TWO JOINT ESTIMATION DECISION METHODS FOR POLLUTANT SOURCE MONITORING : A COMPARISON.

*G . Delmaire , R . Lherbier , E. Carlier**
Laboratoire d'Analyse des Systèmes du Littoral,
195 Avenue du Pasteur Martin Luther King ,
62228 CALAIS BP 649 France
* Laboratoire d'Artois de Mécanique et Habitat (LAMH)
62400 Béthune France
Tel: +33 3 21465664
e-mail : gilles.delmaire@lasl.univ-littoral.fr

Abstract
In this paper, the problem of pollutant source localisation and flow is addressed. It is tackled in a one dimensional context such as river or tunnel with the aid of one remote sensor. N possible sources are considered, one of whom is operating. A specific effort is made upon the impulse responses family which are approximated as a bank of FIR filters with specific delays and orders. In order to reduce the set of solutions, a source flow model is considered which introduces time bounds. Using these features, a joint estimation decision is derived in a bayesian framework and it is compared to the case when source assumptions are dropped. The benefit for this kind of solution is shown practically in terms of localisation quality.

1. INTRODUCTION

Restoration consists in recovering a signal or an image given a corrupted by noise signal and a model. This model is usually linear or linearized around an operating point and measures are corrupted with additive noise. It is usually solved by a functional minimisation [Dem89] [Bla00] [Com84] composed of a term representing the distance to a given model and a regularisation term. This is a case when we consider a spatio-temporal phenomena such as pollutant propagation described by a model H inspected by a few sensors measures at specific locations. The temporal source flow may be recovered according to this framework.

In addition to this estimation problem, the active source has to be chosen among the set of possible ones. Only a few sensors are available so that a whole image of the situation is illusory. In this case, extra a priori information is needed. Two kinds of hypotheses are made on f :

- The flow source is supposed to be punctual that is to say $f(x,nTe) = j(nTe)\,\delta(x-x_0)$
where δ stands for the dirac function and x_0 the source position. x_0 may be chosen inside a compact set or among a finite set of possible positions. This formulation discards the case of convolutive mixtures.
- The activity of the source is time limited between n_u and N_u : $j(nTe)=(\Gamma((n-n_u)Te)-\Gamma((n-N_u)Te))\,u(nTe)$
where Γ accounts for the step function.

A localisation criteria may be derived using one or both previous assumptions in a bayesian context.

The purpose of this work is to present a unified probabilistic framework to the joint estimation problem in the case of simple probability densities. Previous specific assumptions about sources amount to carry out a joint estimation of the position x_0 the time bounds n_u N_u and the input u(t). It 's practically compared to the no source model case in terms of localisation quality.

2 . NOTATIONS

Let $\underline{y}=[y_1; \ldots\ldots y_m]$ be the vector of time measures where m is the size of the vector. The reduction to its significant part will be denoted $\underline{y}_r$.
The same rules apply to $\underline{u}$. M is the size of $\underline{u}$.
A starting sample time of the vector $\underline{y}$ will be denoted n_y whereas the ending sample number uses capital letters N_y. The same rules apply to $\underline{h}$ and $\underline{u}$.

3 . PROBLEM STATEMENT

The direct propagation model may be expressed :

$$\underline{y} = H(d)\,\underline{u} + \underline{b} \qquad (1)$$

where $\underline{b}$ is a random vector representing the noise, $\underline{u}$ is the temporal input flow and $\underline{y}$ the vector of concentration temporal measures. A similar formulation occurs in array processing where DOA have to be recovered [Mar99].

H(d) stands for a Toeplitz matrix made with the components of the impulse response obtained through the resolution of the advection diffusion PDE. d accounts for the source sensor distance and is to be chosen among a set of possible values. The expression of the impulse response [Ser97] (also called for physicians the green function) is :

$$h(t,d,U,D)= \frac{1}{\sqrt{4\,\Pi\,D\,t}}\; e^{-\frac{(d-U\,t)^2}{4\,D\,t}} \qquad (2)$$

where d is the source sensor distance, U the velocity and D the diffusion parameter. In this framework, (U,D) are assumed to be known via a previous identification step. Obviously, the impulse response appears to be infinite. H(d) requires the knowledge of the source sensor distance d.

