H13-234

DESIGNING A MONITORING SYSTEM FOR A SEMI-URBAN SET OF BUILDINGS

Jean-Pierre Issartel and Tristan Gamel

Délégation Générale pour l'Armement, Maîtrise NRBC, Vert le Petit, France

Abstract: A prototype system is being developed for the surveillance of air quality and security at an industrial facility manipulating chemical compounds. This system is based on a recent approach of data assimilation highlighting the importance of an auxiliary geometry, interpreted as an apparent geometry since the close environment of the detectors is maximised as happens in eye's view. The technique allows to optimally design a network that will detect rapidly an accidental release to the atmosphere. It also allows identifying one or several simultaneous accidental releases. This is important for the immediate management of the consequences and also to later discriminate exploiter's responsibility. In addition to the presentation of the theoretical framework of inverse modelling, the work raises several issues with the difficulties of urban dispersion modelling and with the operational constraint of real time computations**.**

Key words: *Industrial security, trace species, inverse problem, source identification.*

INTRODUCTION

In this study, we want to monitor accidental chemical releases in order to ensure the security of a set of industrial buildings. Whereas most studies in assimilation of data have the Earth or continents as a theater, this study corresponds to a local semi-urban area. The work is illustrated with a set of buildings (figure 1) in Aix en Provence (southern France). The meteorology is chosen corresponding to local average conditions corresponding to Pasquill class C stability, with a wind of 3.0 ms^{-1} from SE at 10 m above the ground and away from the buildings. The wind goes across the 600 m of the modeled area within approximately 4 minutes and thus, a time interval of 15 minutes is utilized in the computations. The scattering of a trace species is simulated by the model PANEPR appropriate for the small scale urban areas. Ten detectors are settled among the buildings as indicated on the figures by the red dots and 2 m above the ground. Each of them performs one concentration measurement each thirtieth second. We want to use this sequence of artificial observations to retrieve the origin, time and intensity of a point release. This inverse problem is addressed here with a non Bayesian technique based on a new concept of visibility. A distinction is made between the possible distribution of the emissions and the visibility of these emissions actually provided by the monitoring network in the prevailing meteorological conditions. This visibility is described based on a geometric weighting of the various parts of the environment according to a visibility function introduced by Issartel, J.-P. (2005) and developed by Issartel, J.-P. *et al.* (2007). The artificial data are free from the noises due to the detectors or to the difference between model and reality. The imperfection of the real detectors and the imperfection of the dispersion model are an essential but separate difficulty addressed, for instance, by Sharan, M. *et al.* (2009).

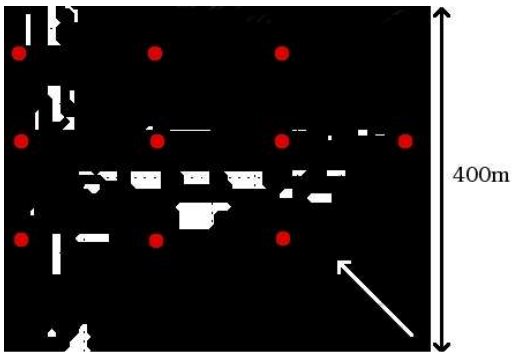

Figure 1. The semi-urban area considered for the illustration of the present study. The industrial buildings are indicated in white colour. The arrow indicates the main wind direction (3 $m.s^{-1}$).

PRIOR ASSUMPTIONS ABOUT POSSIBLE EMISSIONS: THE FUNDAMENTAL GEOMETRY

The most general description of a tracer source may be given as a function σ of the horizontal position x, y, altitude z and time t, $\sigma(x,y,z,t)$ being a rate of release in unit of tracer per kg of air and per second. The mixing ratio of tracer $\chi(x,y,z,t)$ is obtained from advection diffusion equation (1). Each measurement μ_i is obtained as an integral in which the sampling function π_i describes where and when the air of the sample was taken; π_i reduces to a Dirac delta function for an instantaneous measurement at a point.

