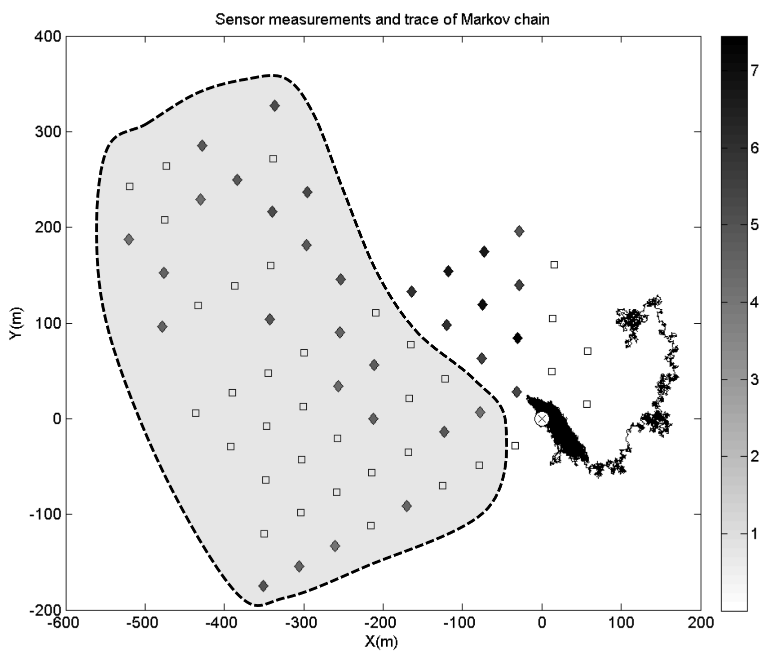

FIGURE 3. Application of SERT model to FFT-07 trial-15 data with continuous dissemination of a tracer gas from a single source. Colored square markers indicate concentration measurement by the sensor positive and blank markers indicate negative (zero) sensor measurements. The trace of the Markov chain is also shown. The actual source is located at (0,0). Sensor measurements of the dispersed plume inside the shaded area are discarded in the calculations, because the forward model is not a suitable plume model for such cases.

Event reconstruction results from Trial-15 of the FFT-07 database are shown in Figures 3 and 4. This particular case was specifically selected because the plume, as can be observed from Figure 3, is dispersed on the field. However, a visual examination of the sensor observations indicates a steady Gaussian plume pattern in the data. The SERT model was not successful in reconstructing the actual release when dispersed sensor information was taken into account. This is not surprising because the Gaussian plume model, which is currently the only forward model in SERT, is not a suitable model for highly dispersed plumes. On the other hand, the non-Gaussian dispersion of the plume, as observed from Figure 3, appears to have taken place during a short period of time. Therefore, when the sensor observations enveloped by the dashed loop in Figure 3 were excluded from the data set, SERT was successful in reconstructing the release. The present approach is definitely ad-hoc, but it highlights the importance of the forward model in event reconstruction. A time-dependent puff model with spatial varying wind field should serve as a better forward model to simulate the whole release in this particular case.

Similar to Figure 2, Figure 4 presents the probability map for the source location and the marginal probability distribution of the emission rate. Note that those sensor observations that correspond to the dispersed plume were ignored. The maximum distance of the probability region from the true source location is approximately 10 m, which is a very good reconstruction for emergency response purposes. Marginal probability distribution of the emission rate shows a wider range of probably values but the maximum in the distribution is in good agreement with the actual emission rate.

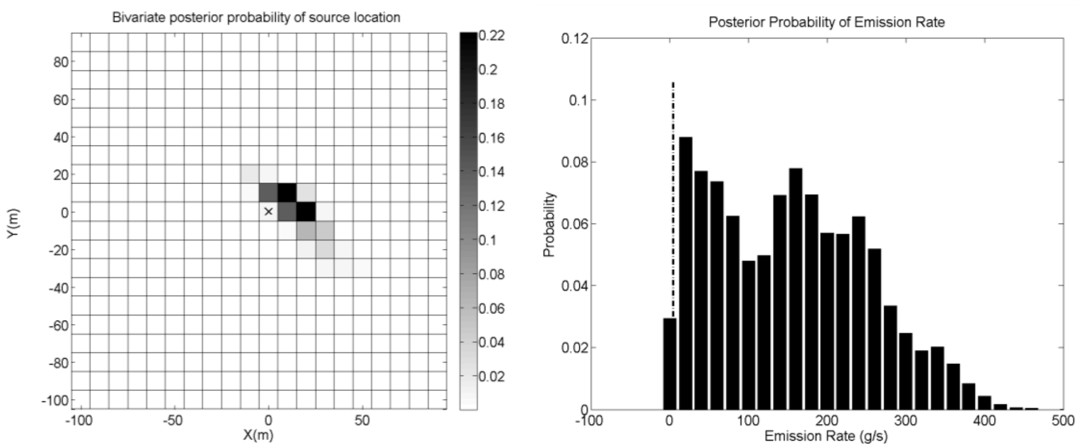

FIGURE 4. Application of SERT model to FFT-07 trial-15 data with continuous dissemination of a tracer gas from a single source. Left: Probability map for source location. The actual source location is located at (0,0). Right: Marginal probability distribution of emission rate. The actual emission rate is represented by the dashed vertical line.