The main problem is to recover simultaneously the distance d and the input flow u(t). A difficulty may be encountered in regularising $\underline{u}$ which disturbs decision making about d. A similar framework is proposed in [Dja97] but the main difference here with this work is that the distance parameter to be restored act on the impulse response so that the matrix H condition number changes with d leading to major difficulties.

In a mixed estimation decision problem, the model family has to be completely described with special emphasis on the time bounds of the model.

3 PROPAGATION MODEL FEATURES

The previous impulse response appears to be infinite. This property affects the computational load, as a result, it would be interesting to approximate these parametric filters as a set of FIR filters. Their major features are :

- they are mass conservative over time for fixed d.
- their maximum value is located at

$$t_M = \frac{1}{U}\left(-\frac{D}{U} + \sqrt{\frac{D^2}{U^2} + d^2} \right)$$

Under the condition D << U.d, t_M is approximately equal to $\frac{d}{U}$.

Then, symmetric finite impulse responses may be derived around t_M . A natural choice is to find the time bounds according to a mass conservation criteria. They can be numerically computed for every distance d and for fixed parameters (U,D).

These time bounds can be modelled as a function of d, conversely d can be expressed as a polynomial function of t : **$d = a\,(t - d/U)^2$** (3)

where a is a constant depending on (U,D) parameters.

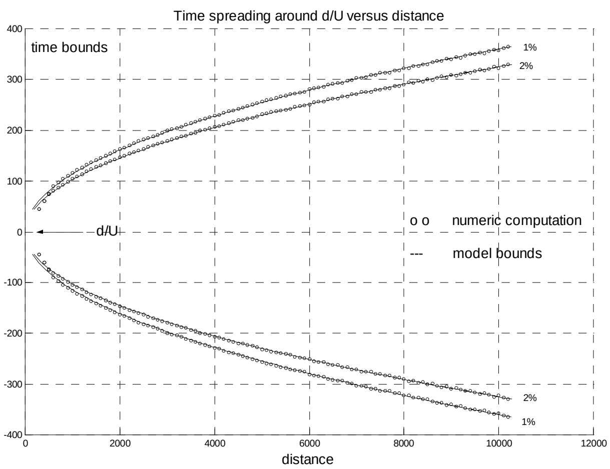

Figure 1: time bounds versus distance

For example, figure 1 shows the time spreading of the pollutant versus distance in the case of 1% and 2% loss of the whole mass. The dotted curve stands for the numeric computation while the solid one accounts for model bounds. Notice that the model sticks very well to numeric computations. Conversely, an expression of the delay and the memory of each filter may be obtained from (3) :

$$t_h(d) = \frac{d}{U} - \sqrt{\frac{d}{a}} \qquad T_h(d) = \frac{d}{U} + \sqrt{\frac{d}{a}} \qquad (4)$$

These results can be used to model approximately the bank of filters for various distances. In discrete time operating conditions, this leads equivalently to find out the delay and the order of each source sensor path.

4 JOINT ACTIVE SOURCE DETECTION AND ESTIMATION

We develop here two methods based on the previous source assumptions. Both rely on a method investigated by Trussel [Tru79] which was initially devoted to restoration and adapted for joint estimation decision. The first one only assumes that the source is punctual whereas the other supposes that both conditions are checked (a source has a cycle of life : a birth time and a death time) [Del00a] [Del00b].

4.1 Without *a priori* on the input signal

Let $\underline{\theta}$ be the vector lumping together the distance d and the input vector : $\underline{\theta} = [d;\underline{u}]$.

