XXIV ENMC
Encontro Nacional de Modelagem Computacional
XII ECTM
Encontro de Ciência e Tecnologia de Materiais

A BAYESIAN INFERENCE APPROACH FOR THE IDENTIFICATION OF MULTIPLE ATMOSPHERIC EMISSIONS WITH UNCERTAINTY QUANTIFICATION

Roseane A.S.Albani[1] - roseanealves75@gmail.com
Vinícius V.L.Albani[2] - v.albani@ufsc.br
Hélio S. Migon[1,3] - migon@im.ufrj.br
Antônio J. Silva Neto[1] - ajsneto@iprj.uerj.br
[1]Universidade do Estado do Rio de Janeiro, Instituto Politécnico - Nova Friburgo, RJ, Brazil
[2]Universidade Federal de Santa Catarina - Florianópolis, SC, Brazil
[3]Universidade Federal do Rio de Janeiro - Rio de Janeiro, RJ, Brazil

Abstract. *This work proposes the use of a Monte Carlo Markov Chain technique to estimate the location and strength of multiple pollutant emissions in the atmosphere. The corresponding dispersion problem is numerically solved by the stabilezed Galerkin/Least-Square finite element method. The sources parameters are estimated simultaneously to the so-called precision, which is related with the uncertainties in concentration measurements, from a set of observed concentrations. Since the number of unknowns is large,* a priori *information is used to reduce the search region. Estimation is carried out by the Metropolis in Gibbs algorithm. The proposed methodology is tested with the Fusion Field Trial 2007 (FFT07).*

Keywords: *Source Identification, Bayesian inference, Air pollution*

1. INTRODUCTION

The proper estimation of effluents from intentional or accidental emissions can be specially useful when a fast response is necessary. For example, we can mention industrial releases of gases to the atmosphere or contaminants to the river currents, the discharge of sewage into rivers and leaking from ships to the ocean.

The source characterization by data assimilation from monitoring data may need a huge number of sensors and the costs could be impractical. A promising option is the estimation of the sources parameters using inverse modeling techniques. In this work we deal with atmospheric releases, so, in this case, firstly, it is necessary to define a dispersion model, or direct problem, that provides a synthetic concentration field and establishes a source-receptor relationship.

Secondly, we define an inverse problem technique. The literature concerning the identification of atmospheric releases presents a wide variety of methodologies, usually based on deterministic and stochastic techniques. In the class of deterministic methods, an objective function relating measured and predicted concentrations is minimized. Several techniques may be used to perform this minimization, such as gradient-based methods or meta-heuristic optimization tools (Albani and Albani, 2020; Albani et al., 2020).

The stochastic methodologies are usually based on Bayesian Inference which considers all the observed and unknown quantities as random variables. In these methodologies, the Monte Carlo Markov Chain (MCMC) methods are used to generate samples for the unknown variables, exploring the so-called full conditional distributions (Gamerman and Lopes, 2006).

Considering source estimation in the atmosphere the aforementioned methodologies were repeatedly employed. See for example, Albani and Albani (2019, 2020); Albani et al. (2020) for deterministic techniques and Addepalli et al. (2011); Albani et al. (2021) for stochastic methods.

The drawbacks arising from the source identification problems stems from both, direct and the inverse problems. Briefly, we can mention the tricky hypothesis involving the simplifications necessary to set up a precise model for the atmospheric dispersion. On the other hand, the inverse problem suffers with uncertainties in the measured datasets and imprecisions in the dispersion model that may result in an ill-posed problem, that is, the inverse problem may not have a solution, the solution may not be unique, or it can be sensitive to small perturbation in the data.

To address such difficult problem, we propose the combination of a Monte Carlo Markov Chain (MCMC) algorithm with *a priori* information and an accurate numerical solution for the dispersion problem. The dispersion model is described by an advection-diffusion equation numerically solved by the Galerkin Least-Square (GLS) FEM formulation, which was proposed in Hughes et al. (1989). The sources parameters and the so-called precision, which is related to the uncertainties in the concentration data, are estimated simultaneously from concentration measurements by the Metropolis in Gibbs MCMC algorithm (Albani et al., 2021). To simplify the estimation procedure by reducing the search region, a series of *a priori* information is introduced based on the distribution of the observed concentrations in the computational domain, the receptors positions, and the wind direction.

