

7.1 INVERSE DISPERSION OF AN UNKNOWN NUMBER OF CONTAMINANT SOURCES

Eugene Yee*

Defence R&D Canada – Suffield, P.O. Box 4000, Medicine Hat, AB, T1A 8K6, Canada

They say that Understanding ought to work by the rules of right reason. These rules are, or ought to be, contained in Logic; but the actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore the true logic for this world is the calculus of Probabilities, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind.

James Clerk Maxwell (1850)

1. INTRODUCTION

An increasingly capable sensing technology for concentration measurements of contaminants [e.g., chemical, biological, or radiological (CBR) agents] released into the turbulent atmosphere, either accidentally or deliberately, has fostered interest in exploiting this information for detection, identification and reconstruction of pollutant (contaminant) sources responsible for the observed concentration. A critical capability gap in current emergency and retrospective management efforts, directed at terrorist incidents involving the covert release of a CBR agent in a densely populated urban center, is the localization of the unknown source(s) following event detection by an array of independent CBR sensors. These sensors are placed at different points in space within a designated region in order to function as “electronic noses” that are able to provide quantitative measurements of the concentration of various air admixtures of contaminants.

For example, the Department of Homeland Security (DHS) has deployed (albeit sparse) arrays of biological agent sensors in 31 cities across the United States as part of the BioWatch program (Shea and Lister, 2003) in order to provide early detection and warning of a covert biological event. The BioWatch program has provided the impetus for recent research efforts directed towards a solution of the source reconstruction problem for inferring the location and emission rate of the source(s) of contamination. Certainly, determination of the characteristics of the unknown source is perhaps the most critical information required by emergency responders for the delineation of hazard zones (toxic corridors) resulting from the contaminant release and for implementation of an appropriate mitigation strategy (e.g., identification of exposed individuals, formulation of decisions for prophylactic treatment in the case of biological agents) required to counter the CBR agent release. Further motivation is provided by a network of 40 radiological detectors that has been set up as a verification tool for the Comprehensive Nuclear Test Ban Treaty (CTBT) in order to provide world wide monitoring of radioactive noble gases that can be used potentially for source localization and characterization of a clandestine nuclear test (Carrigan et al, 1996; Hourdin and Issartel, 2000), following upon the ear-

lier suggestions (Van der Vink and Park, 1994) to supplement seismic monitoring techniques with a monitoring network for atmospheric radionuclides.

Although the problem of the forward prediction of the concentration field resulting from the dispersion of a contaminant released in a complex turbulent flow (with a known source) has received considerable attention from researchers, the “reverse” prediction of the source location and strength using a finite number of noisy concentration data obtained from an array of sensors has received considerably less attention, in spite of the fact that the importance of the solution of this problem for a number of practical applications is obvious. Early work on inverse source determination focussed on the problem of source strength estimation, whereby the location of the source was assumed to be known exactly *a priori*. For example, Hanna et al (1990) used three different types of dispersion models to estimate the emission rate (source strength) for the localized sources used in the Project Prairie Grass experiments. Gordon et al (1988) and Wilson and Shum (1992) applied a forward-time Lagrangian stochastic trajectory model, in conjunction with a mass balance technique, to estimate the rate of ammonia volatilization from field plots (viz., area sources on the ground surface), while Flesch et al (1995) used a backward-time Lagrangian stochastic dispersion model to estimate the emission rate from a sustained surface area source over simple terrain. Kaharabata et al (2000) used an approximate solution to the advection-diffusion equation to determine the source strength from microplots over an open field, using single and multipoint measurements of the concentration downwind of the contaminated field plot. For longer-range dispersion, Robertson and Langner (1998) estimated the emission rates for the European Tracer Experiment (Nodop et al, 1998) using a variational data assimilation procedure to minimize a cost function defined to be the sum of squared differences between the model solution and the concentration measurements over a 24-h assimilation period (window). A different approach was proposed by Siebert and Stohl (1999), who attempted to reconstruct the source strength using Tikhonov regularization. Finally, Skiba (2003) used an adjoint transport model to assess emission rates from various industrial plants in Mexico, whose locations were assumed to be known *a priori*.

A number of researchers have focussed on the problem of recovering the location of an unknown source on the basis of available concentration measurements, given that the source strength is known *a priori*. Khapalov (1994) formulated the source localization problem as a nonlinear identification problem based on the diffusion system modeling involving a distributed system of parabolic type. Another optimization approach was followed by Alpay and Shore (2000) who described an “intelligent” brute-force method for source localization, which is based on the simulation of every possible scenario, with the likely source location being that which minimizes an appropriate norm between the model predictions of concentration for that source location and the measured concentration data. Matthes et al (2005) employed a two-

* Corresponding author address: Hazard Protection Section, Defence R&D Canada – Suffield, P.O. Box 4000, Medicine Hat, AB, T1A 8K6, Canada, e-mail:eugene.yee@drdc-rddc.gc.ca

step approach for source localization for the simple (albeit unrealistic) case of an emitted substance transported by homogeneous advection and isotropic diffusion. This method was designed to overcome convergence difficulties that have been experienced by various researchers using gradient-based optimization methods for formulation of the source localization problem.

The joint estimation of the source location and strength, in the context of dispersion on a global scale, was proposed by Pudykiewicz (1998), who inferred source parameters (location, emission rate, and time of release) by using a superposition of influence functions obtained as a solution of an adjoint tracer transport equation to estimate the source location, in conjunction with the Lagrange duality relation to estimate the strength of the source. Bocquet (2005a, 2005b) applied the principle of maximum entropy in the mean (MEM) to regularize the ill-posed nature of the source inversion. Thomson et al (2006) and Allen et al (2007) proposed, respectively, the application of simulated annealing and a genetic algorithm for solving the optimization problems (involving various cost functions with varying regularization terms) associated with source characterization. In the line of probabilistic (as opposed to optimization) approaches, a Bayesian inferential methodology used in conjunction with the adjoint method for representation of the source-receptor relationship was proposed by Yee (2005, 2006), Yee et al (2006), Keats et al (2006), Keats et al (2007a, 2007b), Yee et al (2007), and Yee (2007) for the simultaneous determination of all source parameters for dispersion of conservative and non-conservative scalars in simple (level, unobstructed terrain) and complex (e.g., urban terrain, complex terrain over continental scales) environments.

In all previous studies cited here, with the exception of Yee (2007), the problem of identification of source parameters was restricted to one source. The determination of the characteristics of multiple localized sources was briefly addressed in Yee (2007), but the approach used there *assumed* that the number of sources was known *a priori*. The problem of reconstruction of an unknown number of sources using a finite number of noisy concentration measurements obtained from an array of sensors is of great interest for current emergency and retrospective management efforts directed at terrorist incidents involving the release of CBR (or, other toxic) agents. In this application, the array of CBR sensors measures the superposition of concentrations arising from multiple interfering sources in which both number of sources and the characteristics (e.g., location, emission rate, release time) of each source are unknown *a priori*. The determination of the number of sources is a difficult problem and involves some form of model selection.

In this paper, we formulate a Bayesian inferential scheme for the joint determination of the number of contaminant sources and the parameters for each source, given a finite number of concentration measurements made by an array of sensors. The Bayesian formulation of a solution for this problem provides a computational challenge owing to the fact that the resulting integrals of the posterior distribution of the source parameters over the hypothesis space are not analytically tractable, and standard numerical methods for integration cannot be applied to give accurate results due to the potential high dimensionality of the hypothesis space (of possible source distribution models). To overcome this problem, we demonstrate how an innovative Metropolis-coupled reversible-jump

Markov chain Monte Carlo method can be used in conjunction with an adjoint representation for the source-receptor relationship (defining the dispersion of the toxic material), in order to provide a fast and reliable framework for Bayesian inference in this context [viz., to estimate the number of unknown sources and the relevant parameters (e.g., emission rate, source location, etc.) for each source].

2. SOURCE-RECEPTOR RELATIONSHIP

2.1 Mean concentration model

To solve the source reconstruction problem using Bayesian probability theory, it is necessary to relate the hypotheses of interest concerning the source distribution to the available concentration data measured by an array of sensors. This is a source-receptor relationship that describes the transport of a substance (e.g., CBR material) through the atmosphere after it has been released from a known source distribution and, as such, provides a prediction of the average concentration of the substance in a small volume centered at any given spatial location during any given time interval. Let the concentration at a spatial location $\mathbf{x} \equiv (x, y, z)$ and at time t be denoted $C(\mathbf{x}, t)$. The mean concentration “seen” by a sensor corresponds to an average of $C(\mathbf{x}, t)$ over the sensor volume and averaging time (centered at \mathbf{x}_r and t_r , respectively) and is given by

$$\begin{aligned} \bar{C}(\mathbf{x}_r, t_r) &\equiv \int_0^T dt \int_{\mathcal{D}} d\mathbf{x} C(\mathbf{x}, t) h(\mathbf{x}, t | \mathbf{x}_r, t_r) \\ &\equiv \langle C, h \rangle(\mathbf{x}_r, t_r), \end{aligned} \quad (1)$$

where $h(\mathbf{x}, t | \mathbf{x}_r, t_r)$ is the spatial-temporal filtering function of (\mathbf{x}, t) for a concentration sensor measurement at (\mathbf{x}_r, t_r) with

$$\int_0^T dt \int_{\mathcal{D}} d\mathbf{x} h(\mathbf{x}, t | \mathbf{x}_r, t_r) = 1, \quad (2)$$

and $\mathcal{D} \times [0, T]$ corresponds to a space-time volume that contains the source and the receptors (sensors).

The concentration $C(\mathbf{x}, t)$ can be determined using either an Eulerian or a Lagrangian approach for the atmospheric diffusion. In this paper we focus on a Lagrangian approach because this is perhaps the most natural and simplest description of turbulent diffusion. This approach involves modeling dispersion through the random walks for “marked” fluid elements. For the source-oriented approach within the Lagrangian description, the concentration $C(\mathbf{x}, t)$ can be determined by releasing “marked” fluid particles from the source distribution $S \equiv S(\mathbf{x}, t)$ and following these particles forward in time using the following stochastic differential equation (Thomson, 1987):

$$\begin{aligned} d\mathbf{X}(t) &= \mathbf{U}(t) dt, \\ d\mathbf{U}(t) &= \mathbf{a}(\mathbf{X}, \mathbf{U}, t) dt + (C_0 \epsilon(\mathbf{X}, t))^{1/2} d\mathbf{W}(t), \end{aligned} \quad (3)$$

where $\mathbf{X} \equiv \mathbf{X}(t) \equiv (X_i(t)) = (X_1(t), X_2(t), X_3(t))$ and $\mathbf{U} \equiv \mathbf{U}(t) \equiv (U_i(t)) = (U_1(t), U_2(t), U_3(t))$ are the (Lagrangian) position and velocity, respectively, of a “marked” fluid element (or, particle) at time t (marked by the source as the fluid element passes through it at some earlier time t'), so (\mathbf{X}, \mathbf{U}) determines the state of the fluid particle at any time t after its initial release from the source distribution S . In Eq. (3), C_0 is the Kolmogorov universal constant (associated with the Kolmogorov similarity hypothesis for the form

of the Lagrangian velocity structure function in the inertial subrange); ϵ is the mean dissipation rate of turbulence kinetic energy; $d\mathbf{W}(t) \equiv (dW_i(t)) = (dW_1(t), dW_2(t), dW_3(t))$ are the increments of a vector-valued (three-dimensional) Wiener process (i.e., these increments have a Gaussian distribution with zero mean and variance dt (infinitesimal time step), and non-overlapping increments are statistically independent); and $\mathbf{a} \equiv (a_i) = (a_1, a_2, a_3)$ is the drift coefficient vector (or, more precisely, the conditional mean acceleration).

Equation (3) is a forward-time Lagrangian stochastic (LS) trajectory simulation model. In this source-oriented approach, “marked” fluid elements with initial space-time coordinates (\mathbf{x}', t') are sampled from a space-time density function proportional to the (prescribed) source distribution $S(\mathbf{x}', t')$. The forward Lagrangian trajectories ($t > t'$) of these “marked” fluid elements, which emanate from the source and move towards the receptor, are determined in accordance to Eq. (3). This algorithmic procedure permits the evaluation of the mean concentration “seen” by a sensor at any arbitrary receptor space-time point (\mathbf{x}_r, t_r) in accordance to Eq. (1). Note that evaluation of the inner product $\overline{C}(\mathbf{x}_r, t_r) = \langle C, h \rangle(\mathbf{x}_r, t_r)$ involves two basic functions: the source distribution $S(\mathbf{x}', t')$ and the detector space-time filtering function $h(\mathbf{x}, t|\mathbf{x}_r, t_r)$. Unfortunately, the source-oriented approach described here is too computationally expensive for use in a Bayesian inferential approach for source reconstruction, because sampling from the posterior distribution for the source parameters (described later in this paper) requires a large number of forward source-receptor relationships to be determined, each of which involves the solution of the stochastic differential equation given by Eq. (3). Therefore, for the Bayesian inversion of concentration data to be practical, fast and efficient techniques are required for the determination of the source-receptor relationship within the context of sampling in the (large) hypothesis space of the source distribution models required for posterior inference.

In view of this, it is much more attractive to apply a receptor-oriented approach¹ for representation of the source-receptor relationship, owing to the fact that this approach provides a fast technique for predicting mean concentration data at a given receptor location for arbitrary hypotheses about the source distribution. As a consequence, we adopt and develop this formulation of the source-receptor relationship, as it provides a fast and reliable framework for Bayesian inference in the context of source reconstruction. To this purpose, we consider a mathematical representation for the source-receptor

¹A rigorous analysis of turbulent diffusion of a single particle in the source-oriented approach is contained in the classical work of Taylor (1921) in which he formulates a fundamental theory for a ‘diffusion by continuous movement’ that addresses explicitly the question of “how far, on average, will a single fluid particle migrate from its point of release in a time t as a result of the turbulent eddying?”. Taylor’s analysis of diffusion in stationary homogeneous turbulence demonstrates clearly the nature of turbulence as a correlated random walk and led to an important advance in our understanding of turbulent diffusion. However, the idea of using a receptor-oriented approach for the computation of turbulent diffusion seems to have been first proposed almost four decades later by Gifford (1959) in the context of the efficient determination of the concentration at a receptor due to multiple sources.

relationship that is dual to the one given by Eq. (1) as follows:

$$\begin{aligned} \overline{C}(\mathbf{x}_r, t_r) &\equiv \int_0^T dt \int_{\mathcal{D}} d\mathbf{x} C(\mathbf{x}, t) h(\mathbf{x}, t|\mathbf{x}_r, t_r) \\ &= \int_{-\infty}^{t_r} dt' \int_{\mathcal{D}} d\mathbf{x}' C^*(\mathbf{x}', t'|\mathbf{x}_r, t_r) S(\mathbf{x}', t') \\ &\equiv \langle C^*, S \rangle(\mathbf{x}_r, t_r), \end{aligned} \quad (4)$$

where $C^*(\mathbf{x}', t'|\mathbf{x}_r, t_r)$ is an adjunct (dual) concentration at space-time point (\mathbf{x}', t') associated with the sensor concentration data at location \mathbf{x}_r and time t_r (with $t' < t_r$). We note that $C^*(\mathbf{x}', t'|\mathbf{x}_r, t_r)$ is explicitly constructed so that it verifies the duality relationship implied by Eq. (4); namely, $\langle C, h \rangle = \langle C^*, S \rangle$.