From the model (1) given in section 3, the likelihood of the measures conditionally to $\underline{\theta}$ is a normal density with variance matrix $R_b = \sigma_b^2\, I$.

$$p(\underline{y}/\underline{\theta}) = (2\,\Pi\,\sigma_b^2)^{-m/2}$$

$$\exp\left(-1/(2\,\sigma_b^2).(\underline{y} - H(d)\,\underline{u})^t\,.\,(\underline{y} - H(d)\,\underline{u})\right)$$

with m=dim($\underline{y}$)

The parameter $\underline{\theta}$ law may be derived using conditional probabilities : $p(\underline{\theta}) = \sum_{i=1}^{N} p_i\; p(\underline{u}/d_i)$

The conditional probability may be expressed as a gaussian density : $p(u/d) \sim \mathcal{N}(\underline{u}/\,\bar{\underline{u}}(d), R_u(d), d)$

Bayes' theorem enables to express the posterior probability density : $p(\underline{\theta}/\underline{y}) \propto p(\underline{y}/\underline{\theta}).p(\underline{\theta})$

This leads to the following expression :

$$-2\ln p(\underline{\theta}/\underline{y}) = (\underline{u} - \bar{\underline{u}}(d))^t.\,R_u(d)^{-1}\,(\underline{u} - \bar{\underline{u}}(d))$$

$$+ 1/\sigma_b^2.(\underline{y} - H(d)\,\underline{u})^t\,(\underline{y} - H(d)\,\underline{u}) - 2\ln p(d) + K$$

Since sources have got equal probabilities, the previous relation may be reduced to the minimisation of the compound criteria :

$$J(d,\underline{u}) = 1/\sigma_b^2.(\underline{y} - H(d)\,\underline{u})^t\,(\underline{y} - H(d)\,\underline{u})$$

$$+(\underline{u} - \bar{\underline{u}}(d))^t.\,R_u(d)^{-1}\,(\underline{u} - \bar{\underline{u}}(d)) \qquad (5)$$

This is a very classical criteria but some difficulties may happen in this specific context :

The variance matrix $R_u(d)$ may be chosen equal for every source but then the restoration quality varies from one source to another since H(d) is more and more ill-conditioned when d grows.

The variance matrix $R_u(d)$ may also be set different from one source to another but this leads to prioritise some sources with respect to the others. The choice for one strategy is somehow difficult.

4.2 MAP solution using source assumptions.

The lack of spatial data suggests us to use an extra *a priori* information about the input source. It is supposed to operate during two time bounds. In a pollution context, this supposes that the source emission is dense. This a reasonable hypothesis since accidental spills are generally continuous from the beginning till the end of the spill or a human intervention. From this hypothesis, it follows that sensor measurements are continuous inside two sample times $[n_y;N_y]$. These times may be computed off line through a change detection algorithm $\hat{n}_y$ and $\hat{N}_y$ [Bas88].

Ideally n_y (resp. N_y) is deduced from $n_y Te = t_h + n_u Te$ (resp. $T_h + n_u Te$) due to the convolution product.

An estimate of n_u (resp Nu) may be obtained using :

$$\hat{n}_u(d) = \hat{n}_y - \left\lfloor \frac{t_h(d)}{Te} \right\rfloor \text{ (resp. } \hat{N}_y - \left\lfloor \frac{T_h(d)}{Te} \right\rfloor \text{)} \quad (6)$$

These relations may be viewed as hard constraints or flexible constraints.

We choose to develop a sequential algorithm described by two stages. The first one is concerned with the change points of the measures. Its results are modelled using some specific probability densities which are used to feed the second stage. This one is concerned with the joint localisation and input source estimation. The only role of the first step is to propose a fair initialisation of the change points. This section just outlines the major steps of the second stage which are detailed in a recent work [Del02].