$$\frac{\partial \chi}{\partial t} + v\nabla\chi + \frac{\partial}{\partial z}\left(\kappa\frac{\partial \chi}{\partial z}\right) = \sigma \qquad \mu_i = \int_{\Omega\times T} \rho\chi\pi_i(x,y,z,t)dxdydzdt = (\chi,\pi_i) \qquad (1)$$

in which v, κ and ρ are the fields of wind, diffusion coefficient and air density. As explained by Issartel, J.-P. and J. Baverel, (2003) the expression for μ_i is a scalar product denoted (,). It may be transformed using an adjoint function or retroplume r_i subject to the retrograde advection diffusion equation:

$$-\frac{\partial r_i}{\partial t} - v\nabla r_i + \frac{\partial}{\partial z}\left(\kappa\frac{\partial r_i}{\partial z}\right) = \pi_i \qquad \mu_i = \int_{\Omega\times T} \rho\sigma r_i(x,y,z,t)dxdydzdt = (\sigma,r_i) \qquad (2)$$

Given the Euclidean geometry of the scalar product (,), the adjoint function describes the sensitivity of the measurement with respect to the various locations and times of the emissions.

It is assumed that the sought source is located on the ground ($z = 0$) and out of the buildings. Then, the emissions are better described by a flux $s(x,y,t)$ in kg m^{-2} s^{-1}. The measurement μ_i becomes an integral through the time interval T and the ground surface Σ_{ext} with the buildings discarded; δ is the Dirac delta function :

$$\sigma(x,y,z,t) = \frac{s(x,y,t)\delta(z)}{\rho(x,y,z,t)} \qquad \mu_i = \int_{\Sigma_{ext}\times T} s a_i(x,y,t)\,dxdydt = (s,a_i)_1 \qquad a_i(x,y,t) = \int_T r_i(x,y,0,t)\,dz \tag{3}$$

In case of an instantaneous point release, the emission function $\sigma(x,y,z,t)$ or $s(x,y,t)$ are proportional to a Dirac delta function. Under the prior hypotheses of superficial outer emissions $s(x,y,t)$, the measurements are accounted for by the new product $(,)_1$. This product is the geometric counterpart of the prior hypotheses and we designate it as the *fundamental product*. This concept, raised in (Issartel J.-P. *et al*. 2007), was formalized by Sharan, M. *et al.* (2009).

EMISSIONS VISIBLE BY THE MONITORING NETWORK: THE RENORMALIZED GEOMETRY

The equations $\mu_i = (s,a_i)_1$ are used to obtain the projection $s_{//} = \sum \lambda_i a_i$ of s on the space spanned by the a_i . The coefficients λ_i are obtained by inverting a matrix **H**:

$$\boldsymbol{\lambda} = \mathbf{H}^{-1}\boldsymbol{\mu} \qquad \text{where} \qquad H_{ij} = (a_i,a_j)_1 \qquad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \\ \lambda_n \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \\ \mu_n \end{bmatrix} \tag{4}$$

This simple method returns an unsatisfactory estimate $s_{//}$ (Issartel J.-P. *et al*. 2007) with artifacts in the form of peaks at detector locations. These peaks are related to singularities of the adjoint functions at detector location. They can be interpreted as artificial information with an excessive attention paid to the surrounding of the detectors. This bias in the visibility of the monitoring network is described through an illumination function E :

$$E(x,y,t) = \boldsymbol{a}(x,y,t)^T \mathbf{H}^{-1}\,\boldsymbol{a}(x,y,t) \qquad \text{with} \qquad \boldsymbol{a}(x,y,t) = \begin{bmatrix} a_1(x,y,t) \\ \\ a_n(x,y,t) \end{bmatrix} \tag{5}$$

Since E is positive and $\int E(x,y,t)\,dx\,dy\,dt = n$ (Issartel, J.-P., 2005), the illumination is interpreted as a density of information. To remove the singularities of E at detector locations, a renormalizing function $f(x,y,t) > 0$ is introduced and the measurements are rewritten as :

$$\mu_i = \int_{\Sigma_{ext}\times T} f s a_{fi}(x,y,t)\,dxdydt = (s,a_{fi})_f \qquad \text{with} \qquad a_{fi}(x,y,t) = \frac{a_i(x,y,t)}{f(x,y,t)} \tag{6}$$

