

Sources Detection and Parameters Estimation of Plume Model Based on Sensor Network Measurements

Chunfeng Huang^{aa}, Tailen Hsing^b, Noel Cressie^c, Auroop R. Ganguly^d,
Vladimir A. Protopopescu^d and Nageswara S. Rao^d

^a*Department of Statistics, Indiana University, Bloomington, IN;* ^b*Department of Statistics, University of Michigan, Ann Arbor, MI;* ^c*Department of Statistics, The Ohio State University, Columbus, OH;* ^d*Oak Ridge National Laboratory, Oak Ridge, TN*

November 9, 2008

Technical Report 08-07
Department of Statistics
Indiana University
Bloomington, IN 47405

^aCorresponding author. Email: huang48@indiana.edu. A preliminary version of this work was presented at 9th ONR/GTRI Workshop on Target Tracking in Sensor Fusion, 2006.

Sources Detection and Parameters Estimation of Plume Model Based on Sensor Network Measurements

Abstract. We consider a network of sensors that measure the intensities of a complex plume composed of multiple absorption-diffusion source components. We address the problem of estimating the plume parameters, including the spatial and temporal source origins and the parameters of the diffusion model for each source, based on a sequence of sensor measurements with additive sensor-measurement errors. The approach not only leads to multiple source detection, but also the characterization and prediction of the combined plume in space and time. The parameter estimation is formulated as a Bayesian inference problem and the solution is obtained using a version of the Markov chain Monte Carlo algorithm. The approach is applied to a simulation study, which shows that accurate parameter estimation is achievable.

Key Words: Bayesian statistics; Markov chain Monte Carlo; partial differential equation; plume model; sensor networks

1. Introduction

Advances in monitoring and communication technologies enable us to address sensitive issues such as the environment and security using sensor networks as evidenced by their deployments in a wide variety of contexts [1], [2]. While the specific configuration of any particular sensor network depends on the problem context, sensor networks typically generate potentially complex and unstructured datasets. These datasets hold keys to understanding the phenomena that the networks monitor, and our ability to analyze and extract useful information crucially determines the effectiveness of the sensor networks.

After the initial explosion/release at a certain time and location, the results of a dirty bomb, a chemical leak, or the release of biological agents take the form of an atmospheric plume that will undergo transport, advection, diffusion, absorption, adsorption, radioactive decay, and delayed re-emission. To detect, study, identify and/or track the plume, we must study its spatio-temporal evolution. In this article we propose a hierarchical Bayesian statistical approach that relies on both physical laws and statistical models in space and time.

Spatially dispersed networks of sensors hold an enormous potential for the detection, identification, tracking and prediction of these phenomena. In particular, the problem of detecting and tracking sources was addressed using sensor networks for chemical dispersions in [3] and [4], and nuclear radiation in [5]. However, there has been limited effort on incorporating the analytical and statistical nature of the plume and sensor models in analyzing the measurements collected by such networks. The paper [6] employed a linear array of detectors to detect a moving radiation source against background radiation. A sensor-network solution for detecting a moving radiological dispersion device was presented using a Bayesian formulation by [5]. Cost benefit analysis of utilizing sensor networks for detecting moving radiation sources was carried out by [7]. These works mainly focus on detecting low-level radiations emitted by sources prior to explosion, and consequently do not explicitly consider the dispersion dynamics.

This paper is structured as follows. Section 2 gives a technical formulation of the problem, and a PDE plume model is introduced in Section 3. Section 4 discusses the Bayesian inferential approach for this model. Section 5 contains a simulation study, and discussion and conclusions are given in Section 6.

2. Formulation of the problem

The objective of this paper is to introduce a statistical approach to systematically analyze and estimate the plume parameters based on sensor measurements. The establishment of such networks is expected to accelerate, and the roles that they play in addressing significant issues will continue to expand in the near future. We consider the following scenario of K pollution sources released into the environment at different times t_{01}, \dots, t_{0K} and spatial locations $\mathbf{s}_{01}, \dots, \mathbf{s}_{0K}$, for some $K > 1$. For simplicity, we assume that the locations are points on the plane. Extensions to higher-dimensional spaces are straightforward. Let

$$u_k(x, y, t) \equiv \text{true plume intensity at time } t \text{ and location } (x, y) \quad (1)$$

due to the k -th source;

clearly $u_k(x, y, t) = 0$ if $t < t_{0k}$. Assume that plume intensities are additive so that the true total plume intensity at location (x, y) and time t is

$$\sum_{k=1}^K u_k(x, y, t).$$

Suppose that sensors are placed at locations $(x_i, y_i), i = 1, 2, \dots, n_s$, and the total plume intensity is observed at times $t_{i,j}, j = 1, 2, \dots, n_i$, for sensor i . The plume intensity is measured by the sensors; we assume that there is an additive random error $\varepsilon_{i,j}$ associated with each measurement so that we observe

$$z_{i,j} = \sum_{k=1}^K u_k(x_i, y_i, t_{i,j}) + \varepsilon_{i,j}, \quad (2)$$

where the $\varepsilon_{i,j}$ may have a nontrivial dependence structure in practice.

An example of u_k is the simple product-form model in [8] and [9]:

$$u_k(x, y, t) = \begin{cases} b e^{-[(x-x_{0k})^2 + (y-y_{0k})^2]^{1/2}} & t \in [t_{0k}, t_{0k} + T] \\ b e^{a(t-t_{0k}-T) - [(x-x_{0k})^2 + (y-y_{0k})^2]^{1/2}} & t > t_{0k} + T, \end{cases}$$

where a, b, T are unknown parameters. Other plume models considered in the literature can be found in [3], [4], [10] and [11]. The product-form model provides a convenient analytic formulation of the spatio-temporal evolution of the plume, but it does not account for propagation or drift due to exogenous factors such as wind.