Figure 1 illustrates explicitly the duality relationship between C , h , C^* , and S . If we interpret S as a vector in V , the vector space of source functions, then C^* can be viewed as the vector dual (or conjugate) to S belonging to the dual vector space V^* (or, conjugate concentration function space) defined to be the space of all linear functionals $C^* : V \rightarrow \mathbb{R}$. There is a one-to-one correspondence between the vector $S \in V$ and the dual vector $C^* \in V^*$, and this correspondence (or, more precisely, isomorphism between V and V^*) can be defined through the scalar product $\langle C^*, S \rangle$ that pairs S with C^* .² Similarly, C and h can be interpreted as dual (or conjugate) vectors lying in the concentration function space U and detector function space U^* , respectively, with U^* being the dual vector space to U with the isomorphism between these two spaces defined through the scalar product $\langle h, C \rangle$ with $C \in U$ and $h \in U^*$. More interestingly, C can be paired with its dual h and C^* can be paired with its dual S such that the duality relationship $\langle C, h \rangle = \langle C^*, S \rangle$ is exactly satisfied. Now, S can be related to C through the mapping G (viz., $C = G(S) \equiv GS$) defined using the forward LS model (source-receptor relationship). However, a mathematically equivalent representation of the source-receptor relationship can be formulated by relating h to C^* through the adjoint mapping G^* ³ (viz., $C^* = G^*(h) \equiv G^*h$) with G^* explicitly constructed so that the duality relation is exactly satisfied: $\langle C^*, S \rangle = \langle G^*h, S \rangle = \langle h, GS \rangle = \langle h, C \rangle = \langle C, h \rangle$, for any source $S \in V$ and any receptor $h \in U^*$.

The adjoint mapping G^* in Fig. 1 can be realized through a backward LS model (the latter of which constitutes a mathematically equivalent representation of the source-receptor relationship). A backward-time Lagrangian trajectory simulation model, that is dual to the forward-time Lagrangian trajectory simulation model given by Eq. (3), can be constructed for the computation of C^* so that it exactly satisfies the duality relationship $\langle C, h \rangle = \langle C^*, S \rangle$. To this end, suppose the backward-time Lagrangian trajectory model is defined as the solution to the following stochastic differential equation:

$$\begin{aligned} d\mathbf{X}^b(t') &= \mathbf{U}^b(t') dt', \\ d\mathbf{U}^b(t') &= \mathbf{a}^b(\mathbf{X}^b, \mathbf{U}^b, t') dt' + (C_0 \epsilon(\mathbf{X}^b, t'))^{1/2} d\mathbf{W}(t'), \end{aligned} \quad (5)$$

where $t' < t_r$ and at any given time t' , $(\mathbf{X}^b, \mathbf{U}^b)$ is a point in the phase space along the backward trajectory of the “marked”

²In fact, the scalar product $\langle C^*, S \rangle$ can be interpreted as an explicit mathematical representation for collection of all linear functionals $C^* : V \rightarrow \mathbb{R}$ that can be defined on V .

³In functional analysis, a linear operator $G^* : U^* \rightarrow V^*$ is called the dual or pull-back of the linear operator $G : V \rightarrow U$ if $\langle G^*h, S \rangle = \langle h, GS \rangle$, $\forall S \in V, h \in U^*$.

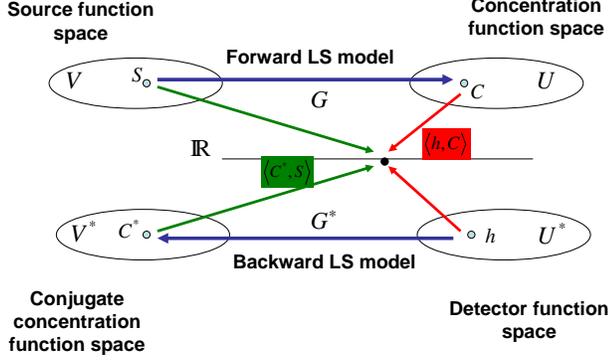


Figure 1: Commutative diagram illustrating the duality relationship $\langle C, h \rangle = \langle C^*, S \rangle$ verified by C, h, C^* , and S , with $C = GS \in U$ ($G : V \rightarrow U$) and $C^* = G^*h \in V^*$ ($G^* : U^* \rightarrow V^*$) $\forall S \in V, h \in U^*$, corresponding to two mathematically equivalent representations for the source-receptor relationship. Each of these two mathematical representations predict that the mean concentration \bar{C} measured by a detector is given by $\bar{C} = \langle C, h \rangle = \langle C^*, S \rangle$. From the perspective of a Lagrangian approach for turbulent diffusion, the mapping G is realized in terms of a forward-time LS model, whereas the associated (adjoint) mapping G^* is realized in terms of a backward-time LS model.

fluid element (here assumed to be marked or tagged as a fluid particle which at time t_r passed through the spatial volume of the detector at location \mathbf{x}_r). The displacement statistics of “marked” fluid elements released from the receptor location \mathbf{x}_r at time t_r can be used to compute $C^*(\mathbf{x}', t' | \mathbf{x}_r, t_r)$ (which is interpreted here as a function of \mathbf{x}' and t'). It can be shown (Thomson, 1987; Flesch et al, 1992) that C^* obtained from Eq. (1) for a detector with the filtering function h and C obtained from Eq. (4) for a release from the source density S is exactly consistent with the duality relationship $\langle C, h \rangle = \langle C^*, S \rangle$ if and only if \mathbf{a}^b in Eq. (5) is related to \mathbf{a} in Eq. (3) through the following “gauge” transformation:

$$a_i^b(\mathbf{x}, \mathbf{u}, t) = a_i(\mathbf{x}, \mathbf{u}, t) - C_0 \epsilon(\mathbf{x}, t) \frac{\partial}{\partial u_i} \ln p_E(\mathbf{u}; \mathbf{x}, t), \quad (6)$$

where $p_E(\mathbf{u}; \mathbf{x}, t)$ is the probability density function of the Eulerian velocity at (\mathbf{x}, t) .⁴ The relationship given by Eq. (6) is derived on the assumption that the conditional acceleration a_i for the forward-time Lagrangian trajectory model satisfies Thomson’s (1987) well-mixed criterion (viz., an initially well-mixed distribution of particles should remain so). As a consequence, Eq. (6) guarantees that the conditional acceleration a_i^b for the backward-time Lagrangian trajectory model also verifies the well-mixed criterion.

2.2. Model for concentration observations

The models described above provide predictions for the

⁴More specifically, the form of the conditional acceleration a_i^b for the backward-time Lagrangian trajectory model given by Eq. (6) ensures that the transition density function $p_L^f(\mathbf{x}, \mathbf{u}, t | \mathbf{x}', \mathbf{u}', t')$ of the forward-time Lagrangian trajectory model of Eq. (3) is related to the transition density function $p_L^b(\mathbf{x}', \mathbf{u}', t' | \mathbf{x}, \mathbf{u}, t)$ of the backward-time Lagrangian trajectory model of Eq. (5) as follows: $p_L^f(\mathbf{x}, \mathbf{u}, t | \mathbf{x}', \mathbf{u}', t') p_E(\mathbf{u}'; \mathbf{x}', t') = p_L^b(\mathbf{x}', \mathbf{u}', t' | \mathbf{x}, \mathbf{u}, t) p_E(\mathbf{u}; \mathbf{x}, t)$. The latter constraint guarantees that the duality relationship $\langle C, h \rangle = \langle C^*, S \rangle$ is satisfied exactly.

“ideal” mean concentration seen by a sensor at the receptor space-time point (\mathbf{x}_r, t_r) . The actual concentration data measured by the sensor will not usually agree with the concentration predicted by the model owing to the noise process imposed on the concentration data, which by its very nature is expected to have a very complicated structure. To this purpose, it is assumed that the actual concentration data available from the sensor array were measured at a finite number of sensor locations and at a finite number of time points at each sensor location. The actual concentration datum $d_{i,j^{(i)}}$ acquired by the sensor at receptor location \mathbf{x}_{r_i} and at time $t_j^{(i)}$ ($i = 1, 2, \dots, N_d$ and $j = 1, 2, \dots, N_t^{(i)}$, where N_d is the number of sensors and $N_t^{(i)}$ is the number of time samples measured at the i -th sensor) is assumed to be the sum of a modeled mean concentration signal $\bar{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta)$ and “noise” $e_{i,j^{(i)}}$, so

$$d_{i,j^{(i)}} = \bar{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta) + e_{i,j^{(i)}}, \quad (7)$$

where Θ is an appropriate parameter vector describing the source distribution S ; and, $\bar{C}(\mathbf{x}_r, t_r; \Theta)$ is the modeled mean concentration at location \mathbf{x}_r and time t_r , determined in accordance to Eq. (4) for a source distribution characterized by parameter vector Θ . For simplicity of notation, the variables in Eq. (7) which are indexed or labelled by $(i, j^{(i)})$ will be ordered in some regular and convenient manner (e.g., lexicographic ordering) and this collection will be indexed by J ($J = 1, 2, \dots, N$, with $N \equiv \sum_{i=1}^{N_d} N_t^{(i)}$ being the total number of measured concentration data). Then, we can write the observational model as follows:

$$d_J = \bar{C}_J(\Theta) + e_J, \quad J = 1, 2, \dots, N, \quad (8)$$

where $\bar{C}_J(\Theta) \equiv \bar{C}(\mathbf{x}_{r_i}, t_j^{(i)}; \Theta)$.

In Eq. (8), e_J is a noise term representing the uncertainty in d_J . In general, e_J consist of errors (e.g, input, stochastic, and measurement) and any real signal in the data that cannot be explained by the model. The random error e_J can be split into four terms as discussed by Rao (2005), so

$$e_J = \eta_J^{(1)} + \eta_J^{(2)} + \eta_J^{(3)} + \eta_J^{(4)}. \quad (9)$$

The first term $\eta_J^{(1)}$ of the error corresponds to model error arising from uncertainties in the representation of various physical processes in the dispersion model used to predict the mean concentration. The second term $\eta_J^{(2)}$ describes the input error arising from uncertainties in the values of empirical parameters and/or specification of the input meteorology (initial and boundary conditions) used by the dispersion model. The third term $\eta_J^{(3)}$ of the error is the stochastic uncertainty arising from the turbulent nature of the atmosphere, which gives rise naturally to random concentration fluctuations in hazardous gas releases. Finally, $\eta_J^{(4)}$ describes the noise inherent in the sensor (essentially measurement or instrument error).

Rao (2005) discusses the nature of these four types of error with respect to characterization of uncertainties in atmospheric dispersion models, and provides a comprehensive review of sensitivity and/or uncertainty analysis methods that have been used to quantify and reduce them. In this paper, all the various error contributions to the noise term are simply lumped together and denoted by e_J [see Eqs. (8) and (9)]. It is assumed that the observer does not have a detailed knowledge of the probability distribution of the noise (aggregate error),

other than that the observer has an estimate for the expected scale of variation of the noise. More specifically, it is assumed that the noise scale parameter associated with e_J (aggregate error of the J -th concentration observation) is provided in the form of a (finite) variance σ_J^2 .

In this paper, we assume *a priori* that the model concentration $\bar{C}_J(\Theta)$ in Eq. (8) results from the release of a contaminant (e.g., CBR material) from N_s transient point sources. It is assumed that the k -th source was located at $\mathbf{x}_{s,k}$ and that this source was activated (turned on) at time T_b^k and deactivated (turned off) at a later time T_e^k . Furthermore, over the time period from T_b^k to T_e^k , the source was assumed to be releasing contaminant at a constant emission rate Q_k ($k = 1, 2, \dots, N_s$). Consequently, the (unknown) source distribution is assumed *a priori* to have the following form:

$$S(\mathbf{x}, t) = \sum_{k=1}^{N_s} Q_k \delta(\mathbf{x} - \mathbf{x}_{s,k}) \left[H(t - T_b^k) - H(t - T_e^k) \right], \quad (10)$$

where $\delta(\cdot)$ and $H(\cdot)$ are the Dirac delta and Heaviside unit step functions, respectively. It is convenient to define a source parameter vector θ_{N_s} corresponding to the source distribution S of Eq. (10) as $\theta_{N_s} \equiv (\mathbf{x}_{s,1}, T_b^1, T_e^1, Q_1, \dots, \mathbf{x}_{s,N_s}, T_b^{N_s}, T_e^{N_s}, Q_{N_s}) \in \mathbb{R}^{6N_s}$. Furthermore, let $\Theta \equiv (N_s, \theta_{N_s})$. If we substitute Eq. (10) into Eq. (4), the model concentration $\bar{C}_J(\Theta)$ “seen” by the sensor at location $\mathbf{x}_{r,J}$ and time $t_{r,J}$ is given explicitly by

$$\bar{C}_J(\Theta) = \sum_{k=1}^{N_s} Q_k \int_{T_b^k}^{\min(t_{r,J}, T_e^k)} C^*(\mathbf{x}_{s,k}, t_s | \mathbf{x}_{r,J}, t_{r,J}) dt_s. \quad (11)$$

It should be noted that Eq. (10) can be interpreted as the mathematical representation⁵ of a general source distribution S (e.g., area or volume source with arbitrary geometry) in terms of the basis functions $b_k(\mathbf{x}, t) = Q_k \delta(\mathbf{x} - \mathbf{x}_{s,k}) \left[H(t - T_b^k) - H(t - T_e^k) \right]$ consisting for elemental (transient) source atoms scattered randomly with positions $\mathbf{x}_{s,k}$ over the domain \mathcal{D} of interest. For a large enough value of N_s , source distributions of arbitrary complexity can be approximated to any prescribed accuracy using the representation of Eq. (10). In this sense, the source reconstruction algorithm described in this paper can be applied to recover an arbitrary source distribution. Nevertheless, the primary focus of this paper is the recovery of the characteristics of a small number of point sources (say $N_s \leq 10$) when the number of sources is unknown *a priori*.