4.2.1 Prior probabilities

From the model (1) given in section 3, the likelihood of the measures conditionally to $\underline{\theta}$ is a normal density with variance matrix $R_b = \sigma_b^2$ I. It may be split into 3 independent parts describing the operating conditions of the source (during life, after death, and before birth). Gathering these results provides a well known result :

$$p(\underline{y}/\underline{\theta}) = (2\,\Pi\,\sigma_b^2)^{-m/2} \exp\left(-1/(2\,\sigma_b^2).(\underline{y}-H(d)\,\underline{u})^t.(\underline{y}-H(d)\,\underline{u})\right)$$

with $m=\dim(\underline{y})$ and $\underline{u} = [0_{n_u-1}\ \ \underline{u}_r^t\ \ 0_{M-N_u}]^t$

The same mechanism (as in section 4.1) may be implemented to compute $p(\underline{\theta})$. The unknown parameter becomes :

$$\underline{\theta} = [d; [\underline{u}_d^t\ \underline{u}_r^t\ \underline{u}_f^t], n_u, N_u, n_y, N_y]$$

The conditional law of $\underline{u}$ may be computed using the model of the input source (see source assumptions). (n_u, N_u) are considered as random variables whose probability densities are a priori known. Hierarchical models can be derived from this setup.

$$p(\underline{u}/d) = p(n_u,N_u/d).p(\underline{u}_r/\bar{\underline{u}}_r,n_u,N_u).p(\underline{u}_d/d,n_u,N_u).p(\underline{u}_f/d,n_u,N_u)$$

where $\underline{u}$ is partitioned into 3 independent parts :

$\underline{u} = [\underline{u}_d^t\ \underline{u}_r^t\ \underline{u}_f^t]^t$ where $\dim(\underline{u}_d) = n_u - 1$, $\dim(\underline{u}_r) = N_u - n_u + 1$, $\dim(\underline{u}_f) = M - N_u$.

According to the assumptions that these sample times are following uniform densities, the conditioned probability of $\underline{u}$ may be condensed after some simple computations in the form :

$$p(\underline{u}/d, \bar{\underline{u}}(d), R_u(d), n_u, N_u) = \mathcal{N}(\underline{u}/\bar{\underline{u}}, R_u(d), n_u, N_u, d)$$

where $R_u(d)$ is a block diagonal matrix made with components of specific size :

$R_u(d) = \text{blockdiag}\{R_{ud}, R_{ur}, R_{uf}\}$ with $R_{ur} = \sigma_{ur}^2$ I

$R_{ud} = \sigma_{ud}^2$ I, $R_{uf} = \sigma_{ud}^2$ I and $\bar{\underline{u}}^t = [0\ \ \bar{\underline{u}}_r^t\ \ 0]^t$.

The size of the variance matrices are related to the size of the associated input vector. The choice of the standard deviations is such that $\sigma_{ur} \gg \sigma_{ud}$ in order to constrain $\underline{u}_d$ and $\underline{u}_f$ to be close to zero.

The whole *a priori* probability may be expressed in the form :

$$p(\underline{\theta}) = p(d)\, 1_{[\hat{n}_u-\delta;\ \hat{n}_u+\delta]}\, 1_{[\min(\hat{N}_u-\delta,\, n_u);\ \hat{N}_u+\delta]}\, \mathcal{N}(\underline{u}/\bar{\underline{u}}(d), R_u(d), n_u, N_u, d) \quad (7)$$

For convenience in the following, the dependence with respect to d is dropped in $R_u(d)$ and $\bar{\underline{u}}(d)$.

4.2.2 MAP solution

The Bayes' rule leads to the result :

$$-2\ln p(\underline{\theta}/\underline{y}) = 1/\sigma_b^2.(\underline{y}-H(d)\,\underline{u})^t(\underline{y}-H(d)\,\underline{u}) - 2\ln p(d) + (\underline{u}-\bar{\underline{u}})^t.R_u^{-1}(\underline{u}-\bar{\underline{u}}) + 2\ln(2\Pi.\sigma_b^2)^{m/2} + 2\ln[(2\Pi)^M.|Ru|]^{1/2} + K$$

The previous expression may be reduced to the optimisation of the following criteria in the case when sources have got equal probabilities:

$$J(n_y, N_y, n_u, N_u, d, \underline{u}) = 1/\sigma_b^2.(\underline{y}-H(d)\,\underline{u})^t(\underline{y}-H(d)\,\underline{u}) + \ln(|Ru|) + (\underline{u}-\bar{\underline{u}})^t.R_u^{-1}(\underline{u}-\bar{\underline{u}}) \quad (8)$$

where indexes are chosen in their parameters range.