Indeed, within this model the consequence of the explosion is instantaneous in the whole space, followed by uniform absorption. On the other hand, one could attempt to simulate each particle and track them through space and time, the sensor measurements providing constants on the simulation. An intermediary step between the extremely simple product model and the realistic, but often cumbersome numerical simulations, is given by analytic PDE solutions, adapted to a statistical modeling framework. These provide an excellent middle ground for evaluating statistical methods, carrying out parameter identification, and refining sensor-fusion techniques.

The main purpose of this paper is to consider the statistical inference of such a PDE plume model based on the spatio-temporal data $z_{i,j}$ described in (2). Due to the large number of unknown parameters, a Bayesian perspective will be adopted and the posterior distribution will be obtained from Markov chain Monte Carlo simulations. This distribution can be used not only to estimate the plume parameters but also to derive their qualitative properties as related to measurements. As mentioned in Section 1, the Bayesian approach has been utilized in detecting moving radiation sources (characterized by simple diffusions) using sensor networks [5], [12] and [6].

3. A PDE Plume model

The most important features of the physical phenomena described in Section 1 are well captured by an absorption-drift-diffusion model. A two-dimensional isotropic, homogeneous version of the model is given by:

$$\frac{\partial u}{\partial t} = -bu + v_1 \frac{\partial u}{\partial x} + v_2 \frac{\partial u}{\partial y} + c^2 \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right), \quad (3)$$

where $u(x, y, t)$ models the plume intensity at the point (x, y) at time t , b is the absorption coefficient, v_1 and v_2 are the components of the (constant) advection velocity, and c is the diffusivity (diffusion coefficient), which corresponds to an isotropic medium. This model assumes that propagation takes place essentially at the surface; the third dimension can also be accounted for in a straightforward manner.

We assume that the medium is infinite (the free atmosphere is a reasonably good realization of this ideal situation) and the initial condition is:

$$u(x, y, t = t_0) = C\delta(x - x_0)\delta(y - y_0), \quad (4)$$

where $\delta(\cdot)$ is the Dirac delta function, which corresponds to a very sharply localized release of intensity C at time $t = t_0$, at the point (x_0, y_0) . Then the solution to (3)

and (4) can be calculated exactly as:

$$u(x, y, t) = \frac{a}{t - t_0} \exp \left\{ -b(t - t_0) - \frac{(x - v_1 t - x_0)^2 + (y - v_2 t - y_0)^2}{4c^2(t - t_0)} \right\}, \quad (5)$$

where a is a normalization constant depending on the intensity of the initial source, C , and on the diffusion constant, c . Formula (5), which represents the Green's function of equation (3) in an infinite medium can be readily generalized to higher dimensions and various anisotropic situations. Notice that it is not of product (time \times space) form.

Propagation in inhomogeneous and/or non-stationary media (i.e., accounting for the presence of buildings, mountains, forests, changing meteorological conditions, etc.) can be described also, by replacing the constant coefficients with appropriate functions of time and position and/or supplementing the evolution equation with appropriate boundary conditions. While in general these situations do not lend themselves to simple analytical solutions [13], numerical approaches are always possible and have been pursued to a great extent [14].

The advantages of the class of models illustrated by equation (3) include (i) their general mathematical properties are very well understood [15] and [13]; (ii) the evolution described by the continuous versions preserves the required physical properties such as positivity and conservation laws [15] and [13]; (iii) due to the parabolic character of the equation, the discretized versions are stable; and (iv) boundary conditions can be accounted for in a systematic and consistent manner.

4. Methodology

From now on we shall focus on the model (5). We assume here that the total number of plume sources K is either known or can be determined from data. An approach to determine K from data will be discussed in Section 5. To simplify notation, write

$$\begin{aligned} & u(x, y, t; \boldsymbol{\tau}_k) \\ = & I(t > t_{0k}) \frac{a_k}{t - t_{0k}} \exp \left\{ -b_k(t - t_{0k}) - \frac{(x - v_1 t - x_{0k})^2 + (y - v_2 t - y_{0k})^2}{c_k(t - t_{0k})} \right\}, \end{aligned}$$

where $\boldsymbol{\tau}_k = (a_k, b_k, c_k, v_{1k}, v_{2k}, t_{0k}, x_{0k}, y_{0k})$ stands for a vector of possible parameter values for the k -th plume source. Let $\tilde{\boldsymbol{\tau}}_k \equiv (\tilde{a}_k, \tilde{b}_k, \tilde{c}_k, \tilde{v}_{1k}, \tilde{v}_{2k}, \tilde{t}_{0k}, \tilde{x}_{0k}, \tilde{y}_{0k})$ denote the corresponding true parameter values. Observe that the true intensity function $u_k(x, y, t)$ defined in (1) can now be written as $u(x, y, t; \tilde{\boldsymbol{\tau}}_k)$. Assume that the errors

$\varepsilon_{i,j}$ in (2) are independent and normally distributed with mean 0 and true variance $\tilde{\sigma}^2$. Thus, the combined parameter of the K plume sources is

$$\boldsymbol{\theta} \equiv (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \dots, \boldsymbol{\tau}_K, \sigma^2),$$

which is a row vector of length $8K + 1$. Also, denote by \mathbf{z} the data $\{z_{i,j}, 1 \leq j \leq n_i, 1 \leq i \leq n_s\}$ defined in (2). Then the likelihood of $\boldsymbol{\theta}$ is

$$p(\mathbf{z}|\boldsymbol{\theta}) \propto (\sigma^2)^{-N/2} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n_s} \sum_{j=1}^{n_i} \left\{ z_{i,j} - \sum_{k=1}^K u(x_i, y_i, t_{i,j}; \boldsymbol{\tau}_k) \right\}^2 \right], \quad (6)$$

where $N = \sum_{i=1}^{n_s} n_i$. The high dimensionality of $\boldsymbol{\theta}$ makes the Bayesian-inference approach attractive.

