22nd International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
10-14 June 2024, Pärnu, Estonia

INVERSE MODELLING FOR SOURCE CHARACTERIZATION IN COMPLEX INDUSTRIAL SITES: DEVELOPMENT OF THE ADJOINT STATE METHOD APPLIED TO A LAGRANGIAN STOCHASTIC DISPERSION MODEL

Jean Salles Loustau[1,2], Lionel Soulhac[1], Ariane Emmanuelli[1], Olivier Duclaux[2], Rami Nammour[3], Chi Vuong Nguyen[1], Ludovic Donnat[2], Catherine Juery[2]

[1]Ecole Centrale de Lyon, CNRS, Universite Claude Bernard Lyon 1, INSA Lyon, LMFA, UMR5509, 69130, Ecully, France
[2]Laboratoire Qualité de l'Air, TotalEnergies, 69360 Solaize, France
[3]E&P Research & Technology USA, TotalEnergies, 77002-5605 Houston (Texas), USA

Abstract: In this paper, a new application of the adjoint state method, to a Lagrangian stochastic (LS) model, in the context of atmospheric dispersion of pollutants, is presented. This method allows to solve the minimization problem with an efficient computation of the objective function gradient, reducing computation cost. The technique is applied to a LS model as it accurately models turbulent dispersion in a complex environment while still providing reasonable computational cost. In a first part, the adjoint method is presented, highlighting its benefits. Then the method is extended to a LS model, described as a Markovian explicit iterative model with least-square misfit. Finally, test cases are discussed.

Key words: *Inverse problem, Adjoint method, Pollutant source characterization, Lagrangian models*

INTRODUCTION

The characterization of atmospheric pollutant sources in industrial sites is a major concern to improve on-site safety and to evaluate environmental impact. However, some of these sources are leaks and diffuse emissions difficult to identify and quantify, especially in real-time. Inverse modelling is a widespread method used to cope with this issue. It relies on sensor measurements downstream of the plume coupled with an inversion algorithm, allowing to retrieve source term characteristics, i.e., emission rate and position. In the literature, one can find different approaches to obtain these characteristics, as grid search "brute force" methods (Ben Salem *et al.*, 2014), consisting in testing a possible source position at every grid point and picking the one that best fits the data. Alternatively, one can solve a minimization problem with optimization methods classically computing the whole Jacobian matrix of the functional (Gill *et al.*, 1981).

Yet, all these techniques are time-consuming, reducing their applicability when time to solution prevails. To address this issue, the adjoint method was developed for the inverse problem theory (Chavent, 1974) to efficiently compute the cost function gradient, allowing for fast local optimization techniques (Keats *et al.*, 2007). Indeed, it avoids computing the whole Jacobian besides being independent of the optimization parameter number. Since then, it has been widely used, including in the atmospheric dispersion field, to find source term characteristics. A literature review highlights applications to Gaussian and Eulerian models (Pudykiewicz, 1998; Giering, 2000) but, to our knowledge, not on a forward LS model, despite being suitable for modelling turbulent dispersion in complex environment while providing reasonable computational cost. Consequently, this study aims at presenting a new application of this method, to a LS model for atmospheric dispersion of pollutants, and its use on some dataset configurations.

In the first section, the adjoint method is introduced, focusing on its benefits compared to other approaches. In the second part, it is extended to a LS model described as a Markovian explicit iterative model with least-square misfit. At last, some numerical verification results obtained by this application are discussed.

ADJOINT STATE METHOD FOR INVERSE PROBLEM SOLVING

Forward model definition

A forward model F_s, depending on model parameters m_s and giving the state (output) variable u_s, defines the generally implicit state equation (Plessix, 2006):

$$\underbrace{F_s}_{n_F \times 1}\left(\underbrace{u_s}_{n_u \times 1}, \underbrace{m_s}_{n_m \times 1}\right) = \underbrace{0}_{n_F \times 1} \tag{1}$$

F_s, u_s and m_s are respectively of size n_F, n_u and n_m with $n_F = n_u$ there being as many outputs as equations. If an explicit relationship f_s exists between u_s and m_s, F_s will stand as $F_s(u_s, m_s) = u_s - f_s(m_s) = 0$. In the atmospheric dispersion context, F_s is a forward transport and dispersion model of pollutants originating from a source s, with u_s a concentration vector (e.g. pollutant species concentrations) provided as the forward model output and m_s a source parameter vector, i.e. source position and strength ($n_m = 4$).

