A bayesian inference procedure based on inverse dispersion modelling for source term estimation in built-up environments

François Septier, Patrick Armand, Christophe Duchenne

Version of Record: https://www.sciencedirect.com/science/article/pii/S1352231020304659
Manuscript_a061f879e3765df8560c6639fd70bafe

A Bayesian inference procedure based on inverse dispersion modelling for source term estimation in built-up environments

François Septier[a,*], Patrick Armand[b], Christophe Duchenne[b]

[a] *Université Bretagne Sud, LMBA UMR CNRS 6205, F-56000 Vannes, France*
[b] *CEA, DAM, DIF, F-91297 Arpajon, France*

Abstract

In atmospheric physics, reconstructing a pollution source is a challenging and important question. It provides better input parameters to dispersion models, and gives useful information to first-responder teams in case of an accidental toxic release. Various methods already exist, but using them requires an important amount of computational resources, especially when the accuracy of the dispersion model increases which is necessary in complex built-up environments. In this paper, a Bayesian probabilistic approach to estimate the location and the temporal emission profile of a pointwise source is proposed. More precisely, an Adaptive Multiple Importance Sampling (AMIS) algorithm is considered and enhanced by an efficient use of a Lagrangian Particle Dispersion Model (LPDM) in backward mode. Twin experiments empirically demonstrate the efficiency of the proposed inference strategy in very complex cases.

Keywords: Bayesian inference, Monte Carlo, STE, inverse dispersion model

1. Introduction

Chemical, radiological, biological, and nuclear (CRBN) releases into the atmosphere may result as a consequences of accidents or criminal activities. In such circumstances, it is essential to have a rapid and efficient identification of the source location as well as the strength of the emission. Indeed, by using these source term parameters as an input of an atmospheric dispersion model, prediction of the pollutant dispersion will provide invaluable information to first-responder teams. The problem consists in obtaining, as quickly as possible, an accurate estimation of the source parameters from noisy observations of concentration levels measured by a network of sensors.

Several strategies have been proposed to solve this challenging source term estimation (STE) problem. The majority of them give a single point estimate of the parameters by solving an optimization problem where a cost function has to be minimized using least squares or genetic algorithms, e.g (Efthimiou et al.,

*Corresponding author
Email address: francois.septier@univ-ubs.fr (François Septier)

2018; Kovalets et al., 2018; Winiarek et al., 2012). Unfortunately, such approaches do not allow us to quantify the uncertainty relative to the given estimates, which could be really problematic in such a context. To overcome this limitation, Bayesian algorithms have been designed to solve an inference problem by aiming at providing the complete probability density function of the parameters of interest given the observed measurements. Owing to the complex nature of the STE model, the exact computation of such a distribution is not feasible in practice, and one has to resort to some approximation techniques based on Monte Carlo methods (Kopka and Wawrzynczak, 2018; Delle Monache et al., 2008; Chow et al., 2008; Keats et al., 2007; Yee et al., 2014). In most of them, the authors propose to use a Markov Chain Monte Carlo (MCMC) kernel to obtain samples from the distribution of interest. However, these MCMC algorithms are known to suffer from several issues, such as the necessary burn-in period required for the convergence of the Markov chain to the correct target distribution, or the choice of its initialization. In this paper, we propose to use an other class of stochastic simulation techniques, based on the principle of importance sampling (IS). More specifically, we focus on an advanced technique called Adaptive Multiple Importance Sampling (AMIS) and recently proposed in (Cornuet et al., 2012). As presented in previous works (Rajaona et al., 2015, 2016), the application of such an adaptive technique on the challenging STE problem allows us to obtain significant gain compared to state-of-the-art algorithms in both synthetic and real data experiments.

Nevertheless, the computational complexity of such Monte Carlo techniques becomes prohibitive when naively applied to STE problem, since the likelihood of every generated samples with respect to the observed measurements has to be computed by running for each of them a forward dispersion model. In complex urban environment, elaborate but time consuming dispersion model which requires the generation of a large number of Lagrangian particles has to be considered. It is therefore of prime interest to design a fast Monte Carlo algorithm in such a challenging context.

In this paper, we present a complete strategy for efficient stochastic simulation techniques, aiming at optimizing the most time-consuming step in the algorithm by using the duality relationship with adjoint model for evaluating concentrations. Moreover, the output of the single initial run of the dispersion model in backward mode is also efficiently utilized both in the initialization step of the adaptive proposal distribution to improve the convergence speed. The rest of this paper is organized as follows. Section 2 describes the statistical model used for the STE problem. Section 3 is devoted to the description of the proposed Bayesian solution based on an adaptive mulitple importance sampling. Numerical experiments are conducted in Section 4. Conclusions are given in Section 5.

2. Problem Formulation

In this section we first present the statistical model of the source term estimation problem, and then develop the Bayesian framework for estimating the characteristics of the source.

2.1. Atmospheric dispersion model

In this study, we consider a point-wise and static source fully characterized by the parameter $\boldsymbol{\theta} = [\boldsymbol{x}_s, \boldsymbol{q}]$ where $\boldsymbol{x}_s = [x_s, y_s]$ is the spatial position of the source and $\boldsymbol{q}$ is the release rate vector resulting from the discretization of the plausible emission time interval into T_s time steps..

The concentration is considered to be observed by N_c sensors deployed over a 2-dimensional monitoring region. The measured concentration acquired by the i-th sensor at time t_j is defined as

$$y_{i,j} = \sum_{n=1}^{T_s} q_n C_{i,j}(\boldsymbol{x}_s, \Delta t_n) + \epsilon_{i,j}, \tag{1}$$

where $j = 1, \ldots, T_c$ with T_c the number of time samples collected by each sensor. Each measurement results from the superposition of the T_s releases at the different time steps $\{\Delta t_n\}_{n=1}^{T_s}$ weighted by their associated emission rates $\{q_n\}_{n=1}^{T_s}$ of the source plus an error term, $\epsilon_{i,j}$. $C_{i,j}(\boldsymbol{x}_s, \Delta t_n)$ corresponds therefore to the mean concentration observed by the i-th sensor at time t_j if a unitary release is made during the time step Δt_n from a source that is located at $\boldsymbol{x}_s$. The random variable term $\epsilon_{i,j}$ encompasses the three classical types of error: the dispersion modeling error, the observation error and the representativeness error due to the interpolation in both time and space of the dispersion model (Koohkan and Bocquet, 2012). As mentioned in (Yee, 2008), the choice of a Gaussian noise is justified by bringing forward the argument of the maximum entropy principle (Jaynes, 2003), which stipulates that such an assumption represents a maximally uninformative state of knowledge. All the measurements obtained at the different time samples of all sensors can be written in the following matrix form:

$$\boldsymbol{y} = \boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{q} + \boldsymbol{\epsilon}, \tag{2}$$

where $\boldsymbol{y} = \begin{bmatrix} y_{1,1} & \cdots & y_{1,T_c} & \cdots & y_{N_c,1} & \cdots & y_{N_c,T_c} \end{bmatrix}^T$ is the vector of observed concentration values and $\boldsymbol{C}(\boldsymbol{x}_s)$, generally called *source-receptor* matrix (Seibert and Frank, 2004), takes the following matrix form

$$\boldsymbol{C}(\boldsymbol{x}_s) = \begin{bmatrix} C_{1,1}(\boldsymbol{x}_s, \Delta t_1) & \cdots & C_{1,1}(\boldsymbol{x}_s, \Delta t_{T_s}) \\ \vdots & \ddots & \vdots \\ C_{1,T_c}(\boldsymbol{x}_s, \Delta t_1) & \cdots & C_{1,T_c}(\boldsymbol{x}_s, \Delta t_{T_s}) \\ \vdots & \ddots & \vdots \\ C_{N_c,1}(\boldsymbol{x}_s, \Delta t_1) & \cdots & C_{N_c,1}(\boldsymbol{x}_s, \Delta t_{T_s}) \\ \vdots & \ddots & \vdots \\ C_{N_c,T_c}(\boldsymbol{x}_s, \Delta t_1) & \cdots & C_{N_c,T_c}(\boldsymbol{x}_s, \Delta t_{T_s}) \end{bmatrix}. \tag{3}$$

As in (Yee, 2009), the likelihood distribution is given using a spatially and temporally independent zero-mean Gaussian multivariate random variable by

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \mathcal{N}\left(\boldsymbol{y}; \boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{q}, \sigma_\epsilon^2 \mathbf{I}_{N_c T_c}\right), \tag{4}$$

where $\mathcal{N}(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ corresponds to the multivariate normal distribution evaluated in $\boldsymbol{y}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ and $\mathbf{I}_{N_c T_c}$ represents the identity matrix of size $(N_c T_c \times N_c T_c)$.

The computation of the source-receptor matrix in Eq. (3) is an important part in an STE procedure as it links the source's characteristics with the measurements and quantifies the predicted concentration value at some location and time from a dispersion model for a given source. As a consequence, in stochastic simulation based inference techniques, this matrix has to be computed for each generated sample (at least several thousands, generally). The computation of this matrix with a Lagrangian particle dispersion model (LPDM) in a forward mode constitutes the most time-consuming step of the algorithm proposed in Rajaona et al. (2015).