The Bayesian approach views the true parameter $\boldsymbol{\theta}$ as a quantity sampled from a probability distribution $p(\boldsymbol{\theta})$, called the prior distribution. The prior distribution describes how likely the possible values of $\boldsymbol{\theta}$ are by taking into account all the information available prior to seeing the data. For instance, if the plume sources are dirty bombs then highly populated areas may be more likely to contain the bombs source locations than lowly populated areas, in which case it makes sense for the prior distribution to reflect this; also, certain values of a_k, b_k, c_k may be more likely than others, as determined by the physical conditions of the environment. However, if no prior information is available, or if one chooses to ignore prior information, then a uniform distribution on $\boldsymbol{\theta}$, often called non-informative prior, could be used.

Bayesian inference focuses on the properties of the posterior distribution $p(\boldsymbol{\theta}|\mathbf{z})$, namely, the conditional distribution of $\boldsymbol{\theta}$ given the data \mathbf{z} . The posterior distribution updates what was known about $\boldsymbol{\theta}$, as contained in the prior distribution, by incorporating the information in the data \mathbf{z} . Once $p(\boldsymbol{\theta}|\mathbf{z})$ is available, inference based on it can be made with regard to the true values of the model parameters from which the data were generated. Also, overall plume intensity at location (x, y) and time t can be predicted using its posterior mean, namely

$$\sum_{k=1}^K \int_{\boldsymbol{\theta}} u(x, y, t; \boldsymbol{\tau}_k) p(\boldsymbol{\theta}|\mathbf{z}) d\boldsymbol{\theta}. \quad (7)$$

The posterior distribution can sometimes be derived in closed form but often has to be computed numerically. In this paper, the posterior distribution will be computed

by Markov Chain Monte Carlo (MCMC) simulations. MCMC methods are widely used in statistics and there are many different approaches. We illustrate how it works with a combination of the Gibbs sampler and the Metropolis algorithm.

For any component θ_i in $\boldsymbol{\theta}$, let $\boldsymbol{\theta}_{[-i]}$ be the vector $\boldsymbol{\theta}$ with θ_i removed, and accordingly let $p(\cdot|\boldsymbol{\theta}_{[-i]}, \mathbf{z})$ be the conditional distribution of θ_i given \mathbf{z} and the other components in $\boldsymbol{\theta}$. The basic Gibbs sampler works as follows.

1. First choose a starting value $\boldsymbol{\theta}$.
2. For each θ_i in $\boldsymbol{\theta}$, generate a random value from $p(\cdot|\boldsymbol{\theta}_{[-i]}, \mathbf{z})$ and update $\boldsymbol{\theta}$ by replacing θ_i with the value generated.
3. Iterate 2 until the empirical distribution of the sampled values of $\boldsymbol{\theta}$ converges.

Theoretically, $p(\boldsymbol{\theta}|\mathbf{z})$ is the limiting distribution, which is estimated with the empirical distribution of sampled values. How many iterations to perform in the Gibbs sampler is, of course, crucial. There is a large volume of literature that addresses this issue. With modern computing capabilities, the number of iterations performed are usually in the range of hundreds of thousands. See [16] for an overview on MCMC and discussions of important issues.

Since

$$p(\theta_i|\boldsymbol{\theta}_{[-i]}, \mathbf{z}) = \frac{p(\boldsymbol{\theta}, \mathbf{z})}{p(\boldsymbol{\theta}_{[-i]}, \mathbf{z})} = \frac{p(\mathbf{z}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int_{\theta_i} p(\mathbf{z}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\theta_i},$$

any normalizing constants will cancel and hence sampling from $p(\cdot|\boldsymbol{\theta}_{[-i]}, \mathbf{z})$ can be easy. However, on occasions, this step may be numerically challenging. One alternative is to use the Metropolis algorithm, explained as follows. Let $J(\theta|\theta')$ be a symmetric bivariate function in θ, θ' and such that $J(\cdot|\theta)$ is a probability distribution for each θ . In step 2 above, whenever sampling from $p(\cdot|\boldsymbol{\theta}_{[-i]}, \mathbf{z})$ is difficult to perform directly, the following steps will be performed instead:

2'(a) Sample a proposal θ_i^* from $J(\cdot|\theta_i)$.

2'(b) Calculate the ratio

$$r = \frac{p(\theta_i^*, \boldsymbol{\theta}_{[-i]}|\mathbf{z})}{p(\boldsymbol{\theta}|\mathbf{z})}.$$

2'(c) Sample u from the uniform $[0, 1]$ distribution; if $u > r$ then replace θ_i by θ_i^* , otherwise keep θ_i .

It can be shown that the combined Gibbs-Metropolis algorithm produces a limiting distribution equal to true posterior distribution under very general conditions. See [16].