With the formulation above, the problem of source reconstruction reduces to the following: given the observed vector of concentration data $\mathbf{D} \equiv (d_1, d_2, \dots, d_N)$, the objective is to estimate N_s and θ_{N_s} or, equivalently, Θ .

3. BAYESIAN PROBABILITY THEORY

To determine the number of sources and reconstruct the characteristics (e.g., location, emission rate, release time) of

⁵Eq. (10) provides a structure-based representation for an arbitrary source distribution, whereby the amount of structure required to approximate the geometrical and physical characteristics of the source can accommodate to any degree of accuracy by varying the number of source atoms used in this representation.

each source (parameters encapsulated in Θ), given the vector of concentration observations \mathbf{D} obtained from an array of sensors, the techniques of Bayesian inference are employed. The Bayesian framework is attractive in the current context because it provides a rigorous mathematical foundation for making inferences about the source parameters and, as a consequence, provides a rigorous basis for quantifying the uncertainties in the estimated source parameters. Bayesian inference can be obtained from the product rule of probability calculus, the latter of which can be derived rigorously starting with the formulation of a small number of desiderata required to define a rational theory of inference as first enunciated by Cox (1947), with a more complete treatment given by Jaynes (2003) in his definitive treatise. This formulation leads to the ordinary rules (sum and product rules) of probability calculus and implies that every allowed (mathematically consistent) theory for inference must be equivalent to probability theory, or else inconsistent (viz., no other calculus is admissible for inference that is consistent with the above mentioned desiderata).⁶

The basic relationship quantifying parameter inference is Bayes’ rule, which in the context of the current problem can be expressed as follows:

$$p(\Theta | \mathbf{D}, I) p(\mathbf{D} | I) = p(\mathbf{D} | \Theta, I) p(\Theta | I). \quad (12)$$

The factors in Eq. (12) are: (1) $p(\Theta | I)$ is the *prior* probability density function (PDF) of the source parameter vector Θ that encapsulates our state of knowledge of the parameters before the receipt of the concentration measurements; (2) $p(\mathbf{D} | \Theta, I)$ is termed the *likelihood function* when considered as a function of Θ , but is known as the sampling distribution when considered as a function of \mathbf{D} (the latter being the probability of observing the concentration data \mathbf{D} when given the source parameters Θ); (3) $p(\mathbf{D} | I)$ is termed the *evidence* (also, frequently referred to as the prior predictive or the marginal likelihood); and, (4) $p(\Theta | \mathbf{D}, I)$ is the *posterior* probability density function of the parameters Θ of interest, that corresponds to the update of $p(\Theta | I)$ incorporating the knowledge gained about Θ after receipt of the concentration observations \mathbf{D} . All the terms here are to be interpreted given the background (contextual) information I (e.g., background meteorology, source-receptor relationship, etc.) that defines the source reconstruction problem. This is the meaning of I after the vertical bar “|” which is used here to denote “conditional upon”.

Using Bayesian inference, all information in the measured concentration data relevant to the problem of estimating the source parameter vector Θ is summarized in the posterior PDF

⁶Cox (1947) demonstrated rigorously that a theory of plausible reasoning that satisfies three desiderata must necessarily lead to a theory for inference that is mathematically equivalent (isomorphic) to probability calculus. These three desiderata are (1) degrees of plausibility can be represented by real numbers; (2) consistency with deductive (Aristotelian) logic in the limit when all propositions are either certainly true or certainly false; and, (3) mathematical (internal) consistency of the schemata in the sense that if a result can be reasoned out in more than one way, every possible way must lead to the same answer. This analysis demonstrates conclusively that Bayesian probability theory (or, probability theory as an extended logic) is the correct mathematical language for inference and, moreover, constitutes the *uniquely* “right” rules for conducting inference (or, plausible reasoning).

$p(\Theta|\mathbf{D}, I)$, which by virtue of Eq. (12), is given by

$$p(\Theta|\mathbf{D}, I) \propto p(\mathbf{D}|\Theta, I)p(\Theta|I). \quad (13)$$

It is noted that for parameter estimation, the evidence $p(\mathbf{D}|I)$ is Θ -independent and simply plays the role of a normalization factor; hence, $p(\Theta|\mathbf{D}, I)$ is expressed simply as a proportionality in Eq. (13). To proceed further in the specification of the posterior PDF, we now need to assign functional forms for the prior PDF $p(\Theta|I)$ and the likelihood function $p(\mathbf{D}|\Theta, I)$.

3.1. Assignment of prior PDF

First, let us consider the assignment of the prior PDF $p(\Theta|I) \equiv p((N_s, \theta_{N_s})|I)$. To proceed, it is necessary to state explicitly what is known about the source parameters before the receipt of the concentration data. To this purpose, we assume the logical independence of the various source parameters; that is to say, knowing the value of the number of sources tells us nothing about the various source characteristics (e.g., location, emission rate, release time), knowing the characteristics of one source tells us nothing about those of another source, knowing the location of a particular source tells us nothing about the emission rate or when the source was turned on/off, etc. In consequence, the prior probability $p((N_s, \theta_{N_s})|I)$ factorizes (by repeated application of the product rule of probability calculus) as follows:

$$p((N_s, \theta_{N_s})|I) = p(N_s|I) \prod_{k=1}^{N_s} p(Q_k|I)p(\mathbf{x}_{s,k}|I) \times p(T_b^k|I)p(T_e^k|T_b^k, I). \quad (14)$$

Let us now consider the assignment of each of the (component) prior distributions in Eq. (14). The prior distribution on N_s (number of sources) is chosen to be a binomial distribution defined on the set $\{1, 2, \dots, N_s^*\}$, where N_s^* is the maximum number of sources. More specifically, the binomial prior for N_s has the form (offset by the minimum number of sources, here assumed without any loss in generality to be one):

$$p(N_s|I) = \frac{(N_s^* - 1)!}{(N_s - 1)!(N_s^* - N_s)!} \times p^{*(N_s-1)}(1 - p^*)^{N_s^* - N_s}, \quad (15)$$

where $p^* \in [0, 1]$ is the binomial rate.

Given N_s , the prior distributions for Q_k , $\mathbf{x}_{s,k}$, T_b^k and T_e^k ($k = 1, 2, \dots, N_s$) are each chosen to be either Bernoulli-uniform or uniform over an appropriate domain of definition. More specifically, the prior distribution for Q_k is assigned a Bernoulli-uniform mixture model:

$$p(Q_k|I) = (1 - \gamma)\delta(Q_k) + \gamma\mathbb{I}_{(0, Q^*)}(Q_k)/Q^*, \quad (16)$$

($k = 1, 2, \dots, N_s$) where γ is an intermittency factor, defined as the probability that the source is turned on (viz., $\Pr(Q_k > 0) = \gamma$, so $\Pr(Q_k = 0) = 1 - \gamma$); and, Q^* is an upper bound for the expected emission rate. Furthermore, $\mathbb{I}_A(x)$ is the indicator function for set A (viz., $\mathbb{I}_A(x) = 1$ if $x \in A$, and $\mathbb{I}_A(x) = 0$ if $x \notin A$). This mixture probability model for Q_k allows for the hypothesis that any given source in the domain can be turned off ($Q_k = 0$) with a finite probability ($= 1 - \gamma$).

The prior distribution for $\mathbf{x}_{s,k}$ is chosen to be uniform over some large region $\mathcal{D} \subset \mathbb{R}^3$:

$$p(\mathbf{x}_{s,k}|I) = \mathbb{I}_{\mathcal{D}}(\mathbf{x}_{s,k})/\mathcal{V}(\mathcal{D}), \quad k = 1, 2, \dots, N_s, \quad (17)$$

where $\mathcal{V}(\mathcal{D})$ is the volume of the region \mathcal{D} .

The prior distributions for T_b^k and T_e^k are assigned uniform distributions with the following forms:

$$p(T_b^k|I) = \mathbb{I}_{(0, T^*)}(T_b^k)/T^*, \quad (18)$$

and

$$p(T_e^k|T_b^k, I) = \mathbb{I}_{(T_b^k, T^*)}(T_e^k)/(T^* - T_b^k), \quad (19)$$

for $k = 1, 2, \dots, N_s$. Here, T^* is an upper bound on the time at which the source was turned on or off.⁷ Note that the time that the source is turned off must necessarily occur after it has been turned on, and this information is encoded in the form of the prior distribution for T_e^k given by Eq. (19), where the distribution is seen to be conditioned on T_b^k .

3.2. Assignment of likelihood function

Next, let us consider the assignment of a functional form for the likelihood function $p(\mathbf{D}|\Theta, I)$. The likelihood function is equivalent to the direct probability for the concentration data \mathbf{D} , given the source parameters Θ . In the absence of a detailed knowledge of the noise distribution e_J [cf. Eq. (9)], other than that it has a finite variance σ_J^2 , the application of the principle of maximum entropy (Jaynes, 2003) informs us that a Gaussian distribution is the most conservative choice for the direct probability of the data \mathbf{D} (or, equivalently, of the noise $\mathbf{e} \equiv (e_1, e_2, \dots, e_N)$).

The entropy of the PDF of the noise is a measure of its information content (viz., it is the asymptotic measure of the size of the basic support set of the distribution or ‘volume’ occupied by the sensibly probable noise values). The principle of maximum entropy is applied to ensure that the PDF representing our ‘state of information’ about the noise values does not encapsulate unwarranted assumptions (e.g., about higher-order moments of the noise which are not available). Choosing a distribution for the noise that provides the largest support set permitted by the information allows the largest range of possible variations in the noise values consistent with the available information (implying the most conservative estimates for these values).

Assigning a Gaussian distribution for the noise using the maximum entropy principle makes no statement about the true (unknown) sampling distribution of the noise (which has a very complicated structure for the current problem). Rather, it simply represents a maximally uninformative state of knowledge, a state of knowledge that reflects what the observer knows about the true noise in the data (namely, the mean and variance of the noise, with all other properties of the noise being irrelevant to the inference since these are unknown to the observer).

From these considerations, the likelihood function for our problem has the following form [in light of Eq. (8)]:

$$p(\mathbf{D}|\Theta, I) = \frac{1}{\prod_{J=1}^N \sqrt{2\pi}\sigma_J} \exp\left(-\frac{1}{2}\chi^2(\Theta)\right), \quad (20)$$

where

$$\chi^2(\Theta) \equiv \sum_{J=1}^N \left(\frac{d_J - \bar{C}_J(\Theta)}{\sigma_J}\right)^2. \quad (21)$$

⁷Different upper bounds can be chosen for T_b^k and T_e^k in the prior PDFs of Eqs. (18) and (19), but for the formulation in this paper we simply used a common upper bound for the source on and off times (with effectively no loss in generality).

3.3. Posterior distribution of source parameters

Inserting Eqs. (14)–(19) and Eqs. (20) and (21) in Eq. (13), the posterior distribution for the source parameters, $p(\Theta|\mathbf{D}, I)$, can be expressed by the following proportionality:

$$\begin{aligned}
p(\Theta|\mathbf{D}, I) &\propto \frac{1}{\prod_{J=1}^N \sqrt{2\pi}\sigma_J} \exp\left(-\frac{1}{2} \sum_{J=1}^N \left(\frac{d_J - \overline{C}_J(\Theta)}{\sigma_J}\right)^2\right) \\
&\times \frac{(N_s^* - 1)!}{(N_s - 1)!(N_s^* - N_s)!} p^{*(N_s - 1)} (1 - p^*)^{N_s^* - N_s} \\
&\times \prod_{k=1}^{N_s} \left[(1 - \gamma)\delta(Q_k) + \gamma \mathbb{I}_{(0, Q^*)}(Q_k) / Q^* \right] \\
&\times \mathbb{I}_{\mathcal{D}}(\mathbf{x}_{s,k}) \mathbb{I}_{(0, T^*)}(T_b^k) \frac{\mathbb{I}_{(T_b^k, T^*)}(T_e^k)}{(T^* - T_b^k)}, \quad (22)
\end{aligned}$$

recalling that the source parameter vector $\Theta \equiv (N_s, \theta_{N_s})$. Here, $\overline{C}_J(\Theta)$ is determined in accordance to Eq. (11) with the adjunct “concentration” field $C^*(\mathbf{x}', t' | \mathbf{x}_r, t_r)$ predicted using a backward-time Lagrangian stochastic model [cf. Eqs. (5) and (6)].