In this study, we propose to use an alternative strategy which consists in using instead the backward mode of a LPDM. Using this backward mode is computationally advantageous if the number of receptors is less than the number of sources considered, which is generally the case in practice. Keats et al. (2007) and Yee et al. (2008) used also a receptor-oriented atmospheric transport model for the prediction of the source-receptor relationship in their Bayesian inference procedure for the rapid computation of $\boldsymbol{C}(\cdot)$.

2.2. A priori knowledge about model parameters

Our belief regarding the characteristics of the unknown state of interest, $\boldsymbol{\theta}$, is encapsulated within the prior probability distributions of the proposed Bayesian model. In this paper, we consider that the release could appear anywhere uniformly in the region of surveillance denoted here by $\Omega \subseteq \mathbb{R}^2$. As a consequence, the following uniform prior distribution is chosen for the position of the source:

$$p(\boldsymbol{x}_s) = \mathcal{U}_\Omega(\boldsymbol{x}_s). \tag{5}$$

Of course, in some scenarios of interest, it could be more appropriate to incorparate a more informative distribution to represent our initial guess about this source location (nuclear plants, industrial sites, etc).

Regarding now the emission rate vector, $\boldsymbol{x}_s$, as in Winiarek et al. (2011), a multivariate normal distribution is considered as a prior information:

$$p(\boldsymbol{q}) = \mathcal{N}\left(\boldsymbol{q}; \boldsymbol{\mu}_q, \boldsymbol{\Sigma}_q\right). \tag{6}$$

As pointed out in Bocquet (2008), this choice is a quite crude approximation since the emission rate cannot take negative values. However, this Gaussian assumption is often used in practice and generally leads to satisfactory performances (Issartel and Baverel, 2003).

2.3. Source term estimation in a Bayesian framework

In this work, a Bayesian solution is considered in order to solve efficiently this challenging problem. Instead of just a point-wise estimation of the source characteristics, $\boldsymbol{\theta}$, we are therefore interested in obtaining the full posterior distribution of the unknown parameters, $p(\boldsymbol{\theta}|\boldsymbol{y})$, which completely characterizes the available information on $\boldsymbol{\theta}$ given the measurements $\boldsymbol{y}$ obtained from all the sensors deployed in the field. With such a quantity, one can obtain all possible quantities of interest about the parameters such as, for example, point estimates or confidence intervals. In this problem, the posterior distribution of interest can be expanded as follows:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{x}_s, \boldsymbol{q}|\boldsymbol{y}) = p(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x}_s)p(\boldsymbol{x}_s|\boldsymbol{y}). \tag{7}$$

Owing to the Gaussian assumption of both the likelihood in Equation (4) and the prior distribution of $\boldsymbol{q}$ in Equation (6), the rule of conjugate priors states that the conditional posterior of the source emission rate $p(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x}_s)$ is therefore Gaussian and can thus be evaluated analytically as

$$p(\boldsymbol{q}|\boldsymbol{x}_s, \boldsymbol{y}) = \mathcal{N}(\boldsymbol{q}; \widetilde{\boldsymbol{\mu}}_{\boldsymbol{q}}, \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}}), \tag{8}$$

where parameters are obtained by:

$$\begin{aligned} \widetilde{\boldsymbol{\mu}}_{\boldsymbol{q}} &= \boldsymbol{\mu}_q + \boldsymbol{K}\left[\boldsymbol{y} - \boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{\mu}_q\right] \\ \widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}} &= \boldsymbol{\Sigma}_q - \boldsymbol{K}\boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{\Sigma}_q, \end{aligned} \tag{9}$$

with:

$$\boldsymbol{K} = \boldsymbol{\Sigma}_q \boldsymbol{C}(\boldsymbol{x}_s)^T \left[\boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{\Sigma}_q \boldsymbol{C}(\boldsymbol{x}_s)^T + \sigma_\epsilon^2 \mathbf{I}_{T_c \times N_c}\right]^{-1}. \tag{10}$$

Unfortunately, the second term $p(\boldsymbol{x}_s|\boldsymbol{y})$ in the complete posteriori distribution of interest in (7) is analytically intractable. Indeed, the dependence of the position of the source in the measurements is highly nonlinear due to the complex structure of the source-receptor matrix $\boldsymbol{C}(\boldsymbol{x}_s)$. By using such a decomposition, instead of having to approximate the full posterior distribution $p(\boldsymbol{x}_s, \boldsymbol{q}|\boldsymbol{y})$, only the posterior marginal distribution $p(\boldsymbol{x}_s|\boldsymbol{y})$ needs finally to be approximated since an analytical expression for $p(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x}_s)$ can be obtained. In this work, we consider efficient stochastic simulation based algorithms to approximate this complex marginal posterior distribution $p(\boldsymbol{x}_s|\boldsymbol{y})$.

3. Proposed Bayesian Algorithm to STE

In this section we first introduce the general principle of the Adaptive multiple importance sampling algorithm (AMIS), then describe the complete Bayesian solution based on AMIS for STE which was originally proposed in Rajaona et al. (2015) and finally we present the proposed strategy to enhance both the convergence speed and the robustness.

3.1. General Principle of AMIS

3.1.1. Importance Sampling

The basic idea of *Importance Sampling* (IS) is to estimate statistical quantities with respect to a specific *target distribution* π, while only having samples drawn from a different distribution $\phi(\cdot)$, called the *proposal distribution* (Robert and Casella, 2004). More specifically, N_p samples $(\boldsymbol{x}^1, \ldots, \boldsymbol{x}^{N_p})$, also called *particles*, are generated from $\phi(\cdot)$ and then, in order to compensate for the fact that we have sampled from a distribution which is not the target one, an *importance* weight is assigned to each particle, as follows for $i = 1, \ldots, N_p$:

$$w^i = \frac{\pi(\boldsymbol{x}^i)}{\phi(\boldsymbol{x}^i)}, \tag{11}$$

As a consequence, the target distribution can be approximated by this principle of IS with the following empirical measure:

$$\pi(\boldsymbol{x}) \approx \sum_{i=1}^{N_p} \widetilde{w}^i \delta_{\boldsymbol{x}^i}(d\boldsymbol{x}), \tag{12}$$

and any expectation of some function $h(\cdot)$ with respect to the target distribution is estimated by

$$\mathbb{E}_\pi\left[h(\boldsymbol{x})\right] = \int h(\boldsymbol{x})\pi(\boldsymbol{x})d\boldsymbol{x} \approx \sum_{i=1}^{N_p} \widetilde{w}^i h(\boldsymbol{x}^i), \tag{13}$$

where the $\widetilde{w}^i$ are the normalized importance weights (i.e. $\widetilde{w}^i = w^i[\sum_{j=1}^{N_p} w^j]^{-1}$) and $\delta(\cdot)$ is the Dirac function. In this paper, we propose to approximate the complex marginal posterior distribution of the source location, $p(\boldsymbol{x}_s|\boldsymbol{y})$, by an IS procedure:

$$p(\boldsymbol{x}_s|\boldsymbol{y}) \approx \sum_{i=1}^{N_p} \widetilde{w}^i \delta_{\boldsymbol{x}_s^i}(d\boldsymbol{x}_s), \tag{14}$$

with $\left\{\boldsymbol{x}_s^i, \widetilde{w}^i\right\}_{i=1}^{N_p}$ being the random weighted samples from an IS-based algorithm. By plugging this apprxiamtion into Eq. (7), the complete posterior distribution of interest will therefore be estimated as:

$$p(\boldsymbol{x}_s, \boldsymbol{q}|\boldsymbol{y}) \approx \sum_{i=1}^{N_p} \widetilde{w}^i p(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x}_s^i)\delta_{\boldsymbol{x}_s^i}(d\boldsymbol{x}_s). \tag{15}$$

IS requires to choose a proposal distribution that we can easily sample from. This choice is of course crucial as poor performances can be easily obtained when the proposal distribution is not appropriately chosen. To obtain a good empirical approximation of the target distribution, the proposal distribution should be close to the target distribution. Indeed, a high discrepancy between the target density and the importance density will result to an increasing variability of the importance weights which may strongly affect the accuracy of the final estimate. A careful design of this proposal distribution is therefore necessary but generally difficult to do.

3.1.2. The AMIS algorithm: an adaptive method

To overcome the difficulty of designing an appropriate proposal distribution, some adaptive procedures have been proposed in the context of IS. Originally referred to as *Population Monte Carlo algorithms* (PMC) in (Cappé et al., 2004), the objective consists in iteratively adapting the proposal distribution by learning some information from the samples that have been generated. More specifically, PMC methods are iterative IS-based sampling techniques: 1) at each step, samples from a proposal distribution are generated and a weight is assigned to each of them according to the IS identity, and 2) the proposal distribution is thus adapted using these random weighted samples in order to make it closer to the target distribution. Since the seminal paper (Cappé et al., 2004), several variants have been proposed, see (Bugallo et al., 2017) for a detailed overview. Recently, a sophisticated adaptive importance sampler, named *Adaptive Multiple Importance Sampling* (AMIS), has been proposed in (Cornuet et al., 2012). The main novelty of the AMIS is in the use of a recycling mechanism that allow us to use all the samples that have been generated until now to improve both the adaptivity of the proposal distribution and the variance of the resulting estimator. In the classical PMC algorithm, only the particles drawn during the last iteration are used for the adaptation procedure. By using such a recycling

strategy, a significant improvement could be obtained by the AMIS as shown in (Cornuet et al., 2012).