5. Numerical results

In this section, we present some numerical results for the statistical inference on θ from the PDE plume model. We consider two scenarios, of having $K = 2$ and $K = 3$ plume sources. For convenience, we shall make the simplification that, for the advection velocities, v_2 is equal to 0, and $v_1 = v$ is the same for all sensors. This assumption depicts the real-world situation where the wind is blowing at more or less constant velocity in the direction of the x -axis in the area of the plume. Note that this direction can be replaced by any other known direction, and that we do not assume the wind speed is known. The locations of the plume sources (unknown) and sensors (known) in our examples are given by the configurations in Figure 1, where sensor locations are denoted by “ \circ ” and source locations are also denoted by “ \times ”.

[Figure 1 about here]

First we consider an example for two sources (i.e., $K = 2$). Assume that the true PDE plume model is specified by the parameters in Table 1.

Table 1: Parameters of simulated two-source model

	\tilde{a}_k	\tilde{b}_k	\tilde{c}_k	\tilde{v}_{1k}	\tilde{v}_{2k}	\tilde{t}_{0k}	$(\tilde{x}_{0k}, \tilde{y}_{0k})$
$k = 1$	3.0	1.0	1.0	0.6	0.0	0.4	$(-0.2, 0.0)$
$k = 2$	3.0	1.0	1.0	0.6	0.0	0.5	$(0.1, 0.0)$

The data $\mathbf{z} = \{z_{i,j}\}$ are observed at 20 equally-spaced time points from 0 to 2 at each of the 9 sensor locations shown in Figure 1. In the simulations, we assumed that the distribution of the measurement error $\varepsilon_{i,j}$ is normal with mean 0 and standard deviation $\tilde{\sigma} = .05$. We first conducted a single simulation run; see Figure 2, which shows the data $z_{i,j}$ observed at the 9 sensors at 20 time points.

[Figure 2 about here]

Based on the data, we conducted statistical inference using the Bayesian approach described in the previous section. We assumed that the prior distributions for each of the parameters were mutually independent, where non-informative priors were used for a_i, b_i, c_i, t_i, v , and an inverse χ^2 distribution with scale parameter .01 and degrees of freedom 1 was used for σ^2 . All of the parameters except σ^2 were updated using the Metropolis algorithm. MCMC simulations were implemented with 100,000 iterations, where the last 50,000 simulation results were used to compute the empirical posterior distributions. The distributions so obtained are summarized in the plots in Figure 3, where the bar charts are relative frequency histograms and the superimposed smooth curves are the corresponding density estimates obtained by kernel smoothing. It can be observed that the estimated posterior distributions are fairly tight and include the true parameters in their ranges.

[Figure 3 about here]

We then used (7) to predict the true plume intensity at time $t = 2.5$ over the spatial region $[-1, 1] \times [-1, 1]$. Note that this time point and a portion of the spatial region are beyond the coverage of the sensors as reflected by the data \mathbf{z} . The true levels are displayed in the left plot of Figure 4, and the predicted levels are displayed in the right plot of Figure 4. Overall, the plume levels are predicted quite well over the entire region.

[Figure 4 about here]

Next we conducted 50 simulation runs, that is, we simulated the data $z_{i,j}, 1 \leq i \leq 20, 1 \leq j \leq 9$, fifty independent times, and for each run we went through the MCMC simulations just described to obtain the posterior distributions. For each parameter θ_ℓ , we computed the posterior mean $\int \theta_\ell p(\boldsymbol{\theta}|\mathbf{z}) d\boldsymbol{\theta}$. Consequently, we obtained 50 posterior means from the 50 simulations. The posterior mean θ_ℓ are presented in Figure 5. It can be seen that the posterior means are all in a tight range of the corresponding true values.

[Figure 5 about here]

Next we considered an example for three plume sources. Assume that the true PDE plume model is specified by the parameters in Table 2.

The data $z_{i,j}$ are observed at 40 equally-spaced time points from 0 to 2.5 at each of

Table 2: *Parameters of simulated three-source model*

	\tilde{a}_k	\tilde{b}_k	\tilde{c}_k	\tilde{v}_{1k}	\tilde{v}_{2k}	\tilde{t}_{0k}	$(\tilde{x}_{0k}, \tilde{y}_{0k})$
$k = 1$	3.0	1.0	1.0	0.6	0.0	0.4	$(-0.2, 0.1)$
$k = 2$	3.0	1.0	1.0	0.6	0.0	0.6	$(-0.3, 0.0)$
$k = 3$	3.0	1.0	1.0	0.6	0.0	0.8	$(-0.1, 0.0)$

the 25 sensor locations as described by the right plot in Figure 1. We again assumed that the distribution of the measurement error $\varepsilon_{i,j}$ is normal with mean 0 and standard deviation $\tilde{\sigma} = .05$. The estimated posterior distributions for the parameters obtained by MCMC simulations are described by the plots in Figure 6. Observe that, even with the additional plume source, the data collected by the 25 sensors lead to much tighter posterior distributions than those seen for the two-source example in Figure 3 using only 9 sensors.

[Figure 6 about here]

5. Discussion

In this paper, we present a statistical approach for identifying the parameters of the PDE plume model based on sensor data. We showed that even with a moderate amount of data, the identification of the model can be achieved using a Bayesian MCMC approach. The estimated model can then be used for spatial and temporal prediction of plume level beyond the sensor range.