CONCLUSIONS

The Stochastic event reconstruction tool (SERT) [12] is further tested against the FFT-07 trial data. Unlike previous test cases, FFT-07 trial data represents short range atmospheric releases. Results have indicated that the SERT model is able to reconstruct dispersion events that appear to be steady on average. Estimates of source release location, based on maximum probability region were found to be within approximately 10 m of the true locations. Results significantly benefits from the data-driven approach [12] adopted in the Gaussian plume model. For highly dispersed plumes, the constant wind Gaussian plume based forward model in SERT is not a suitable model. However, SERT performed well when observations that detect the dispersed plume were ignored in an ad-hoc fashion. Future work will focus on incorporating additional forward models into SERT to reconstruct different release scenarios.

REFERENCES

1. Government Accountability Office, "Homeland security: First Responders' Ability to Detect and Model Hazardous Releases in Urban Areas Is Significantly Limited," Report No: GAO-08-180, 2008
2. Shea, D.A., and Lister, S.A., "The BioWatch Program: Detection of Bioterrorism," Congressional Research Service Report No. RL 32152, 2003
3. N. Platt and D. Deriggi, "Comparative Investigation Of Source Term Estimation Algorithms Using Fusion Field Trial 2007 Data," 8th Conference on Artificial Intelligence Applications to Environmental Sciences at AMS Annual Meeting, Atlanta, GA, Jan 17-21, 2010
4. D. P. Storwold, "Detailed Test Plan for the Fusing Sensor Information from Observing Networks (FUSION) Field Trial 2007 (FFT-07)," Meteorology Division, West Desert Test Center, U.S. Army Dugway Proving Ground, September 2007
5. W. R. Hogan, G. F. Cooper, M. M. Wagner, and G. L. Wallstrom, "An Inverted Gaussian Plume Model for Estimating the Location and Amount of Release of Airborne Agents from Downwind Atmospheric Concentrations," Technical Report, University of Pittsburgh, 2004
6. V. Akcelik, G. Biros, O. Ghattas, K. R. Long, and B. B. Waanders, "A Variational Finite Element Method for Source Inversion for Convective Diffusive Transport," *Finite Elements in Analysis & Design*, **39**, 683-705, (2003)
7. C. T. Allen and S. E. Haupt, "Source Characterization With A Genetic Algorithm-Coupled Dispersion-Backward Model Incorporating SCIPUFF," *J. Appl. Meteorol. Climatol*, **41**, 465–479, (2007)
8. G. Johannesson, B. Hanley and J. Nitao, "Dynamic Bayesian models via Monte Carlo-An Introduction with Examples," Lawrence Livermore National Laboratory, UCRL-TR-207173, 2004
9. F. K. Chow, B. Kosovic and S.T. Chan, "Source inversion for contaminant plume dispersion in urban environments using building-resolving simulations," *J. Appl. Meteorol. Climatol*, **47**, 1553-1572, (2008)
10. S. Neumann, L. Glascoe, B. Kosovic, K. Dyer, W. Hanley and J. Nitao, "Event Reconstruction For Atmospheric Releases Employing Urban Puff Model UDM With Stochastic Inversion Methodology," 6th Symposium on the Urban Environment, American Meteorological Society, GA, 2006
11. Y. M. Marzouk, H. N. Najm, L. A. Rahn, "Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems," *J. Comput. Phys*, **224**, 560–586, (2007)
12. I. Senocak, N. Hengartner, M. B. Short and B. Daniel, "Stochastic Event Reconstruction of Atmospheric Contaminant Dispersion using Bayesian Inference," *Atmos. Environ.*, **42**, 7718–7727, (2008)
13. A. Tarantola, *Inverse Problems Theory and Methods for Model Parameter Estimation*, SIAM, 2005, pp. 342.
14. H. A. Panofsky and J. A. Dutton, *Atmospheric Turbulence*, John Wiley, 1984
15. W. R. Gilks, S. Richardson and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 1996.
16. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.* **21**, 1087–1092 (1953)