As in (Rajaona et al., 2015), the AMIS algorithm will be used to approximate the marginal posterior distribution of the source location

$$\pi(\boldsymbol{x}_s) = p(\boldsymbol{x}_s|\boldsymbol{y}) \propto p(\boldsymbol{y}|\boldsymbol{x}_s)p(\boldsymbol{x}_s), \tag{16}$$

which therefore corresponds to the target distribution. The prior distribution is defined in Equation (5) and the marginal likelihood is obtained using Equations (4) and (6) as follows:

$$\begin{aligned} p(\boldsymbol{y}|\boldsymbol{x}_s) &= \int p(\boldsymbol{y}|\boldsymbol{x}_s, \boldsymbol{q})p(\boldsymbol{q})d\boldsymbol{q} \\ &= \mathcal{N}\left(\boldsymbol{y}; \boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{\mu}_q, \boldsymbol{C}(\boldsymbol{x}_s)\boldsymbol{\Sigma}_q\boldsymbol{C}(\boldsymbol{x}_s)^T + \sigma^2_{\boldsymbol{\epsilon}}\mathbf{I}_{T_c \times N_c}\right). \end{aligned} \tag{17}$$

In the AMIS algorithm, A parametric proposal distribution has to be chosen. This distribution has to be flexible enough and easy to sample from in order to obtain satisfactory performances. In (Rajaona et al., 2015) following the idea presented in (Cappé et al., 2008), a mixture of D distributions, with parameters adapted at each iteration, was considered. In this work, we introduce an additional "defensive" component which will remain unchanged through the iteration of the algorithm

$$\phi(\boldsymbol{x}_s|\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \alpha^{(0)}\phi^{(0)}(\boldsymbol{x}_s) + (1 - \alpha^{(0)}) \sum_{d=1}^{D} \alpha^{(d)}\phi^{(d)}(\boldsymbol{x}_s|\boldsymbol{\mu}^{(d)}, \boldsymbol{\Sigma}^{(d)}). \tag{18}$$

The aim of the static component, $\alpha^{(0)}\phi^{(0)}(\boldsymbol{x}_s)$, is to guarantee that the importance function remains bounded by $\pi(\boldsymbol{x}_s)/(\alpha^{(0)}\phi^{(0)}(\boldsymbol{x}_s))$ whatever happens during the adaptation, thus guaranteeing a finite variance. It is preferable to keep $\alpha^{(0)}$ as low as possible (e.g. $\alpha^{(0)} = 0.1$) to not limit the performances achievable by the adaptation procedure.

In this paper, each component of the mixture is considered to be a bivariate Gaussian distribution. The adaptation of the proposal distribution will thus consists in updating the following parameters:

- the vector of mixture weights $\boldsymbol{\alpha} = \{\alpha^{(1)}, \ldots, \alpha^{(D)}\}$ such as $\sum_{d=1}^{D} \alpha^{(d)} = 1$,
- the parameters of each D bivariate Gaussian distribution, i.e. their mean vector and covariance matrix: $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \left\{(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}^{(1)}), \ldots, (\boldsymbol{\mu}^{(D)}, \boldsymbol{\Sigma}^{(D)})\right\}$.

An initial tunning of these parameters is required before running the AMIS algorithm. Unfortunately, this starting distribution generally has a major impact on the performances of such an adaptive algorithm. In our previous work (Rajaona et al., 2015), the parameters of the proposal was set up so that the generated samples cover roughly all the surveillance area since the source was considered to be located a priori uniformly in the area. In this work, we propose to use a novel efficient strategy in order to initialize this proposal distribution, so that the algorithm converges more rapidly to the region of interest. This method of initializing the parameters as well as the choice of $\phi^{(0)}(\cdot)$ will be described in Section 3.2.

In summary, the following steps are performed during the k-th iteration of the AMIS:

1. N_p particles, $\boldsymbol{x}_{s,k}^i$, are sampled from the current proposal ϕ_k,
2. An importance weight w_k^i is computed for each of these samples,
3. A recycling mechanism is performed on the previous importance weights, $\boldsymbol{w}_{0:k-1}$, by using a correction that take into account the current form of the proposal,
4. The adaptation of the proposal distribution is done by updating its parameters $(\boldsymbol{\alpha}_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ into $(\boldsymbol{\alpha}_{k+1}, \boldsymbol{\mu}_{k+1}, \boldsymbol{\Sigma}_{k+1})$ using all the collection of weighted samples $\left\{\widetilde{w}_n^i, \boldsymbol{x}_{s,n}^i\right\}_{n=0,\ldots,k}^{i=1,\ldots,N_p}$.

In order to perform the adaption of the proposal distribution between two successive iterations, the most popular choice of measure to quantify some distance between two probability density functions is the *Kullback-Leibler* (KL) divergence Cappé et al. (2008), defined as:

$$\mathcal{D}_{KL}(\pi \parallel \phi) = \int \log\left(\frac{\pi(\boldsymbol{x}_s)}{\phi(\boldsymbol{x}_s)}\right) \pi(\boldsymbol{x}_s) d\boldsymbol{x}_s. \tag{19}$$

By minimizing this divergence through the iterations of the algorithm, the parameters of the proposal distribution will be set, so that this distribution will become closer and closer to the target distribution. In adaptive IS algorithms, the integral in Eq. (19) is approximated using the population of samples obtained until now. As a consequence, when the proposal distribution used is the one defined in Eq. (18) with Gaussian distributions, the updating mechanism is given at the end of the k-iteration and for $d = 1, \ldots, D$, by:

$$\begin{aligned}
\alpha_{k+1}^{(d)} &= \frac{\widetilde{\alpha}_{k+1}^{(d)}}{\sum_{l=1}^{D} \widetilde{\alpha}_{k+1}^{(l)}} \quad \text{with} \quad \widetilde{\alpha}_{k+1}^{(d)} = \sum_{n=1}^{k} \sum_{i=1}^{N_p} \widetilde{w}_n^i \rho_{n,i}^{(d)} \\
\boldsymbol{\mu}_{k+1}^{(d)} &= \frac{\sum_{n=1}^{k} \sum_{i=1}^{N_p} \widetilde{w}_n^i \rho_{n,i}^{(d)} \boldsymbol{x}_{s,n}^i}{\widetilde{\alpha}_{k+1}^{(d)}} \\
\boldsymbol{\Sigma}_{k+1}^{(d)} &= \frac{\sum_{n=1}^{k} \sum_{i=1}^{N_p} \widetilde{w}_n^i \rho_{n,i}^{(d)} (\boldsymbol{x}_{s,n}^i - \boldsymbol{\mu}_{k+1}^{(d)})(\boldsymbol{x}_{s,n}^i - \boldsymbol{\mu}_{k+1}^{(d)})^T}{\widetilde{\alpha}_{k+1}^{(d)}},
\end{aligned} \tag{20}$$

where $\rho_{n,i}^{(d)}$ is an intermediary value corresponding to the probability of a given particle belonging to the d-th component of the mixture and is given by:

$$\rho_{n,i}^{(d)} = \frac{(1-\alpha^{(0)})\alpha_k^{(d)} \phi_k^{(d)}(\boldsymbol{x}_{s,n}^i | \boldsymbol{\mu}_k^{(d)}, \boldsymbol{\Sigma}_k^{(d)})}{\alpha^{(0)}\phi^{(0)}(\boldsymbol{x}_s) + (1-\alpha^{(0)}) \sum_{m=1}^{D} \alpha_k^{(m)} \phi_k^{(m)}(\boldsymbol{x}_{s,n}^i | \boldsymbol{\mu}_k^{(m)}, \boldsymbol{\Sigma}_k^{(m)})}. \tag{21}$$

Since the collection of samples obtained after K iterations of the proposed AMIS approximates the posterior marginal distribution of the source location as

$$p(\boldsymbol{x}_s | \boldsymbol{y}) \approx \sum_{n=1}^{K} \sum_{i=1}^{N_p} \widetilde{w}_n^i \delta_{\boldsymbol{x}_{s,n}^i}(d\boldsymbol{x}_s), \tag{22}$$

the complete posterior distribution of interest including both the position and the release-rate parameters, is thus approximated by:

$$p(\boldsymbol{x}_s, \boldsymbol{q} | \boldsymbol{y}) \approx \sum_{n=1}^{K} \sum_{i=1}^{N_p} \widetilde{w}_n^i p(\boldsymbol{q} | \boldsymbol{x}_{s,n}^i, \boldsymbol{y}) \delta_{\boldsymbol{x}_{s,n}^i}(d\boldsymbol{x}_s), \tag{23}$$

with $p(\boldsymbol{q}|\boldsymbol{x}^i_{s,n}, \boldsymbol{y})$ is the distribution defined in Eq. (8). The AMIS algorithm is summarized in Algo 1.

