



**HAL**  
open science

## Identifiability and identification of a pollution source in a river by using a semi-discretized model

Nathalie Verdière, Ghislaine Joly-Blanchard, Lilianne Denis-Vidal

► **To cite this version:**

Nathalie Verdière, Ghislaine Joly-Blanchard, Lilianne Denis-Vidal. Identifiability and identification of a pollution source in a river by using a semi-discretized model. *Applied Mathematics and Computation*, 2013, 221, pp.1-9. 10.1016/j.amc.2013.06.022 . hal-01993267

**HAL Id: hal-01993267**

**<https://hal.science/hal-01993267>**

Submitted on 4 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Identifiability and identification of a pollution source in a river by using a semi-discretized model

Nathalie Verdière<sup>1</sup>, Ghislaine Joly-Blanchard<sup>1</sup>, Lilianne Denis-Vidal<sup>1</sup>

<sup>a</sup>University of Le Havre, BP 1123, 76063 Le Havre Cedex

<sup>b</sup>University of Technology of Compiègne, BP 20 529, 60205 Compiègne, France

<sup>c</sup>University of Sciences and Technologies of Lille, 59650 Villeneuve d'Ascq. France.

---

## Abstract

This paper is devoted to the identification of a pollution source in a river. A simple mathematical model of such a problem is given by a one-dimensional linear advection-dispersion-reaction equation with a right hand side spatially supported in a point (the source) and a time varying intensity, both unknown. There exist some identifiability results about this distributed system. But the numerical estimation of the unknown quantities require the introduction of an approximated model, whose identifiability properties are not analyzed usually. This paper has a double purpose: - to do the identifiability analysis of the differential system considered for estimating the parameters, - to propose a new numerical global search of these parameters, based on the previous analysis. Another consequence of this approach is to give the unknown pollution intensity directly as the solution of a differential equation. Lastly, the numerical algorithm is described in detail, completed with some applications.

*Keywords:* Semi-discretized model, Identifiability, Parameter Identification, Application to Pollution.

---

## 1. Introduction

The quality of the water is of a crucial interest in our society. It can be estimated by measuring, for example, the quantity of organic matters contained in the water. This paper is concerned with the determination of the location and the intensity of a pollution source from the measurements of a pollutant concentration linked to organic matters. The model, given here, corresponds to a river portion of length  $L$  surrounded by factories or other possible pollution sources. A first simple model is given by a linear advection-dispersion-reaction equation (1) (see [O] for details) where  $u$ ,  $D$ ,  $V$ , and  $R$  denote the substrate concentration, the dispersion coefficient, the transport velocity and the reaction coefficient respectively. The right hand side member depends on two unknown parameters:  $\lambda \in \mathcal{L}$  (the set of functions  $\mathcal{L}$  will be specified in the remainder), the flow rate of the pollutant, and  $a \in ]0, L[$ , the location of the source. The

function  $g$  is assumed to be smooth, at least in  $C^2(0, L)$ , and  $\delta(x - a)$  is the Dirac mass at the point  $a$ .

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) - D \frac{\partial^2 u}{\partial x^2}(x, t) + V \frac{\partial u}{\partial x}(x, t) + Ru(x, t) \\ \quad = \lambda(t)\delta(x - a), & (x, t) \in ]0, L[ \times ]0, T[ \\ u(x, 0) = g(x), & x \in ]0, L[, \\ u(0, t) = 0, & t \in ]0, T[, \quad \frac{\partial u}{\partial x}(L, t) = 0, & t \in ]0, T[, \end{cases} \quad (1)$$

The pollution is assumed to be known at the initial time and the boundary conditions translate the unidirectional nature of the transport. Indeed there is no significant transport upstream, therefore the concentration is assumed to be zero at some point situated upstream ( $x = 0$ ). On the other hand the downstream point  $x = L$  is far enough from the source so that a zero gradient can be considered.

It is well-known that the problem (1) has a unique solution that is smooth enough in a neighborhood of  $L$  since  $a \in ]0, L[$ . So, it makes sense to define  $u(L, t)$  as the observation function  $y(t) = u(L, t)$ ,  $t \in (0, T)$ . In order to test a new approach, a first study was done when  $\lambda$  is constant ([V]). It was proved that, when the initial conditions are considered known, the observation function  $y$  is sufficient to obtain the identifiability of the pollution source in the model (1) whereas when they are unknown, a second point of observation has to be added. In order to estimate the parameters  $\lambda$  and  $a$ , an approximated system was used and its global identifiability was proved by using an elimination approach. From the identifiability analysis, a new numerical procedure of parameter estimation was elaborated and tested.