In the analysis above, we assumed that the true number of sources is known. In practice, the number may have to be determined from data. There are many model-selection criteria that can be used for problems of this nature, the most widely used being Akaike’s information criterion (AIC [17], [18]), Bayesian information criterion (BIC [19]), and deviance information criterion (DIC [20]). We demonstrate the use of BIC here. In general, BIC picks the optimal model from among those under consideration by minimizing

$$BIC = -2 \log(\text{max likelihood}) + p \log n,$$

where n is sample size, max likelihood is the maximum value of the likelihood function based on the assumption that the number of parameters is p . Note that in our example in Section 5, the number of parameters is equal to $6K + 2$ if the number of sources

is K . Based on the data of the two-source example in Figure 2, the BIC values are computed for $K = 1, 2, 3, 4$ and displayed in Table 3, where the maximized likelihood is approximated by plugging in the parameters' posterior means of the parameters into the likelihood function. This worked very well for the problem on hand. Clearly, BIC picked the correct model, $K = 2$.

Table 3: BIC values for $K = 1, 2, 3, 4$. The smallest BIC value is attained at $K = 2$, the true value of K .

K	1	2	3	4
BIC	196.54	-800.50	-771.44	-740.86

A number important issues require further research and some of which are discussed in the following.

- (i) Various computational issue are crucial in the problem. The MCMC simulations that we presented for the two-source example took about 90 seconds to implement on a Linux workstation (Dual Pentium 4 Xeon at 3GHz with 4GB RAM) using Fortran codes. The computational speed could be substantially improved using a more efficient computing platform. The more interesting question is to perform the computations in real time. In other words, as new plume evolution information becomes available, the estimation/prediction can be updated based on the previous calculations without restarting from scratch. One possible way to achieve this goal is to utilize sequential Monte Carlo algorithms as in [21], which requires further investigation.

- (ii) We assumed that the measurement errors $\varepsilon_{i,j}$ are independent. This assumption can be relaxed. For convenience, in (6), let \mathbf{w} denote the vector of $z_{i,j} - \sum_{k=1}^K u(x_i, y_i, t_{i,j}; \boldsymbol{\tau}_k)$, $1 \leq i \leq n_s$, $1 \leq j \leq n_i$. Then

$$p(\mathbf{z}|\boldsymbol{\theta}) \propto (\sigma^2)^{-N/2} \exp \left\{ -\frac{1}{2\sigma^2} \mathbf{w}^T \mathbf{w} \right\}.$$

Suppose that the $\varepsilon_{i,j}$ are dependent having the covariance matrix Σ . Then the likelihood function becomes

$$p(\mathbf{z}|\boldsymbol{\theta}) \propto |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} \mathbf{w}^T \Sigma^{-1} \mathbf{w} \right\},$$

where the parameters of Σ are now included in θ . Although the MCMC procedure can, in principle, be carried out as before given any Σ , in reality Σ needs to be modeled parsimoniously using only a small number of parameters. A simple example of this is to assume that the measurement errors for different sensors are independent but within each sensor the measurement errors follow an autoregressive process of order one. In this case, an extra parameter, the autoregressive coefficient, will be needed for each sensor. Ultimately, the complexity of the model will have to be balanced by the amount of data available.

- (iii) In our simulations, we assumed that the sensors are placed on a grid. This is because our trial runs revealed that this scheme of sensor placement leads to the most satisfactory statistical inference on average. An important question is if we have a fixed number of sensors with limited query capabilities, how should we design the network so as to optimize the inference. Spatial design of this type is discussed in [22].

Acknowledgements

This work was supported by the SensorNet Program at the Oak Ridge National Laboratory under UT-Battelle LLC contract 4000042084.

References

- [1] Brooks, R. R., and Iyengar, S. S. *Frontiers of Distributed Sensor Networks*, CRC Press (2004).
- [2] Culler, D., Estrin, D., and Srivastava, M. Overview of sensor networks. *IEEE Computer* (2004), 41-49.
- [3] Chen, Y., Moore, K., and Song, Z. Diffusion boundary determination and zone control via mobile actuator-sensor networks (mas-net): challenges and opportunities. In *Proceedings of SPIE: Intelligent Computing: Theory and Applications*, 5421 (2004), 102-113.
- [4] Ishida, H., Nakamoto, T., Moriizumi, T., Kikas, T., and Janata, J. Plume-tracking robots: A new application of chemical sensors. *Biological Bulletin*, 200 (2001), 222-226.

- [5] Brennan, S. M., Mielke, A. M., Torney, D. C., and Maccabe, A. B. Radiation detection with distributed sensor networks. *IEEE Computer* (2004), 57-59.
- [6] Nemzek, R. J., Dreicer, J. S., Torney, D. C., and Warnock, T. T. Distributed sensor networks for detection of mobile radioactive sources. *IEEE Trans. on Nuclear Science*, 51, 4 (2004), 1693-1700.
- [7] Stephens, D. L., and Peurrung, A. J. Detection of moving radioactive sources using sensors networks. *IEEE Trans. on Nuclear Science*, 51, 5 (2004), 2273-2278.
- [8] Rao, N. S. V. Identification of class of simple product-form plumes using sensor networks. In *Innovations and Commercial Applications of Distributed Sensor Networks Symposium* (2005a).
- [9] Rao, N. S. V. Identification of simple product-form plumes using networks of sensors with random errors. Manuscript (2005b).
- [10] Li, W., Farrell, J. A., and Carde, R. Tracking of fluid-advected odor plumes: strategies inspired by insect orientation to pheromone. *Adaptive Behavior* (2001), 143-170.
- [11] Sykes, R. I., Lewellen, W. S., and Parker, S. F. A Gaussian plume model of atmospheric dispersion based on second-order closure. *Journal of Climate and Applied Meteorology*, 25 (1986), 322-331.
- [12] Brennan, S. M., Mielke, A. M., and Torney, D. C. Radiation source detection by sensor networks. *IEEE Trans. on Nuclear Science*, 52, 3 (2005), 813-819.
- [13] Morse, P. M., and Feshbach, H. *Methods of Mathematical Physics*, McGraw Hill (1953).
- [14] Sykes, R. I. *PC-Scipuff, Version 1.3*, A.R.A.P. Report no. 72 (2000).
- [15] Greenberg, W., van der Mee, C. V. M., and Protopopescu, V. *Boundary Value Problems in Abstract Kinetic Theory*, Birkhauser (1987).
- [16] Gelman, A., Karlin, J. B., Stern, H. S., and Rubin, D. B. *Bayesian Data Analysis*, Chapman and Hall (2003).