Algorithm 1: The AMIS algorithm

Initialization:

Generate N_p particles $\left\{\boldsymbol{x}^i_{s,0}\right\}$ from initial proposal $\phi_0(\cdot|\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$

for $i = 1 : N_p$ **do**

$$\delta^i_0 = N_p \phi_0(\boldsymbol{x}^i_{s,0}|\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

$$w^i_0 = \frac{p(\boldsymbol{y}|\boldsymbol{x}^i_{s,0})p(\boldsymbol{x}^i_{s,0})}{\phi_0(\boldsymbol{x}^i_{s,0}|\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)}$$

end

Normalize the importance weights: $\widetilde{w}^i_0 = w^i_0[\sum_{j=1}^{N_p} w^j_0]^{-1}$

Update the proposal's parameters using Equation (20) to obtain $(\boldsymbol{\alpha}_1, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$

Iterations:

for $k = 1 : K$ **do**

Generate N_p particles $\left\{\boldsymbol{x}^i_{s,k}\right\}$ from the proposal $\phi_k(\cdot|\boldsymbol{\alpha}_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$

for $i = 1 : N_p$ **do**

$$\delta^i_k = N_p \phi_0(\boldsymbol{x}^i_{s,k}|\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) + \sum_{l=1}^{k} N_p \phi_l(\boldsymbol{x}^i_{s,k}|\boldsymbol{\alpha}_l, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)$$

$$w^i_k = (k+1)N_p \frac{p(\boldsymbol{y}|\boldsymbol{x}^i_{s,k})p(\boldsymbol{x}^i_{s,k})}{\delta^i_k}$$

end

for $l = 0, \ldots, k-1$ *and* $i = 1, \ldots, N_p$ **do**

Update the importance weights that were previously generated:

$$\delta^i_l = \delta^i_l + N\phi_k(\boldsymbol{x}^i_{s,l}|\boldsymbol{\alpha}_k, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$

$$w^i_l = (k+1)N_p \frac{p(\boldsymbol{y}|\boldsymbol{x}^i_{s,l})p(\boldsymbol{x}^i_{s,l})}{\delta^i_l}$$

end

Normalize the importance weights, $\forall\ l = 0, \ldots, k$ and $i = 1, \ldots, N_p$: $\widetilde{w}^i_l = w^i_l[\sum_{m=1}^{k}\sum_{j=1}^{N_p} w^j_k]^{-1}$

Update the proposal's parameters using Equation (20) to obtain $(\boldsymbol{\alpha}_{k+1}, \boldsymbol{\mu}_{k+1}, \boldsymbol{\Sigma}_{k+1})$

end

Output: The entire collection of weighted particles $\left\{\widetilde{w}^i_n, \boldsymbol{x}^i_{s,n}\right\}_{n=0,\ldots,K}^{i=1,\ldots,N_p}$ to approximate the posterior marginal distribution of the source location as described in Equation (22)

3.2. Efficient initialization

In this section, we propose an efficient procedure to automatically set the initial parameters of the adaptive proposal distribution of the AMIS in Eq. (18). As already mentioned by Cornuet et al. (2012), the starting distribution has clearly a major impact on the resulting performances of such adaptive sampling algorithms. Indeed, it is quite difficult to recover from a poor starting sample since the adaptivity is only based on the visited regions of the simulation space.

The initialization consists in setting the parameters of the D bivariate Gaussian distributions $(\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ as well as choosing the "defensive" distribution $\phi^{(0)}(\cdot)$. In this work, we propose to use the results of pollutant concentration levels obtained from the run in backward mode of the dispersion model. Indeed, the runs of receptor-oriented atmospheric transport model allows us to obtain $N_c \times T_c \times T_s$ maps of concentration levels on the domain of interest which is discretized spatially on a 2-D mesh, denoted by $\mathcal{X}$. An histogram is obtained by normalizing an aggregated map that is the the average of all these intermediate concentration maps. The complete procedure is detailed in Algo 2.

Algorithm 2: Construction of the backward two-dimensional concentration map $\mathcal{B}(\cdot)$

Inputs: $\rho_{\text{detec}} > 0$, $\rho_{\min} > 0$, $\gamma \in [0, 1]$

$\forall \boldsymbol{x}_c \in \mathcal{X}, n = 1, \ldots, T_s$:

for $j = 1, \ldots, T_c$ **do**

  for $i = 1, \ldots, N_c$ **do**

    if $y_{i,j} > \rho_{detec}$ **then**

      $B_{i,j}(\boldsymbol{x}_c, \Delta t_n) = \begin{cases} 1 & \text{if } C_{\boldsymbol{x}_c, \Delta t_n}(\boldsymbol{x}_i, \Delta t_j) > \rho_{\min} \\ 0 & \text{otherwise} \end{cases}$

    else

      $B_{i,j}(\boldsymbol{x}_c, \Delta t_n) = -\alpha \frac{\rho_{\text{detec}} - y_{i,j}}{\rho_{\text{detec}}}$

    end

  end

end

Output: $\mathcal{B}(\boldsymbol{x}_c) = \frac{\sum_{i=1}^{N_c} \sum_{j=1}^{T_c} \sum_{n=1}^{T_s} B_{i,j}(\boldsymbol{x}_c, \Delta t_n)}{\sum_{\boldsymbol{x}_c \in \mathcal{X}} \sum_{i=1}^{N_c} \sum_{j=1}^{T_c} \sum_{n=1}^{T_s} B_{i,j}(\boldsymbol{x}_c, \Delta t_n)}$

Now, in order to initialize the parameters of the D bivariate Gaussian distribution, we propose to find the ones that define the mixture with the minimum of Kullback-Leibler distance with the histogram $B(\cdot)$ obtained with Algo 2. By defining,

$$\phi_{Adapt}(\boldsymbol{x}_s; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \sum_{d=1}^{D} \alpha^{(d)} \phi^{(d)}(\boldsymbol{x}_s | \boldsymbol{\mu}^{(d)}, \boldsymbol{\Sigma}^{(d)}), \tag{24}$$

the parameters are obtained by:

$$\begin{aligned}
(\boldsymbol{\alpha}_0, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) &= \underset{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}}{\operatorname{argmin}} \, \mathcal{D}_{KL}\left(B(\boldsymbol{x}_s) || \phi_{Adapt}(\boldsymbol{x}_s; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})\right) \\
&= \underset{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}}{\operatorname{argmin}} \int_{\Omega} \log\left(\frac{\mathcal{B}(\boldsymbol{x}_s)}{\phi_{Adapt}(\boldsymbol{x}_s; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})}\right) \mathcal{B}(\boldsymbol{x}_s) d\boldsymbol{x}_s \\
&= \underset{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}}{\operatorname{argmax}} \int_{\Omega} \log\left(\phi_{Adapt}(\boldsymbol{x}_s; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma})\right) \mathcal{B}(\boldsymbol{x}_s) d\boldsymbol{x}_s \\
&= \underset{\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Sigma}}{\operatorname{argmax}} \int_{\Omega} \log\left(\sum_{d=1}^{D} \alpha^{(d)} \phi^{(d)}(\boldsymbol{x}_s | \boldsymbol{\mu}^{(d)}, \boldsymbol{\Sigma}^{(d)})\right) \mathcal{B}(\boldsymbol{x}_s) d\boldsymbol{x}_s.
\end{aligned} \tag{25}$$

This optimization cannot be solved analytically but an iterative procedure (in the same spirit as the Expectation-Maximization) can be used to find a

solution of this problem. By considering that the mesh is relatively fine, the parameters of the d-th component can be updated at the n-th iteration as follows:

$$\begin{aligned}
\alpha_n^{(d)} &= \sum_{\boldsymbol{x}_c \in \mathcal{X}} \mathcal{B}(\boldsymbol{x}_c)\gamma_{\boldsymbol{x}_c,n}^{(d)} \\
\boldsymbol{\mu}_n^{(d)} &= \frac{\sum_{\boldsymbol{x}_c \in \mathcal{X}} \mathcal{B}(\boldsymbol{x}_c)\gamma_{\boldsymbol{x}_c,n}^{(d)} \boldsymbol{x}_c}{\alpha_n^{(d)}} \\
\boldsymbol{\Sigma}_n^{(d)} &= \frac{\sum_{\boldsymbol{x}_c \in \mathcal{X}} \mathcal{B}(\boldsymbol{x}_c)\gamma_{\boldsymbol{x}_c,n}^{(d)} \left(\boldsymbol{x}_c - \boldsymbol{\mu}_n^{(d)}\right)\left(\boldsymbol{x}_c - \boldsymbol{\mu}_n^{(d)}\right)^T}{\alpha_n^{(d)}},
\end{aligned} \tag{26}$$

with

$$\gamma_{\boldsymbol{x}_c,n}^{(d)} = \frac{\alpha_{n-1}^{(d)} \mathcal{N}\left(\boldsymbol{x}_c; \boldsymbol{\mu}_{n-1}^{(d)}, \boldsymbol{\Sigma}_{n-1}^{(d)}\right)}{\sum_{l=1}^{D} \alpha_{n-1}^{(l)} \mathcal{N}\left(\boldsymbol{x}_c; \boldsymbol{\mu}_{n-1}^{(l)}, \boldsymbol{\Sigma}_{n-1}^{(l)}\right)}, \tag{27}$$

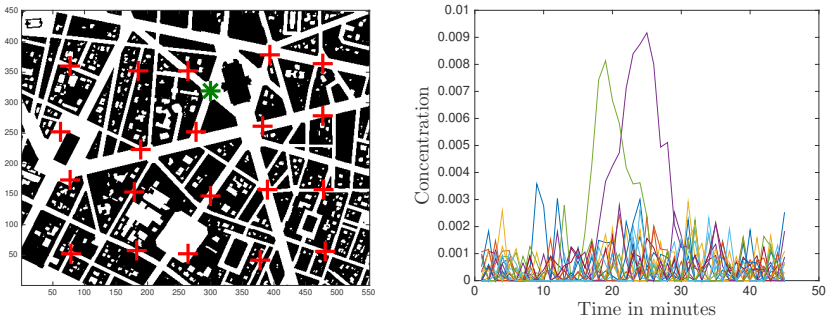
These updates are repeated until convergence. Finally, we propose to simply define the "defensive" component of the proposal distribution by:

$$\phi^{(0)}(\boldsymbol{x}_s) = \mathcal{U}_{\mathcal{X}_{NB}}(\boldsymbol{x}_s), \tag{28}$$

which thus consists in considering an uniform distribution over all plausible cells from the mesh of the area under study ($\mathcal{X}_{NB} \subset \mathcal{X}$).