3.4. Summary statistics for inferred source parameters

The posterior distribution for Θ provides the full solution for the multiple source reconstruction problem. Inferences on the values of the source parameters are based on this posterior distribution.⁸ The posterior distribution may be summarized by various statistics of interest such as the posterior mean of each source parameter, say $\theta_{N_s}^i$ (which can be an emission rate Q_k , or source location $\mathbf{x}_{s,k}$, or time at which the source was turned on (off) T_b^k (T_e^k) for the k -th source for a source distribution consisting of N_s sources with $k = 1, 2, \dots, N_s$):

$$\overline{\theta_{N_s}^i} = \mathcal{E}[\theta_{N_s}^i | \mathbf{D}] \equiv \int \theta_{N_s}^i p((N_s, \theta_{N_s}) | \mathbf{D}, I) d\theta_{N_s}, \quad (23)$$

where $\mathcal{E}[\cdot]$ denotes mathematical expectation. A measure of the uncertainty of this estimate of $\theta_{N_s}^i$ is the posterior standard deviation $\sigma(\theta_{N_s}^i)$:

$$\begin{aligned}
\sigma^2(\theta_{N_s}^i) &= \mathcal{E}[(\theta_{N_s}^i - \overline{\theta_{N_s}^i})^2 | \mathbf{D}] \\
&\equiv \int (\theta_{N_s}^i - \overline{\theta_{N_s}^i})^2 p((N_s, \theta_{N_s}) | \mathbf{D}, I) d\theta_{N_s}. \quad (24)
\end{aligned}$$

⁸The posterior distribution also forms the basis for predictions of the expected concentration arising from the (unknown) source distribution at space-time points in the domain where there are no detectors. The Bayesian approach to prediction is based on the predictive probability density function, $p(\mathbf{D}' | \mathbf{D}, I)$, given by

$$\begin{aligned}
p(\mathbf{D}' | \mathbf{D}, I) &= \int_{\mathbb{R}^{6N_s}} p(\mathbf{D}' | \Theta, I) p(\Theta | \mathbf{D}, I) d\Theta \\
&= \int_{\mathbb{R}^{6N_s}} p(\mathbf{D}' | \Theta, I) p(\Theta | \mathbf{D}, I) d\Theta,
\end{aligned}$$

where \mathbf{D}' is the vector of expected concentration values (at various space-time locations in the domain) that will be predicted. Note that the predictive density is determined by the convex hull spanned by the possible hypotheses for the source distribution models Θ , with the boundaries (or “vertices”) for the convex hull determined by the likelihood function $p(\mathbf{D}' | \Theta, I)$ for \mathbf{D}' . Finally, with perfect knowledge of the source (so $\Theta = \Theta^*$ exactly), $p(\Theta | \mathbf{D}, I) = \delta(\Theta - \Theta^*)$ and $p(\mathbf{D}' | \mathbf{D}, I) = p(\mathbf{D}' | \Theta^*, I)$ (viz., the predictive density in this special case is simply the sampling distribution for \mathbf{D}' with a fixed value of $\Theta = \Theta^*$).

Alternatively, a $p\%$ credible [or, highest posterior density (HPD)] interval that contains the source parameter θ^i with $p\%$ probability, with the lower and upper bounds of the interval specified such that the probability density within the interval is everywhere larger than that outside it, can be used as a measure of the uncertainty in the determination of θ^i . Finally, an estimate for the number of sources can be obtained from the maximum a posteriori estimate as follows:

$$\hat{N}_s = \operatorname{argmax}_{N_s} p(N_s | \mathbf{D}, I), \quad (25)$$

where $p(N_s | \mathbf{D}, I)$ is the posterior probability of the number of sources, given the concentration data and the contextual background information. The posterior probability for the number of sources is a marginal posterior probability, where all the parameters we are not interested in (e.g., θ_{N_s}) are removed, through the process of integrating over these parameters. This process is referred to as marginalization.

In general, Bayesian probability theory never advocates the minimization, maximization, or optimization of any objective or cost function (in sharp contrast to the underlying basis for the application of regularization procedures for source reconstruction which attempts to find an “optimal” solution by setting a balance between the importance of quality (regularization) and of fitting the data, with sometimes arbitrary choices for the functional used to represent the regularization in the problem). Rather, the rules of Bayesian probability theory demand that we sum or integrate over unknown quantities, so that the effect is to average over all plausible values of these quantities. The underlying philosophy of Bayesian probability theory for source reconstruction is to find and explore all regions in the hypothesis space (of source distribution models) of reasonably large plausibility, and not simply to find the highest point of maximum posterior probability.⁹ This procedure allows a rigorous assessment of the uncertainty in our inferences of the source parameters.

4. BAYESIAN COMPUTATION AND MARKOV CHAINS

A perusal of Eq. (22) shows that the posterior distribution is highly nonlinear in the source parameters associated with location $(\mathbf{x}_{s,k})$ and with the times the source was turned on (T_b^k) or off (T_e^k), and that explicit evaluations of Eqs. (23) and (24) (and similar integrals or functionals arising in Bayesian analysis) are impossible. In view of this, we apply posterior sampling for evaluation of these integrals, which is implemented using a Markov chain Monte Carlo (MCMC) algorithm (Gilks et al, 1996; Gelman et al, 2003). A MCMC algorithm can be used to generate samples of source distribution models (characterized by Θ) from the posterior distribution in Eq. (22). Towards this objective, a Markov chain $\{\Theta^{(t)}\} \equiv \{(N_s^{(t)}, \theta_{N_s}^{(t)})\}$ is constructed whose stationary (or, invariant) distribution is the posterior distribution $p(\Theta | \mathbf{D}, I)$ of the parameters $\Theta = (N_s, \theta_{N_s})$.

⁹Indeed, in a number of cases of source reconstruction encountered by the author [including some cases of source inversion for dispersion over simple terrain such as for Project Prairie Grass (Yee, 2005)], the posterior distribution of the source parameters is multimodal and highly asymmetrical, and the highest peak in this distribution is very narrow and located far from the bulk of the distribution where the probability mass is concentrated.

All quantities of interest, such as posterior means of various source parameters and various marginal posterior distributions, can be estimated by sample path averages of the Markov process $\{\Theta^{(t)}\}$. For example, the marginal posterior distribution $p(N_s|\mathbf{D}, I)$, which is required in Eq. (25) for the inference of N_s , can be estimated as follows:

$$\begin{aligned}\widehat{p}(N_s|\mathbf{D}, I) &= \frac{1}{N'} \#\{t: N_s^{(t)} = N_s\} \\ &= \frac{1}{N'} \sum_{t=1}^{N'} \mathbb{I}_{\{N_s\}}(N_s^{(t)}),\end{aligned}\quad (26)$$

where N' is the number of samples $\Theta^{(1)}, \Theta^{(2)}, \dots$ drawn from a sampled realization of the Markov chain and $\mathbb{I}_{\{N_s\}}(\cdot)$ is used to select those samples $\{\Theta^{(t)}, t = 1, 2, \dots, N'\}$ drawn from the Markov chain that correspond to source distribution models containing exactly N_s sources (viz., those samples of source distribution models with $N_s^{(t)} = N_s$). Similarly, from the samples of source distributions generated from such a Markov chain, the posterior means of various source parameters $\theta_{N_s}^i$ [see Eq. (23)] can be estimated by

$$\widehat{\mathcal{E}}[\theta_{N_s}^i|\mathbf{D}] = \frac{\sum_{t=1}^{N'} \theta_{N_s}^{i(t)} \mathbb{I}_{\{N_s\}}(N_s^{(t)})}{\sum_{t=1}^{N'} \mathbb{I}_{\{N_s\}}(N_s^{(t)})}.\quad (27)$$

The difficulty in the construction of a Markov chain for sampling source distribution models from the posterior distribution of Eq. (22) resides in the fact that N_s is an unknown, requiring the design of a Markov chain that can simultaneously explore both parameter and model spaces. More specifically, the length of the parameter vector $\theta_{N_s} \in \mathbb{R}^{6N_s}$ is variable because it depends on the value of N_s (number of source atoms) which is not known *a priori*. This can result in algorithmic complications. Furthermore, if the expected number of source atoms in the source distribution is large, then so too is the corresponding hypothesis space which must be sampled. In the current problem, we need to consider a set of source distribution models $\mathcal{M} = \{M_{N_s}\}_{N_s=1}^{N_s^*}$ indexed by the integer parameter $N_s \in \mathcal{N}_s \equiv \{1, 2, \dots, N_s^*\}$, each characterized by a source parameter vector $\theta_{N_s} \in \mathbb{R}^{6N_s}$. Here, model M_{N_s} corresponds to all permissible source distributions having exactly N_s sources. The simultaneous exploration of a number of candidate source models involving differing numbers of sources can be realized using the reversible-jump MCMC algorithm that was originally introduced by Green (1995) in the context of a Bayesian model determination problem. The remainder of this section will describe the design of a reversible-jump MCMC algorithm for multiple source reconstruction, in the case where the number of sources is unknown *a priori*. Actually, this algorithm can be used also for the reconstruction of an arbitrary source distribution of unknown geometry (e.g., volume source whose geometry is unknown *a priori*) by simply allowing N_s to be very large, or even infinite (if no upper bound N_s^* is imposed on the allowable number of sources N_s in the source distribution).¹⁰

¹⁰In the case in which N_s is allowed to be infinite, the prior distribution for N_s given by Eq. (15) must be replaced by a discrete probability distribution having support over the natural numbers \mathbb{N} (e.g., Poisson distribution, geometric distribution, etc.). Since there is no upper limit on the number of source atoms in the source distribution, the overall hypothesis space is formally a countably

4.1. Propagation moves

The objective of MCMC is to construct a Markov chain whose stationary distribution is one that coincides exactly with the target probability density function that we are trying to sample from [e.g., in our case that target probability distribution is the posterior distribution $p(\Theta|\mathbf{D}, I)$ given in Eq. (22)]. In this subsection, we consider the problem of construction of a MCMC algorithm that samples from $p(\Theta|\mathbf{D}, I)$ for fixed N_s . The basic MCMC algorithm consists of two components: (1) a transition or proposal distribution function (or transition kernel) $T(\Theta'|\Theta)$; and, (2) an acceptance probability $\alpha(\Theta, \Theta')$. These two components are related as follows: given a chain in the current state $\Theta^{(t)} = \Theta$ at iteration t , a proposed new state $\Theta' = \mathcal{P}(\Theta) = \mathcal{P}((N_s, \theta_{N_s})) = (N_s, \theta'_{N_s})$ is drawn from some proposal (transition) distribution $T(\Theta'|\Theta)$ and this new point is accepted as the new state of the chain at iteration $(t+1)$ with the standard Metropolis-Hastings (M-H) acceptance probability (Gelman et al, 2003) given by

$$\alpha(\Theta, \Theta') = \min \left\{ 1, \frac{p(\Theta'|\mathbf{D}, I)T(\Theta|\Theta')}{p(\Theta|\mathbf{D}, I)T(\Theta'|\Theta)} \right\}.\quad (28)$$

Note that $\mathcal{P}(\cdot)$ is a pure *propagation* operation (move) that “translates” $\theta_{N_s} \in \mathbb{R}^{6N_s}$ to $\theta'_{N_s} \in \mathbb{R}^{6N_s}$, while keeping N_s fixed (update move in a fixed-dimensional hypothesis space). If the proposal for this propagation move is accepted, then the new state at iteration $(t+1)$ is $\Theta^{(t+1)} = \Theta'$; otherwise, $\Theta^{(t+1)} = \Theta$. Finally, it is noted that the M-H algorithm does not require the normalization of $p(\Theta|\mathbf{D}, I)$ be known [cf. Eq. (28)].

To proceed further, it is necessary to specify appropriate forms for the proposal density $T(\Theta'|\Theta)$. For the propagation move, N_s is fixed and the update occurs only for θ_{N_s} . To this purpose, it is useful to partition θ_{N_s} into two blocks of parameters as follows: $\theta_{N_s} = (\theta^1, \theta^2)$ where $\theta^1 \equiv (Q_1, Q_2, \dots, Q_{N_s}) \in \mathbb{R}^{N_s}$ and $\theta^2 = (\mathbf{x}_{s,1}, T_b^1, T_e^1, \dots, \mathbf{x}_{s,N_s}, T_b^{N_s}, T_e^{N_s}) \in \mathbb{R}^{5N_s}$. This particular partitioning of the source parameter vector distinguishes parameters that are related linearly (θ^1) and nonlinearly (θ^2) to the model concentration data as specified in Eq. (11). With this partitioning of the source parameter vector, a cycle of two different types of “blocked” component updates (Gibbs step and M-H step) involving θ^1 and θ^2 (respectively) are combined in sequence to form a single iteration of the MCMC sampler for generation of the updated state of the Markov chain.

First, let us update the emission rates Q_k of the various sources ($k = 1, 2, \dots, N_s$), which we have collected together in θ^1 . For the first part of the iteration $(t+1)$, we fix the source parameters in $\theta^2 = \theta^{2(t)}$ (e.g., source locations, activation times, deactivation times) to the values obtained in the previous iteration t and focus on the re-sampling of Q_k ($k = 1, 2, \dots, N_s$). We choose here to sample the emission rates one-at-time using Gibbs sampling. The Gibbs sampler updates Q_k as a direct draw from the univariate full conditional posterior distribution $p(Q_k|\theta_{-k}^1, \theta^{2(t)}, \mathbf{D}, I)$, where θ_{-k}^1 is used to denote a subvector containing all the components of θ^1 with the exception of Q_k . More specifically, the re-sampling of Q_k ($k = 1, 2, \dots, N_s$) for iteration $(t+1)$ proceeds using

infinite union of subspaces $\theta = \bigcup_{N_s=1}^{\infty} \theta_{N_s}$, where $\theta_{N_s} \in \mathbb{R}^{6N_s}$ denotes the $6N_s$ -dimensional hypothesis space corresponding to the collection of source distribution models with exactly N_s source atoms.

a systematic sweep Gibbs sampling strategy as follows (viz., we sample $Q_k^{(t+1)}$ in accordance to the conditional probability distributions indicated below):¹¹

$$\begin{aligned} Q_1^{(t+1)} &\sim p(Q_1|Q_2^{(t)}, Q_3^{(t)}, \dots, Q_{N_s}^{(t)}, \theta^{2(t)}, \mathbf{D}, I); \\ Q_2^{(t+1)} &\sim p(Q_2|Q_1^{(t+1)}, Q_3^{(t)}, \dots, Q_{N_s}^{(t)}, \theta^{2(t)}, \mathbf{D}, I); \\ &\vdots \\ Q_{N_s}^{(t+1)} &\sim p(Q_{N_s}|Q_1^{(t+1)}, Q_2^{(t+1)}, \dots, Q_{N_s-1}^{(t+1)}, \theta^{2(t)}, \mathbf{D}, I); \end{aligned} \quad (29)$$

which yields $\theta^{1(t+1)} = (Q_1^{(t+1)}, Q_2^{(t+1)}, \dots, Q_{N_s}^{(t+1)})$ after N_s cycles. It is noted that Gibbs sampling can be applied to re-sample Q_k because the univariate conditional distribution $p(Q_k|\theta_{-k}^1, \theta^{2(t)}, \mathbf{D}, I)$ has a particularly simple form from which it is possible to obtain Q_k as a direct draw. Indeed, a perusal of Eqs. (22), (11), and (16) shows that $p(Q_k) \equiv p(Q_k|\theta_{-k}^1, \theta^{2(t)}, \mathbf{D}, I)$ possesses a simple Bernoulli-Gaussian (truncated) distribution (which can be sampled from directly):

$$\begin{aligned} p(Q_k) \propto \exp \left(-Q_k \left(\sum_{J=1}^N \left[\sum_{\substack{r=1 \\ r \neq k}}^{N_s} Q_r f_r^J(\theta^2) - d_J \right] f_k^J(\theta^2) \right) \right. \\ \left. - \frac{1}{2} Q_k^2 \sum_{J=1}^N f_k^J(\theta^2) \right) \left[(1 - \gamma) \delta(Q_k) \right. \\ \left. + \gamma \mathbb{1}_{(0, Q^*)}(Q_k) / Q^* \right], \end{aligned} \quad (30)$$

where for simplicity of notation we have defined

$$f_k^J(\theta^2) \equiv \int_{T_b}^{\min(t_{r,J}, T_e^k)} C^*(\mathbf{x}_{s,k}, t_s | \mathbf{x}_{r,J}, t_{r,J}) dt_s. \quad (31)$$