Methodology for inverse problem solving

Generally, as an explicit analytical inversion of equation (1) does not exist, the inverse problem must be considered as a minimization one. The goal is to minimize the difference between the observation data d, i.e. the concentration observations in this context, and the output data u_s provided by the forward model (modelled concentrations). It amounts to minimize a cost function J, depending on model parameters m_s:

$$\underbrace{J}_{1 \times 1}\left(\underbrace{m_s}_{n_m \times 1}\right) = \underbrace{E}_{1 \times 1}\left(\underbrace{u_s}_{n_u \times 1}, \underbrace{m_s}_{n_m \times 1}\right) \tag{2}$$

with E the error functional of the differences between the model data u_s and the observations d. To find the values of m_s minimizing the cost function J, its gradient must be computed. The minimization problem in then solved iteratively in practise, updating the m_s values at each iteration of the optimization algorithm.

Introduction of the adjoint state equation to compute the gradient

Classically, the computation of ∇J requires the whole Jacobian matrix, i.e. the matrix of the derivatives relatively to model parameters m_s. Providing the full expression of the gradient:

$$\underbrace{\nabla J}_{n_m \times 1} := \left(\frac{dJ}{dm_s}\right)^T = \left(\frac{\partial E}{\partial u_s}\frac{du_s}{dm_s} + \frac{\partial E}{\partial m_s}\right)^T = \underbrace{\left(\frac{du_s}{dm_s}\right)^T}_{n_m \times n_u}\underbrace{\left(\frac{\partial E}{\partial u_s}\right)^T}_{n_u \times 1} + \underbrace{\left(\frac{\partial E}{\partial m_s}\right)^T}_{n_m \times 1} \tag{3}$$

the matrix $\frac{du_s}{dm_s}$ appears as the bottleneck term as u_s does not normally depend explicitly on m_s, preventing the differentiation of u_s. Moreover, this term, which needs to be computed for each perturbation dm_s, i.e. typically on each grid point, depends on the number of parameters. At the industrial scale, the number of grid points can exceed many thousands, leading to large computation cost. The adjoint method proposes an alternate, quicker way to compute ∇J. It formally derives adjoint equations from transport models (Pudykiewicz, 1998), providing the sensitivity of model output (e.g. concentration) to input variables (e.g. emission rate and location). From the state equation (1) specifying the state variable u_s, the adjoint state equation is obtained such that:

$$\underbrace{\left(\frac{\partial F_s}{\partial u_s}\right)^T}_{n_u \times n_F}\underbrace{\lambda_s}_{n_F \times 1} = \underbrace{\left(\frac{\partial E}{\partial u_s}\right)^T}_{n_u \times 1} \tag{4}$$

specifying the adjoint state λ_s, of dimensions $n_F \times 1$. Equation (4) is nothing more than a linear system to solve, where λ_s is not a matrix but a vector this time. The adjoint state does not depend on the number of optimization parameters anymore, reducing computation time.

Gradient computing

Once equation (4) is solved, i.e. λ_s found, the gradient ∇J determined by each component can be computed:

$$\underbrace{\nabla J}_{n_m\times 1} \coloneqq \left(\frac{\mathrm{d}J}{\mathrm{d}m_s}\right)^T = \underbrace{\left(\frac{\partial E}{\partial m_s}\right)^T}_{n_m\times 1} - \underbrace{\left(\frac{\partial F_s}{\partial m_s}\right)^T}_{n_m\times n_F} \underbrace{\lambda_s}_{n_F\times 1} \tag{5}$$