4. Numerical Experiments

As shown in Rajaona et al. (2015), the performances of the core AMIS algorithm have been validated using synthetic and real concentration data in the framework of the Fusion Field Trials 2007 (FFT-07) experiment. In this paper, the proposed enhancements of the inference algorithm are tested using twin experiments. More precisely, as shown in Figure 1a, an urban area of 1.1km $\times$ 0.9km $\times$ 1.6km meshed at an horizontal and vertical resolution of 2 meters which corresponds to the Opéra quarter in Paris is considered. In this example of Fig. 1, the fictitious source and sensors are represented by a green asterisk and red crosses, respectively. For simplicity, both are assumed to be located at the same level from the ground. Even if the source term localization is considered in these experiments to be a two-dimensional problem, all the computations from the LPDM are performed in three-dimensional space. Unit releases emitted each minute from the 20 sensors over a 45-minute period were simulated in a backward mode, during which weather conditions varied gradually from a west-northwest wind to a north-northeast wind. These backward computations allow us to obtain the source-receptor matrix in Eq. (3) for every source term position. The different measures of concentration as a function of time are depicted in Fig. 1b with a different color for each sensor.



(a) **(b)**

Figure 1: Scenario under study – Left: 20 sensors (red) & 1 source (green). Right: Measurements obtained every minute by the 20 sensors from 09:00 to 09:45.

The dispersion simulations were carried out using Parallel-Micro-SWIFT-SPRAY (PMSS). Originally, Micro-SWIFT-SPRAY (MSS) (Tinarelli et al., 2013) was developed in order to provide a simplified, but rigorous CFD solution of the flow and dispersion in built-up environments in a limited amount of time. MSS is constituted by the local scale, high resolution, versions of the SWIFT and SPRAY models. SWIFT is a 3D terrain-following mass-consistent diagnostic model taking account of the buildings and providing the 3D fields of wind, turbulence, and temperature. SPRAY is a 3D Lagrangian Particle Dispersion Model able to account for the presence of buildings. Both SWIFT and SPRAY can deal with complex terrains and evolving meteorological conditions as specific features of the release like heavy or light gases. More recently, SWIFT and SPRAY have been efficiently parallelized in time, in space, and in numerical particles leading to the PMSS system (Oldrini et al., 2017). PMSS has been systematically validated against numerous wind tunnel and in-field experimental campaigns for short and prolonged releases (Trini Castelli et al., 2018). In all configurations, PMSS results comply with statistical acceptance criteria defined by Hanna and Chang (2012) commonly accepted to validate dispersion models in built-up environments. Furthermore, the SPRAY dispersion model can be run in both direct mode (from the source to a number of sensors) and retrograde mode (from sensors where detections are possibly made to areas indicating the possible locations of sources) providing respectively concentration and conjugate concentration data (Armand et al., 2013). The use of the backward mode has the advantage of reducing quite drastically the number of computations required to obtain the required source-receptor matrix $\boldsymbol{C}(\cdot)$. Indeed, only $N_c \times T_c$ (N_c: number of sensors, T_c: the number of time samples collected by each sensor) computations are needed whereas using the forward mode $N_p \times K \times T_s$ (N_p: number of algorithmic particles, K: number of AMIS iterations, T_s: number of intervals considered for the release) computations are necessary. Since in most cases $N_p \times K >> N_c$, the use of the backward mode is clearly preferable in order to reduce the overall computational time.

Regarding the parameters of the statistical model describing the prior knowledge of the release rate vector $\boldsymbol{q}$ in Eq. (6), the following covariance matrix is

considered to ensure a certain smoothness:

$$\mathbf{\Sigma}_q = \sigma_q^2 \begin{bmatrix} \kappa(1,1) & \kappa(1,2) & \cdots & \kappa(1,T_s) \\ \vdots & \ddots & & \vdots \\ \kappa(T_s,1) & \cdots & \kappa(T_s,T_s-1) & \kappa(T_s,T_s) \end{bmatrix}, \tag{29}$$

with $\kappa(\cdot,\cdot)$ a squared exponential kernel (Rasmussen and Williams, 2006):

$$\kappa(i,j) = \exp(-\frac{|i-j|^2}{l}). \tag{30}$$

The proposed AMIS algorithm detailed in Algo 1 was tested with $N_p = 100$ particles and $K = 20$ iterations. The adaptive part of the proposal distribution is a mixture of $D = 9$ bivariate Gaussian distributions which are equally spaced in the surveillance area as shown in Figure 2c and in Figure 3c. Figures 2 and 3 show the results obtained for the proposed initialization procedure of the adaptive part of the proposal distribution for two different fictitious source positions. With such a procedure, the mixture of *Gaussian* distributions is efficiently fitted to the backward map $\mathcal{B}(\cdot)$ obtained using Algo 2. Figure 4 highlights the large benefit of using this initialization strategy for the AMIS. The use of such a procedure based on the output of the backward LPDM allows us to sample particles in region of high interest directly in the first iterations of the algorithm, thus leading to a more rapid convergence to the correct solution.

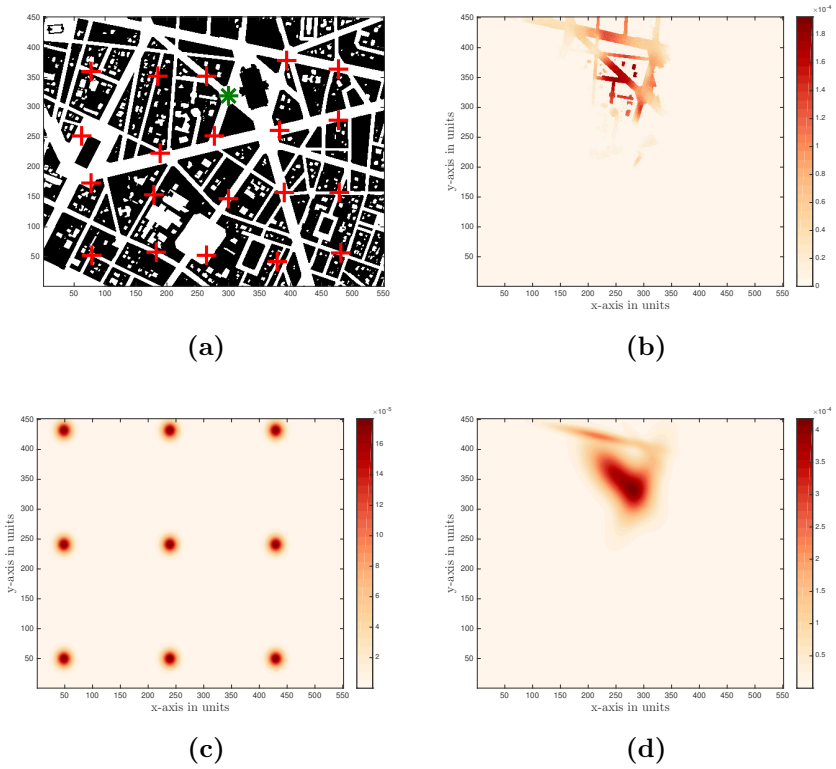


Figure 2: Initialization of the importance distribution using our proposed strategy in the presence of a prompt source - (a) Position of the source in green and sensors in red - (b) 2D map $\mathcal{B}(\cdot)$ obtained using Algo 2 - (c) Initial mixture of Gaussian distributions - (d) Final mixture of Gaussian distributions that minimizes the KL distance with the backward map $\mathcal{B}(\cdot)$ in (b)