Therefore, the contribution of this work stems from the combination of these techniques to solve a short-range source identification problem from multiple emissions.

2. SOURCE PARAMETERS ESTIMATION MODEL

This section describes the models and algorithms and their combination to estimate the set of parameters of a multiple point source problem in the atmospheric boundary layer (ABL). In the sequence, the proposed MCMC adaptive algorithm is presented followed by the case study we deal in this work and the dispersion model or direct problem employed to describe it.

2.1 The Estimation Procedure

Let us consider an n-dimensional vector containing the measured concentrations of a given pollutant emitted from unknown sources. Denote such vector by $\mathbf{c}^{\text{obs}}$. We want to determine the location of the emission sources and the amount of pollutant emitted. In other words, we want to determine the entries of

$$\mathbf{w}_s = [x_{s,1}, y_{s,1}, z_{s,1}, Q_{s,1}, \dots, x_{s,N}, y_{s,N}, z_{s,N}, Q_{s,N}]^T$$

which is the $4 \times N$-dimensional vector, where N is the number of emissions, containing the location of the sources $(x_{s,j}, y_{s,j}, z_{s,j}, \quad j = 1, \dots, N)$ and their corresponding strengths $Q_{s,j} j = 1, \dots, N$.

The relation between the emission sources and the receptors are established by the numerical solution of the dispersion problem in Eq (6)–(8) evaluated at the receptors' locations, and denoted by $\mathbf{c}(\mathbf{w}_s)$ and the observed concentrations as follows:

$$\mathbf{c}(\mathbf{w}_s) = \mathbf{c}^{\text{obs}}. \tag{1}$$

Since there are uncertainties in the concentration measurements and in the direct problem modeling, Eq. (1) does not hold in general, which means that it must be reformulated. Based on the hypotheses that concentrations are non-negative quantities whose discrepancy can be asymmetric, the relation between $\mathbf{c}(\mathbf{w}_s)$ and $\mathbf{c}^{\text{obs}}$ is given as follows:

$$\ln\left(\mathbf{c}(\mathbf{w}_s)\right) - \ln\left(\mathbf{c}^{\text{obs}}\right) = \varepsilon, \tag{2}$$

where ε is the random noise, which is n-dimensional and normal-distributed, with mean $\mathbf{0}$ and covariance matrix $p^{-1}I$, where p is the precision, which is unknown and related to the uncertainties concerning the concentrations measurements. Thus, $\varepsilon \sim N(\mathbf{0}, p^{-1}I)$. In other words, the quantities that must be estimated are the precision p and the vector with the sources parameters $\mathbf{w}_s$. It is worth noticing that, we are assuming that the observed concentrations are independent.

The likelihood function is then given by

$$P\left(\mathbf{c}^{\text{obs}}\,|p, \mathbf{w}_s\right) \propto p^{\frac{n}{2}} \exp\left(-\frac{p}{2}\left\|\ln\left(\mathbf{c}(\mathbf{w}_s)\right) - \ln\left(\mathbf{c}^{\text{obs}}\right)\right\|_{\ell_2}^2\right). \tag{3}$$

The unknowns quantities p and $\mathbf{w}_s$ have independent prior distributions. The precision p has as the prior distribution Gamma $\left(\frac{n_0}{2}, \frac{d_0}{2}\right)$. The sources parameters $\mathbf{w}_s$ have a uniform distribution. The set where such uniform distribution is defined takes into consideration a series of assumptions based on the concentration measurements, wind direction, and receptors location. More details shall be given latter in the text.

