19th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
3-6 June 2019, Bruges, Belgium

APPLICATION OF THE BAYESIAN APPROACH AND INVERSE DISPERSION MODELLING TO SOURCE TERM ESTIMATES IN BUILT-UP ENVIRONMENTS

François Septier[1], Christophe Duchenne[2] and Patrick Armand[2]

[1] Université Bretagne Sud, Lab-STICC UMR CNRS 6285, F-56000 Vannes, France
[2] CEA, DAM, DIF, F-91297 Arpajon, France

Abstract: This paper presents an improved version of a STE adaptive algorithm based on probabilistic Bayesian inference (AMIS) that estimates the complete distribution of the source term parameters in case of an adverse (accidental or malevolent) atmospheric release. After introducing the STE problem in a probabilistic manner, we describe the proposed Bayesian inference algorithm with the use of the Lagrangian Particle Dispersion Model (LPDM) in backward mode. The enhanced procedure is illustrated with a synthetic example, using Retro-SPRAY, the backward implementation of SPRAY, the LPDM of the PMSS suite.

Key words: *Source term estimate, Bayesian approach, Adaptive Importance Sampling, backward LPDM.*

INTRODUCTION

Hazardous releases into the atmosphere may result from diverse circumstances, most often accidents or, less frequently, criminal activities. In some cases, the releases may happen insidiously without obvious trace (like smoke or an explosion). However, surreptitious releases might be detected by a network of sensors or people in trouble. Therefore, the quick and efficient identification of the source location and strength of the emission is of prime importance for the plants operators and rescue teams. Atmospheric dispersion modelling and mathematical methods can help in reconstructing the source term parameters, given a set of measurements coming from sensors. While the problem is challenging due to its ill-posed nature, several methods have been developed in the field of source term estimate (STE). Most of them rely on an optimization problem where a cost function has to be minimized using least squares or genetic algorithms, e.g. (Winiarek, 2012).

Another possibility is the probabilistic Bayesian approach which yields several upsides as it allows the incorporation of both model and observational errors and the use of potential prior information about the source. The framework provided by Bayesian analysis reformulates the point estimation problem as the search of the posterior probability density function of the source parameters. Several related examples exist in the literature, that use the notorious Markov Chain Monte Carlo (MCMC) algorithm, such as (Delle Monache et al., 2008), (Keats et al., 2007) or (Yee et al., 2014). Nevertheless, these MCMC techniques are prone to several issues, regarding the inherent *burn-in* time necessary before the convergence, or the choice of how to initialize properly the Markov chain. In this study, we focus on an alternative Bayesian method called Adaptive Multiple Importance Sampling (AMIS). As presented in (Rajaona et al., 2015), adding an advanced adaptive layer to the classical IS scheme allows us to obtain promising results for STE problems in a correct amount of time, compared to state-of-the-art methods.

Unfortunately, the computational time of such stochastic simulation-based techniques are highly dependent of the dispersion model used. Indeed, for each generated sample of the location of the source, a novel forward run has to be performed to evaluate the likelihood of the measurements. In complex urban environment, elaborate model, typically based on the use of a large number of Lagrangian particles, has to be used to have an accurate evaluation of the dispersion which is therefore time-consuming to run several times during the procedure.

This paper presents an improvement of the original method, aiming at optimizing the most time-consuming step in the algorithm by using the duality relationship with adjoint models for evaluating concentrations.

Moreover, the output of the dispersion model in backward mode is also efficiently utilized both in the initialization step and the "defensive" component of the adaptive proposal distribution to improve the convergence speed and the robustness of the proposed inference approach, respectively.

PROBLEM FORMULATION

Atmospheric dispersion model

In this study, we consider a point-wise and static source fully characterized by the parameter $\boldsymbol{\theta} = (\boldsymbol{x_s}, \boldsymbol{q})$ with $\boldsymbol{x_s} = (x_s, y_s)$ corresponds to 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 $j = 1, \dots, T_c$ with T_c the number of time samples collected by each sensor. Each measurement results from the superposition of the T_s releases on 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 modelling 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 $\boldsymbol{y} = \left[y_{1,1}, \dots, y_{1,T_c}, \dots, y_{N_c,1}, \dots, y_{N_c,T_c}\right]^T$ 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:

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

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

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

where $p(\boldsymbol{y}|\boldsymbol{\theta}) = \boldsymbol{\mathcal{N}}(\boldsymbol{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ corresponds to the multivariate normal distribution evaluated in $\boldsymbol{y}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\mu}$ and $\boldsymbol{I}_{N_c T_c}$ 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. It quantifies the predicted concentration value at some location and time from a dispersion model for a given source. 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). As in (Keats et al., 2007) and (Yee et al., 2008), we use 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 $\boldsymbol{C}(\cdot)$.