Extension to a Markovian explicit iterative model with least-square misfit

To apply the adjoint to a LS model, the method has been extended to an explicit iterative model case, with a Markov process and a least-square error functional. In this case, $\frac{\partial F_s}{\partial u_s}$ becomes the identity matrix in (4) and λ_s now equals the source term $\left(\frac{\partial E}{\partial u_s}\right)^T$, such that:

$$\underbrace{\lambda_s}_{n_F\times 1} = \underbrace{\left(\frac{\partial E}{\partial u_s}\right)^T}_{n_u\times 1} = \underbrace{\left(\frac{u_s - d}{\bar{d}*\bar{d}}\right)}_{n_u\times 1} \tag{6}$$

Moreover, E does not depend explicitly on m_s, $\frac{\partial E}{\partial m_s} = 0$ in equation (5). Finally, combining equation (5), the derivative of the forward explicit model (1) and equation (6), ∇J can be computed such that:

$$\underbrace{\nabla J}_{n_m\times 1} = -\underbrace{\left(\frac{\partial F_s}{\partial m_s}\right)^T}_{n_m\times n_F} \underbrace{\lambda_s}_{n_F\times 1} = \underbrace{\left(\frac{\partial f_s}{\partial m_s}\right)^T}_{n_m\times n_F} \underbrace{\left(\frac{u_s - d}{\bar{d}*\bar{d}}\right)}_{n_F\times 1} \tag{7}$$

Considering an iterative model of N total iterations, the outputs are now computed by successive applications of a sub-model of f_s. Moreover, in the atmospheric dispersion context, the model includes a Markov process. Hence, only the first sub-model f_s^1 depends on m_s. Besides, the global output vector u_s is now the concatenation of the N sub-vectors u_s^i. Finally, the global forward model vector F_s is defined as:

$$\underbrace{F_s}_{n_u\times 1}\left(\underbrace{u_s}_{n_u\times 1}, \underbrace{m_s}_{n_m\times 1}\right) = \left[\underbrace{[u_s^1 - f_s^1(m_s)]^T}_{1\times n_1} \dots \underbrace{[u_s^i - f_s^i(u_s^{i-1})]^T}_{1\times n_i} \dots \underbrace{[u_s^N - f_s^N(u_s^{N-1})]^T}_{1\times n_N}\right]^T = \underbrace{0}_{n_u\times 1} \tag{8}$$

where u_s^i is the sub-vector of u_s of dimension n_i, f_s^i is the corresponding explicit sub-model of f_s at the i^{th} iteration and $n_u = \sum_{i=1}^N n_i = n_F$. In (8), only the first member $[u_s^1 - f_s^1(m_s)]^T$ depends on m_s. Hence, the Jacobian matrix $\left(\frac{\partial F_s}{\partial m_s}\right)^T$ is null everywhere except at its terms involving f_s^1. Besides, under the assumption of only final iteration observations, one can solve iteratively the adjoint state equation (4), which is now:

$$\underbrace{\begin{bmatrix} 1 & -\left(\frac{\partial f_s^2}{\partial u_s^1}\right)^T & 0 & \cdots & 0 \\ 0 & \ddots & -\left(\frac{\partial f_s^3}{\partial u_s^2}\right)^T & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & -\left(\frac{\partial f_s^N}{\partial u_s^{N-1}}\right)^T \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}}_{n_u\times n_u} \underbrace{\begin{bmatrix} \lambda_s^1 \\ \lambda_s^2 \\ \vdots \\ \lambda_s^{N-1} \\ \lambda_s^N \end{bmatrix}}_{n_u\times 1} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \left(\frac{\partial E}{\partial u_s^N}\right)^T \end{bmatrix}}_{n_u\times 1} \tag{9}$$

$$\text{with } \underbrace{\lambda_s^i}_{n_i\times 1} = \underbrace{\prod_{j=i}^{N-1}\left(\frac{\partial f_s^{j+1}}{\partial u_s^j}\right)^T}_{n_i\times n_N} \underbrace{\left[\left(\frac{\partial E}{\partial u_s^N}\right)^T\right]}_{n_N\times 1} \overset{\text{by (6)}}{=} \underbrace{\prod_{j=i}^{N-1}\left(\frac{\partial f_s^{j+1}}{\partial u_s^j}\right)^T}_{n_i\times n_N} \underbrace{\left[\frac{u_s^N - d^N}{\bar{d}^N * \bar{d}^N}\right]}_{n_N\times 1} \tag{10}$$