The full conditionals for the precision p and the sources parameters $\mathbf{w}_s$ are given by:

$$P\left(p\,\middle|\mathbf{w}_s, \mathbf{c}^{\text{obs}}\right) \propto P\left(\mathbf{c}^{\text{obs}}\,|p, \mathbf{w}_s\right) P(p), \tag{4}$$

$$P\left(\mathbf{w}_s\,\middle|p, \mathbf{c}^{\text{obs}}\right) \propto P\left(\mathbf{c}^{\text{obs}}\,|p, \mathbf{w}_s\right) P(\mathbf{w}_s), \tag{5}$$

where the functions $P(\mathbf{w}_s)$ and $P(p)$ denote the prior densities for $\mathbf{w}_s$ and p, respectively. Note that, since $P(p)$ is a Gamma density and the likelihood function is a Gaussian density, then the full conditional for the precision p is proportional to a Gamma density. In other words, p given $\mathbf{w}_s$ and $\mathbf{c}^{\text{obs}}$ has random distribution given by Gamma $\left(\frac{\tilde{n}}{2}, \frac{\tilde{d}}{2}\right)$, with the parameters $\tilde{n} = n_0 + n$ and $\tilde{d} = d_0 + \|\ln(\mathbf{c}^{\text{obs}}) - \ln(\mathbf{c}(\mathbf{w}_s))\|_{\ell_2}^2$. Such distribution can be easily sampled by a Gibbs sampler. On the other hand, the samples for the full conditional of $\mathbf{w}_s$ cannot be obtained directly. Thus, to obtain samples for both p and $\mathbf{w}_s$, the Metropolis in Gibbs Monte Carlo Markov chain (MCMC) is used (Gamerman and Lopes, 2006; Müller, 1994; Robert et al., 2010; Albani et al., 2021). Algorithm 1 presents the pseudo-code of the method.

The covariance matrix Z in the Gaussian random walk distribution in Algorithm 1 is defined to turn the algorithm adapted to the size of the search region, i.e., it is a diagonal matrix, where the entries in the diagonal are given by the step 10^{-3} times the length of the search domain in the corresponding direction, i.e., the length of intervals with the values the source parameters can assume.

Algorithm 1 Metropolis in Gibbs MCMC algorithm.

```
procedure MCMC ALGORITHM
    $j = 0$
    $p^j = p, \mathbf{w}_s^j = \mathbf{w}_s$
    while $j \leq$ MaxIter do
        Draw $\mathbf{u} \sim N(\mathbf{w}_s^j, Z)$
        Evaluate $\tilde{n}$ and $\tilde{d}$
        Draw $p^{j+1} \sim$ Gamma $\left(\frac{\tilde{n}}{2}, \frac{\tilde{d}}{2}\right)$
        Evaluate $P\left(\mathbf{w}_s^j \,\middle|\, p^{j+1}, \mathbf{c}^{\text{obs}}\right)$ and $P\left(\mathbf{u} \,\middle|\, p^{j+1}, \mathbf{c}^{\text{obs}}\right)$
        Evaluate $\beta = \min\left(1, \frac{P\left(\mathbf{u} \,\middle|\, p^{j+1}, c^{\text{obs}}\right)}{P\left(\mathbf{w}_s^j \,\middle|\, p^{j+1}, \mathbf{c}^{\text{obs}}\right)}\right)$
        Draw $l \sim U[0, 1]$
        if $l < \beta$ then
            Accept: $\mathbf{w}_s^{j+1} = \mathbf{u}$
        else
            Reject: $\mathbf{w}_s^{j+1} = \mathbf{w}_s^j$
        end if
        $j = j + 1$
    end while
end procedure
```

Prior Density and Prior Information for the Source Parameters Since the number of estimated quantities is large, *a priori* information can play an important role in the process. The search region can be considerably reduced by using physical information concerning the experiment and its setup. Based on the wind direction, the receptors locations, and the distribution of measured concentrations in the domain, it is possible to say if emission sources are close to each other or far apart, and the most probable region in the domain where the emissions are located.

Based on the concentration distribution given by an isopleth showing the concentration distribution in the xy-plane, we assume that the emissions are located close to each other. We also assume that they have similar strength. In other words, after ordering the sources, we let the parameters of the first one to vary inside the following set

$$A = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}] \times [z_{\min}, z_{\max}] \times [Q_{\min}, Q_{\max}].$$

The parameters of the remaining sources are defined in terms of the first one. Thus, if the location and strength of the first source are denoted by the vector $[x_1, y_1, z_1, Q_1]^T$, the parameters of the other sources are defined as

$$[x_j, y_j, z_j, Q_j]^T = [x_1 + s_{x,j}, y_1 + s_{y,j}, z_1 + s_{z,j}, Q_1 + s_{Q,j}]^T, \quad j = 1, \ldots, N,$$

with $s_{x,j} \in [s_{x,\min}, s_{x,\max}]$ $j = 1, \ldots, N$, and similarly for $s_{y,j}$, $s_{z,j}$, and $s_{Q,j}$ with $j = 1, \ldots, N$.