It is the same idea which is developed in this paper but the variation in time of the pollution intensity increases the complexity of the solution. For the model (1), results of identifiability, obtained by using a decomposition on a basis of eigenfunctions, and parameter identification were given in [E] with two points of observation, one upstream, the other downstream from the source provided the flow rate vanishes before reaching the final control time  $T$  i.e, there exists  $T^* \in (0, T)$  such that  $\lambda(t) = 0$  for  $t \geq T^*$ . The identification method gives the source position explicitly provided the knowledge of  $u(T^*)$  which is done in using least-squares and determine the intensity function by solving a deconvolution problem.

In this paper, an approximate system is obtained from an approximation of the Dirac mass by a smooth function and a semi-discretization in space. It consists of a system of differential equations, the size of which corresponds to the number of discretization points of the interval  $[0, L]$ . First, the paper is focused on the proof of the identifiability of the so-obtained model, based on an elimination approach. The initial conditions are integrated in this study and the work is positioned in the case of an existing tolerated pollution. Then it is shown how the previous identifiability analysis leads to a numerical global search of the localization of the pollution source and the pollution intensity. Indeed the unknown function  $\lambda(t)$  is computed from the solution of a differential

equation directly, without truncating its decomposition on a basis function as it is done usually. This method is close to the one proposed by [L] or [O].

The paper is presented as follows: Section 2 introduces the approximation differential system, Section 3 gives identifiability results, section 4 explains in detail the numerical algorithm and presents some numerical results.

## 2. The semi-discretized model

The model used for the parameter numerical estimation is introduced. Firstly, since regular functions are needed to use differential algebra, the Dirac mass  $\delta(x - a)$  is approximated by a Gaussian function:

$$w(x) = \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{(x-a)^2}{\sigma^2}}, \quad (2)$$

where the coefficient  $\sigma$  will be chosen so that the error between the solution of (1) and the solution of the continuous model with  $\delta(x - a)$  replaced by  $w(x)$  be smaller than the error of the semi-discretized scheme. Indeed, if  $\check{u}$  is the solution of (1) with the Dirac mass substituted by (2), the following estimation can be easily shown:

$\forall \epsilon > 0, \exists \sigma^* > 0$ , such as if  $0 < \sigma < \sigma^*$  :

$$\|u(t) - \check{u}(t)\|_2 \leq \frac{\lambda_{max}\epsilon}{\kappa} \sqrt{1 - e^{-\kappa T}} e^{-\nu L},$$

with  $\nu = -\frac{V}{2D}$ ,  $\kappa$  an ellipticity constant and  $\lambda_{max}$  the maximal value of the flow rate on  $(0, T)$ .

Then the system (1) is discretized in space by a centered difference scheme. If  $(N + 1)$  discretization points  $(x_i)_{0 \leq i \leq N}$  in space are considered, it leads to:

$$\begin{cases} \dot{u}(x_1, t) - \beta u(x_1, t) - \gamma u(x_2, t) \\ \quad = \lambda(t)w(x_1) + h\epsilon_1(t), \\ \dot{u}(x_i, t) - \alpha u(x_{i-1}, t) - \beta u(x_i, t) - \gamma u(x_{i+1}, t) \\ \quad = \lambda(t)w(x_i) + h\epsilon_i(t), i = 2, \dots, N - 1, \\ \dot{u}(x_N, t) - (\alpha + \gamma) u(x_{N-1}, t) - \beta u(x_N, t) \\ \quad = \lambda(t)w(x_N) + h\epsilon_N(t). \end{cases} \quad (3)$$

where  $h = \frac{L}{N}$ ,  $x_i = ih$  for  $i = 0, \dots, N$ ,

$$\alpha = \frac{V}{h} + \frac{D}{h^2}, \quad \beta = -\left(\frac{V}{h} + 2\frac{D}{h^2} + R\right), \quad \gamma = \frac{D}{h^2},$$

and

$$\epsilon_i(t) = -\frac{1}{2}V \frac{\partial^2 u}{\partial x^2}(x_i, t) + \frac{h}{6} \left( \varphi_{1,i} D \frac{\partial^4 u}{\partial x^4}(x_i + \varphi_{2,i} h, t) - V \frac{\partial^3 u}{\partial x^3}(x_i + \varphi_{1,i} h, t) \right) \text{ with } |\varphi_{k,i}| < 1, k = 1, 2.$$

Consequently, the flow rate  $\lambda$  is assumed to be  $C^{N-1}([0, T])$ , whatever is the value of  $N$ . The value  $w(x_i)$ , defined by (2), is explicitated as:

$$w(x_i) = \frac{1}{\sigma\sqrt{\pi}} e^{-\frac{a^2}{\sigma^2}} e^{-\frac{(ih)^2}{\sigma^2}} \left( e^{h\frac{2a}{\sigma^2}} \right)^i.$$