This new expression is associated with the renormalized product $(,)_f$ weighted by f . In the new weighted geometry, the source is estimated according to its projection $s_{//f} = \Sigma\, \lambda_i\, a_{fi}$ with now coefficients: $\boldsymbol{\lambda} = \mathbf{H}_f^{-1}\boldsymbol{\mu}$ where $H_{f\,i\,j} = (a_{fi},a_{fj})_f$. The illumination becomes:

$$E_f(x,y,t) = f(x,y,t)\,\boldsymbol{a}_f(x,y,t)^T\,\mathbf{H}_f^{-1}\,\boldsymbol{a}_f(x,y,t) \tag{7}$$

Again, E_f is positive and $\int E_f(x,y,t)\,dx\,dy\,dt = n$. The optimal weights φ obey the renormalising condition that for all x, y, t :

$$\varphi(x,y,t) = E_\varphi(x,y,t) \qquad \text{equivalent to} \qquad \boldsymbol{a}_\varphi(x,y,t)^T\,\mathbf{H}_\varphi^{-1}\,\boldsymbol{a}_\varphi(x,y,t) = 1 \tag{8}$$

The figure 2 shows the optimal renormalizing function φ and figure 3 shows renormalized inversion clearly improved compared to the more classical one.

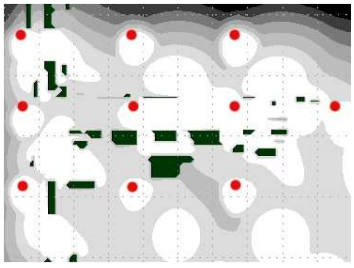

Figure 2. The renormalizing function φ is shown at a given time with an arbitrary colour scale, the grey colour levels correspond to a factor 10. Notice that φ decreases away from the monitoring network and indicates a visibility mainly in the upwind direction.

The function $s_{//\varphi}$ is continuously distributed through space and time. Thus, in general it cannot be used to estimate a point release directly. However, $s_{//\varphi}$ has an essential property emphasised by the figure 3: in case the observed measurements are produced from a point release, $s_{//\varphi}$ is maximum for the location and date of this release. This allows to identify it from the observations. Indeed, the estimate associated with observations $\boldsymbol{\mu}$ may be written: $s_{//\varphi}(x,y,t) = \boldsymbol{\mu}^T \mathbf{H}_{\varphi}^{-1} \boldsymbol{a}_{\varphi}(x,y,t)$. If the release is from location and time (x_0, y_0, t_0) with the intensity q_0, the measurements are proportional to the local value of the sensitivity functions, i.e. $\boldsymbol{\mu}_0 = q_0 \boldsymbol{a}(x_0, y_0, t_0) = q_0 \varphi(x_0, y_0, t_0) \boldsymbol{a}_{\varphi}(x_0, y_0, t_0)$. From these observations $\boldsymbol{\mu}_0$ we would obtain, writing $C_0 = q_0 \varphi(x_0, y_0, t_0)$, a renormalized estimate:

$$s_{//\varphi 0}(x,y,t) = C_0 \boldsymbol{a}_{\varphi}(x_0, y_0, t_0)^T \mathbf{H}_{\varphi}^{-1} \boldsymbol{a}_{\varphi}(x,y,t) \leq C_0 \qquad (9)$$