Using these simplifications, the prior density for the sources parameters $\mathbf{w}_s$ is defined as the uniform distribution in the set $A \times B^{N-1}$, with

$$B = [s_{x,\min}, s_{x,\max}] \times [s_{y,\min}, s_{y,\max}] \times [s_{z,\min}, s_{z,\max}] \times [s_{Q,\min}, s_{Q,\max}],$$

where the power $N-1$ represents the Cartesian product taken $N-1$ times. In the steps of the MCMC algorithm, we use truncation to enforce that all the proposals for the sources parameters stay inside the set A.

2.2 The Direct Problem

The numerical concentration **c** in Eq. (1) is given as the solution of the dispersion model, described by the Advection-diffusion Partial Differential Equation (PDE),

$$\mathbf{u} \cdot \nabla \mathbf{c} - \nabla \cdot (\mathbf{K} \nabla \mathbf{c}) = \sum_{j=1}^{N} \mathbb{S}_j \text{ in } \Omega \tag{6}$$

where Ω denotes the simulation domain, that is a slice of the ABL. The wind vector velocity is represented by **u**, the ABL turbulence effect is incorporated by the tensor **K**, which is a diagonal matrix with the non-zero components K_x, K_y, and K_z, representing the turbulent diffusion in the x, y and z directions, respectively. The right-hand side of the equation corresponds to the contribution of the set of point source emissions, where n represents the total number of sources, and $\mathbb{S}_j$ is the jth source, defined as

$$\mathbb{S}_j(x, y, z, t) = Q_{s,j} \delta(x - x_{s,j}) \delta(y - y_{s,j}) \delta(z - z_{s,j}),$$

wherein $[x_{s,j}, y_{s,j}, z_{s,j}]^T$ are the source coordinates, $Q_{s,j}$ is the emission rate, $j = 1...N$, where N is the number of sources, and δ is the Dirac delta distribution.

The dispersion model (6) can be associated to the following boundary conditions:

$$\mathbf{n} \cdot \nabla \mathbf{c} = 0, \tag{7}$$

with the vector **n** denoting the outward normal on the boundaries of Ω at $z = z_0$ and $z = H$, and

$$\mathbf{c} = 0 \quad \text{elsewhere}. \tag{8}$$

Wherein z_0 represents the surface roughness length and H denotes the upper boundary of the computational domain Ω. The boundary condition (7) assumes a homogeneous and flat surface, with no gradient, or flux, at the boundary. Finally, condition (8) means that the concentration goes to zero far enough from the sources.

Equation (6) must be solved for all the iterations of the MCMC Algorithm 1. This procedure could be prohibitive depending on the considered problem and simulations setups. Thus, considering the linearity of the Eq. (6), it is possible to establish a direct relationship between the sources and the sensors through the so-called adjoint state PDE. This procedure is known to reduce significantly the computational cost of the simulations, since in this case, the direct problem can be solved only once for all the MCMC iterations.

Representing the observed concentration at the kth sensor by $\mathbf{c}^{obs}(x_k, y_k, z_k)$, it follows that

$$\mathbf{c}^{obs}(x_k, y_k, z_k) = \sum_{j=1}^{N} \int_{\Omega} \mathbf{c}_k^* \mathbb{S}_j d\Omega = \sum_{j=1}^{N} \langle \mathbf{c}_k^*, \mathbb{S}_j \rangle, \tag{9}$$

where $\mathbf{c}_k^*$ is the solution of the following adjoint-state PDE (Mamonov and Tsai, 2013):

$$-\mathbf{u} \cdot \nabla \mathbf{c}^* - \nabla \cdot (\mathbf{K} \nabla \mathbf{c}^*) = \mathcal{S}_k, \tag{10}$$

wherein $\mathcal{S}_k$ stands for the kth sensor, defined as

$$\mathcal{S}_k(x, y, z) = \delta(x - x_k)\delta(y - y_k)\delta(z - z_k),$$

where (x_k, y_k, z_k) is the spatial coordinates of the kth sensor for n sensors. The boundary conditions associated to Eq. (10) are the following:

