

B.-E. AINSEBA

J.-P. KERNEVEZ

R. LUCE

**Identification de paramètres dans les problèmes
non linéaires à données incomplètes**

M2AN - Modélisation mathématique et analyse numérique, tome
28, n° 3 (1994), p. 313-328

http://www.numdam.org/item?id=M2AN_1994__28_3_313_0

© AFCET, 1994, tous droits réservés.

L'accès aux archives de la revue « M2AN - Modélisation mathématique et analyse numérique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>



IDENTIFICATION DE PARAMÈTRES
DANS LES PROBLÈMES NON LINÉAIRES
À DONNÉES INCOMPLÈTES (*)

par B.-E. AINSEBA ⁽¹⁾, J.-P. KERNEVEZ ⁽²⁾ et R. LUCE ⁽³⁾

Communiqué par J.-L. LIONS

Résumé. — *Dans cet article, on s'intéresse à l'identification de certains paramètres intervenant dans le second membre d'équations différentielles ou aux dérivées partielles non linéaires. Les modèles étudiés sont incomplets dans le sens où l'on connaît mal les données initiales. Deux approches sont possibles, l'une classique est la méthode des moindres carrés où l'on cherche aussi bien les paramètres à identifier que les termes manquants, la seconde est la méthode des sentinelles (J.-L. Lions [6], [7]) plus adaptée à notre situation. Deux problèmes ont été examinés sur lesquels on a comparé les sentinelles et les moindres carrés.*

Abstract. — *The aim of this paper is to look at the identification of some parameters which appear in the expression of right-hand term of non linear Differential Equations and P.D.E.'s. The studied models are incompleted since the initial condition are not well known. Two approaches are possible. One is classical and uses the least square method, this approach needs to compute the parameters and the missing terms. The other is the Sentinels Method more suited to our situation. We compare the both methods on two different examples.*

1. CADRE GÉNÉRAL

$$\begin{cases} \dot{Y} + \mathcal{A}Y + F(Y) = \mathcal{B}Y + \mathcal{B}\lambda \\ Y(0) = Y^0 + \tau_0 \bar{Y}^0 \end{cases} \quad (1.1)$$

où Y est un vecteur de n composantes, \mathcal{A} est un opérateur différentiel ou aux dérivées partielles, $F(Y)$ est une fonction non linéaire de Y , \mathcal{B} est un

(*) Manuscrit reçu le 25 mai 1993.

⁽¹⁾ Université de Bordeaux 2, UFR, MI2S, BP 69, 33076 Bordeaux Cedex, France et U.R.A., CNRS 226 CeReMaB, Université de Bordeaux 1, France.

⁽²⁾ Université de Technologie de Compiègne, BP 649, 60200 Compiègne, France.

⁽³⁾ Laboratoire de Mathématiques Appliquées, Université de Pau et U.R.A., CNRS 1204, I.P.R.A., Av. de l'Université, 64000 Pau, France.

opérateur linéaire, V est un vecteur connu de m composantes et λ est une perturbation à déterminer. $\tau_0 \bar{Y}^0$ est une perturbation supposée petite de la condition initiale.

L'observation du système est $\mathcal{C}Y$ où \mathcal{C} est un opérateur linéaire. Le but est d'identifier le vecteur λ (ou certaines composantes de ce vecteur).

1.1. Méthode des sentinelles

Il s'agit de construire une fonctionnelle S sensible aux paramètres à identifier et insensible aux termes manquants :

$$S(\tau_0, \lambda) = \int_0^T w \mathcal{C}Y dt$$

où w est un vecteur à déterminer de telle façon que :

$$(P) \quad \left\{ \begin{array}{l} \frac{\partial S}{\partial \tau_0}(0, 0) = 0, \quad \frac{\partial S}{\partial \lambda_1}(0, 0) = \alpha_1 \neq 0, \quad \frac{\partial S}{\partial \lambda_i}(0, 0) = 0 (i \neq 1) \\ w \text{ de norme minimale.} \end{array} \right.$$

Il est facile de voir que trouver w solution de (P) revient à déterminer w qui mène le système (1.2) vers (1.3) (J.-L. Lions [6], [7]).

$$\begin{cases} -\dot{P} + \mathcal{A}^* P + F'(\bar{Y}) P = \mathcal{C}^* w \\ P(T) = 0 \end{cases} \quad (1.2)$$

où \bar{Y} est solution de (1.1) pour $(\tau_0, \lambda) = 0$, \mathcal{A}^* et \mathcal{C}^* désignant les opérateurs adjoints de \mathcal{A} et \mathcal{C} respectivement.

$$P(0) = 0, \quad \int_0^T \mathcal{B}P(t) dt = (\alpha_1, 0, \dots, 0). \quad (1.3)$$

Trouver w revient à résoudre un problème linéaire ce qui nous ramène au cas traité dans [2]. L'estimation de λ_1 est obtenue par un développement limité à l'ordre 1 de S au voisinage de $(0, 0)$. Le développement n'étant pas exact dans ce cas ; on peut donc envisager des sentinelles d'ordre supérieur à 1 (voir paragraphe 2).