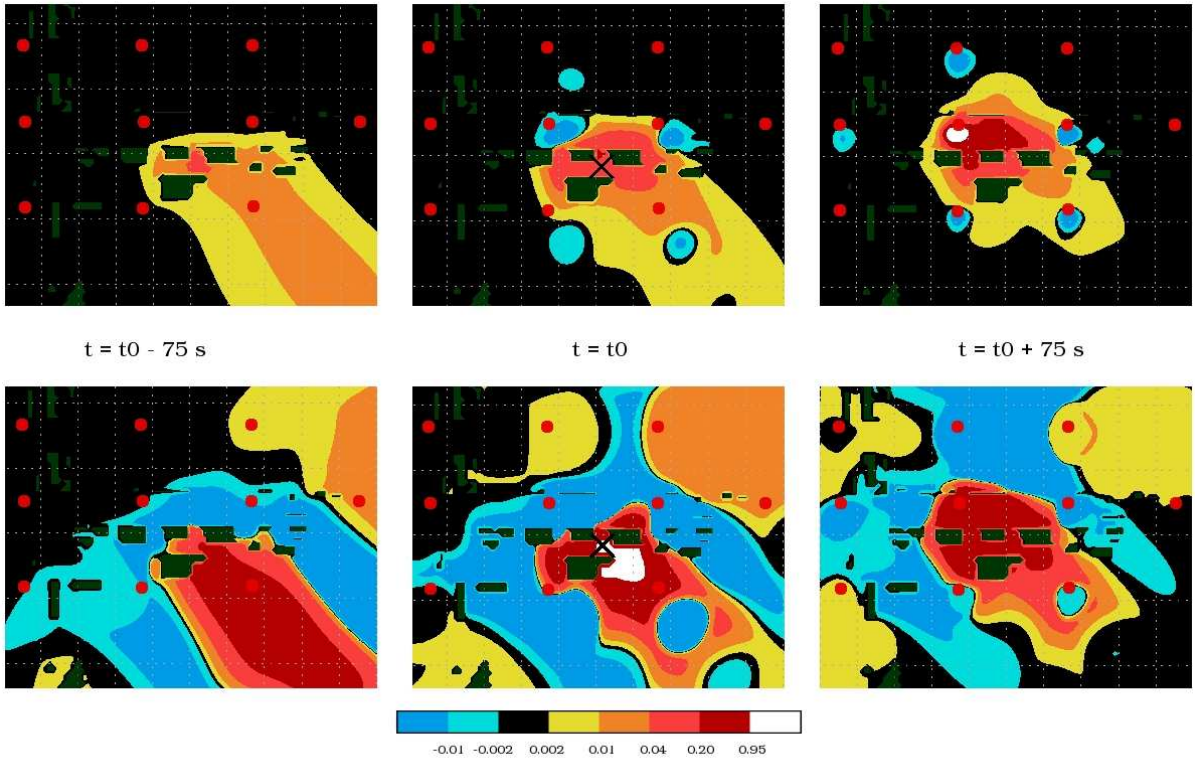

Figure 3. Non-renormalized estimation $s_{//}$ (top panels) and renormalized estimation $s_{//\varphi}$ (bottom panels) at times t_0-75s, t_0 and t_0+75s, obtained from 100 measurements generated from a point release (black cross) at t_0. The non renormalized estimation is maximum at detector location whereas the renormalized estimation is maximum at release location. The colour levels (arbitrary units) are separated by a factor 5, blue levels representing negative values

The Cauchy-Schwarz inequality, $(\boldsymbol{a}^T \mathbf{H}_{\varphi}^{-1} \boldsymbol{b})^2 \leq (\boldsymbol{a}^T \mathbf{H}_{\varphi}^{-1} \boldsymbol{a})(\boldsymbol{b}^T \mathbf{H}_{\varphi}^{-1} \boldsymbol{b})$, and the renormalizing condition (equation 8) imply that $s_{//\varphi 0}$ becomes maximum exactly for the sought location and time (x_0, y_0, t_0). The renormalized assimilation thus provides a unique possibility to determine the origin of an accident from remotely observed concentrations. The quality of this identification is better, with a sharper maximum, if the release happens in a well illuminated region (Sharan, M. *et al.*, 2009). This is a criterion for network design.

NUMBER OF MEASUREMENTS REQUIRED FOR IDENTIFYING SIMULTANEOUS RELEASES

In order to obtain criteria for designing a monitoring network, it is necessary to determine the arrangement and number of detectors required. In particular, it is important to discriminate several small releases simultaneously detected by various detectors from a bigger one detected by the whole network. The theory for identification of single point releases can be extended for the identification of several simultaneous point releases. Figure 4 shows that, when a set of measurements $\boldsymbol{\mu}_0$ has been generated from several point releases, the estimated function $s_{//\varphi}$ displays local maxima close to the location of each of the releases. This property is supported by theoretical arguments and may be exploited to identify the location and intensity of the releases from efficient computations and with accuracy limited only to the noise in the data.