$$\mathbf{n} \cdot \nabla \mathbf{c}^* = 0 \quad \text{on the boundaries of } \Omega \text{ at } z = z_0 \text{ and } z = H, \tag{11}$$

and

$$\mathbf{c}^* = 0 \quad \text{elsewhere.} \tag{12}$$

Again, **n** denotes the outward normal on the boundary of Ω. In the adjoint-state PDE, the sensors work as sources. Thus, the concentration at each sensor will be obtained by the scalar product between the source term and the solution of the PDE (10) according to Eq. (9).

Meteorological variables affect considerably atmospheric dispersion, which is described by an advection-diffusion PDE. When the ABL flow is advection-dominated, the numerical methodologies are challenged to solve the atmospheric dispersion equation (6), or equivalently Eq. (10). In this work the Galerkin-Least/Squares formulation from Hughes et al. (1989) is applied to prevent likely non-physical oscillations, usually observed in the standard FEM solutions when dealing with advection-diffusion problems. Here this formulation is not presented since it is performed in the author's previous works such as Albani and Albani (2019, 2020).

2.3 The case study

In this work we evaluate the proposed methodology to estimate multiple point source parameters, namely, the spatial coordinates and emission rates against the Fusion Field Trial (FFT07) tracer experiment (Storwald, 2007). This experiment is just briefly described in this work. To go into a more detailed description, see Albani and Albani (2020); Albani et al. (2020). Many trials were undertaken during FFT07 experiments, under several atmospheric conditions and considering a different number of point sources. In this work, we consider the datasets from experiment 55. During FFT07 trial 55, the tracer gas propylene (C_3H_6) was released during 15 minutes at a constant rate from four point sources 2 meters above the ground level. After released, the (C_3H_6) was collected over a total of 100 tracer sampling units, arranged over a rectangular grid sampling of 475 m $\times$ 450 m. The sensors were displayed keeping the distance of 50 m from each other. Figure 1 (left) shows the sensors and sources disposition during the FFT07 experiments.

The FFT07 provides substantial database regarding meteorological and concentration measurements to evaluate short-range source identification techniques. The meteorological variables were measured over different positions at 2, 4, 8, 16 and 32 meters above the ground level. The measured meteorological variables were employed to calculate the turbulent fluxes of sensible heat and momentum to further compute the Monin-Obukhov length (L), the friction velocity (u_*) and the surface roughness length (z_0). Considering Trial 55, the computed values for these parameters are $-39.7(m)$, $0.26(m/s)$ and $0.013(m)$ respectively.

Atmospheric dispersion is strongly influenced by the meteorological conditions, including turbulence, the wind intensity and direction. In this work, since we address the source parameter estimation of the FFT07 experiment, we can take advantage of the homogeneity and flatness of the terrain by using functional profiles to describe the wind field and turbulence. This procedure is used to avoid solving the Navier-Stokes PDE problem to obtain the wind field, saving computational time. The parametric profiles for the wind intensity and turbulent diffusion coefficients considered in the present work to solve the dispersion model in Eq. (10) are based on the Monin-Obukhov similarity theory (MOST). For the sake of the size limitations, we omit the description of these profiles. However, they can be found in Albani and Albani (2020). We also assume that $K_x = K_y$.

3. NUMERICAL RESULTS

This section performs an experimental evaluation of the Algorithm 1 using the FFT07 dataset. Although the numerical evaluation of the direct problem (10) is an important step in the code evaluation process, we omit it here for the sake of space and refer the reader to the previous work Albani and Albani (2020).

In order to appropriately setup the *a priori* density of the vector of source parameters $\mathbf{w}_s$, we consider the concentration distribution in the xy-plane given by the isopleth in Figure 1 (right), the wind direction, and the sensors distribution in Figure 1. Notice that, the resulting searching domain contains the true locations of the sources. Thus, the search region is given by $A \times B^3$, since $N = 3$, where

$$A = [-100, 100] \times [50, 600] \times [0.013, 50] \times [0, 30]$$

and

$$B = [-25, 25] \times [-250, 250] \times [-1.5, 1.5] \times [-5, 5],$$

with the spatial parameters defined in meters and the source strengths defined in g/s.