Finally, one can compute ∇J, using λ_s^1 found solving iteratively equation (9) with equation (10), the sparse Jacobian matrix $\left(\frac{\partial F_s}{\partial m_s}\right)^T$ with Markov process and the expression of ∇J in equation (7):

$$\underbrace{\nabla J}_{n_m\times 1} = \underbrace{\left(\frac{\partial f_s^1}{\partial m_s}\right)^T}_{n_m\times n_1} \underbrace{\prod_{j=1}^{N-1}\left(\frac{\partial f_s^{j+1}}{\partial u_s^j}\right)^T}_{n_1\times n_N} \underbrace{\left[\frac{u_s^N - d^N}{\bar{d}^N * \bar{d}^N}\right]}_{n_N\times 1} \tag{11}$$

Equation (11) is a clever way to compute the gradient, by the chain rule application. ∇J is obtained computing only a product iteratively, by simply calculating the partial derivatives of f_s^i iteratively. Neither the a priori knowledge of the final output u_s^N nor the final observations d^N are necessary, being only required as global product terms. Knowing a priori the number of iterations N is not necessary either.

APPLICATION TO A LAGRANGIAN STOCHASTIC PARTICLE DISPERSION MODEL

To present the methodology, it is applied to a simplified LS model, with Gaussian steady isotropic homogeneous turbulence and diagonal Reynolds stresses. A p particle originating from an instantaneous

point source position $\boldsymbol{x_s}$ is transported at each time step δt through a stochastic advection process. Its path is described at each time t by its Lagrangian position $X_{p,i}$ and Lagrangian fluctuating velocity $U'_{p,i}$:

$$\begin{cases} X_{p,i}(t+\delta t) = X_{p,i}(t) + \left(\overline{u_i}(\boldsymbol{X_p}, t) + U'_{p,i}(t)\right)\,\delta t \\ U'_{p,i}(t+\delta t) = \left(1 - \frac{\delta t}{T_L}\right) U'_{p,i}(t) + \sigma_u \sqrt{\frac{2}{T_L}}\,\delta\xi_{p,u_i} \end{cases} \quad \text{with } U'_{p,i}(t_0) = \xi_{p,u_i}(t_0)\sigma_u \text{ and } X_{p,i}(t_0) = x_{s,i} \quad (12)$$

with i the spatial component at time t. The evolution of $U'_{p,i}$ is given by the Langevin stochastic equation (Thomson, 1987). $\overline{u_i}$ is the mean velocity, while σ_u is the standard deviation of velocity fluctuations and T_L the Lagrangian time. $\delta\xi_{p,u_i}$ is a random variable with Gaussian p.d.f of 0 mean value and δt variance. Moreover, p owns a pseudo-mass M_p such that, under the assumption of the absence of decay process:

$$\forall t,\ M_p(t+\delta t) = M_p(t) \ \text{ with } \ M_p(t_0) = \frac{M_s}{N_p} \quad (13)$$

with M_s and N_p respectively the mass of pollutant and the number of particles released from the source s. Finally, the average concentration $\overline{C_s}(\boldsymbol{x_r},\ t_N)$ at a sensor position $\boldsymbol{x_r}$ for final time t_N is computed using a density kernel approach, i.e. as the sum of contributions from all particles in the computational domain:

$$\overline{C_s}(\boldsymbol{x_r},\ t_N) = \sum_{p=1}^{N_p} M_p(t_N)\, K\left(\boldsymbol{X_p}(t_N) - \boldsymbol{x_r}, h_r\right) \quad (14)$$

with K a kernel function, modelling the detector response function of a sensor r, acting as a spatial filter centred in $\boldsymbol{x_r}$ and h_r the smoothing radius of the chosen kernel function.