Let us introduce  $\ell(t) = \frac{\lambda(t)}{\sigma\sqrt{\pi}}e^{-\frac{a^2}{\sigma^2}}$ ,  $k_i = e^{-\frac{(ih)^2}{\sigma^2}}$ ,  $Q = e^{h\frac{2a}{\sigma^2}}$ ,  $g_h = (g(x_i))_{i=1,\dots,N}^T$ ,  $u_h(t) = (u(x_i, t))_{i=1,\dots,N}^T$ ,  $\epsilon_h(t) = (\epsilon_i(t))_{i=1,\dots,N}^T$  ( $^T$  designs the transpose vector) and

$$A_h = \begin{pmatrix} \beta & \gamma & 0 & \dots & 0 \\ \alpha & \beta & \gamma & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \alpha & \beta & \gamma \\ 0 & \dots & 0 & \alpha + \gamma & \beta \end{pmatrix}, b_h(t) = \begin{pmatrix} k_1\ell(t)Q \\ k_2\ell(t)Q^2 \\ \vdots \\ k_N\ell(t)Q^N \end{pmatrix}.$$

Thus, the system (3) can be rewritten

$$u_h'(t) = A_h u_h(t) + b_h(t) + h\epsilon_h(t).$$

Neglecting the derivatives of order greater than two and considering the initial conditions, one gets:

$$v_h'(t) = A_h v_h(t) + b_h(t), v_h(0) = g_h, \quad (4)$$

where  $v_h$  is an approximated solution of  $u_h$ . Indeed, since  $A_h = (a_{i,j})$  is a real tridiagonal matrix whose coefficients verify  $a_{k,k+1} \times a_{k+1,k} > 0$  for  $k = 1, \dots, N$ , it is similar to a Hermitian matrix and its eigenvalues are real. Besides, according to the Gerschgorin theorem, these eigenvalues are negative. Hence, according to the expressions of  $u_h$  and  $v_h$  and the Gerschgorin theorem, one gets:

$$\exists C > 0, \|u_h(t) - v_h(t)\|_\infty \leq \frac{hCK_h}{R}$$

where  $K_h$  is the condition number of  $A_h$ . In the simulations,  $N$  has been chosen so that  $K_h$  is the smallest possible, that is  $N + 1 = 150$ .

### 3. Identifiability of the semi-discretized model

The numerical estimation of the parameters is based on (4), consequently its identifiability analysis has to be performed. First results about the identifiability concept can be found in [W]. In the 90's, [D], [F], and [O] proposed a new approach of identifiability based on differential algebra, and which does not require the existence of a control. [F] have introduced the concept of characteristic presentation (or decomposition), leading to an algorithm of differential elimination called "Rosenfeld-Groebner algorithm" and implemented in the package Diffalg in Maple ([F]). This identifiability approach ignores the initial conditions of the system. But the initial conditions can play a crucial role ([L], [S]) and a software based on "Rosenfeld-Groebner algorithm" and which takes them into account has been written ([L], [L]).

A very important repercussion of the identifiability analysis is the elaboration of differential polynomials linking outputs, parameters and inputs if any, in the most cases allowing the obtention of parameter estimates without any a priori knowledge. In the context of the considered semi-discretized differential model, the unknown parameters are not only constants but also functions and there is no input. Then, the classical approaches based on differential algebra for

identifiability analysis ([L], [S]) cannot be used. Let us begin by recalling some basic definitions and giving some useful notations.

### 3.1. Basic concepts of differential algebra

In the following, a differential polynomial  $p(x)$  is a polynomial in some variables, functions of time,  $x_1, \dots, x_n, y_1, \dots, y_m$  and a finite number of its derivatives with coefficients in the field  $\mathbb{Q}$  or a field extension of  $\mathbb{Q}$ . The differential polynomials generate a ring in the indeterminates  $x_1, \dots, x_n, y_1, \dots, y_m$  with their derivatives up to any order. It is called a *differential ring*  $K\{x_1, \dots, x_n, y_1, \dots, y_m\}$ .

A *differential ideal* of a differential ring  $\mathcal{R}$  is an ideal of  $\mathcal{R}$  stable under derivations. The differential ideal  $I$  generated by  $p_1, \dots, p_s$  some elements of  $\mathcal{R}$  is denoted  $I = [p_1, \dots, p_s]$ . The *radical differential ideal*  $J = \{p_1, \dots, p_s\}$  generated by the  $p_i, i = 1, \dots, s$  is the set of all element of  $\mathcal{R}$  a power of which belongs to  $I = [p_1, \dots, p_s]$ . Considering a set  $S$  of differential polynomials, a differential polynomial  $p$  vanishes on all the zeros of  $s$  if and only if  $p$  belongs to the radical differential ideal generated by  $S$  ([F], [F]). Any radical differential ideal can be decomposed into an intersection of some differential ideals (said *characterizable*) and the obtained representation is called a *characteristic decomposition*. These ideals are defined by sets of differential polynomials, differential *characteristic sets* which are kinds of canonical forms ([R], [S]).