Algorithm 1 is initialized with an uniformly-distributed sample for $\mathbf{w}_s$ in $A \times B^3$. The initial step for p is generated by the very vague prior distribution Gamma$(\frac{n_0}{2}, \frac{d_0}{2})$ with $n_0 = 10^{-3}$ and $d_0 = 10^{-3}$. The Markov Chain has 50 thousand states, with the first 10 thousand set as the burn-in set. From the remaining states, we select 800 states using a step-size of 50 states. The histograms of the resulting chains of the source parameters and the precision p can be found in Figure 2. The corresponding summary statistics is shown in Table 1.

The histograms of the unknowns in Fig. 2 illustrate the potential stabilization of the Markov Chains. Moreover, the values provided by Gelman-Rubin test and presented in Tab. 1 are close to the ideal value 1.00, confirming the convergence of the chains. As Tab. 1 show, the sample regions contain the major part of the true source parameters. In addition, whenever the sample regions failed to contain the true parameters, they were in the neighborhood of such values. Although the overall accuracy of the estimations are satisfactory, since the present inverse problem is highly ill-posed, the method must be further refined, since, in some cases, it failed to find regions containing some of the true parameters of the sources.

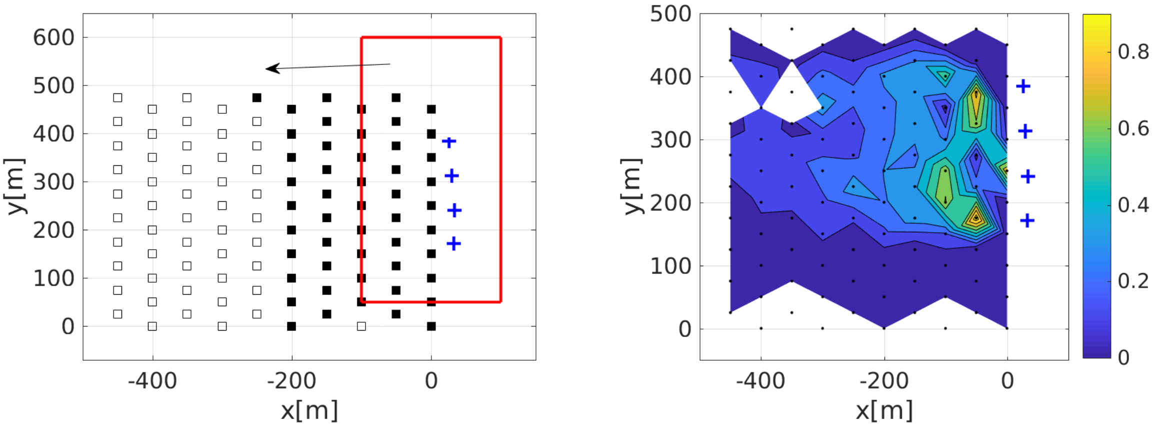

Figure 1- Left: The sensors distribution in the xy-plane of the computational domain. The arrow indicates the wind direction and the squares are the true locations of the sources. The sensors used in the simulations are the filled squares. The rectangle is the resulting search region. Right: Isopleth showing the concentration distribution in the computational domain. Both: Crosses represent the sources locations.