The adjoint method is applied to the forward LS Markovian explicit iterative model with least-square misfit described in the previous section. The matrices $\frac{\partial f_s^1}{\partial m_s}$, $\frac{\partial f_s^{j+1}}{\partial u_s^j}$ and $\frac{\partial f_s^N}{\partial u_s^{N-1}}$ in equation (11) then become:

$$\underbrace{\left(\frac{\partial f_s^1}{\partial m_s}\right)^T}_{n_m \times 1} = \left[\underbrace{P_{1,s}^T}_{n_m \times 7} \cdots \underbrace{P_{p,s}^T}_{n_m \times 7} \cdots \underbrace{P_{N_p,s}^T}_{n_m \times 7}\right] \quad (15)$$

$$\underbrace{\left(\frac{\partial f_s^{j+1}}{\partial u_s^j}\right)^T}_{n_j \times n_{j+1}} = \operatorname{diag}\left(\underbrace{Q_{1,s}^{j+1^T}}_{7\times 7}, \cdots, \underbrace{Q_{p,s}^{j+1^T}}_{7\times 7}, \cdots, \underbrace{Q_{N_p,s}^{j+1^T}}_{7\times 7}\right) \quad (16)$$

$$\underbrace{\left(\frac{\partial f_s^N}{\partial u_s^{N-1}}\right)^T}_{n_{N-1} \times n_N} = \left[\underbrace{R_{1,s}}_{n_N \times 7} \cdots \underbrace{R_{p,s}}_{n_N \times 7} \cdots \underbrace{R_{N_p,s}}_{n_N \times 7}\right]^T \quad (17)$$

For each particle p, the terms $P_{p,s}$, $Q_{p,s}^{j+1}$ and $R_{p,s}$ correspond, for the LS model, to the sensitivity matrixes of, respectively, the seven characteristics of p ($\boldsymbol{X_p}$, $\boldsymbol{U'_p}$ and M_p) with respect to source parameters m_s at release time, the characteristics of p at the $(j+1)^{th}$ iteration with respect to the j^{th} one (transport steps), and the concentration of particles at the sensor r with respect to the characteristics of p at iteration N. The gradient ∇J in equation (11) is now, using equations (15), (16) and (17):

$$\underbrace{\nabla J}_{n_m \times 1} = \underbrace{\left(\sum_{t=1}^{N}\sum_{p=1}^{N_p} O_{p,s}\right)}_{n_m \times n_N} \underbrace{\left[\frac{u_s^N - d^N}{d^N * d^N}\right]}_{n_N \times 1} \text{ with } O_{p,s} = \underbrace{P_{p,s}^T \left[\prod_{j=1}^{N-2} Q_{p,s}^{j+1^T}\right] R_{p,s}^T}_{n_m \times n_N} \quad (18)$$

The term $O_{p,s}$ in (18), of size $4 * n_N$ here, can be computed independently for each particle, allowing parallelization hence reducing restitution time. As previously stated, the product is calculated iteratively while the forward model is running, which avoids particle data storage at each time step and excessive memory use. Lastly, this work deals with observations taken at final iteration only. Yet, in a more general case, if measurements occur at multiple times, or with an unsteady source moving and/or emitting at various instants, $O_{p,s}$ can easily be adapted by summing equivalent terms corresponding to these cases.

APPLICATION RESULTS ON A NUMERICAL TEST CASE

A CH_4 concentration field and its adjoint ones are computed, emitted by a continuous release from a point source located 10 meters above the ground, in a steady configuration with Gaussian isotropic homogeneous turbulence and diagonal Reynolds stresses, for a neutral atmosphere. The computational domain is a test case power plant at the centre of a 600-meter radius disk. The flow part is precomputed from a CFD database and the LS model runs the dispersion part ($\delta_t = 5\text{s}$) until it converges to a steady state. A smoothing stage

is applied to the concentration field. This direct field and its adjoint ones like the emission and the Y-position ones, shown respectively in Figures 1-a), 1-b) and 2, are used to compute sensitivities on the ground, using (18). These gradients represent the opposite directions of source displacement which reduce the concentration error between the model and the observations, i.e. the cost function, allowing to approach source true characteristics. Several results and sensitivity studies will be discussed to illustrate the application of the adjoint state method to the LS model.