The notion of *ranking* is fundamental in elimination methods. It is a well-ordering over the indeterminates and their derivatives. An elimination ranking denoted " $\prec$ " has been used, which is such that  $u_i^{(k)} \prec u_i^{(k+p)}$  and  $u_i^{(k)} \prec u_j^{(l)} \implies u_i^{(k+p)} \prec u_j^{(l+m)}$ , for any indeterminates  $u_i$  and  $u_j$  and arbitrary integers  $k, l, m$  and  $p$ . On the other hand, an orderly ranking has been also used, which satisfies, with the same notations ( $p > 0$ )  $u_i^{(k)} \prec u_j^{(k+p)}$ . The notation used in Diffalg to set an orderly ranking is  $[u_1, \dots, u_n]$  (leftmost elements are greater than rightmost ones and the derivatives are ordered by an orderly ranking). The highest ranking variable or derivative of a variable in a differential polynomial is called the *leader* of the polynomial.

### 3.2. Identifiability approach by elimination

The system (4) can be rewritten as a differential polynomial system that is completed with the observation  $y = v_N$  and with  $\dot{Q} = 0$  and  $\dot{k}_i = 0$  ( $i = 1, \dots, N$ ) since  $Q$  and  $k_i$  are constant. The resulting system can be described by the following differential polynomials ( $Sd_N$ ):

$$\begin{cases} \dot{v}_1 = (\beta v_1 + \gamma v_2 + k_1 \ell Q), \\ \dot{v}_i = (\alpha v_{i-1} + \beta v_i + \gamma v_{i+1} + k_i \ell Q^i), i = 2, \dots, N-1 \\ \dot{v}_N = ((\alpha + \gamma)v_{N-1} + \beta v_N + k_N \ell Q^N), \\ y = v_N, \dot{Q} = 0, \\ \dot{k}_i = 0, k_i \neq 0 \text{ for } i = 1, \dots, N. \end{cases}$$

It is assumed that the constants  $\alpha, \beta, \gamma$  are not solutions of algebraic equations. Then, the differential ideal generated by the equations of ( $Sd_N$ ) can be considered in the differential ring  $K\{v_1, \dots, v_N, y, \ell, Q, k_1, \dots, k_N\}$  generated

by the field of constants  $K = \mathbb{Q}(\alpha, \beta, \gamma)$ , the states  $(v_i)_{1 \leq i \leq N}$ , the output  $y$ , the unknown function  $\ell$ , the unknown parameter  $Q$  and the constants  $k_i$  for  $i = 1, \dots, N$ . Thus, this section is devoted to the identifiability of the function  $\ell$  and the parameter  $Q$ . In the following definition,  $\mathcal{U} = \mathcal{L} \times ]0, T[$  represents the admissible set of the parameters.

**Definition 3.1.** *The model  $(Sd_N)$  is globally identifiable at  $(\ell, Q) \in \mathcal{U}$  if there exists a finite time  $t_1 > 0$  such that, for all  $t \in [0, t_1]$ , if  $y[(\tilde{\ell}(t), \tilde{Q}); t] = y[(\ell(t), Q); t]$  with  $(\tilde{\ell}, \tilde{Q}) \in \mathcal{U}$  and  $\ell, \tilde{\ell}$  are not equal to zero then  $(\tilde{\ell}, \tilde{Q}) = (\ell, Q)$ .*

The main result is given in the theorem 3.1 whose proof is based on an induction argument ([V]).

**Theorem 3.1.** *The radical of the ideal generated by  $(Sd_N)$  endowed with the ranking*

$$[Q, \ell, y, k_1, \dots, k_N], \prec [v_1, \dots, v_N]$$

*admits a characteristic decomposition  $\mathcal{C}$  which contains the polynomials  $\dot{Q}, \dot{k}_1, \dots, \dot{k}_N, P(y, \ell, Q, k_1, \dots, k_N), R_i(y, \ell, Q, v_1, k_1, \dots, k_N)$  for  $i = 1, \dots, N$  and the polynomial  $P$  is equal to*

$$f_N(y) + \sum_{i=0}^{N-1} c_{N,i}(Q)\ell^{(i)}.$$

*The leader of the polynomial  $P$  is  $y$ , and the leader of the polynomial  $R_j$  is  $v_j$  for  $j = 1, \dots, N$ .*