Parameter	True	Min.	Q1	Median	Q3	Max.	GR Test
$x_{s,1}$ [m]	33.0	-59.2	-26.9	-21.8	-15.2	5.76	1.00
$y_{s,1}$ [m]	171	137	222	244	266	344	1.00
$z_{s,1}$ [m]	2.00	0.01	7.57	8.87	10.2	17.3	1.02
$Q_{s,1}$ [g/s]	11.4	10.2	13.53	14.38	15.2	19.9	1.01
$x_{s,2}$ [m]	33.8	-54.6	-25.3	-19.55	-13.4	8.61	1.00
$y_{s,2}$ [m]	241	183	257	279	299	372	1.00
$z_{s,2}$ [m]	2.00	1.36	8.85	10.2	11.6	18.6	1.02
$Q_{s,2}$ [g/s]	11.4	6.61	9.76	10.6	11.5	15.4	1.01
$x_{s,3}$ [m]	30.0	-37.1	-5.32	0.00	6.19	28.0	1.00
$y_{s,3}$ [m]	313	222	285	301	316.78	394.92	1.00
$z_{s,3}$ [m]	2.00	0.80	8.42	9.79	11.1	18.3	1.02
$Q_{s,3}$ [g/s]	4.65	6.15	9.06	9.93	10.7	15.1	1.00
$x_{s,4}$ [m]	26.0	-32.6	-4.05	1.60	7.38	27.6	1.00
$y_{s,4}$ [m]	384	368	495	518	537	600	1.00
$z_{s,4}$ [m]	2.00	2.04	9.83	11.16	12.54	19.7	1.02
$Q_{s,4}$ [g/s]	11.4	1.25	4.60	5.46	6.25	10.8	1.00
p	-	0.01	0.02	0.02	0.02	2.61	1.0003

Table 1- True values (True) and summary statistics of the chains, i.e., minimum value (Min.), the first quartile (Q1), the median value (Median), the third quartile (Q3), the maximum value (Max.) and the value of the Gelman-Rubin convergence test (GR Test) for each variable.

4. CONCLUSIONS

To solve the identification of multiple emission sources, we proposed the use of a Bayesian Inference technique, based on the Metropolis in Gibbs MCMC algorithm, in combination with an accurate numerical solution of the dispersion problem, which is based on the stabilized GLS/FEM formulation. Due to the large number of unknowns and the well-known ill-posedness