The term involving |Ru| is depending on the size of the reduced input vector. The dimension of the reduced input vector appears explicitly in this term :

$$J(n_y,N_y,n_u,N_u,d,\underline{u}) = 1/\sigma_b^2.(\underline{y}-H(d)\,\underline{u})^t(\underline{y}-H(d)\,\underline{u}) + (N_u-n_u+1).\ln(\sigma_{ur}^2/\sigma_{ud}^2) + (\underline{u}-\bar{\underline{u}})^t.R_u^{-1}(\underline{u}-\bar{\underline{u}}) + k \quad (9)$$

This is the main difference with the previous criteria given in (5) and more generally with classical methods. Indeed, classical methods use a constant variance matrix which results in the removal of this term in a usual criteria.

Using the specific form of the variance matrix, it turns out that the criteria penalises long time spills more than short ones ($\dim(\underline{u}_r)$ is greater, |Ru| is bigger) and as a result enables far sources to be as well considered as near sources. Indeed, some computations may show that the input size includes a term of the kind $-\frac{1}{Te}\sqrt{\frac{d}{a}}$ which roughly penalises near sources.

5. SIMULATION RESULTS

To assess the performances of both methods, we investigate an example where the pollutant is moving in a river . The current velocity and the diffusion parameter are supposed to be known and equal to : U= 1m/s and D= $1m^2/s$. Five sources are considered, they are located every four kilometres from x= 0 km to x=16 km, they are numbered from 1 to 5. The remote sensor is situated at x=20 km.

Measures are corrupted with a zero mean gaussian noise with variance matrix σ^2 I with σ = 4E-3. A random choice of the active source is done at each experiment. Both methods are compared on N=50 runs. The first one is used with $R_u = \sigma_u^2$ I σ_u =5E-3 while the other is implemented with σ_{ur} =5E-3 and a ratio $\sigma_{ur}^2/\sigma_{ud}^2 = 5$.

The range parameter of input indexes is chosen equal to δ = 3. The criteria is evaluated for each source and sample time. The restoration of the input is performed iteratively according to Trussel scheme [Tru79]. The localisation quality is carried out using a rate of right decision. This rate (figure2) turns out to be very good for the second method (100%) while the other gives very poor results(24%). In fact, the method 1 remains stuck to the choice of the nearer source. Moreover, histograms of sample times are presented for the second method showing a great similitude with the expected ones. Finally, the second method seems very efficient in this context.