*The function  $f_N(y)$  is a linear function depending on  $y$  and its derivatives and whose the highest derivative with respect to  $y$  is  $y^{N-1}$ . Besides, the each term of the sequence  $(c_{N,i})_{i=1, \dots, N}$  is a constant depending on  $Q$  and the parameters of the initial system  $\alpha, \beta, \gamma$ .*

**Remark 3.1.** *In this context, the systems  $(Sd_N)$  and  $(\mathcal{C})$  obtained after canceling all the polynomials of  $\mathcal{C}$  are equivalent systems (by using the radical of the ideal generated by  $(\mathcal{C})$  endowed with the ranking  $[Q, \ell, y, v_1, \dots, v_N, k_1, \dots, k_N]$ ).*

**Remark 3.2.** *Up to now, initial conditions have not yet been considered in the semi-discretized system, they are now integrated in the following identifiability study. In system (1), the substrate concentration at the initial time is assumed to be known on  $(0, T)$ . For the semi-discretized system, it is supposed to be known at  $x_{N-2}, x_{N-1}$  and  $x_N$ . Moreover, it is supposed not equal to zero at the initial time which corresponds to the case of a residual, tolerated and well-known pollution. Finally, the speed of the substrate concentration, at the initial time, will be assumed not to vary near the observation point, that is  $\dot{v}_{N-1}(0)$  will be supposed equal to  $\dot{v}_N(0)$  and consequently it is assumed to be known. Indeed, the numerical estimation of parameters has been done with this assumption.*

**Theorem 3.2.** *Suppose that  $v_{N-2}(0), v_{N-1}(0), v_N(0)$  and  $\dot{v}_{N-1}(0)$  are known. Suppose too that  $\ell$  is not equal to zero on  $[0, t_1]$ ,  $t_1 > 0$ . Then, the model  $(Sd_N)$  is globally identifiable at  $(\ell, Q) \in \mathcal{U}$ .*

*Proof* - From the theorem 3.1, one gets:

$$f_N(y) + \sum_{i=0}^{N-1} c_{N,i}(Q)\ell^{(i)} = 0. \quad (5)$$

This equation is a linear differential equation of order  $N - 1$  ( $c_{N,N-1} \neq 0$ ). In order to show the model identifiability, it would be enough to show the identifiability of the parameter  $Q$  which implies the identifiability of the parameters  $c_{N,i}$ . Then, the identifiability of  $\ell$  will be deduced from the resolution of the differential equation (5).

- *Identifiability of  $Q$  and  $\ell(0)$  by using  $(Sd_N)$*

From the two last equations of  $(Sd_N)$  taken at 0, an explicit expression of  $k_N\ell(0)Q^N$  and  $k_{N-1}\ell(0)Q^{N-1}$  depending on the observation and  $v_{N-2}(0)$ ,  $v_{N-1}(0)$  and  $v_N(0)$  is obtained so  $Q$  and  $\ell(0)$  too. Thus, the latter are identifiable. Therefore, since the coefficients  $(c_{N,i})_{i=1,\dots,N}$  are function of  $Q$ , they are identifiable.

- *Identifiability of  $\ell$  by using  $(\mathcal{C})$*

Now, it remains to establish that the initial conditions of (5) can be uniquely determined from the observation. The following property is proved, for  $i = 2, \dots, N$ , by using polynomials  $R_1, \dots, R_N$  and an induction argument: " $\ell^{(i-1)}(0)$  and  $(v_{N-j-1}^{(i-j-1)}(0))_{j=-1,0,\dots,i-2}$  are uniquely determined from  $y$  and  $(v_i(0))_{i=1,\dots,N}$ ."

The proof is based on an elimination method and gives the way to obtain an explicit expression. It can be obtained by Maple (section 4.1).

**Remark 3.3.** Since  $Q = e^{\frac{2ah}{\sigma^2}}$ , the unknown parameter  $a$  is identifiable and is given by:

$$a = \frac{N\sigma^2}{2L} \log \left[ \frac{k_{N-1}}{k_N} \frac{\dot{y}(0) - (\alpha + \beta)v_{N-1}(0) - \beta y(0)}{\dot{v}_{N-1}(0) - (\alpha v_{N-2}(0) + \beta v_{N-1}(0) + \gamma y(0))} \right]. \quad (6)$$

#### 4. Numerical parameter identification

For estimating the numerical values of the unknown parameters  $(\lambda, a)$ , the measurement is done at discrete times  $(t_i)_{1 \leq i \leq M}$ , which induces a numerical evaluation of observation derivatives involved in the parameter-output polynomial, and, consequently, a numerical error. The choice of the derivatives evaluation will be also crucial. The method proposed by [F] will be used in the numerical applications. This method does not necessitate any knowledge *a priori* of statistical properties of the signals. Furthermore, the estimators lay on explicit formulas. The previous identifiability's study will give a numerical procedure for getting explicit formulas of the unknown parameters which allow a first approximation of them, without *a priori* knowledge, as it is explained below.