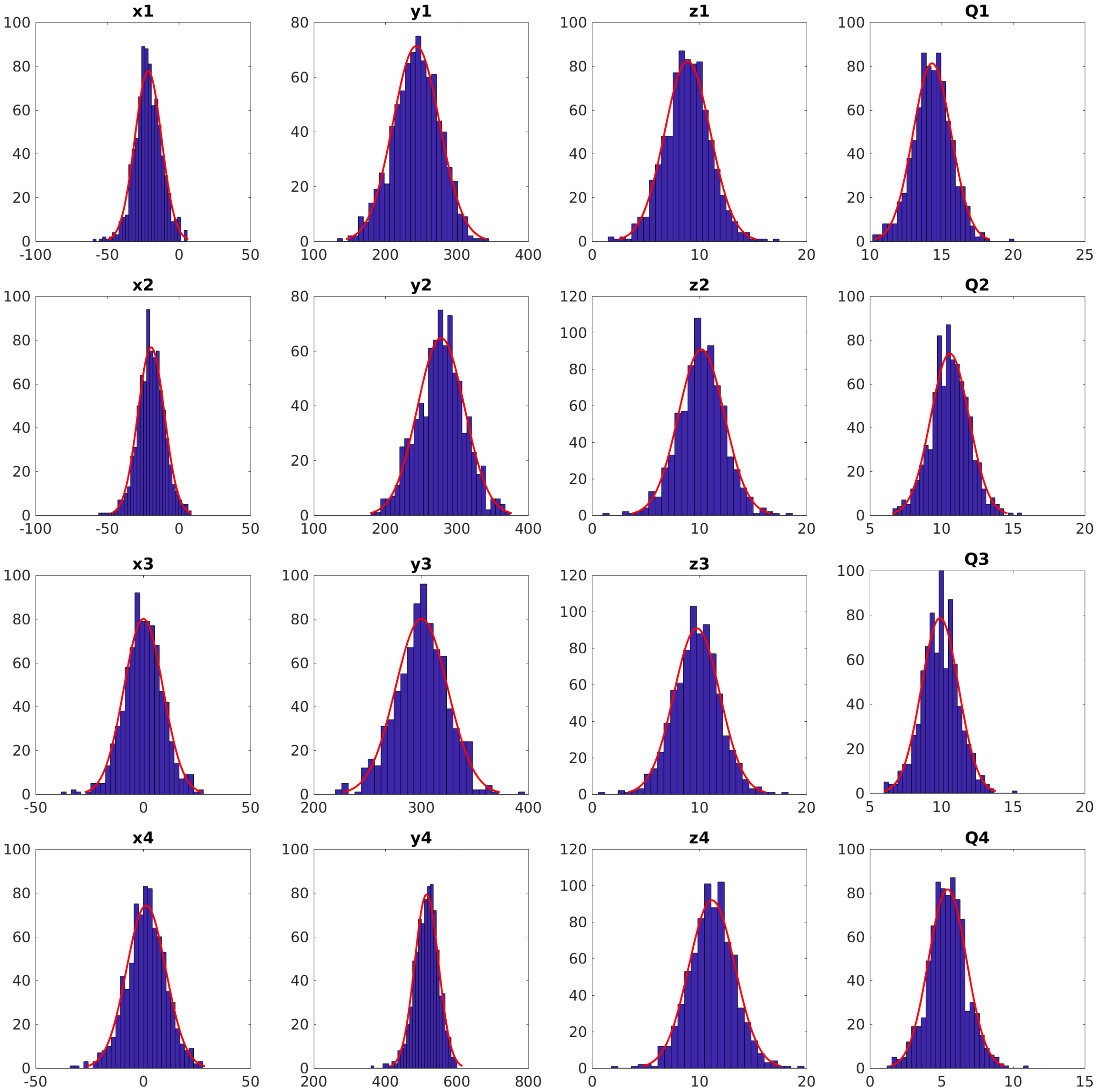

Figure 2- Histograms of the chains of the parameters $x_{s,j}$, $y_{s,j}$, $z_{s,j}$, and $Q_{s,j}$, $j = 1, \ldots, 4$.

of the present inverse problem, we included *a priori* information by considering the experiment setup and the distribution of the measured concentration in the computational domain. In addition, the Metropolis in Gibbs algorithm allowed the estimation of the so-called precision, which is related to the uncertainties in the concentration measurements. The combination of all these features led to satisfactorily accurate estimations of the parameters of the sources. However, the method shall be further refined in a forthcoming work.

Acknowledgements

RA thanks Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) for the financial support through the grants E-26/202.932/2019 and E-26/202.933/2019. AJSN acknowledges the financial support provided by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) through the grants 88887.311757/2018-00 and 88887.194804/2018-00, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) through the grant 308958/2019-5, and the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) through the grant E-26/202.878/2017. HM thanks PAPD/Rio de Janeiro State University for the financial support through the grant E-26/007/10667/2019.

REFERENCES

Addepalli, B., Sikorski, K., Pardyjak, E., and Zhdanov, M. (2011). Source characterization of atmospheric releases using stochastic search and regularized gradient optimization. *Inverse Probl. Sci. Eng.*, 19(8):1097–1124.

Albani, R. and Albani, V. (2019). Tikhonov-Type Regularization and the Finite Element Method Applied to Point Source Estimation in the Atmosphere. *Atmos. Environ.*, 211:69–78.

Albani, R. and Albani, V. (2020). An Accurate Strategy to Retrieve Multiple Source Emissions in the Atmosphere. *Atmos. Environ.*, 233:117579.

Albani, R., Albani, V., Migon, H., and Silva Neto, A. (2021). Uncertainty quantification and atmospheric source estimation with a discrepancy-based and a state-dependent adaptative MCMC. *Environ Pollut.*, 290:118039.

Albani, R., Albani, V., and Silva Neto, A. (2020). Source characterization of airborne pollutant emissions by hybrid metaheuristic/gradient-based optimization techniques. *Environ Pollut.*, 267:115618.

Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC Press.

Hughes, T., Franca, L., and Hulbert, G. (1989). A new Finite Element Formulation for Computational Fluid Dynamics: VIII. The Galerkin/Least-Square Method for Advective-Diffusive Equations. *Comput. Methods Appl. Mech. Engrg.*, 73:173–189.

Mamonov, A. and Tsai, Y.-H. R. (2013). Point source identification in nonlinear advection-diffusion-reaction systems. *Inverse Problems*, 29(3):035009.

Müller, P. (1994). Metropolis based posterior integration schemes. In *Numerical Recipes in Fortran (2nd Edition*. Citeseer.

Robert, C. P., Casella, G., and Casella, G. (2010). *Introducing Monte Carlo Methods with R*, volume 18. Springer.

Storwald, D. P. (2007). Detailed test plan for the fusing sensor information from observing 696 networks (fusion) field trial (FFT-07). Technical report.