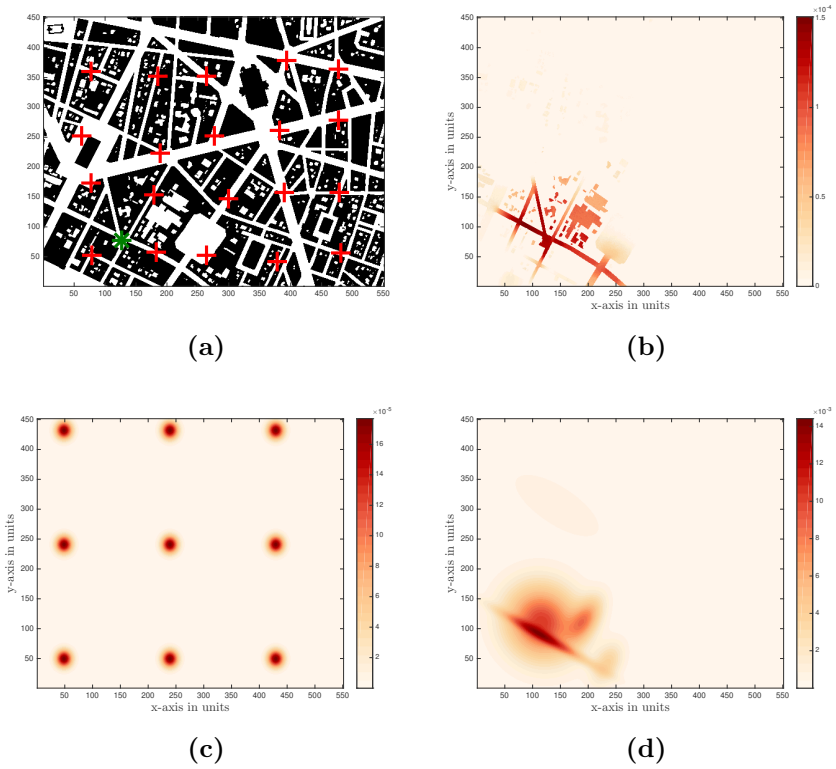


Figure 3: Initialization of the importance distribution using our proposed strategy in the presence of a prompt source - (a) Position of the source in green and sensors in red - (b) 2D map $\mathcal{B}(\cdot)$ obtained using Algo 2 - (c) Initial mixture of Gaussian distributions - (d) Final mixture of Gaussian distributions that minimizes the KL distance with the backward map $\mathcal{B}(\cdot)$ in (b)

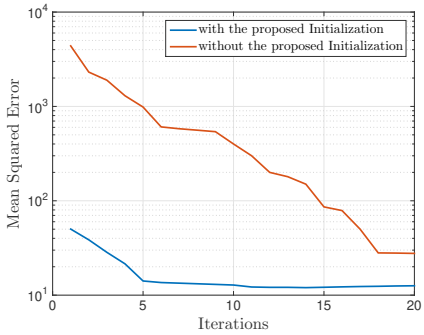


Figure 4: Comparison of the mean squared error between the true source position and the approximated posterior mean at the different iterations of the AMIS with and without the proposed initialization strategy.

Figures 5 to 8 give the output obtained by a single run of the proposed AMIS algorithm in the case of different release rate vector and with a source

located at (300;320) as in Fig. 1a. In each of these four figures, along with the ground truth values, the estimation of the marginal posterior distribution of the source position, $p(x_s|\boldsymbol{y})$ and $p(y_s|\boldsymbol{y})$, as well as the two first moments of the conditional posterior distribution $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$ are depicted. Whatever the fictive release considered - a prompt, a continuous or a more complex time-varying - the AMIS is able to make a correct estimation of both the source position and the release rate with a good accuracy in all cases. Moreover, the representation of the estimation using a probability distribution function provides an explicit representation of the uncertainty around the estimation, emphasizing how useful the Bayesian inference framework is when compared to results provided by other methods such as optimization-based ones in which only a point-estimate is available.

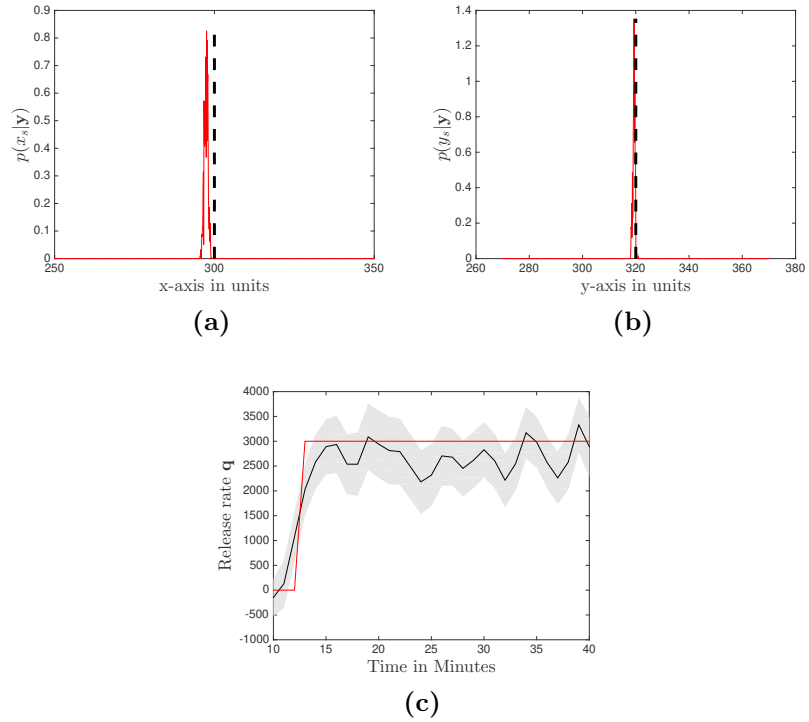


Figure 5: Results of the AMIS for STE with a continuous release [variance of the observation uncertainty $\sigma_\epsilon^2 = 10^{-6}$ in Eq. (4)]. Estimation in red of (a) $p(x_s|\boldsymbol{y})$ and (b) $p(y_s|\boldsymbol{y})$ and the true value in dashed black. (c): Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$ (black) and $\pm 2\text{diag}(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red), where $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{q}}$ is covariance matrix of the conditional posterior distribution of $\boldsymbol{q}$ obtained in Eq. (9).

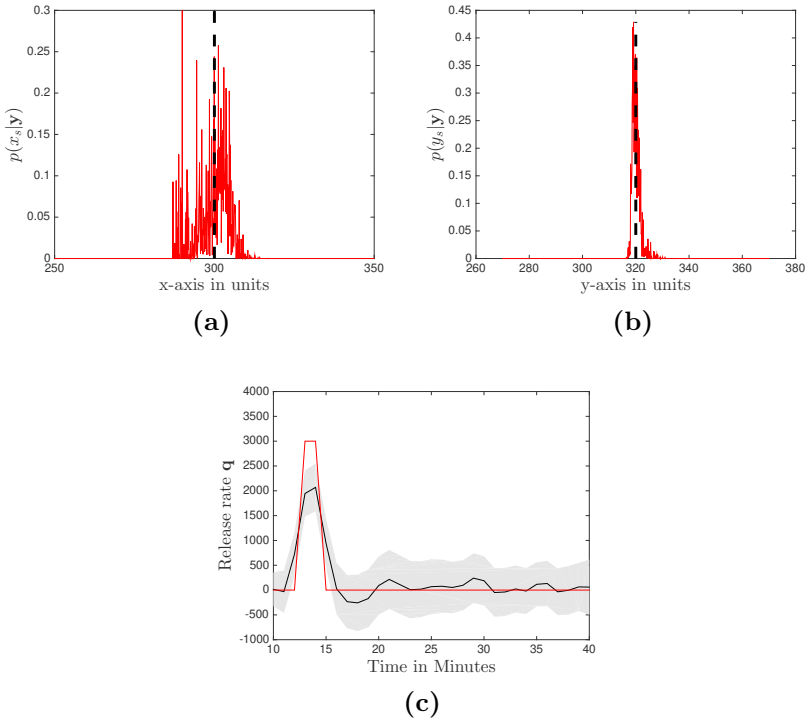


Figure 6: Results of the AMIS for STE with a prompt release [$\sigma_\epsilon^2 = 10^{-6}$]. Estimation in red of (a) $p(x_s|\boldsymbol{y})$ and (b) $p(y_s|\boldsymbol{y})$ and the true value in dashed black. (c): Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$ (black) and $\pm 2\text{diag}(\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red).

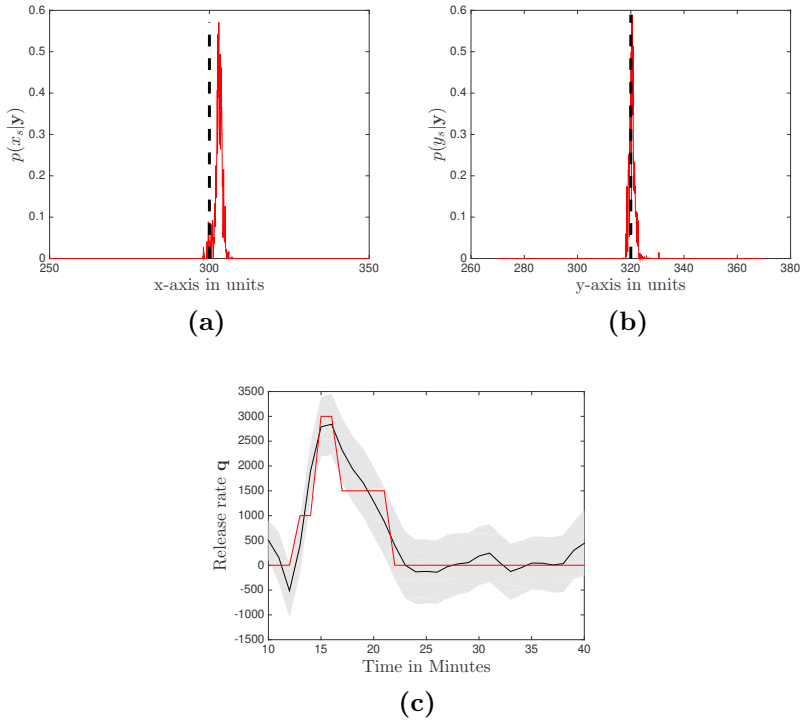


Figure 7: Results of the AMIS for STE with a time-varying release rate [$\sigma_\epsilon^2 = 10^{-6}$]. Estimation in red of (a) $p(x_s|\boldsymbol{y})$ and (b) $p(y_s|\boldsymbol{y})$ and the true value in dashed black. (c): Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$ (black) and $\pm 2\text{diag}(\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red).