#### 4.1. Presentation of the method

- *Expression of  $\ell$  and its derivatives in zero*

The estimation of  $a$  comes from (6). From (5), one can deduced that he flow rate is the solution of a differential system of first order  $\dot{L}(t) = AL(t) + F(t)$  and where  $A$ , a compagnon matrix and  $F$  are easily found. In order to determinate the initial conditions of this system, that is  $L(0)$ , the package Diffalg is used. For example, with 11 points, the ranking  $[v_1, \dots, v_{10}, k_1, \dots, k_{10}, y, q, \ell]$  is applied to the equations  $(Sd_N)$ . It leads to polynomials which contain  $\ell$  and its derivatives of order  $n$  ( $n = 1, \dots, 9$ ),  $y$  and its derivatives and  $v_1, \dots, v_{10}$ . For example, it leads to the following expression of  $\dot{\ell}$ :

$$\begin{aligned} \dot{\ell} = & (\ddot{v}_{10} - (\alpha + \gamma)\alpha v_8 - (\alpha + \gamma)\beta v_9 \\ & - (\alpha + \gamma)\gamma v_{10} - (\alpha + \gamma)k_9 Q^9 \ell - \beta \dot{v}_{10}) / k_{10} / Q^{10}, \end{aligned}$$

hence  $\ell'(0)$ . The other derivatives are deduced by the same way.

The matrix  $A$  is diagonalizable since it can be proved that if the characteristic polynomial of  $A$  has a multiple root then  $\alpha$ ,  $\beta$  and  $\gamma$  are solutions of an algebraic equation which has been excluded. Therefore, the linear differential system is solved by a direct solution using the diagonalization of the matrix  $A$ , hence  $\ell$  and consequently  $\lambda$  evaluated by:  $\lambda(t) = \ell(t)e^{\frac{a^2}{\sigma^2}t} \sigma \sqrt{\pi}$ .

- *Evaluation of the observation derivatives*

Some observation derivatives of important order have to be estimated. To do that, the method proposed by [F] is used in the numerical applications. First, the derivatives in 0 will be given. Let  $x$  a real function, analytic around 0. The Taylor expansion of  $x(t)$  is  $x(t) = \sum_{n \geq 0} x^{(n)}(0) \frac{t^n}{n!}$  and it can be approximated on an interval  $(0, \epsilon)$  by  $x_p(t) = \sum_{n=0}^p x^{(n)}(0) \frac{t^n}{n!}$  ( $x_p^{(n)}(0) = x^{(n)}(0)$ ) which verifies  $\frac{d^{p+1}}{dt^{p+1}} x_p = 0$ . In the operational domain, the previous equality gives:

$$s^{p+1} x_p - s^p x_p(0) - s^{p-1} \dot{x}_p(0) \dots - x_p^{(p)}(0) = 0.$$

The derivatives in zero,  $x_p^{(i)}(0) = \frac{d^i}{dt^i} x_p(t) |_{t=0}$  are also obtained from the solution of the system of linear equations ( $m = 0, \dots, p$ ,  $\nu \geq p + 1$ ):

$$\begin{aligned} & s^{-\nu} \frac{d^m}{ds^m} \left\{ x_p^{(p)}(0) + x_p^{(p-1)}(0)s + \dots + x_p(0)s^p \right\} \\ & = s^{-\nu} \frac{d^m}{ds^m} \left\{ s^{p+1} x_p \right\}. \end{aligned}$$

The formula in the time domain are deduced by the well-known rules of the operational calculus and by replacing  $x_p$  by  $x$ .

#### 4.2. Numerical results

For checking the validity of our approach, an accidental pollution has been simulated in a river during 4 hours with:

$L = 1000m$ ,  $V = 0.66m/s$ ,  $R = 10^{-5}/s$ ,  $D = 5m^2/s$ .  $\lambda$  is built from the function

$$\lambda(t) = \sum_{i=1}^3 3\alpha_i e^{-\beta_i(t-\tau_i)^2},$$

with  $\alpha_1 = 1.2$ ,  $\alpha_2 = 0.4$ ,  $\alpha_3 = 0.6$ ,  $\beta_1 = 10^{-6}s^{-2}$ ,  $\beta_2 = 5.10^{-6}s^{-2}$ ,  $\beta_3 = 10^{-6}s^{-2}$ ,  $\tau_1 = 4500s$ ,  $\tau_2 = 6500s$ ,  $\tau_3 = 9000s$ .