Following the update of Q_k for iteration $(t+1)$ in accordance to Eq. (29), the second part of this iteration involves updating the nonlinear source parameters collected in θ^2 from $\theta^{2(t)}$ to $\theta^{2(t+1)}$. To accomplish this part of the iteration, the emission rate parameters are assumed to be fixed at $\theta^1 = \theta^{1(t+1)}$. For the updating of these parameters, it is no longer possible to use a Gibbs sampler because the resulting univariate full conditional distribution of these parameters cannot be sampled from directly. In view of this, we need to perform a M-H step to update these parameters with an appropriately selected proposal distribution. Towards this objective, we sample each of the nonlinear parameters (locations, activation times, and deactivation times of the various sources) using a M-H step with proposal distribution $T(\theta_l^{2'}|\theta_l^2)$ given as follows:

$$\begin{aligned} \theta_l^{2'}|\theta_l^2 &\sim N(\theta_l^2, \beta_l^2) \\ &= \frac{1}{\sqrt{2\pi\beta_l^2}} \exp \left(-\frac{[\theta_l^{2'} - \theta_l^2]^2}{2\beta_l^2} \right), \end{aligned} \quad (32)$$

($l = 1, 2, \dots, 5N_s$) for valid $\theta_l^{2'}$ (e.g., if the prior information encoded in the prior distribution requires the parameter to lie in a certain interval, then a proposal for this parameter that lies outside this interval is rejected). In Eq. (32), θ_l^2 ($l = 1, 2, \dots, 5N_s$) denotes the l -th component of θ^2 . The proposal distribution here provides a candidate $\theta_l^{2'}$ that is a

perturbation of the current value of θ_l^2 , which is obtained by drawing from a Gaussian distribution with variance β_l^2 . Each β_l determines the scale of the proposal steps for the nonlinear parameter θ_l^2 , and this scale needs to be carefully selected in order to ensure an appropriate acceptance probability for the proposed (propagation) move. We use proposal distributions that are mixtures of seven Gaussian distributions centered on θ_l^2 , each with the form given by Eq. (32). However, each Gaussian distribution of this mixture has a different standard deviation β , with the standard deviations chosen in such a manner as to cover (usually) several orders of magnitude.

Equations (29) and (32) completely specify one iteration involving the sampling of a new candidate source distribution model θ_{N_s} for fixed N_s (propagation move). In order to ensure that the Markov chain is reversible, the Gibbs and M-H samplers of Eqs. (29) and (32) are used systematically to update each parameter in θ^1 and θ^2 , respectively, in the forward direction for even iterations and in the reverse direction for odd iterations. More specifically, for odd iterations, the procedure for updating θ^1 in Eq. (29) is applied backwards in the order $k = N_s, N_s - 1, \dots, 1$ (and, similarly, for the update of θ^2 for odd iterations).

The propagation move described here only involved the update of θ_{N_s} for a fixed number of sources N_s . In order to move between configurations involving different numbers of sources, we need to use a reversible-jump MCMC procedure that allows updates between states of different dimensions in the hypothesis space (e.g., such as those associated with the creation of a new source or the annihilation of an existing source in the source distribution model).

4.2. Creation and annihilation moves

A reversible-jump MCMC sampling algorithm (which we use to accommodate between-model moves such as, for example, a change in the number of sources) was first introduced by Green (1995) in the context of the development of a methodology for addressing the model selection problem. The reversible-jump MCMC algorithm is very appealing in that it can be considered to be a natural generalization of the standard MCMC algorithm, with the generalization allowing not only moves in a parameter space of fixed dimension, but also ‘‘jumps’’ between model spaces M_{N_s} of different dimensions (viz., involving different numbers of sources N_s). More specifically, suppose a dimension-changing (or, between-model) move of type m is proposed, and the new state Θ' is generated by a deterministic invertible function $g(\Theta, \mathbf{v})$, where \mathbf{v} is a random vector with distribution $f(\mathbf{v})$ [so, $\mathbf{v} \sim f(\mathbf{v})$]. In other words, g is an operator that maps state Θ and the random vector \mathbf{v} into the new state Θ' . Green (1995) demonstrated that the acceptance probability of this proposed between-model move m has the following form:

$$\alpha(\Theta, \Theta') = \min \left\{ 1, \frac{p(\Theta'|\mathbf{D}, I)r_{m'}(\Theta')}{p(\Theta|\mathbf{D}, I)r_m(\Theta)f(\mathbf{v})} \left| \frac{\partial g(\Theta, \mathbf{v})}{\partial(\Theta, \mathbf{v})} \right| \right\}, \quad (33)$$

where m' denotes the reverse move to m and $r_m(\Theta)$ denotes the probability of choosing a move of type m in the state Θ . The final term in the ratio of Eq. (33) is the Jacobian which results from the change in variables associated with the dimension-changing move.

We consider two types of dimension-changing moves: namely, a creation move that results in the addition of a single new source to the current source distribution, and an

¹¹The symbol ‘‘ \sim ’’ is used to denote ‘‘is distributed as’’.

annihilation move that results in the removal of a single existing source from the current source distribution. To be more specific, a creation operator \mathcal{C} generates a between-model move from $(N_s, \theta_{N_s}) \in \mathbb{R}^{1+6N_s}$ in model M_{N_s} to $(N_s + 1, \theta_{N_s+1}) \in \mathbb{R}^{1+6N_s+6}$ in model M_{N_s+1} , so $\Theta' = (N_s + 1, \theta_{N_s+1}) \equiv \mathcal{C}(\Theta) = \mathcal{C}((N_s, \theta_{N_s}))$. Similarly, the reverse move, associated with the annihilation operator \mathcal{C}^\dagger , results in an existing source being deleted: so, $\Theta' = (N_s - 1, \theta_{N_s-1}) \equiv \mathcal{C}^\dagger(\Theta) = \mathcal{C}^\dagger((N_s, \theta_{N_s}))$ where $(N_s - 1, \theta_{N_s-1}) \in \mathbb{R}^{1+6N_s-6}$.

To move from model M_{N_s} to M_{N_s+1} using the creation operator \mathcal{C} , we propose the generation of a new source at a location \mathbf{x}_{s, N_s+1} with source strength Q_{N_s+1} , emitting material between the activation and deactivation times $T_b^{N_s+1}$ and $T_e^{N_s+1}$, respectively. The ‘‘coordinates’’ of the new source are generated by drawing random samples from some proposal density which we choose to be the prior density for each coordinate: so, from Eqs. (16) to (19) we sample $\mathbf{x}_{s, N_s+1} \sim p(\mathbf{x}_{s, N_s+1}|I)$, $Q_{N_s+1} \sim p(Q_{N_s+1}|I)$, $T_b^{N_s+1} \sim p(T_b^{N_s+1}|I)$, and $T_e^{N_s+1} \sim p(T_e^{N_s+1}|T_b^{N_s+1}, I)$. Now, let us assemble the parameters describing the new source as $\psi = (\mathbf{x}_{s, N_s+1}, T_b^{N_s+1}, T_e^{N_s+1}, Q_{N_s+1})$, so $\theta_{N_s+1} = \mathcal{C}(\theta_{N_s}) = (\theta_{N_s}, \psi)$ for random vector $\psi \in \mathbb{R}^6$ whose components are sampled as described above.

Suppose that at any iteration, we propose a creation move with probability p_C , with the reverse (annihilation) move having the probability p_{C^\dagger} . The proposal probability for the creation move is then¹²

$$\begin{aligned} q_C(\Theta, \Theta') &= p_C p(\mathbf{x}_{s, N_s+1}|I) p(T_b^{N_s+1}|I) \\ &\quad \times p(T_e^{N_s+1}|T_b^{N_s+1}, I) p(Q_{N_s+1}|I) \\ &= p_C \frac{1}{\mathcal{V}(\mathcal{D})} \frac{1}{T^*} \frac{1}{(T^* - T_b^{N_s+1})} p(Q_{N_s+1}|I). \end{aligned} \quad (34)$$

The probability $q_{C^\dagger}(\Theta', \Theta)$ for the reverse (annihilation) move is equal to the probability of choosing this move (p_{C^\dagger}) times the probability of picking a particular source from the $N_s + 1$ available sources for annihilation, so¹³

$$q_{C^\dagger}(\Theta', \Theta) = p_{C^\dagger} \frac{1}{N_s + 1}. \quad (35)$$

In view of Eq. (22), the ratio of posterior distributions in Eq. (33) for $\Theta' = \mathcal{C}(\Theta)$ has the following form:

$$\begin{aligned} \frac{p(\Theta'|\mathbf{D}, I)}{p(\Theta|\mathbf{D}, I)} &= \exp\left(-\frac{1}{2}(\chi^2(\Theta') - \chi^2(\Theta))\right) \\ &\quad \times \frac{p(N_s + 1|I)}{p(N_s|I)} \frac{1}{\mathcal{V}(\mathcal{D})} \frac{1}{T^*} \frac{1}{(T^* - T_b^{N_s+1})} \\ &\quad \times p(Q_{N_s+1}|I). \end{aligned} \quad (36)$$

Finally, if we substitute Eqs. (34), (35), and (36) in Eq. (33), and note that for the creation operator \mathcal{C} the Jacobian term in Eq. (33) is simply unity, the acceptance probability $\alpha(\Theta, \Theta')$ for the creation move $\Theta' = \mathcal{C}(\Theta)$ is given by

$$\begin{aligned} \alpha(\Theta, \Theta') &= \min\left\{1, \frac{p(N_s + 1|I)}{p(N_s|I)} \frac{p_{C^\dagger}}{p_C} \frac{1}{N_s + 1}\right. \\ &\quad \left. \times \exp\left(-\frac{1}{2}(\chi^2(\Theta') - \chi^2(\Theta))\right)\right\}. \end{aligned} \quad (37)$$

As demonstrated by Green (1995), a sufficient condition to ensure reversibility of the trans-dimensional Markov chain is for the acceptance ratio for the reverse move to be given by the reciprocal of that for the forward move. In the present context, the condition of detailed balance required for Markov chain reversibility implies that the acceptance probability, associated with the annihilation move $\Theta' = \mathcal{C}^\dagger(\Theta)$ for the removal of a source (from a source distribution containing N_s sources), must have the following form:

$$\begin{aligned} \alpha(\Theta, \Theta') &= \min\left\{1, \frac{p(N_s - 1|I)}{p(N_s|I)} \frac{p_C}{p_{C^\dagger}} N_s\right. \\ &\quad \left. \times \exp\left(-\frac{1}{2}(\chi^2(\Theta') - \chi^2(\Theta))\right)\right\}. \end{aligned} \quad (38)$$

It only remains now to specify the probabilities p_C and p_{C^\dagger} for creation and annihilation moves, respectively. Of course, with probabilities specified for the creation and annihilation moves, the probability p_P for the remaining propagation move (source parameters updated for a hypothesis space of fixed dimension) is determined as $p_P = 1 - p_C - p_{C^\dagger}$. Furthermore, it is useful to allow the probabilities p_C and p_{C^\dagger} to depend on the number of sources N_s in the current state. To indicate explicitly this dependence on N_s , we will augment the notation and express the probability of a creation and annihilation move as $p_C^{N_s}$ and $p_{C^\dagger}^{N_s}$, respectively. For $N_s = 1$, $p_{C^\dagger}^{N_s} = 0$ because at least one source must be responsible for the concentration measured by an array of sensors. Also, for $N_s = N_s^*$, $p_C^{N_s} = 0$ for otherwise the preassigned maximum number of sources that may be responsible for the measured concentration will be exceeded. For all other cases, the probabilities for creation and annihilation moves will be specified as follows:

$$\begin{aligned} p_C^{N_s} &= \frac{1}{2} \min\left\{1, \frac{p(N_s + 1|I)}{p(N_s|I)}\right\}, \\ p_{C^\dagger}^{N_s+1} &= \frac{1}{2} \min\left\{1, \frac{p(N_s|I)}{p(N_s + 1|I)}\right\}. \end{aligned} \quad (39)$$

Finally, it should be noted that with the dependence of p_C and p_{C^\dagger} on N_s as in Eq. (39), the ratios of these two probabilities in Eqs. (37) and (38) need to be interpreted as follows: namely, $p_{C^\dagger}/p_C \rightarrow p_{C^\dagger}^{N_s+1}/p_C^{N_s}$ in Eq. (37) and $p_C/p_{C^\dagger} \rightarrow p_C^{N_s-1}/p_{C^\dagger}^{N_s}$ in Eq. (38).

4.3. Parallel tempering and Metropolis-coupled MCMC

To overcome problems associated with a Markov chain that is slowly mixing in the hypothesis space and which can cause the chain to be stuck in a local minimum, we have implemented a form of parallel tempering to obtain a Metropolis-coupled MCMC algorithm that improves the speed at which the hypothesis space is explored. The Metropolis-coupled MCMC algorithm has been described by Geyer (1991) and is based on the idea of using a series of transition densities $T_1(\Theta'|\Theta), \dots, T_r(\Theta'|\Theta)$ with corresponding (unnormalized) invariant (stationary) distributions $p_1(\Theta|\mathbf{D}, I), \dots, p_r(\Theta|\mathbf{D}, I)$. In our current application of Metropolis-coupled MCMC, we run r Markov chains in parallel to give samples $\Theta_{[i]}^{(t)}$ ($i = 1, 2, \dots, r$), where chain i has the stationary distribution given by

$$p_i(\Theta|\mathbf{D}, I) = p(\Theta|I) p^{\lambda_i}(\mathbf{D}|\Theta, I), \quad i = 1, 2, \dots, r, \quad (40)$$

¹²This term should be identified with $r_m(\Theta)f(\mathbf{v})$ in Eq. (33).