- [17] Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (1974), 716-723.
- [18] Sakamoto, Y., Ishiguro, M., and Kitagawa, G. *Akaike Information Criterion Statistics*, D. Reidel Publishing Company (1986).
- [19] Schwarz, G. Estimating the dimension of a model. *Annals of Statistics*, 6 (1978), 461-464.
- [20] Spiegelhalter, D., Best, N. G., Carlin, B. P., and van der Linde, A. Bayesian measures of model complexity and fit. *Journal of Royal Statistical Society B*, 64 (2002), 583-639.
- [21] Wikle, C. K., and Berliner, L. M. A Bayesian tutorial for data assimilation. *Physica D* (2007), 1-16.
- [22] Cressie, N. *Statistics for Spatial Data*, Wiley (1993).

Figure Captions

Figure 1: The left/right plot contains the source locations (\times) and sensor locations (\circ) for the examples with two/three sources.

Figure 2: The plots contain the data collected by the 9 sensors.

Figure 3: Estimated posterior distributions of the parameters based on a single MCMC run for the two-source example.

Figure 4: The left plot contains the true plume levels based on the estimated posterior distributions of the parameters. The right plot contains the predicted plume levels.

Figure 5: Estimated posterior means of the parameters based on 50 MCMC runs.

Figure 6: Estimated posterior distributions of the parameters based on one MCMC run for the three-source example.

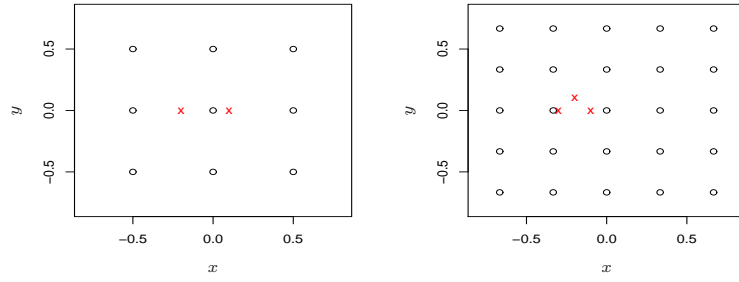


Figure 1: The left/right plot contains the source locations (\times) and sensor locations (\circ) for the examples with two/three sources.

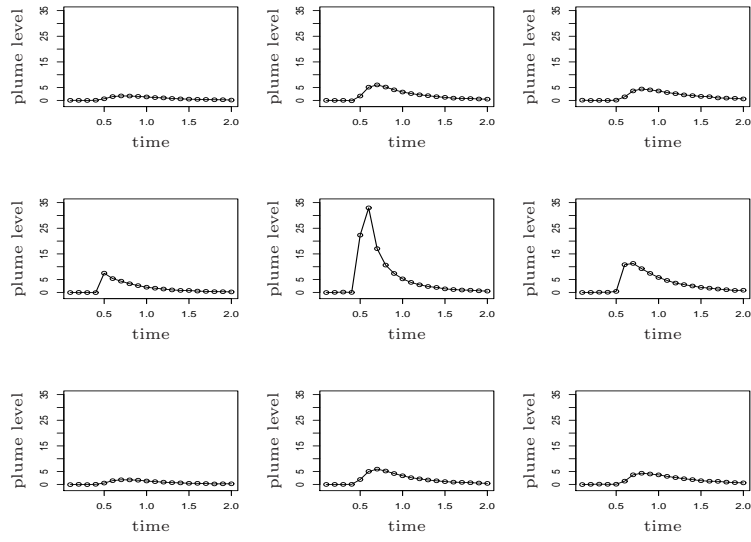


Figure 2: The plots contain the data collected by the 9 sensors.