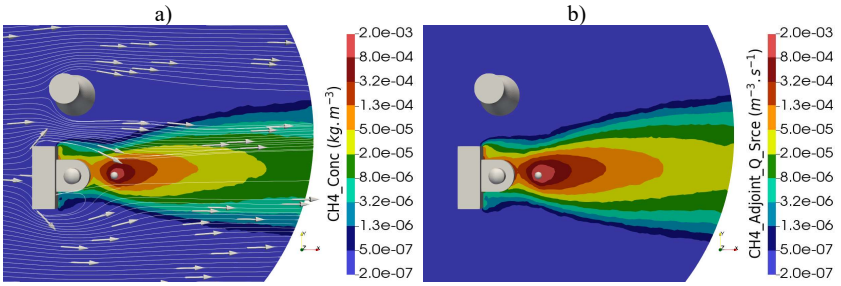

Figure 1. In log. scale: a) Concentration field with wind streamlines; b) Adjoint emission field.

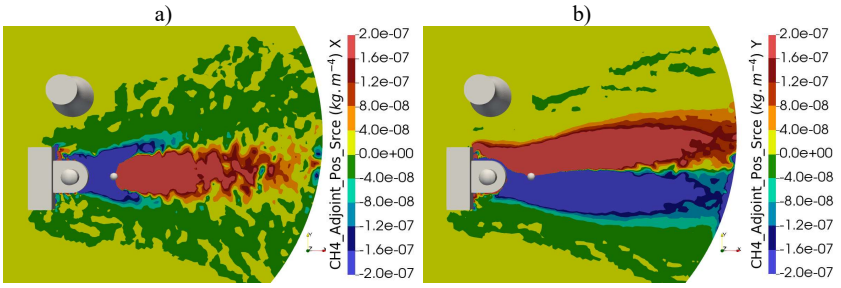

Figure 2. In linear scale: a) Adjoint X-position field; b) Adjoint Y-position field.

CONCLUSION

In this work, a new application of the adjoint method has been presented for a LS model. Use of the adjoint to solve the minimization problem avoids the computation of the whole Jacobian matrix and dependence on the number of optimization parameters, saving significant computation time. In this study, the benefits of the approach have been highlighted and its use extended to a LS model, well suited to turbulent dispersion modelling in a complex environment. Hence, the combined use of the adjoint method with a LS model is well suited to source characterization with real time constraint in complex industrial sites.

REFERENCES

Ben Salem, N., L. Soulhac, P. Salizzoni and M. Marro, 2014: Pollutant source identification in a city district by means of a street network inverse model. *Int. J. Environ. and Pollution,* **55** (1-4), 50-57.

Chavent, G., 1974: Identification of functional parameters in partial differential equations. *Identifications of parameter distributed systems*, 31-48, Goodson, RE. & Polis, USA.

Giering, R., 2000: Tangent linear and adjoint biogeochemical models. *Geophys. Monogr. Ser.,* **114**, 33-48.

Gill, P.E., W. Murray and M.H. Wright, 1981: Practical Optimization, Academic Press Ltd, USA, 401 pp.

Keats, A., E. Yee and F.-S. Lien, 2007: Bayesian inference for source determination with application to a complex urban environment. *Atmos. Env.,* **41** (3), 465-479.

Plessix, R.-E., 2006: A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys. J. Int.,* **167**, 495-503.

Pudykiewicz, J.A., 1998: Application of adjoint tracer transport equations for evaluating source parameters. *Atmos. Env.,* **32** (17), 3039-3050.

Thomson, D. J., 1987: Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *J. Fluid Mech.*, **180**, 529–556.