1.2. Méthode des moindres carrés

Il s'agit de minimiser la fonction coût J suivante :

$$J(\tau_0 \bar{Y}^0, \lambda) = \int_0^T \|\mathcal{C}Y - Z_{\text{obs}}\|^2 dt$$

où Z_{obs} est l'observation.

On cherche au même titre les paramètres à identifier et les termes manquants.

2. IDENTIFICATION DU PARAMÈTRE DE RAYLEIGH-BÉNARD DANS L'ÉQUATION DE LORENZ

2.1. Présentation du problème

On considère l'équation de Lorenz présentant des solutions chaotiques (C. Sparrow [12], R. Luce [10]) :

$$\begin{cases} \dot{x} = Pr (y - x) \\ \dot{y} = -xz + rx - y \\ \dot{z} = xy - bz \end{cases} \tag{2.1}$$

où Pr, b sont des constantes connues.

Notre but est d'identifier le paramètre r , sachant que les conditions initiales sont mal connues. Sous la forme (2.1) le problème n'entre pas dans le cadre général que l'on s'est posé au paragraphe 1. On va donc effectuer le changement de variable $z' = r - z$ et ne considérer par la suite que l'équation (2.2), on suppose que $r = \bar{r} + \lambda_1$ avec \bar{r} connu et λ_1 est à déterminer.

$$\begin{cases} \dot{x} = Pr (y - x) \\ \dot{y} = xz' - y \\ \dot{z}' = -xy - bz' + b\bar{r} + b\lambda_1 \end{cases} \quad \text{avec} \quad \begin{cases} x(0) = x_0 + \tau_1 \\ y(0) = y_0 + \tau_2 \\ z'(0) = z'_0 + \tau_3 \end{cases} \tag{2.2}$$

où les τ_i sont des perturbations des valeurs initiales.

2.2. Identification par la méthode des moindres carrés

Il s'agit de minimiser la fonction coût J suivante :

$$J(\tau_i, \lambda_1) = \int_0^T \| \mathcal{E}Y - Z_{\text{obs}} \|^2 dt \tag{2.3}$$

avec $Y = (x, y, z')^T$.

La minimisation de la fonctionnelle J nécessite la connaissance de son gradient par rapport aux différents paramètres que l'on obtient en utilisant une méthode Lagrangienne. On note $Q = (Q_1, Q_2, Q_3)^T$ la variable adjointe ; le système d'optimalité est donné par (2.2), (2.3), (2.4) l'équation adjointe et (2.5) le gradient.

$$\begin{cases} -\dot{Q} = A^T Q + \mathcal{E}Y - Z_{\text{obs}} \\ Q(T) = 0 \end{cases} \quad \text{avec} \quad A = \begin{pmatrix} -\text{Pr} & \text{Pr} & 0 \\ z' & -1 & x \\ -y & -x & -b \end{pmatrix} \quad (2.4)$$

où A est la matrice jacobienne de (2.2).

$$\frac{\partial J}{\partial \lambda_1} = b \int_0^T Q_3(t) dt, \quad \frac{\partial J}{\partial \tau_i} = Q_i(0). \quad (2.5)$$

La nullité du gradient n'est qu'une condition nécessaire qui caractérise le minimum de J , elle n'est pas suffisante et il peut y avoir des minima locaux.

2.3. Identification par la méthode des sentinelles

On cherche une sentinelle

$$S(\tau_i, \lambda_1) = \int_0^T w \mathcal{E}Y dt.$$

Un développement limité à l'ordre 2 de S au voisinage de 0 donne :

$$\begin{aligned} S(\tau_i, \lambda_1) &= S(0,0) + \sum_{i=1}^3 \tau_i \frac{\partial S}{\partial \tau_i}(0,0) + \lambda_1 \frac{\partial S}{\partial \lambda_1}(0,0) \\ &+ \frac{\lambda_1^2}{2} \frac{\partial^2 S}{\partial \lambda_1^2}(0,0) + \sum_{i=1}^3 \frac{\tau_i^2}{2} \frac{\partial^2 S}{\partial \tau_i^2}(0,0) \\ &+ \sum_{i=1}^3 \sum_{j=1 \atop (j \neq i)}^3 \tau_i \tau_j \frac{\partial^2 S}{\partial \tau_i \partial \tau_j}(0,0) + \sum_{i=1}^3 \tau_i \lambda_1 \frac{\partial^2 S}{\partial \tau_i \partial \lambda_1}(0,0) \\ &+ \text{termes d'ordre 3.} \end{aligned}$$

S définit une sentinelle d'ordre 1 si w est de norme minimale et si

$$\frac{\partial S}{\partial \tau_i}(0,0) = 0, \quad \frac{\partial S}{\partial \lambda_1}(0,0) = \alpha_1.$$

Si de plus

$$\frac{\partial^2 S}{\partial \tau_i \partial \tau_j}(0,0) = 0 \quad (i, j = 1, \dots, 3), \quad \frac{\partial^2 S}{\partial \tau_i \partial \lambda_1}(0,0) = 0 \quad (i = 1, \dots, 3),$$

$$\frac{\partial^2 S}{\partial \lambda_1^2}(0, 0) = 0 \text{ (ou } \neq 0 \text{)} (*)$$

alors S définit une sentinelle d'ordre 2.