Specification of prior knowledge about parameters

Specifying the prior consists in representing our belief about the unknown state $\boldsymbol{\theta}$ using probability distributions before obtaining any observations. In this work, we consider that a source can be anywhere uniformly in some spatial surveillance area, denoted by $\Omega \subseteq \mathbb{R}^2$, i.e. $p(\boldsymbol{x_s}) = \boldsymbol{\mathcal{U}}_\Omega(\boldsymbol{x_s})$. Depending on the scenario, more informative distribution could be used in order to take into account that the source is more likely to be on some specific area (nuclear plants, industrial sites, etc.).

As in (Winiarek et al., 2011), a multivariate Gaussian distribution is considered as prior for the emission rate vector:

$$p(\boldsymbol{q}) = \boldsymbol{\mathcal{N}}(\boldsymbol{q}; \boldsymbol{\mu_q}, \boldsymbol{\Sigma_q}) \tag{5}$$

Bocquet (2008) points out that this is a crude approximation as the source emission rate vector should not have negative values, yet this simple assumption is often used in atmospheric dispersion inverse modelling with reasonable performances (Issartel and Baverel, 2003).

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 interested can be expanded as follows:

$$p(\boldsymbol{\theta}|\boldsymbol{y}) = p(\boldsymbol{x_s}, \boldsymbol{q}|\boldsymbol{y}) = \boldsymbol{p}(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x_s})p(\boldsymbol{x_s}|\boldsymbol{y}) \quad (6)$$

Owing to the Gaussian assumption of both the likelihood in Equation (4) and the prior distribution of $\boldsymbol{q}$ in Equation (5), the rule of conjugate priors states that the conditional posterior of the source emission rate $\boldsymbol{p}(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x_s})$ is therefore Gaussian and can thus be evaluated analytically.

Unfortunately, the second term $p(\boldsymbol{x_s}|\boldsymbol{y})$ in the complete posteriori distribution of interest in (6) 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 $\boldsymbol{p}(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x_s})$ can be obtained. In this work, we consider efficient stochastic simulation-based algorithms to approximate this complex posterior distribution $p(\boldsymbol{x_s}|\boldsymbol{y})$.

PROPOSED BAYESIAN ALGORITHM TO STE

An adaptive method: the AMIS algorithm

To approximate the marginal posterior distribution $p(\boldsymbol{x_s}|\boldsymbol{y})$ in Eq. (6), we resort to an adaptive version of the Importance Sampling algorithm (Robert and Casella, 2004), which consists, at the t-th iteration, in:

1. Drawing a population of N_p samples, $\left\{\boldsymbol{x}_{s,t}^1, \dots, \boldsymbol{x}_{s,t}^{N_p}\right\}$, from a proposal distribution $\phi(\boldsymbol{x_s}; \varphi_t)$
2. Computing the *importance weights* $w_t^i = \frac{p(\boldsymbol{x}_{s,t}^i|\boldsymbol{y})}{\phi(\boldsymbol{x}_{s,t}^i;\varphi_t)}$
3. Adapting the parameters φ of the proposal distribution $\phi(\boldsymbol{x_s}; \varphi_t)$ using the generated random weighted samples $\left\{\boldsymbol{x}_{s,t}^i, w_t^i\right\}_{i=1}^{N_p}$ so that it tends to fit the posterior distribution we try to approximate $p(\boldsymbol{x_s}|\boldsymbol{y})$.

The target distribution of interest, $p(\boldsymbol{x_s}|\boldsymbol{y})$ here, is thus approximated by an empirical measure, thus leading to the following approximation of the full posterior distribution defined in Eq. (6):

$$p(\boldsymbol{x_s}, \boldsymbol{q}|\boldsymbol{y}) \approx \sum_{\boldsymbol{t=1}}^{\boldsymbol{T}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{N_p}} \overline{w}_t^i \boldsymbol{p}\left(\boldsymbol{q}|\boldsymbol{y}, \boldsymbol{x}_{s,t}^i\right) \delta_{\boldsymbol{x}_{s,t}^i}(\boldsymbol{x_s}) \quad (7)$$

where the $\overline{w}_t^i$ are the normalized importance weights (i.e. $\overline{w}_t^i = w_t^i \left[\sum_{t=1}^{T} \sum_{j=1}^{N_p} w_t^j\right]^{-1}$) and $\delta(\cdot)$ is the Dirac function. The main novelty between both traditional adaptive IS and the AMIS is the use in the latter of all the collection of particles drawn at the previous iterations by using a recycling scheme to improve **both** the learning of the proposal and the accuracy of the approximation of the target distribution (in the PMC only the particles form the last iteration are used). Using every single simulated particle at every step of this iterative algorithm could therefore lead to a significant improvement as shown in (Cornuet et al., 2012).