The observation corresponds to  $y(t) = u(L, t)$ . The true signal  $\bar{y}$  was obtained by solving (1) with the classical finite element scheme presented at section 2 with 150 discretization points, the pollution location not being a mesh point. Therefore, the Dirac mass is approximated by putting the whole mass in this point. The measured observation  $y$  was simulated from the true signal  $\bar{y}$  and is supposed to follow a random law with  $\bar{y}$  mean and  $(s\bar{y})^2$  variance.

The coefficient  $s$  is computed so that the relative error has a maximum value of 0.05 with an error probability less than 0.003. Thus, if  $y = \bar{y} + s\bar{y}\eta_0$  with  $\eta_0$  following the reduced normal law,  $s$  is chosen so that  $P\left(\left|\frac{y-\bar{y}}{\bar{y}}\right| \leq 0.05\right) = 0.997$ . According to the table of the reduced normal law,  $s \approx 0.017$ .

*Algorithm* - The numerical estimation of the unknown parameters is based on the semi-discretized scheme (4) whose equations are dimensionless. The factories (the pollution sources) are assumed to be located between  $400m$  and  $900m$ . The differential polynomial  $P$  has been obtained with  $N + 1 = 11$  discretization points, which is a very small number for such a distributed system but which allows the symbolic computation of (5) with maple. Besides, let us recall that the aim of this work is to make a first evaluation of the parameters.

The derivatives of the observation at the initial time have been computed by applying the method introduced in the previous section, with  $p = 1$  and  $\nu = 2$  for the first derivative and  $p = 9$  and  $\nu = 12$  for higher derivatives. The numerical parameter identification have been performed on the time interval  $[2000, 14400]$ , i.e. the initial time corresponds to 2000. For this initial time, there exists a residual pollution of order  $10^{-3}$ . Before giving numerical results, notice that the expressions giving  $a$  and  $\lambda(t)$  depend on the parameter  $\sigma$  which intervenes in the approximation of the Dirac mass. In theory, it has to be taken very small, but it is not the case owing to the numerical errors introduced by estimating the derivatives for example. The numerical procedure will consist in varying  $\sigma$  in the interval  $[0.01; 0.5]$  and for each given value  $\tilde{\sigma}$  of  $\sigma$  in this interval, the identified parameters  $(a_{\tilde{\sigma}}, \lambda_{\tilde{\sigma}})$  are estimated according to the formulas (6) and  $\lambda(t) = \ell(t)e^{\frac{a}{\tilde{\sigma}^2}}\tilde{\sigma}\sqrt{\pi}$  where  $\ell(t)$  is obtained in solving the differential system  $\dot{L}(t) = AL(t) + F(t)$  defined at section 4.1. Afterwards, the corresponding  $y_{\tilde{\sigma}}$  is obtained by solving (4) and the following relative error is computed:  $e_{\tilde{\sigma}} = 0.5 \frac{\text{norm}(y - y_{\tilde{\sigma}})}{\text{norm}(y)}$ . Consequently, the value of  $\sigma$  is chosen such that it gives the smallest relative error and an estimate of the parameters corresponds to this choice.

The results obtained by the algorithm are summed up at Figures 1, 2, 3 and 4. The results are quite satisfactory, due to the search of the parameter  $\sigma$  in the approximated system. However, the flow rate found in the first case is not as good as in the others. Indeed, since  $a$  is located far from the observation point, a degradation of the substrate, which is traduced by the coefficient  $R$ ,

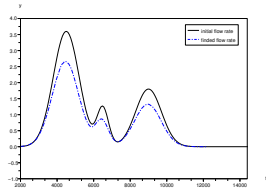


Figure 1: For  $a = 442m$ ,  $\tilde{a} = 441.6m$ ,  $e_\sigma = 0.25$ .

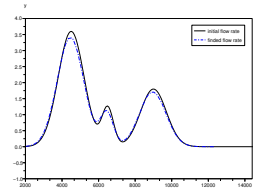


Figure 2: For  $a = 547m$ ,  $\tilde{a} = 546.5m$ ,  $e_\sigma = 0.04$ .

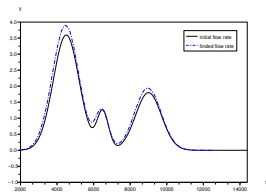


Figure 3: For  $a = 621m$ ,  $\tilde{a} = 623.6m$ ,  $e_\sigma = 0.099$ .

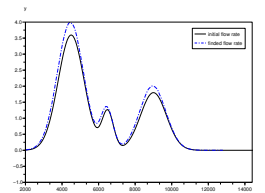


Figure 4: For  $a = 853m$ ,  $\tilde{a} = 849.2m$ ,  $e_\sigma = 0.120$ .

has happened.

Finally, the sensitivity of the identified parameters with respect to the introduction of noises on the measures  $y(t)$  has been studied in describing their relative errors for different noise intensities in the case of  $a = 621$  (figure 5).