¹³This term is identified with $r_{m'}(\Theta')$ in Eq. (33).

where $\lambda_i \in [0, 1]$ ($i = 1, 2, \dots, r$) is an increasing sequence (viz., $\lambda_i < \lambda_j$ for $i < j$) with $\lambda_1 = 0$ and $\lambda_r = 1$.

The parameter λ_i in Eq. (40) is interpreted as a tempering parameter. This parameter is used to raise the likelihood function to the λ_i power to give a modified posterior proportional to $p(\Theta|I)p^{\lambda_i}(\mathbf{D}|\Theta, I)$. It is noted that $\lambda_r = 1$ is associated with the desired target posterior probability distribution [cf. Eq. (22)] that we want to sample from. The other simulations correspond to a “ladder” of modified posterior distributions indexed by i . As the tempering parameter increases from zero to one, the effects of the concentration data are introduced slowly through the “softened” likelihood function $p^{\lambda_i}(\mathbf{D}|\Theta, I)$ [$\lambda_i \in [0, 1]$ for $i = 1, 2, \dots, r - 1$]. In particular, $\lambda_i < 1$ implies that the corresponding modified posterior distribution is broader (or flatter) than the actual posterior distribution, allowing the states to move more freely in the hypothesis space.

Let us denote the state of the i -th Markov chain at iteration t by $\Theta_{[i]}^{(t)}$ ($i = 1, 2, \dots, r$). In this form of Metropolis-coupled MCMC using parallel tempering, the r Markov chains are run simultaneously and at each iteration there is a prescribed probability (or, equivalently, on average every N_{iter} iterations) that a pair of adjacent chains (say, i and $i+1$ with $i = 1, 2, \dots, r-1$) is randomly selected, and a proposal is made to swap the states of these two chains. More specifically, if at iteration t , a proposed swapping operation between chains i and $i+1$ is accepted, then we swap the states of the chains $\Theta_{[i]}^{(t)} \rightarrow \Theta_{[i+1]}^{(t)}$ and $\Theta_{[i+1]}^{(t)} \rightarrow \Theta_{[i]}^{(t)}$ with an acceptance probability for this swap given by

$$\alpha_{\text{swap}} = \min \left\{ 1, \frac{p_i(\Theta_{[i+1]}^{(t)}|\mathbf{D}, I)p_{i+1}(\Theta_{[i]}^{(t)}|\mathbf{D}, I)}{p_i(\Theta_{[i]}^{(t)}|\mathbf{D}, I)p_{i+1}(\Theta_{[i+1]}^{(t)}|\mathbf{D}, I)} \right\}. \quad (41)$$

This swap enables the exchange of information across the population of r parallel simulations. More specifically, this allows the chain associated with the desired target posterior distribution (viz., that corresponding to $\lambda_r = 1$) to sample from remote regions of the posterior distribution, which in turn facilitates chain mobility in the hypothesis space and ensures a more reliable relaxation of the chain into the (potentially exponentially tiny) regions of this space where the posterior probability mass is concentrated. We apply the Metropolis-coupled MCMC algorithm described here with $r = 21$. The tempering parameters λ_i used on this ladder of parallel simulations are uniformly spaced between 0 and 1.

5. TESTS WITH SYNTHETIC CONCENTRATION DATA

In this section, we present the results of the application of our proposed algorithm for multiple source reconstruction for two cases: namely, case 1 involves two unknown continuous sources and case 2 involves two unknown instantaneous sources. In each of these cases, simulated concentration data will be generated and the Bayesian inference scheme described above will be applied to determine the number of unknown sources and for each of these sources to reconstruct the unknown source parameters.

We consider the simulation of concentration data sets corresponding to dispersion over a level and unobstructed terrain. Owing to the horizontal homogeneity of the terrain, the mean wind flow and turbulence statistics will be assumed to be horizontally homogeneous and stationary. Given the short times

and distances over which we will be modeling dispersion, this assumption is quite acceptable (and, indeed, at these distances the dispersion is assumed to occur entirely within the atmospheric surface layer). To simulate the concentration seen by the array of sensors, we use the forward-time LS model given by Eq. (3) with the drift coefficient vector \mathbf{a} corresponding to Thomson’s (1987) well-mixed three-dimensional forward LS model for Gaussian turbulence (viz., the background atmospheric velocity is assumed to be described by a multivariate Gaussian probability density function). The choice of the Kolmogorov constant $C_0 = 4.8$ is used for our simulations, the latter value having been obtained by calibrating the LS model against concentration data measured during a benchmark field experiment Project Prairie Grass (Wilson et al, 2001). We consider dispersion in a neutral wall shear layer, with the wind flow statistics required by the LS model prescribed in accordance with well-known surface-layer relations based on Monin-Obukhov theory (Stull, 1988).

For the two examples used to test the Bayesian inference algorithm for multiple source reconstruction, the concentration measured by a particular sensor in the array will result from the superposition of two or more partially overlapping plumes (or, clouds) produced by continuously-emitting (or, instantaneous) sources. It is noted that for multiple instantaneous point sources $T_e^k \rightarrow T_b^k$, whereas for multiple continuous point sources $T_b^k \rightarrow -\infty$ and $T_e^k \rightarrow \infty$ [for $k = 1, 2, \dots, N_s$; cf. Eqs. (10) and (11)]. In consequence, for a source distribution model consisting of multiple instantaneous point sources, the only relevant parameters are the source location $\mathbf{x}_{s,k}$, the release mass Q_k , and the release time $T_b^k = T_e^k \equiv T_s^k$ for each source ($k = 1, 2, \dots, N_s$); whereas, for a source distribution model consisting of multiple continuous point sources, the only relevant parameters are the source location $\mathbf{x}_{s,k}$ and the emission rate Q_k for each source ($k = 1, 2, \dots, N_s$). Finally, for the simulations, it is assumed that all the sources are emitting at ground level ($z = 0$) and that this is known *a priori* (viz., this knowledge is considered to be part of the background information I). As a result, the unknown location parameters for each source are its alongwind (x_s) and crosswind (y_s) positions, only.

5.1. Example 1: two unknown continuous sources

For the first example, we synthesized artificial concentration data for the case of two continuously-emitting sources which were located upwind of an array consisting of 42 detectors arranged as shown in Fig. 2. Each detector was placed at a height of 1.5 m above ground level. The two ground level continuously-emitting sources were located at $(x_s, y_s) = (-50, 0)$ m and at $(-250, 0)$ m and each source had an emission rate of $Q \equiv q_s = 1.0 \text{ g s}^{-1}$. The flow statistics for the neutrally-stratified atmospheric surface layer (through which the dispersion occurred) require the specification of two surface layer parameters: friction velocity $u_* = 0.25 \text{ m s}^{-1}$ and aerodynamic roughness length $z_0 = 0.015 \text{ m}$. The concentration data generated using the forward-time LS model (with these flow statistics as input) were embedded within white and normally distributed noise with a standard deviation equal to 10% of the true concentration amplitude.

We applied our proposed algorithm for multiple source reconstruction to the simulated concentration data. We perform Metropolis-coupled MCMC sampling with $r = 21$ chains with

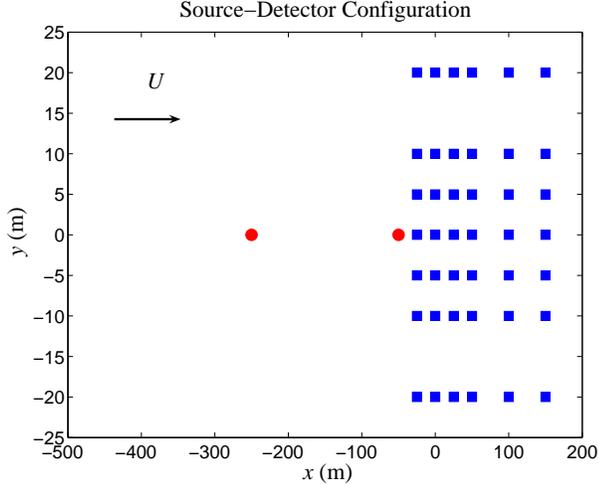


Figure 2: Two point sources located upwind of an array of 42 concentration detectors arranged as shown. A solid dot denotes a source and a solid square denotes a detector.

swaps between these chains attempted at every $N_{\text{iter}} = 25$ iterations on average. The proposed algorithm randomly initializes all unknown parameters (source location, emission rate) in accordance to the prescribed prior distributions for these parameters [cf. Eqs. (16) and (17)] with $Q^* = 100 \text{ g s}^{-1}$ and $\mathcal{D} = [-2000, -25] \text{ m} \times [-500, 500] \text{ m}$ providing the prior bounds on the emission rate q_s and on the source location (x_s, y_s) , respectively. The initial number of sources was randomly assigned in accordance to the prior distribution for N_s given by Eq. (15), where $N_s^* = 4$ (maximum allowable number of sources) and the hyperparameter $p^* = 1/3$ is chosen so that the expected number of sources is 1 (hence, the prior distribution favors the wrong choice for the number of sources).¹⁴

¹⁴From Eq. (15), the expected number of sources $\langle N_s \rangle = (N_s^* - 1)p^*$, so with $N_s^* = 4$ the choice $p^* = 1/3$ gives $\langle N_s \rangle = 1$.

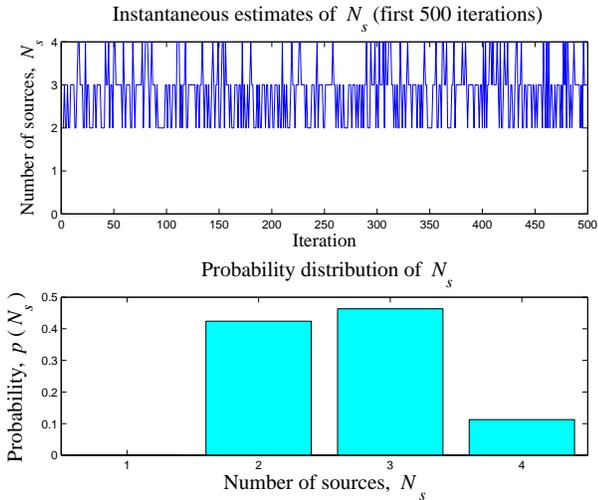


Figure 3: Instantaneous estimates of N_s for the first 500 post burn-in iterations of the reversible jump MCMC sampler (top) and the posterior probability distribution for the number of sources (bottom) estimated using the 50,000 post burn-in samples.

We have chosen the hyperparameter $\gamma = 1.0$ [cf. Eq. (16)] in the prior distribution for the emission rate (implying that no source in the source distribution model can be turned off). For this example, the MCMC algorithm was run for 100,000 iterations, with the first 50,000 iterations corresponding (conservatively) to the burn-in with the result that these samples were discarded in the subsequent analysis. The remaining 50,000 post burn-in samples were used for the posterior inference.

Figure 3 (top) displays changes in N_s (number of sources) as a function of the iteration number for the first 500 post burn-in iterations of the reversible-jump MCMC sampler. Note the dimension-changing moves involving creation of a new source (e.g., transitions from $N_s = 2$ to $N_s = 3$ or from $N_s = 3$ to $N_s = 4$) or annihilation of an existing source (e.g., transitions from $N_s = 4$ to $N_s = 3$ or from $N_s = 3$ to $N_s = 2$). Interestingly, after convergence of the Markov chain to the stationary distribution, moves across models from $N_s = 2$ to $N_s = 1$ (and, vice-versa) do not occur. Indeed, Fig. 3 (bottom) which exhibits the probability distribution, $p(N_s) \equiv p(N_s | \mathbf{D}, I)$, for the number of sources estimated from the 50,000 post burn-in samples, shows that the hypothesis $N_s = 1$ source is excluded. The simulations settle in a distribution which favors equally (approximately or better) the hypotheses $N_s = 2$ or 3, with a smaller probability for $N_s = 4$.

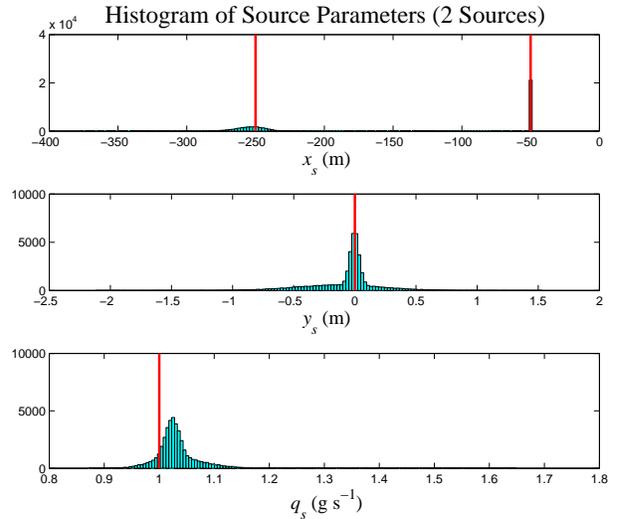


Figure 4: Histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel) obtained from the post burn-in samples of source distribution models with $N_s = 2$ (viz., from the samples of source distribution models having exactly two source atoms). The vertical lines in each panel mark the true values for the source parameters.