On efficient use of the LPDM in backward mode

As mentioned previously, the computation of the source-receptor matrix in Eq. (3) with a Lagrangian particle dispersion model (LPDM) in a forward mode constitutes the most time-consuming step of the AMIS algorithm proposed in (Rajaona et al., 2015). As a consequence, in this study, we propose to use the duality relationship mentioned in (Keats et al., 2007) by running a backward LPDM to fill the source-receptor matrix, thus drastically reducing the computational complexity of the overall Bayesian inference procedure.

Moreover, we propose to use the outputs of the backward LPDM runs to design an efficient procedure to automatically set the initial parameters, φ_0, of the adaptive proposal distribution of the AMIS. As already remarked in (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. In this paper, we propose

to use a mixture of D normal distributions and an additional "defensive" component which will remain unchanged through adaptive procedure, i.e.:

$$\phi(\boldsymbol{x_s};\varphi_t) = \alpha^{(0)}\phi^{(0)}(\boldsymbol{x_s}) + (1-\alpha^{(0)})\sum_{d=1}^{D}\alpha_t^{(d)}\mathcal{N}(\boldsymbol{x_s};\boldsymbol{\mu}_t^{(d)},\boldsymbol{\Sigma}_t^{(d)}) \quad (8)$$

The aim of the static component is to guarantee that the importance function remains bounded whatever happens during the adaptation, thus guaranteeing a finite variance. However, 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. As a summary, the parameters φ_t to be adapted during the t-th iteration of the AMIS consists in $\left\{\alpha_t^{(d)},\boldsymbol{\mu}_t^{(d)},\boldsymbol{\Sigma}_t^{(d)}\right\}_{d=1}^{D}$.

The proposed initialization consists in fitting the adaptive components of the mixture to a spatial map obtained from the backward runs of the LPDM. More precisely, this spatial map is resulting from the cumulative weighted sum of binary retro-propagation maps (obtained after thresholding using a plausible level of associated release rate) from all the sensors with a weight of (+1) and (-1) for those that have been activated or not activated, respectively, during the considered period of time.

NUMERICAL EXPRIMENTS

The dispersion simulations are 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 LPDM 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). Furthermore, the SPRAY dispersion model can be run in direct mode (from the source to a number of sensors) but also in the retrograde mode (from sensors where detections are possibly made to areas indicating the possible locations of sources) which has been used in this paper. As illustrated in **Figure 1**, we considered an urban area of 1.1km × 0.9km × 1.6km meshed at an horizontal and vertical resolution of 2 meters. 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.

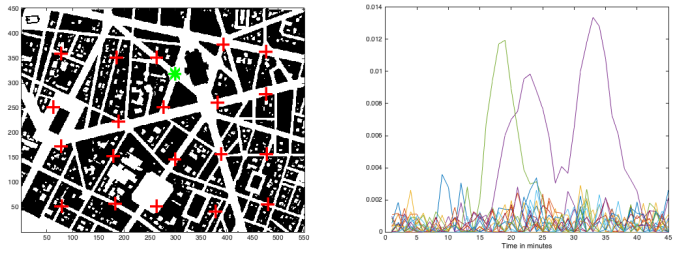

Figure 1. Scenario under study – *Left*: 20 sensors (red) & 1 source (green).
Right: Measurements obtained every minute by the 20 sensors from 09:00 to 09:45.

To assess the performances of the inference procedure, $N_p = 100$ particles are drawn during $T = 20$ iterations and the proposal distribution is composed of a mixture of $D = 9$ adaptive components that are initialized using our proposed fitting procedure of the spatial map from the backward LPDM. As seen in **Figure 2**, the results illustrate good estimation performances for both the source location and the release rate.

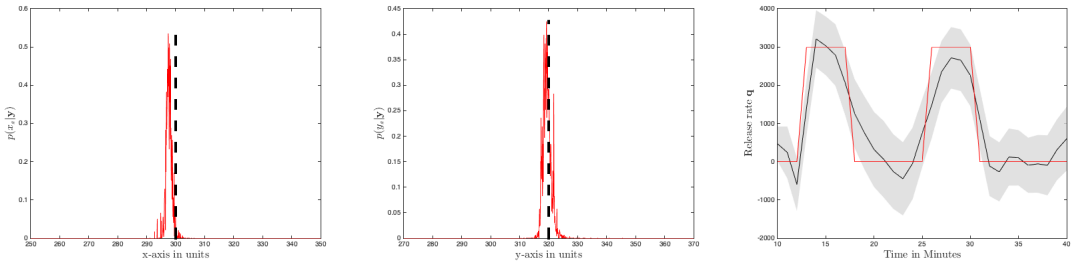

Figure 2. Results of the AMIS for STE. *Left and middle*: estimation of $p(x_s|\boldsymbol{y})$ and $p(y_s|\boldsymbol{y})$ (red) and the true value (dashed black). *Right*: Mean of $p(\boldsymbol{q}|\boldsymbol{y},\hat{\boldsymbol{x}}_s)$ (black) and ±2σ confidence interval (grey) compared to the ground truth (red).