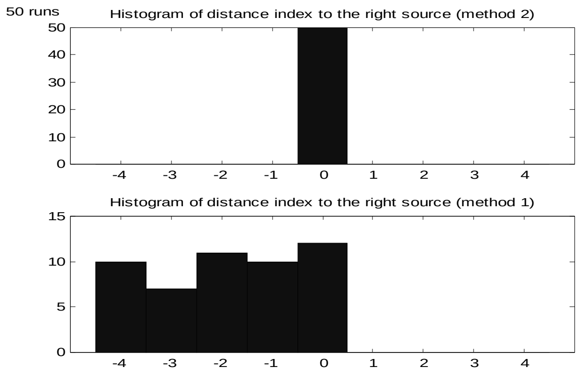

Figure 2 : Comparison of the localisation quality

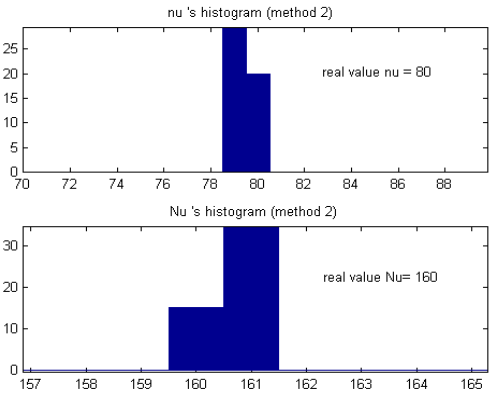

Figure 3 : Histogram of the time bounds

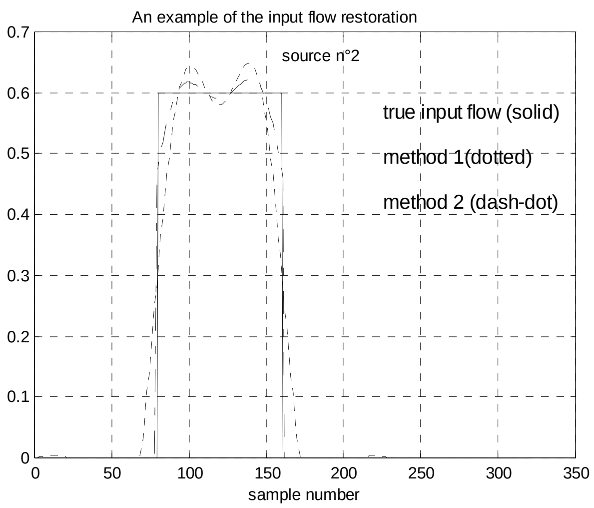

Figure 4 : An example of the input restoration (method 1 and method 2).

6. CONCLUSION

In this paper, the problem of joint localisation and estimation of the source flow is addressed. It is assumed that only one source is active during the pollution event. Two methods are investigated : the first one uses no assumption about the input source while the other relies on a specific input model described by two time bounds (a birth time and a death time). A MAP criteria enables to carry out a decision in each case and to recover in the same time the input flow. The main interest for the second method lies in a term of the criteria which depends on the size of the reduced input vector.

Practically, the first method is shown to often decide for a nearer source while the other seems to be very efficient in this context. This encourages to take advantage of *a priori* information about input sources.

REFERENCES:

Basseville M. (1988). Detecting Changes in Signals and Systems. A survey, *Automatica*, **vol 24. n°3**, pp 309-326.

Blanc-Feraud L., (2000). Sur quelques problèmes inverses en traitement d’images. Habilitation à diriger des Recherches. July 7 2000.

Commenges D (1984). The Deconvolution Problem : Fast Algorithms Including the Preconditioned Conjugate Gradient to Compute a MAP Estimator. *IEEE Trans. A.C.*, **Vol A-C 29, n° 3**, pp 229-243.

Delmaire G (2000), M. Benjelloun. Estimation of pollutant source specifications using a conditional deconvolution method, IFAC Sysid 2000, June 21-23 2000, Santa Barbara, USA.

Delmaire G. (2000), M. Benjelloun. Failure isolation and estimation using a conditional deconvolution approach for pollution monitoring, IFAC Safeprocess 2000, June 14-16 2000, Budapest, Hongrie.

Delmaire G (2002)., A joint estimation decision algorithm for pollutant source characterisation, just proposed at ISIT2002, June 30 –July 5 2002, Lausanne, Switzerland.

Demoment G. (1989), Image Reconstruction : Overview of Common Estimation Structures and Problems, *IEEE Trans. On Acoustics Speech and Signal Processing*, **Vol. 37, n° 12,**.pp 2024-2036.

Djafari A.M. (1997), Une Méthode Bayesienne pour la Localisation et la Segmentation de Sources de formes connues, *in Proc. GRETSI*, Grenoble, France, pp 136-138.

Marcos S. (1999), Méthodes Haute Résolution : Traitement D'antennes et Analyse Spectrale, HERMES.

Serrano S. E.(1997) The Hydrology of river pollution. In Hydrology for engineers Geologists and Environmental professionals. Hydroscience Inc,Lexington [ED], pp 341-368.

Trussel H.J. and B.R. Hunt (1979). Improved Methods of Maximum A Posteriori Restoration. *IEEE Trans. On Computers*, **Vol C-27**, n° 1, pp 57-62.