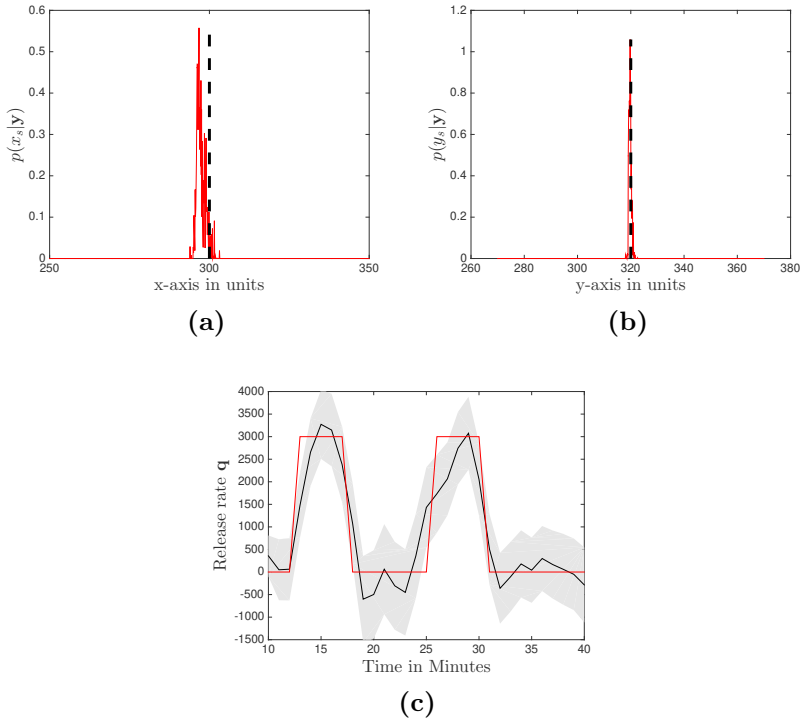
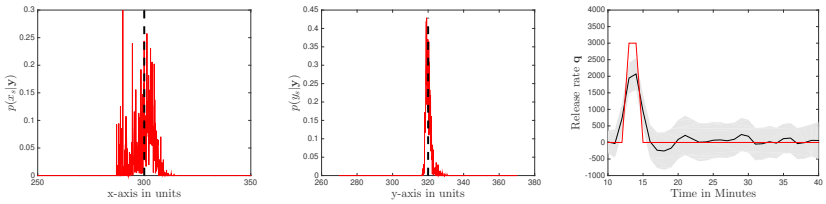
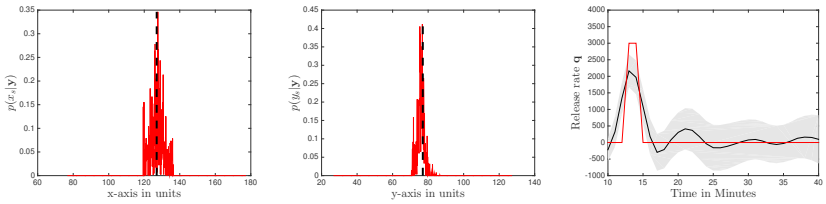


Figure 8: Results of the AMIS for STE with a time-varying release rate [$\sigma_\epsilon^2 = 10^{-6}$]. Estimation in red of (a) $p(x_s|\boldsymbol{y})$ and (b) $p(y_s|\boldsymbol{y})$ and the true value in dashed black. (c): Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$ (black) and $\pm 2\text{diag}(\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red).

In Figure 9, the same results in the case of a fictitious explosive release located at two different positions show the ability of the proposed inference algorithm to provide accurate answer in both situations.



(a)



(b)

Figure 9: Results of the AMIS for STE with an explosive release [$\sigma_\epsilon^2 = 10^{-6}$] with two positions of the source at (127;77) (a) and (300;320) (b). Left and middle: estimation of $p(x_s|\boldsymbol{y})$ and $p(y_s|\boldsymbol{y})$ (red) and the true value (dashed black). Right: Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$(black) and $\pm 2\text{diag}(\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red).

In Figure 10, the sensitivity of the approach with respect to the value σ_q^2 in the covariance matrix in Eq. (29) is studied in a time-varying setting as the one presented in Figure 7. The overall accuracy of the posterior mean provided by the AMIS could be impacted by a poor value of σ_q^2 but it is reasonable to assume that in practice this value could be set with a prior knowledge on plausible release ranges for the problem of interest and the area under surveillance.

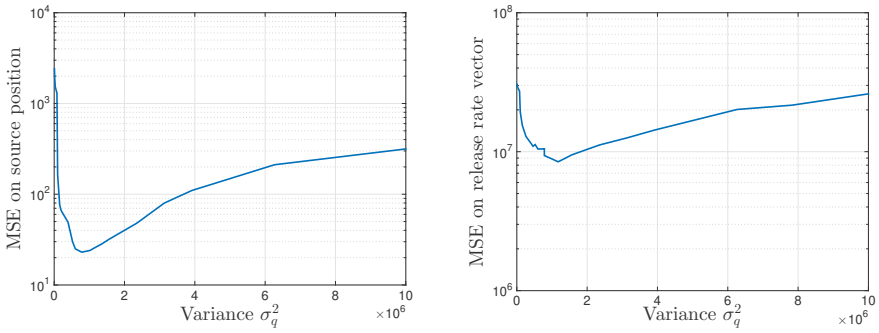


Figure 10: Mean squared error on the position (left) and release vector (right) obtained using different values of the variance in the prior

Finally in Figure 11, the impact of the noise variance level, σ_ϵ^2 in the observation equation, Eq. (4), on the estimation results provided by the proposed AMIS is studied in the case of a time-varying release rate as the one presented in Figure 7. As discussed in Section 2.1, this increase of variance could result for example from a larger error on the dispersion model, or on the sensor level due to poorer electronic components. As expected, the uncertainty is increasing with the noise variance level. More importantly, even in an extremely noisy environment, the proposed AMIS is able to provide an accurate estimation of the source position which is quite remarkable as the true detections in the measurements are clearly indistinguishable from the noise, e.g. Figure 11e.