Figure 3 clearly highlights the large benefit of using the proposed initialization strategy for the AMIS. The use of such a procedure based on the output of the backward LPDM allow 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.

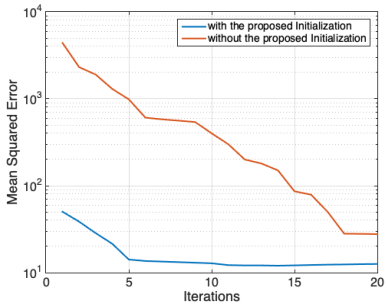

Figure 3. Comparison of the mean squared error between the true source position and the approximated posterior mean at the different iterations of the AMIS with and without the proposed initialization strategy.

CONCLUSION

In this report, 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. More precisely, 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 the algorithm. 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. As a consequence, the algorithm is faster to converge since the proposal is better initialized.

REFERENCES

M. Bocquet. Inverse modelling of atmospheric tracers: non-Gaussian methods and second-order sensitivity analysis. *Nonlin. Processes Geophys.*, 15:127–143, 2008.

O. Cappé, A. Guillin, J.-M. Marin, and C. P. Robert. Population monte carlo. *J. Comput. Graph. Stat.*, 13(4):907–929, 2004.

J.-M. Cornuet, J.-M. Marin, A. Mira, and C. P. Robert. Adaptive Multiple Importance Sampling. *Scand. Jour. of Stat.*, 2012.

L. Delle Monache, J. K. Lundquist, B. Kosovic, G. Johannesson, K. M. Dyer, R. D. Aines, F. K. Chow, R. D. Belles, W. G. Hanley, S. C. Larsen, G. a. Loosmore, J. J. Nitao, G. a. Sugiyama, and P. J. Vogt. Bayesian Inference and Markov Chain Monte Carlo Sampling to Reconstruct a Contaminant Source on a Continental Scale. *Journal of Applied Meteorol. and Clim.*, 47(10):2600–2613, October 2008.

J. Issartel and J. Baverel. Inverse transport for the verification of the Comprehensive Nuclear Test Ban Treaty. *Atmospheric Chemistry and Physics*, (3):475–486, 2003.

E. Jaynes. Probability Theory: *The Logic of Science*. Cambridge University Press, 2003. ISBN 0521592712.

A. Keats, E. Yee, and F. S. Lien. Bayesian inference for source determination with applications to a complex urban environment. *Atmospheric Environment*, 41(3):465–479, January 2007.

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

O. Oldrini, P. Armand, C. Duchenne, C. Olry and G. Tinarelli. 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*, 17(3):1-18, 2017.

H. Rajaona, F. Septier, P. Armand, Y. Delignon, C. Olry, A. Albergel, and J. Moussafir. An adaptive Bayesian inference algorithm to estimate the parameters of a hazardous atmospheric release. *Atmospheric Environment*, 122:748–762, December 2015.

C. P. Robert and G. Casella. Monte Carlo statistical methods. Springer, 2004.

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

G. Tinarelli, L. Mortarini, S. Trini-Castelli, G. Carlino, J. Moussafir, C. Olry, P. Armand and D. Anfossi. Review and validation of Micro-Spray, a Lagrangian particle model of turbulent dispersion. In Lagrangian Modeling of the Atmosphere, Geophysical Monograph, Volume 200, American Geophysical Union (AGU), 311-327, 2013.

V. Winiarek, J. Vira, M. Bocquet, M. Sofiev, and O. Saunier. Towards the operational estimation of a radiological plume using data assimilation after a radiological accidental atmospheric release. *Atmospheric Environment*, 45(17):2944–2955, 2011.

V. Winiarek, M. Bocquet, O. Saunier, and A. Mathieu. 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*, 117(D5):D05122, March 2012.

E. Yee. Theory for Reconstruction of an Unknown Number of Contaminant Sources using Probabilistic Inference. *Boundary-Layer Meteorology*, 127(3):359–394, March 2008.

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