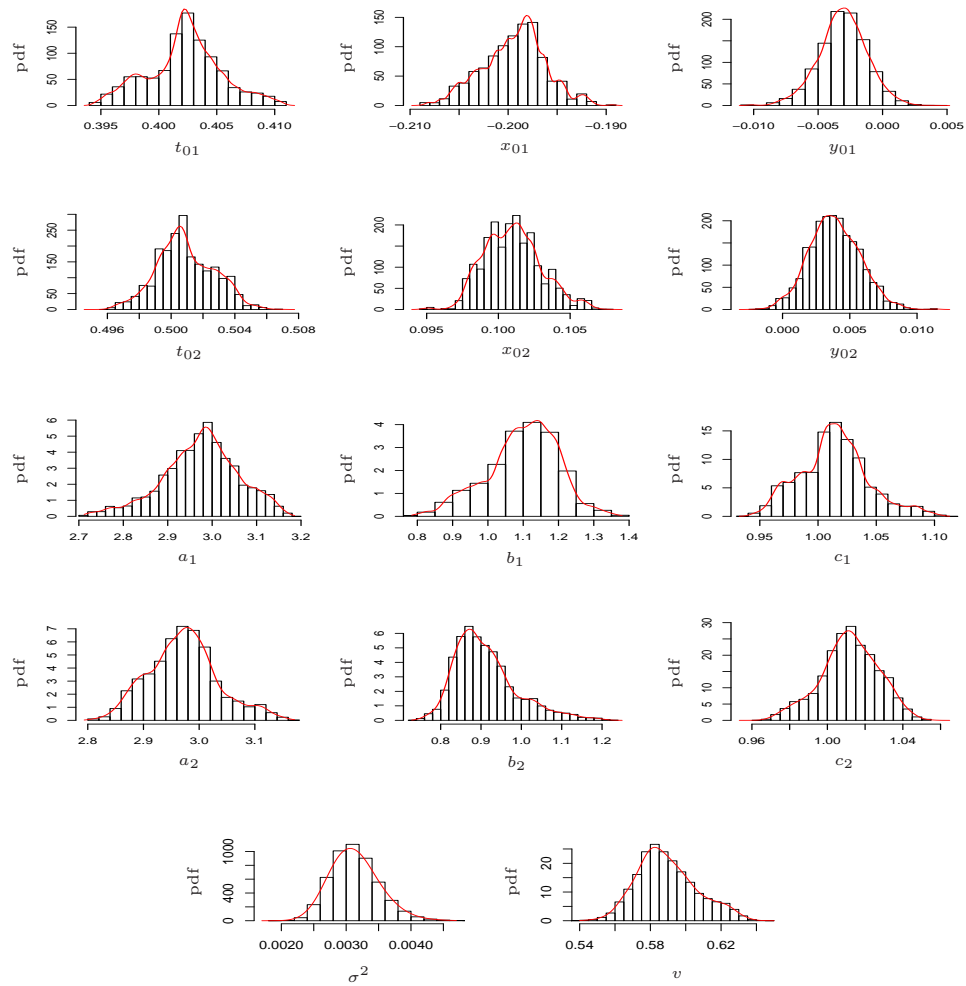


Figure 3: Estimated posterior distributions of the parameters based on a single MCMC run for the two-source example.

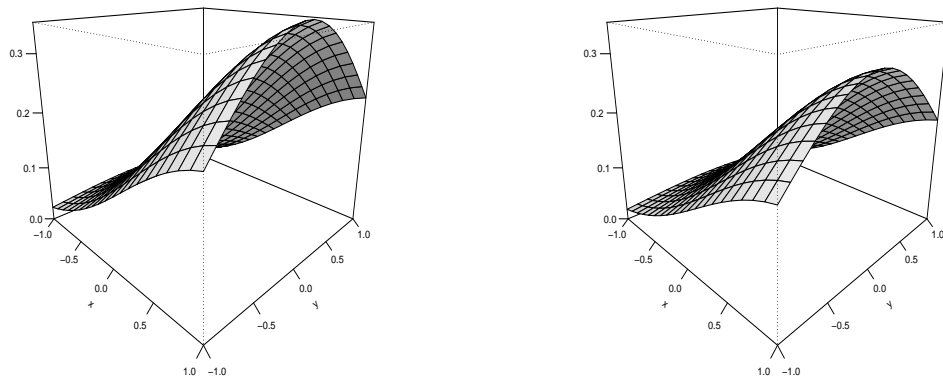


Figure 4: The left plot contains the true plume levels based on the estimated posterior distributions of the parameters. The right plot contains the predicted plume levels.