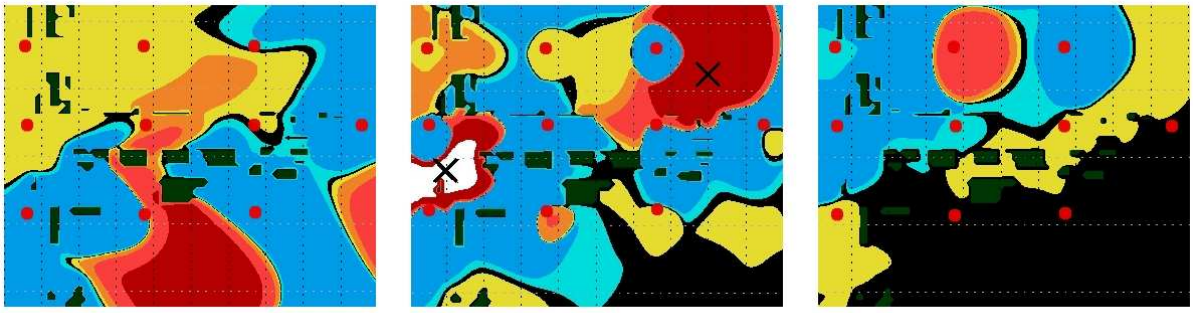

Figure 4. The renormalized estimation $s_{//\varphi}$ having two local maxima at location and time of two simultaneous releases (black crosses) at t_0.

The computations with artificial model data emphasize a surprising result about the minimum number of concentration measurements required for identifying $m > 1$ simultaneous point releases. In order to reduce the computational volume, time dimension has been suppressed by considering point sources $s(x,y)$ (figure 4 is built without this simplification, i.e. with $s=s(x,y,t)$). This brings some minor modification in the fundamental geometry corresponding to the Euclidean geometry of the ground only. At first glance, the minimum number of measurements required for identifying the various releases should be $3m$ because each release is characterized by two parameters for horizontal location and intensity. However, we have verified in synthetic experiments that a lesser number is sufficient. We are able to retrieve two simultaneous releases from at least five synthetic measurements and three simultaneous releases from at least six synthetic measurements. The pair or triplet of locations and intensities are retrieved exactly in all of many repeated trials except when one of the releases is placed in a weakly seen region. This surprising result may be partly understood. When n measurements are performed, it seems that the adjoint vectors $\boldsymbol{a}(x_1, y_1)$, $\boldsymbol{a}(x_2, y_2)$,…, $\boldsymbol{a}(x_n, y_n)$ computed at n locations should be, in general, linearly independent.

Notice that m releases of intensity $q_1, q_2,\ldots, q_m$ at locations (ξ_1, ζ_1), (ξ_2, ζ_2),…, (ξ_m, ζ_m) generate a measurement vector ideally free from noises :

$$\boldsymbol{\mu}_0 = q_1 \boldsymbol{a}(\xi_1, \zeta_1) + q_2 \boldsymbol{a}(\xi_2, \zeta_2)+\ldots+ q_m \boldsymbol{a}(\xi_m, \zeta_m) \qquad (10)$$

The identification of the releases amounts to identify the above decomposition of the measurement vector. We argue that this decomposition is unique, in general, provided that $n \geq 2m$. Indeed, if this decomposition is not unique, we can write two alternative decompositions:

$$\boldsymbol{\mu}_0 = q_1 \boldsymbol{a}(\xi_1, \zeta_1) + q_2 \boldsymbol{a}(\xi_2, \zeta_2)+\ldots+ q_m \boldsymbol{a}(\xi_m, \zeta_m) = q'_1 \boldsymbol{a}(\xi'_1, \zeta'_1) + q'_2 \boldsymbol{a}(\xi'_2, \zeta'_2)+\ldots+ q'_m \boldsymbol{a}(\xi'_m, \zeta'_m) \qquad (11)$$