Figure 4 exhibits histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel). These histograms were constructed from the subset of post burn-in samples consisting of two source objects (i.e., samples with $N_s = 2$). The bimodality of the histogram of x_s with mode locations at $x_s \approx -50 \text{ m}$ and $x_s \approx -250 \text{ m}$ indicates that the alongwind positions of the two sources have been correctly identified. Because both sources are located at $y_s = 0 \text{ m}$ and have an emission rate $q_s = 1 \text{ g s}^{-1}$, the single mode in the histograms at $y_s \approx 0 \text{ m}$

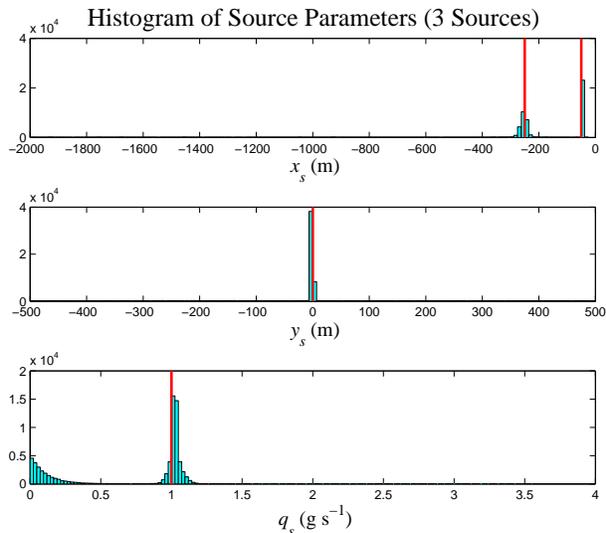


Figure 5: Histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel) obtained from the post burn-in samples of source distribution models with $N_s = 3$ (viz., from the samples of source distribution models having exactly three source atoms). The vertical lines in each panel mark the true values for the source parameters.

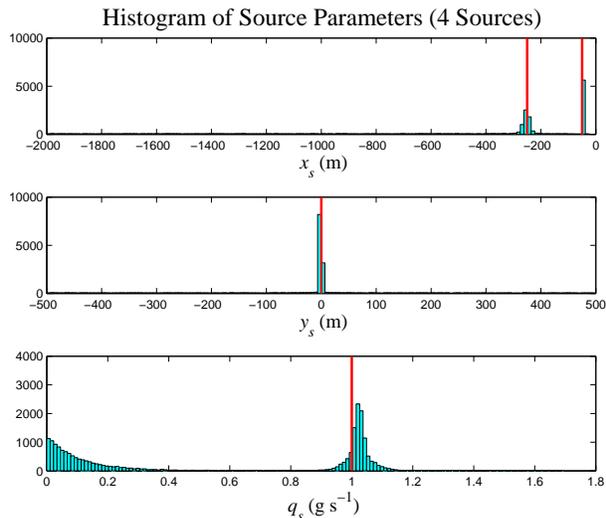


Figure 6: Histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel) obtained from post burn-in samples of source distribution models with $N_s = 4$ (viz., from the samples of source distribution models having exactly four source atoms). The vertical lines in each panel mark the true values for the source parameters.

and $q_s \approx 1.0 \text{ g s}^{-1}$ indicates that these source parameters have been correctly inferred for the two sources.

Figures 5 and 6 display histograms of the alongwind location x_s (top panel), crosswind location y_s (middle panel), and emission rate q_s (bottom panel). These histograms were obtained, respectively, from the subset of post burn-in source distribution model samples consisting of exactly three ($N_s = 3$)

Table 1: The posterior mean, posterior standard deviation (Std Dev.), and lower and upper bounds of the 95% HPD interval of the parameters $x_{s,k}$ (m), $y_{s,k}$ (m) and $q_{s,k}$ (g s^{-1}) for $k = 1, 2$ calculated from the totality of samples for each of the two identified clusters in Figs. 4, 5 and 6 for $N_s = 2, 3$, and 4, respectively. These are the parameters that characterize the two source atoms ($k = 1, 2$) associated with the two clusters of samples.

	Mean	Std Dev.	95% HPD
$k = 1$			
x_s	-254	11	(-276, -234)
y_s	-0.13	0.37	(-0.85, 0.58)
q_s	1.04	0.05	(0.95, 1.14)
$k = 2$			
x_s	-50	0.2	(-50.5, -49.6)
y_s	-0.00195	0.036	(-0.07, 0.07)
q_s	1.02	0.01	(0.99, 1.05)

and four ($N_s = 4$) source atoms. Note from both these figures that there really only exists two modes in the histograms of x_s , and it is seen that these modes correspond to the correct alongwind locations of the two sources. Furthermore, the extra source atom in the histograms in Fig. 5 and the two extra source atoms in the histograms of Fig. 6 are randomly distributed in \mathcal{D} (viz., are not concentrated into a significant cluster or clusters of points). Interestingly, these extra source atoms have emission rates associated with the mode in the histogram of q_s at $\approx 0 \text{ g s}^{-1}$ (viz., corresponding to “dim” source atoms). The mode in the histogram of q_s at $\approx 1.0 \text{ g s}^{-1}$ is associated with the modes in the histogram of x_s at $\approx -250 \text{ m}$ and $\approx -50 \text{ m}$ and the mode in the histogram of y_s at $\approx 0 \text{ m}$. Hence, even the source reconstruction based on samples of source distribution models with $N_s = 3$ or $N_s = 4$ only give two main clusters of points in \mathcal{D} located at the true positions of the two sources. These two main clusters in the histogram of x_s are associated with the modes in the histogram of y_s and q_s at $\approx 0 \text{ m}$ and $\approx 1.0 \text{ g s}^{-1}$, respectively (corresponding as such to the true crosswind positions and emission rates of the two sources).

From Figs. 4, 5, and 6 corresponding to the case $N_s = 2, 3$ and 4, respectively, we see that there exists two clusters along x_s in which the samples are concentrated. These two clusters are associated with two source atoms. We have gathered all samples of source distribution models for $N_s = 2, 3$, and 4 for each of these two clusters and plotted histograms of the parameters $\{x_s, y_s, q_s\}$ for each detected source atom in Figs. 7 and 8. The posterior mean, posterior standard deviation, and lower and upper bounds for the 95% HPD interval of the parameters for each detected source object were calculated from the samples for $N_s = 2, 3$, and 4 contained in each of the two identified clusters and the results are summarized in Table 1.¹⁵ From this information, we see that the two source

¹⁵It should be stressed that labels used for the source atoms in Table 1 [e.g., source atom 1 is identified to be the source at $(x_s, y_s) = (-250, 0) \text{ m}$, whereas source atom 2 is the source at $(x_s, y_s) = (-50, 0) \text{ m}$] are completely arbitrary here because the posterior probability distribution of the source parameters

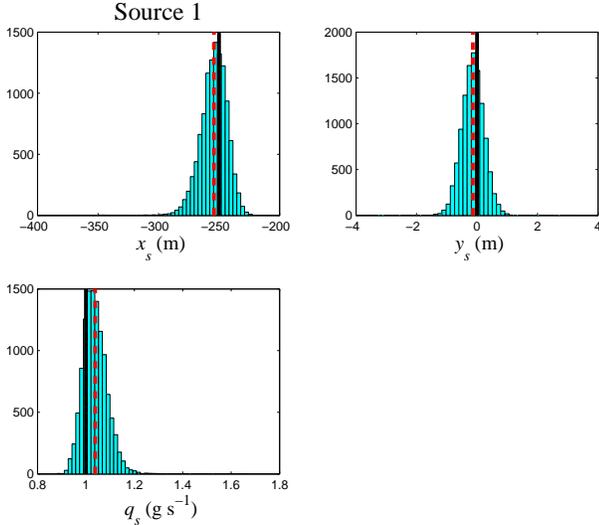


Figure 7: Histograms for the three parameters, namely along-wind location x_s (top left frame), crosswind wind location y_s (top right frame), and emission rate q_s (bottom left frame) that characterize source atom 1. The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

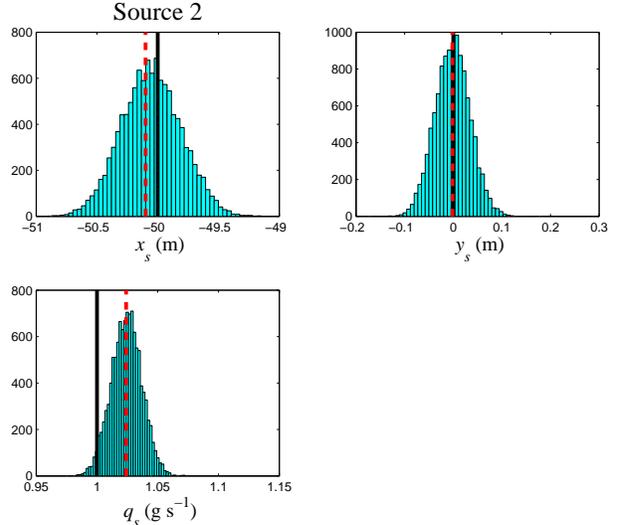


Figure 8: Histograms for the three parameters, namely along-wind location x_s (top left frame), crosswind wind location y_s (top right frame), and emission rate q_s (bottom left frame) that characterize source atom 2. The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the posterior mean of the associated marginal posterior distribution.

objects have been characterized to a reasonable accuracy.

5.2. Example 2: two unknown instantaneous sources

For our second example, we synthesized artificial concentration data using a forward-time LS model for the case of two instantaneous sources which were located upwind of an array consisting of 42 detectors arranged as shown in Fig. 2. This example is exactly the same as Example 1, except for the fact that the two continuous ground-level sources have been replaced with two instantaneous ground-level sources at $(x_s, y_s) = (-50, 0)$ m and at $(-250, 0)$ m. These two sources instantaneously released $q_s = 1$ g and 5 g of material at release times of $T_s = 10$ s and 2 s relative to an arbitrary time origin, respectively. The synthetic concentration-time data generated using the forward-time LS model were corrupted with white and normally distributed noise with a standard deviation equal to 10% of the true concentration amplitude.

We applied our proposed algorithm for multiple source reconstruction to this simulated concentration-time data set. The Metropolis-coupled reversible-jump MCMC algorithm used the same parameters as described previously for Example 1, with the following exceptions. The intermittency factor, γ , in Eq. (16) which is defined as the probability that the source is turned on (viz., $Q_k > 0$) is chosen to be 0.2, rather than 1.0 used in Example 1. This choice allows for the hypothesis that any given source in the domain can be turned off ($Q_k = 0$) with a finite probability ($= 1 - \gamma$) [viz., the given source does not release any material into the atmosphere and, hence, does not contribute to the concentration signal measured by the

[cf. Eq. (22)] is invariant under a reordering (relabelling) of the identifiers used for each source atom. This degeneracy simply corresponds to different (but equivalent) identifications of what is meant by source atom 1, source atom 2, etc.

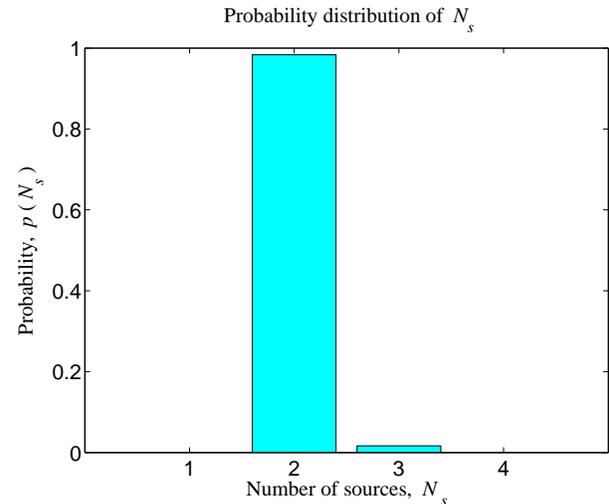


Figure 9: The posterior probability distribution for the number of sources estimated using the 50,000 post burn-in samples.

array of detectors]. Furthermore, for this example, the upper bounds for the release mass Q^* and release time T^* used in the definition of the prior uniform distributions for these two quantities [cf. Eqs. (16) and (18)] were chosen to be 100 g and 25 s, respectively. As in Example 1, the MCMC algorithm was run for 100,000 iterations, with the first 50,000 iterations corresponding to the burn-in with the result that these samples were discarded from the subsequent analysis. The remaining 50,000 post burn-in samples were used for the posterior inference, the results of which will be presented below.

Figure 9 displays the probability distribution for the number of sources $p(N_s) \equiv p(N_s | \mathbf{D}, I)$ for this example. The

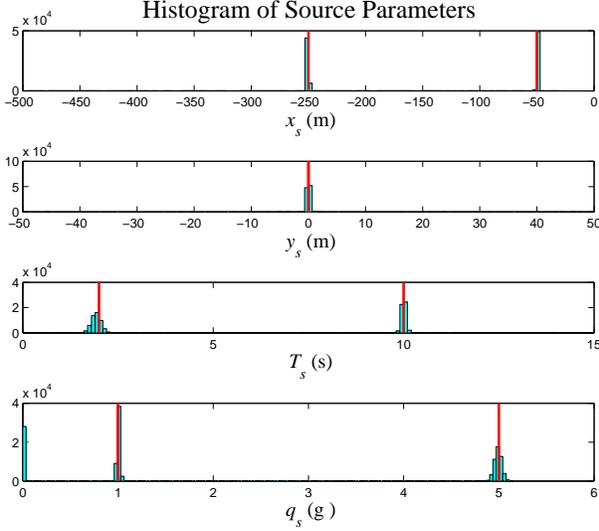


Figure 10: Histograms of the alongwind location x_s (top panel), crosswind location y_s (second panel), release time T_s (third panel), and release mass q_s (bottom panel) obtained from all the post burn-in samples of source distribution models extracted for Example 2. The vertical lines indicate the true values of the parameters for the two instantaneous sources.