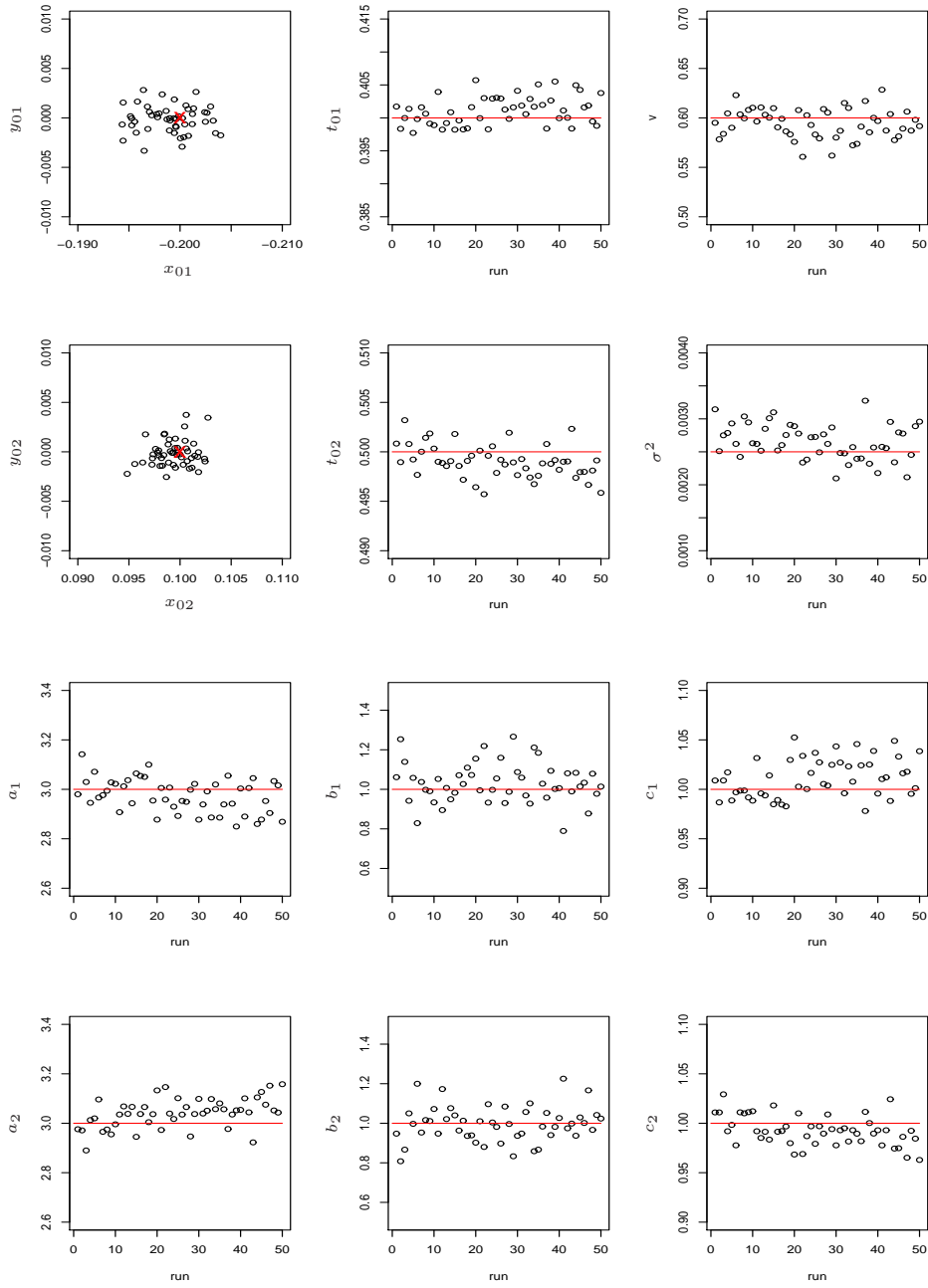


Figure 5: Estimated posterior means of the parameters based on 50 MCMC runs.

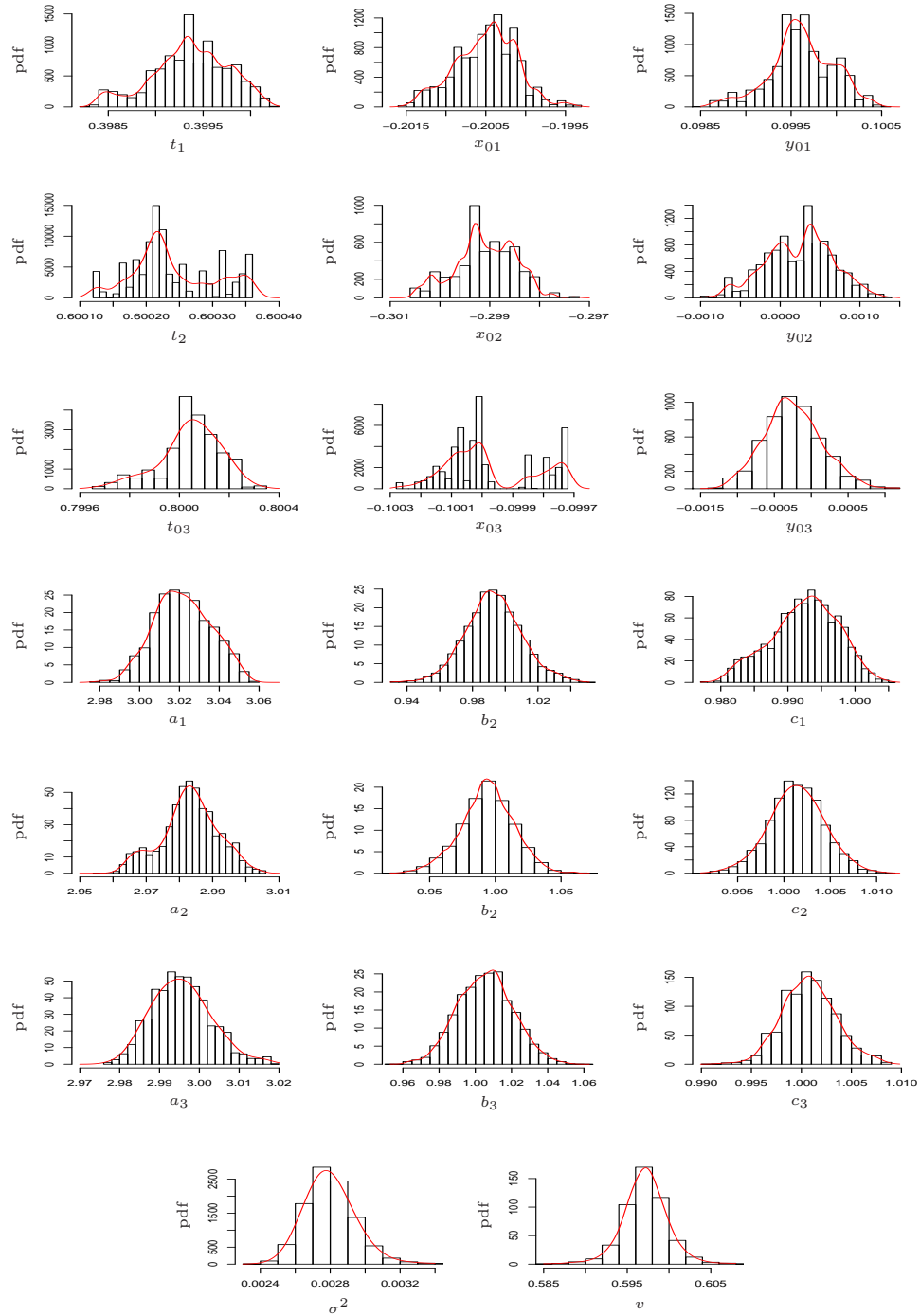


Figure 6: Estimated posterior distributions of the parameters based on one MCMC run for the three-source example.