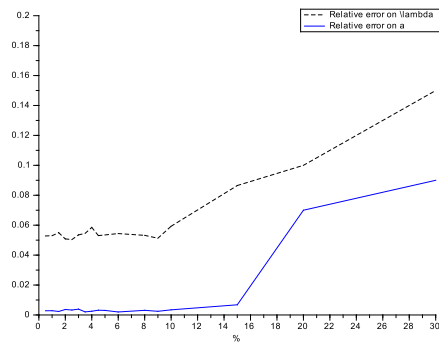


Figure 5: Relative errors on  $a$  and  $\lambda$  for different noise intensities when  $a = 621$ .

## 5. Conclusion

In this paper an original method for estimating time-varying parameters in a distributed system without using a basis of special functions has been tested. It consists in elaborating an approximated system whose identifiability has been proved by an elimination procedure whatever the mesh refinement. The identifiability analysis leads to the development of numerical algorithms of global optimization. Of course, since it is an elimination approach, some derivatives of important order appeared. However, the method developed by M. Fliess and H.Sira-Ramirez proposed an effective way to estimate them without any knowledge *a priori* of statistical properties of the signals. Finally, some numerical simulations are confirming the interest of the proposed approach.

- [F] F. Boulier, D. Lazard, F. Ollivier, M. Petitot (1995), "Representation for the radical of finitely generated differential ideal", *Proc. ISSAC'95*, pp 158-166, Montréal, Canada
- [F] F. Boulier and F. and Lemaire(2000), "Computing canonical representatives of regular differential ideals", *Proc. ISSAC 2000*, Scotland.
- [L] L. Denis-Vidal, G. Joly-Blanchard, C. Noiret(2001a), "Some effective approaches to check the identifiability of uncontrolled nonlinear systems", *Math. Comp. in Simulation* 57, pp 35-44.
- [L] L. Denis-Vidal, G. Joly-Blanchard, C. Noiret, M. Petitot(2001b), "An algorithm to test identifiability of non-linear systems", *Proc. IFAC NOLCOS*, CD-289, Russia .
- [D] S. Diop, M. Fliess(1991), "Nonlinear observability, identifiability, and persistent trajectories", *Proc. 30th CDC*, Brighton, pp 714-719.
- [E] A. El Badia, T. Hua Duong, A Hamdi(2005), "Identification of a point source in a linear advection -dispersion-reaction equation: application to a pollution source problem", *Inverse Problems* 21, pp 1-17.
- [F] M. Fliess, S. T. Glad(1993), "An algebraic approach to linear and nonlinear control", *Essays on control: perspectives in the theory and its application*, Cambridge, MA, Birkhäuser, 7, pp 223-267 .
- [F] Michel Fliess, Mamadou Mboup, Hugues Mounier, Hebertt Sira-Ramirez, "Questioning some paradigms of signal processing via concrete examples", *Proc. Conf. Diff. Alg. Meth Flatness Sign. Proc. Estim. Mexico*, Novembre 2003.
- [L] B. Laroche, Ph. Martin and P. Rouchon (2000), "Motion planing for the heat equation". *Int. Journal of Robust and Nonlinear Control*, vol 10, 629-643.
- [O] Okubo(1980), "Diffusion and Ecological Problems: Mathematical Models", *Springer Verlag*.

- [O] F. Ollivier(1997), "Identifiabilité des systèmes", *Technical Report*, 97-04, GAGE, Ecole polytechnique, juin 1997.
- [O] F. Ollivier and Sedoglavic(2001), A generalization of flatness to nonlinear systems of partial differential equations. Application to the command of a flexible rod., *Proc. of the 5th IFAC Symposium, "NCS"*, Russia 2001, vol. 1, Elsevier, pp. 196-200.
- [R] J.F. Ritt(1950), "Differential Algebra", *Providence:RI: American Mathematical Society*.
- [S] Saccomani, M.P., Audoly, S., D'Angiό. L.(2003) "Parameter identifiability of nonlinear systems: the role of initial conditions", *Automatica*, 39, pp 619–632.
- [V] N. Verdière, G. Joly-Blanchard, L. Denis-Vidal(2004), "Identifiability and identification of a pollution source: a semi-discretization method", *Proc. MTNS04*, CD-MP10, Louvain, Belgium.
- [V] N. Verdière(2005), "Identifiabilité de systèmes d'équations aux dérivées partielles semi-discrétisées et applications à l'identifiabilité paramétrique de modèles en pharmacocinétique et en pollution", PhD thesis, Université Technologique de Compiègne, 2005. <http://www.lmac.utc.fr/preprints>.
- [W] E. Walter(1982), "Identifiability of state space models", *Lecture Notes Biomath.* Vol. 46.