Then, by defining $(x_i, y_i) = (\xi_i, \zeta_i)$ and $(x_{m+i}, y_{m+i}) = (\xi'_i, \zeta'_i)$ for $i = 1, 2,\ldots, m$, we obtain a family $\boldsymbol{a}(x_1, y_1)$, $\boldsymbol{a}(x_2, y_2)$,…, $\boldsymbol{a}(x_{2m}, y_{2m})$ of less than n adjoint vectors corresponding to different locations. According to equation (11) they are linearly dependent. However, this is in contradiction with the aforementioned generic property. The decomposition is unique and this implies that it may be identified (for instance by looping over all m-tuples of locations). If the number of measurements is sufficiently large, i.e. $n \geq 2m$, the m releases and their locations can be identified. The computations show that the six parameters corresponding to the location and intensity of three releases may be identified from five measurements. For identifying the six parameters associated with two point releases, at least five measurements are required in the computations. This is less than the number of parameters but does not coincide exactly with the condition $n \geq 2m$. Further investigations are required to fully understand the numerical facts. This will help in designing monitoring networks with the least number of samplers.

CONCLUSIONS

The model PANEPR for fluid advection diffusion was used with winds calculated around the buildings for the prescribed boundary condition corresponding to a Pasquill class C stability and a free wind of 3 ms^{-1} from SE. The adjoint functions were obtained from a backward integration of the retrograde transport equation on a grid of 5.0×5.0 m^2 and a time step of 5s. Due to the constant wind conditions, the adjoint functions of the successive measurements at each given detector are identical except for a time shift. The computations made with artificial measurements produced free from noises illustrate the usefulness of the proposed non-Bayesian framework for monitoring atmospheric contaminations. Contaminations due to industrial accidents, most of the time, may be regarded as point releases. The proposed technique allows to identify the origin of such releases. The advantages of the proposed technique, compared to earlier ones, are many. First, the computations are fully and simply realizable without summoning any arbitrary simplifying assumptions. Second, the visibility functions φ clearly indicates the regions well or badly seen by the monitoring network and the contribution of each detector to the global performance. Third, the technique turns the set of measurements into visual information showing the structure of the emissions. This visualization is limited by the resolution capability of the monitoring network. In particular, several simultaneous contaminations may be discriminated. In an industrial context, this is important to discriminate several minor releases from a major accident. In addition, it is possible to clarify responsibilities.

This study raises issues of practical and theoretical importance. The effect of the noise is not described in the present abstract but a more extensive study (Sharan, M. *et al.*, 2009) shows acceptable performance of the noisy inversion. In principle, the proposed technique is based on the simultaneous processing of the whole set of measurements available at all times. However, this is not practically realizable because the continuous flow of measurements is rapidly enormous. Accordingly, it

will be necessary to investigate the fading correlations between measurements separated by large time intervals. Finally, a surprising result is observed numerically: the number of measurements required for identifying several simultaneous point releases is strictly less than the number of parameters required to describe them. A tentative explanation has been proposed. This result, relying on the global properties of the informational system is contrary to all classical descriptions. Its complete understanding is left for further investigations but will be probably helpful in optimizing the design and exploitation of monitoring networks.

REFERENCES

Issartel, J.-P. and Baverel, J., 2003: Adjoint backtracking for the verification of the CTBT, *Atmospheric Chemistry and Physics*, 3, 475-486.

Issartel, J.-P., 2005: Emergence of a tracer source from air concentration measurements: a new strategy for linear assimilation, *Atmospheric Chemistry and Physics*, **5**, 249-273.

Issartel, J.-P., Sharan, M., Modani, M., 2007: An inversion technique to retrieve the source of a tracer with an application to synthetic satellite measurements, *Proceedings of the Royal Society*, **463**, 2863-2886.

Sharan, M., Issartel, J.-P, Singh, S. K., Kumar, P., 2009: An inversion technique for the retrieval of single point emissions from atmospheric concentration measurements, *Proceedings of the Royal Society*, **465**, 2069-2088.