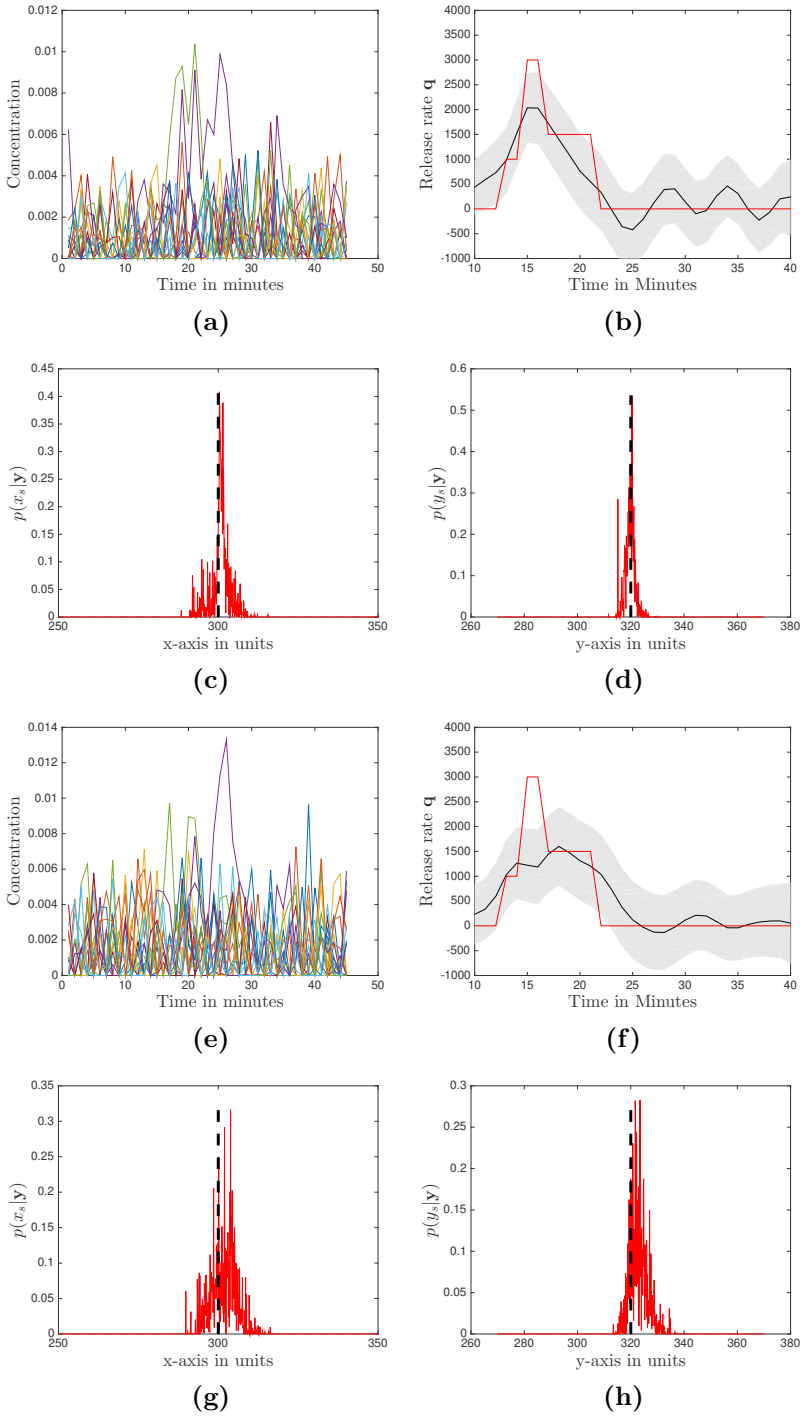


Figure 11: Impact of the noise variance level in the observation equation on the AMIS performances [(a-d): $\sigma_\epsilon^2 = 5.10^{-6}$, (e-h): $\sigma_\epsilon^2 = 10^{-5}$] in the presence of source with a varying release rate. In (a or e), measurements obtained every minute by the 20 sensors from 9:00 to 9:45. Estimation of $p(x_s|\boldsymbol{y})$ (c or g) and $p(y_s|\boldsymbol{y})$ (d or h) in red and the true value (dashed black). In (b or f) Mean of $p(\boldsymbol{q}|\boldsymbol{y}, \widehat{\boldsymbol{x}_s})$(black) and $\pm 2\text{diag}(\widetilde{\boldsymbol{\Sigma}}_{\boldsymbol{q}})$ confidence interval (grey) compared to the ground truth (red).

5. Conclusion

In this paper, a Bayesian solution is proposed to solve the source term estimation problem. More precisely, an enhanced version of the adaptive algorithm based on probabilistic Bayesian inference originally proposed in Rajaona et al. (2015) that estimates the parameters of the source term in case of an atmospheric release is described. In particular, we firstly propose to use the backward mode of the dispersion model in order to avoid multiple forward runs, a time-consuming task within the iterations of such a simulation-based algorithm. Using the dispersion model in backward mode, $N_c \times T_c$ computations are needed where N_c is the number of sensors and T_c the number of time samples collected by each sensor. In the test-case presented in the paper, N_c and T_c are equal respectively to 20 and 45. One can notice that in a practical situation, these figures would be lower. Each computation has a maximum computational time of 30 minutes as the urban simulation domain is of a limited extent. While the run of the adaptive algorithm is very quick, the simulations with the dispersion model are the major part of the computational burden associated to the source term estimate. As the dispersion runs may be carried out independently and, for instance, distributed to the cores of a large cluster, it is important to point out that the computational needs of our method are thus fully tractable regarding the present computational capabilities.

Then, we propose to also use the output of this backward run in order to efficiently design the initial parameters of the adaptive proposal distribution, which is crucial for the first exploration of the algorithm. Numerous twin experiments allow us to empirically demonstrate the efficiency of the proposed enhancements in complex built-up environments, both in terms of estimation accuracy and computation time since the algorithm is faster to converge owing to a better initialization of the proposal distribution and accuracy as shown by the promising results obtained by the proposed scheme.

The numerical experiments illustrate that the proposed inference solution could be sensitive to the parameters value chosen for the distributions used in the statistical model. As a consequence, it could be interesting to consider, as a potential extension of this work, the estimation of such hyperparameters jointly with the characteristics of the source. Finally, a future research direction of interest would be to design a statistical procedure to optimize the location of each sensor to increase the efficiency of inference algorithms to rapidly detect and identify the source characteristics.

References

Armand P, Olry C, Albergel A, Duchenne C, Moussafir J. Development and application of retro-spray, a backward atmospheric transport and dispersion model at the regional and urban scale. In: 15th International conference on Harmonisation within Atmospheric dispersion Modelling for Regulatory Purposes, Harmo'15. 2013. p. 789–93.

Bocquet M. Inverse modelling of atmospheric tracers: non-Gaussian methods and second-order sensitivity analysis. Nonlin Processes Geophys 2008;15:127–43.

Bugallo MF, Elvira V, Martino L, Luengo D, Miguez J, Djuric PM. Adaptive Importance Sampling: The past, the present, and the future. IEEE Signal Processing Magazine 2017;34(4):60–79.

Cappé O, Douc R, Guillin A, Marin JM, Robert CP. Adaptive Importance Sampling in General Mixture Classes. Statistics and Computing 2008;18(4):447–59.

Cappé O, Guillin A, Marin JM, Robert CP. Population monte carlo. Journal of Computational and Graphical Statistics 2004;13(4):907–29.

Chow FK, Kosovic B, Chan ST. Source inversion for contaminant plume dispersion in urban environments using building-resolving simulations. Journal of Applied Meteorology and Climatology 2008;47(6):1553–72.

Cornuet JM, Marin JM, Mira A, Robert CP. Adaptive Multiple Importance Sampling. Scandinavian Journal of Statistics 2012;.

Delle Monache L, Lundquist JK, Kosović B, Johannesson G, Dyer KM, Aines RD, Chow FK, Belles RD, Hanley WG, Larsen SC, Loosmore Ga, Nitao JJ, Sugiyama Ga, Vogt PJ. Bayesian Inference and Markov Chain Monte Carlo Sampling to Reconstruct a Contaminant Source on a Continental Scale. Journal of Applied Meteorology and Climatology 2008;47(10):2600–13.

Efthimiou GC, Kovalets IV, Argyropoulos CD, Venetsanos A, Andronopoulos S, Kakosimos KE. Evaluation of an inverse modelling methodology for the prediction of a stationary point pollutant source in complex urban environments. Building and Environment 2018;143:107–19.

Hanna S, Chang J. Acceptance criteria for urban dispersion model evaluation. Meteorological Atmospheric Physics 2012;116:133–46.

Issartel J, Baverel J. Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. Atmospheric Chemistry and Physics 2003;(3):475–86.

Jaynes E. Probability Theory: The Logic of Science. Cambridge University Press, 2003.

Keats A, Yee E, Lien FS. Bayesian inference for source determination with applications to a complex urban environment. Atmospheric Environment 2007;41(3):465–79.

Koohkan MR, Bocquet M. Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes. Tellus 2012;64:1–17.

Kopka P, Wawrzynczak A. Framework for stochastic identification of atmospheric contamination source in an urban area. Atmospheric Environment 2018;195:63–77.

Kovalets IV, Efthimiou GC, Andronopoulos S, Venetsanos AG, Argyropoulos CD, Kakosimos KE. Inverse identification of unknown finite-duration air pollutant release from a point source in urban environment. Atmospheric Environment 2018;181:82–96.

Oldrini O, Armand P, Duchenne C, Olry C, Tinarelli G. Description and preliminary validation of the PMSS fast response parallel atmospheric flow and dispersion solver in complex built-up areas. J of Environmental Fluid Mechanics 2017;17(3):1–18.

Rajaona H, Septier F, Armand P, Delignon Y, Olry C, Albergel A, Moussafir J. An adaptive Bayesian inference algorithm to estimate the parameters of a hazardous atmospheric release. Atmospheric Environment 2015;122:748–62. doi:10.1016/j.atmosenv.2015.10.026.

Rajaona H, Septier F, Delignon Y, Armand P, Olry C, Albergel A. A Bayesian approach of the Source Term Estimate coupling retro-dispersion computations with a Lagrangian Particle Dispersion Model and the Adaptive Multiple Importance Sampling. In: 17th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes (HARMO 17). Budapest, Hungary; 2016. p. 1–5.

Rasmussen C, Williams C. Gaussian processes for machine learning. The MIT Press, 2006.

Robert CP, Casella G. Monte Carlo statistical methods. Springer, 2004.

Seibert P, Frank A. Source-receptor matrix calculation with a Lagrangian particle dispersion model in backward mode. Atmospheric Chemistry and Physics 2004;4(1):51–63.

Tinarelli G, Mortarini L, Trini Castelli S, Carlino G, Moussafir J, Olry C, Armand P, Anfossi D. Description and preliminary validation of the PMSS fast response parallel atmospheric flow and dispersion solver in complex built-up areas. American Geophysical Union (AGU) 2013;200(A):311–27.

Trini Castelli S, Armand P, Tinarelli G, Duchenne C, Nibart M. Validation of a Lagrangian particle dispersion model with wind tunnel and field experiments in urban environment. Atmos Env 2018;193:273–89.

Winiarek V, Bocquet M, Saunier O, Mathieu A. Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant. Journal of Geophysical Research 2012;117(D5):D05122.

Winiarek V, Vira J, Bocquet M, Sofiev M, Saunier O. Towards the operational estimation of a radiological plume using data assimilation after a radiological accidental atmospheric release. Atmospheric Environment 2011;45(17):2944–55.

Yee E. Theory for Reconstruction of an Unknown Number of Contaminant Sources using Probabilistic Inference. Boundary-Layer Meteorology 2008;127(3):359–94.

Yee E. Validation of a Sensor-Driven Modeling Paradigm for Multiple Source Reconstruction with FFT-07 Data. Technical Report; Suffield; 2009.

Yee E, Hoffman I, Ungar K. Bayesian Inference for Source Reconstruction: A Real-World Application. International Scholarly Research Notices 2014;2014(1):1–12.

Yee E, Lien FS, Keats A, D'Amours R. Bayesian inversion of concentration data: Source reconstruction in the adjoint representation of atmospheric diffusion. Journal of Wind Engineering and Industrial Aerodynamics 2008;96(10-11):1805–16.