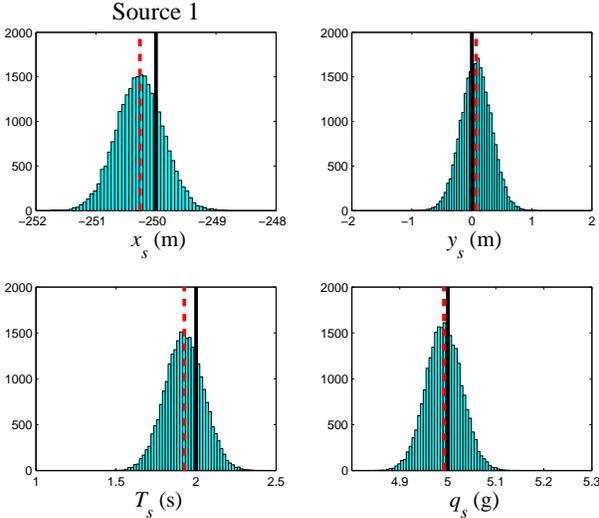


Figure 11: Histograms for the four parameters, namely alongwind location x_s (top left frame), crosswind wind location y_s (top right frame), release time T_s (bottom left frame), and release mass q_s (bottom right frame) that characterize source atom 1. The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

algorithm predicts the correct number of sources in this case (namely, $N_s = 2$) with a probability greater than about 0.98. The information embodied in the concentration data was sufficient to move the simulations towards a source distribution model having the correct number of sources. Interestingly, we found that for the prior specification of release mass with $\gamma = 0.2$, many of the samples for more complex models involv-

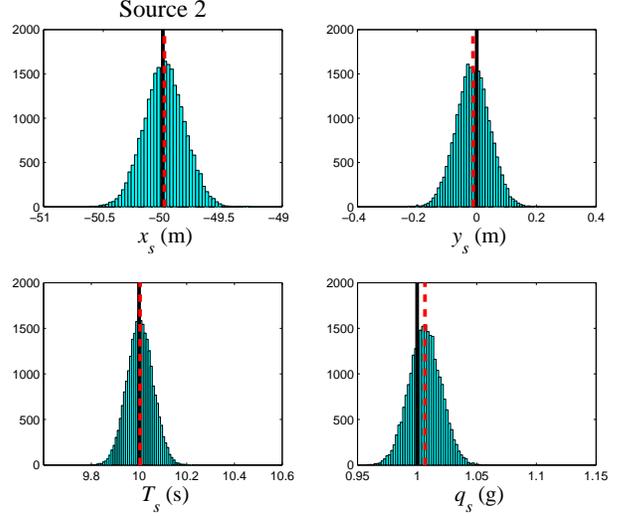


Figure 12: Histograms for the four parameters, namely alongwind location x_s (top left frame), crosswind wind location y_s (top right frame), release time T_s (bottom left frame), and release mass q_s (bottom right frame) that characterize source atom 2. The solid vertical line indicates the true value of the parameter and the dashed vertical line corresponds to the best estimate of the parameter obtained as the posterior mean of the associated marginal posterior distribution.

ing $N_s > 2$ sources had the additional sources turned off (viz., $Q_k = 0$). We note that for samples involving N_s sources, but with N_d of these sources turned off, the source model here is considered to have $N_s - N_d$ sources, and not N_s sources (owing to the fact that sources that are turned off do not contribute to the model concentration seen at the detectors). In consequence, some of the samples associated with the mode of $p(N_s) \equiv p(N_s | \mathbf{D}, I)$ in Fig. 9 at $N_s = 2$ actually corresponds to more complex models involving more than 2 sources, but with the additional sources turned off.

Figure 10 displays histograms of the source parameters $\{x_s, y_s, T_s, q_s\}$ constructed from all post burn-in samples of source distribution models extracted for Example 2. We note that Fig. 10 demonstrates that the samples drawn from the posterior distribution tend to cluster into two regions along x_s and T_s , which can be identified with two source atoms located at two different alongwind positions, with two distinct release times. Note that in the histogram of q_s , there are three distinct clusters. Two of these clusters correspond to two source atoms having two different release masses. The third cluster corresponds to a source atom with zero release mass. This cluster is associated with the additional source atoms in the samples of source distribution models having $N_s = 3$ or 4 source atoms, but with the additional source atoms “turned off” (viz., these sources did not release any mass, and hence did not contribute to the concentration measured by the detectors of the array).

Figures 11 and 12 exhibit histograms of the source parameters $\{x_s, y_s, T_s, q_s\}$ for the two detected source atoms in this example. Table 2 summarizes the recovered source parameters for each of these two sources in terms of the posterior mean, posterior standard deviation, and lower and upper bounds for the 95% HPD interval of the source parameters extracted from the marginal posterior distributions for the various param-

Table 2: The posterior mean, posterior standard deviation (Std Dev.), and lower and upper bounds of the 95% HPD interval of the parameters $x_{s,k}$ (m), $y_{s,k}$ (m), T_s^k (s) [measured relative to an arbitrary time origin], and $q_{s,k}$ (g) for $k = 1, 2$ calculated from samples of source distribution models with $N_s = 2$ (the latter corresponding to the most probable number of sources as inferred from Fig. 9).

	Mean	Std Dev.	95% HPD
$k = 1$			
x_s	-250.3	0.4	(-251.0, -249.5)
y_s	0.071	0.25	(-0.55, 0.43)
T_s	1.93	0.12	(1.69, 2.15)
q_s	4.99	0.04	(4.91, 5.06)
$k = 2$			
x_s	-50.0	0.17	(-50.3, -49.7)
y_s	-0.012	0.056	(-0.12, 0.095)
T_s	10.0	0.055	(9.9, 10.1)
q_s	1.00	0.013	(0.98, 1.03)

ters. The results shown here suggest that in this example the multiple source reconstruction algorithm was very effective in the sense that the correct number of sources was estimated, and for each of these sources the parameters characterising each source were recovered with very good accuracy.

6. CONCLUSIONS

In this paper, we have presented a Bayesian inference approach for multiple source reconstruction from a limited number of noisy concentration data obtained from an array of sensors for the case where the number of sources is unknown *a priori*. In this approach, we use a model that relates the source distribution to the concentration data (source-receptor relationship), and then apply Bayesian probability theory to formulate the posterior density function for the source parameters including the number of sources. The evaluation of the posterior density function and of its features of interest requires a numerical procedure. To this end, the computational algorithm required here is implemented using a Metropolis-coupled reversible-jump Markov chain Monte Carlo method to draw samples from the posterior density function. We showed how to design creation and annihilation moves that allow the Markov chain to jump between hypothesis spaces corresponding to different numbers of sources in the source distribution. The algorithm can be applied to problems involving a large number of source parameters (viz., for problems with N_s expected to be very large) implying that the method can be used for the reconstruction of volume (or area) sources with arbitrary *a priori* unknown geometry.¹⁶

¹⁶The source reconstruction technology described in this paper, which uses Bayesian probability theory and concentration sensor observations to answer the questions “How many discrete sources are present?” and “Given that there are N_s discrete sources, what are the best estimates of the locations, emission rates, release times, etc. of each source?”, can be viewed as a particular realization of the scientific method and a quantitative formulation of Occam’s razor. The methodology presented here utilizes the full

The new methodology has been successfully applied to simulated concentration data corresponding to continuously-emitting and instantaneous sources. It is shown that the proposed method performs well: the number of sources is correctly identified using the procedure, and the parameters (e.g., location, emission rate, release time) that characterize each identified source are reliably estimated. In addition, the methodology provides a rigorous determination of the uncertainty (e.g., standard deviation, credible intervals) in the inference of the source parameters, hence extending the potential of the methodology as a tool for quantitative source reconstruction.

7. ACKNOWLEDGEMENTS

The author wishes to acknowledge support from the Chemical Biological Radiological Nuclear Research and Technology Initiative (CRTI) Program under project number CRTI-02-0093RD.

REFERENCES

- Allen, C. T., G. S. Young, and S. E. Haupt, 2007: Improving pollutant source characterization by better estimating wind direction with a genetic algorithm. *Atmos. Environ.*, **41**, 2283–2289.
- Alpay, M. E. and M. H. Shor, 2000: Model-based solution techniques for the source localization problem. *IEEE Trans. Control Sys. Tech.*, **8**, 895–904.
- Bocquet, M., 2005a: Reconstruction of an atmospheric tracer source using the principle of maximum entropy. I: Theory. *Quart. J. Roy. Meteorol. Soc.*, **131**, 2191–2208.
- Bocquet, M., 2005b: Reconstruction of an atmospheric tracer source using the principle of maximum entropy. II: Applications. *Quart. J. Roy. Meteorol. Soc.*, **131**, 2209–2223.
- Carrigan, C. R., R. A. Heinle, G. B. Hudson, J. J. Nitao, and J. J. Zucca, 1996: Trace gas emissions on geological faults as indicators of underground nuclear testing. *Nature*, **382**, 528–531.
- Cox, R. T., 1946: Probability, frequency, and reasonable expectation. *Am. J. Phys.*, **14**, 1–13.
- Flesch, T. K., J. D. Wilson, and E. Yee, 1995: Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *J. Appl. Meteorol.*, **34**, 1320–1332.
- Gelman, A., J. Carlin, H. Stern, and D. Rubin, 2003: *Bayesian Data Analysis*, Chapman and Hall, CRC Press, Boca Raton, Florida, 668 pp.
- Gifford, F. A., 1959: Computation of pollution from several sources. *Int. J. Air Pollution*, **2**, 109–110.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter, 1996: *Markov Chain Monte Carlo in Practice*, Chapman and Hall, CRC Press, Boca Raton, Florida, 486 pp.
- Gordon, R., M. Leclerc, P. Schuepp, and R. Brunke, 1988: Field estimates of ammonia volatilization from swine manure by a simple micro-meteorological technique. *Can. J. Soil Sci.*, **68**, 369–380.

Bayesian probability theory to calculate the posterior probability for a particular source distribution model structure and, consequently, is fully general in the sense that it can be applied to reconstruct any arbitrary source distribution.

- Green, P. J., 1995: Reversible-jump Markov chain Monte Carlo computation and bayesian model determination. *Biometrika*, **82**, 711–732.
- Hanna, S. R., J. Chang, and D. G. Strimaitis, 1990: Uncertainties in source emission rate estimates using dispersion models. *Atmos. Environ.*, **24A**, 2971–2980.
- Hourdin, F. and J.-P. Issartel, 2000: Sub-surface nuclear tests monitoring through the CTBT xenon network. *Geophys. Res. Lett.*, **27**, 2245–2248.
- Kaharabata, S. K., P. H. Schuepp, and R. L. Desjardins, 2000: Source strength determination of a tracer gas using an approximate solution to the advection-diffusion equation for microplots. *Atmos. Environ.*, **34**, 2343–2350.
- Jaynes, E. T., 2003: *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 727 pp.
- Keats A., F.-S. Lien, E. Yee, 2006: Source determination in built-up environments through Bayesian inference with validation using the MUST array and Joint Urban 2003 tracer experiments. *Proceedings of the 14th Annual Conference of the Computational Fluid Dynamics Society of Canada*, Vol. 14, 8 pp.
- Keats A., E. Yee, F.-S. Lien, 2007a: Bayesian inference for source determination with applications to a complex urban environment. *Atmos. Environ.*, **41**, 465–479.
- Keats A., E. Yee, F.-S. Lien, 2007b: Efficiently characterizing the origin and decay rate of a nonconservative scalar using probability theory. *Ecological Modelling*, **205**, 437–452.
- Khapalov, A. V., 1994: Localization of unknown sources for parabolic systems on the basis of available observations. *Int. J. Systems Sci.*, **25**, 1305–1322.
- Matthes, J., L. Gröll, L. and H. B. Keller, 2005: Source localization by spatially distributed electronic noses for advection and diffusion. *IEEE Trans Signal Processing*, **53**, 1711–1719.
- Nodop, K., R. Connolly, and F. Girardi, 1998: The field campaigns of the European Tracer Experiment (ETEX): overview and results. *Atmos. Environ.*, **32**, 4095–4108.
- Rao, K. S., 2005: Uncertainty analysis in atmospheric dispersion modeling. *Pure Appl. Geophys.*, **162**, 1893–1917.
- Robertson, L. and J. Langner, 1998: Source function estimate by means of variational data assimilation applied to the ETEX-I tracer experiment. *Atmos. Environ.*, **32**, 4219–4225.
- Seibert, P. and A. Stohl, 1999: Inverse modeling of the ETEX-1 release with a Lagrangian particle model. Third GLOREAM Workshop, Ischia, Italy, 10 pp.
- Shea, D. A. and S. A. Lister, 2003: The BioWatch program: detection of bioterrorism. Congressional Research Service Report Number RL 32152. Available at <http://www.fas.org/sgp/crs/terror/RL32152.html>.
- Skiba, Y. N., 2003: On a method of detecting the industrial plants which violate prescribed emission rates. *Ecological Modelling*, **159**, 125–132.
- Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 666 pp.
- Taylor, G. I., 1921: Diffusion by continuous movements. *Proc. London Math. Soc.*, Ser. 2, **20**, 196–211.
- Thomson, D. J., 1987: Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.*, **180**, 529–556.
- Thomson, L. C., B. Hirst, G. Gibson, S. Gillespie, P. Jonathan, K. D. Skeldon, and M. J. Padgett, 2006: An improved algorithm for locating a gas source using inverse methods. *Atmos. Environ.*, **41**, 1128–1134.
- Van der Vink, G. E. and J. Park, 1994: Nuclear test ban monitoring: new requirements, new resources. *Science*, **263**, 634–635.
- Wilson, J. D. and W. K. N. Shum, 1992 A re-examination of the integrated horizontal flux method for estimating volatilization from circular plots. *Agric. Forest Meteorol.*, **57**, 281–295.
- Wilson, J. D., T. K. Flesch, and L. A. Harper, 2001: Micro-meteorological methods for estimating surface exchange with a disturbed windflow. *Agricultural and Forest Meteorology*, **107**, 207–225.
- Yee, E., 2007: Probabilistic inference: an application to the inverse problem of source function estimation. *The Technical Cooperation Program (TTCP), Chemical and Biological Defence Group, Technical Panel 9 Annual Meeting*, Defence Science and Technology Organization, Melbourne, Australia, 31 January – 4 February, 2005.
- Yee, E., 2006: A Bayesian approach for reconstruction of the characteristics of a localized pollutant source from a small number of concentration measurements obtained by spatially distributed “electronic noses”. *Russian-Canadian Workshop on Modeling of Atmospheric Dispersion of Wasp Agents*, Karpov Institute of Physical Chemistry, Moscow, Russia.
- Yee E., 2007: Bayesian probabilistic approach for inverse source determination from limited and noisy chemical or biological sensor concentration measurements. Chemical and Biological Sensing VIII (Augustus W. Fountain III, ed.), *Proc of SPIE*, Vol. 6554, 65540W, doi:10.1117/12.721630, 12 pp.
- Yee, E., F.-S. Lien, A. Keats, K. J. Hsieh, and R. D’Amours, 2006: Validation of Bayesian inference for emission source distribution reconstruction using the Joint Urban 2003 and European Tracer Experiments. *CWE2006: Proceedings of the Fourth International Symposium on Computational Wind Engineering*, **4**, 837–840.
- Yee, E., F.-S. Lien, A. Keats, and R. D’Amours, 2007: Bayesian inversion of concentration data: source reconstruction in the adjoint representation of atmospheric diffusion. *J. Wind Eng. Ind. Aerodyn.*, accepted for publication.