PROPOSITION 2.1 : Déterminer une sentinelle d'ordre 1 revient à trouver w de norme minimale qui mène le système (2.6) vers le système (2.7)

$$\begin{cases} -\dot{P} = A^T(\bar{Y})P + \mathcal{C}^T w \\ P(T) = 0 \end{cases} \tag{2.6}$$

$$P(0) = 0, \quad \int_0^T bP_3(t) dt = \alpha_1. \tag{2.7}$$

Si en plus de (2.7) on a (2.8)

$$\int_0^T G(Y_\gamma, Y_\nu)^T P(t) dt = 0 \quad \forall \gamma, \nu \in \{\tau_i, \lambda_1\} \tag{2.8}$$

$$\text{où } G(Y_\gamma, Y_\nu) = \begin{pmatrix} 0 \\ z'_\nu x_\gamma + z'_\gamma x_\nu \\ -x_\nu y_\gamma - x_\gamma y_\nu \end{pmatrix}$$

et où, à titre d'exemple x_ν désigne la dérivée de x , l'une des composantes de la solution de (2.2), par rapport au paramètre ν . Alors S est une sentinelle d'ordre 2.

Preuve : Il suffit de dériver S et (2.2) par rapport aux paramètres, de multiplier (2.6) par Y_ν et d'intégrer par parties pour obtenir (2.7). Pour obtenir (2.8), on multiplie (2.6) par $Y_{\nu, \gamma}$ et on intègre par parties.

Remarques :

— Dans ce cas en dimension finie, montrer l'existence d'un tel w n'est pas une chose facile. Il faut montrer que l'opérateur :

$$\begin{aligned} \mathcal{L} : L^2(]0, T[) &\mapsto \mathbb{R}^2 \\ w &\mapsto \left(P(0), \int_0^T bP_3(t) dt \right) \end{aligned}$$

(*) Si $\frac{\partial^2 S}{\partial \lambda_1^2} = 0$ alors l'estimation de λ_1 est donnée par la formule obtenue pour les sentinelles d'ordre 1. Si $\frac{\partial^2 S}{\partial \lambda_1^2} = \alpha_2 \neq 0$ alors l'estimation de λ_1 est donnée par la solution de norme la plus petite de l'équation $\frac{\alpha_2}{2} \lambda_1^2 + \alpha_1 \lambda_1 + S(0, 0) - S(\tau_i, \lambda_1) = 0$.

est surjectif (pour les sentinelles d'ordre 1). Une condition suffisante est que la matrice $\mathcal{C}^T \mathcal{C}$ soit de rang maximum, mais cette condition est beaucoup trop restrictive.

— On ne peut pas appliquer les sentinelles d'ordre 2 sur un problème en dimension infinie, sauf si la perturbation de la condition initiale peut se décomposer sur une base de dimension finie. En effet le calcul de $G(Y_{\tau_0}, Y_\nu)$ nécessite le calcul de Y_{τ_0} qui dépend de \bar{Y}^0 inconnue. Si la condition initiale se décompose sur une base finie, on exprimera l'insensibilité par rapport à chacune des composantes de la base, de la même manière que nous avons procédé pour le cas que nous étudions.

2.4. Résultats numériques

La minimisation de la fonction coût (2.3) du problème de moindres carrés est obtenue en utilisant un algorithme de gradient conjugué. L'état (1.1) est discrétisé par un schéma implicite d'ordre 1. Le système d'équations non linéaires obtenu est résolu à chaque pas de temps par une méthode de Newton. Le système d'optimalité est calculé à partir de l'équation d'état et de la fonction coût discrétisées.

Pour la détermination de w on introduit la fonction coût suivante :

$$J(w) = \frac{1}{2} \left((p(0))^2 + \left(\int_0^T bP_3(t) dt - \alpha_1 \right)^2 \right) \quad (2.9)$$

pour les sentinelles d'ordre 1 et

$$J(w) = \frac{1}{2} \left((p(0))^2 + \left(\int_0^T bP_3(t) dt - \alpha_1 \right)^2 + \sum_{\gamma, \nu \in \{\tau_i, \lambda_1\}} \left(\int_0^T G(Y_\gamma, Y_\nu)^T P(t) dt \right)^2 \right) \quad (2.10)$$

pour les sentinelles d'ordre 2.

Pour trouver le contrôle w minimisant la fonction coût, on considère une méthode d'optimisation utilisant un algorithme de gradient conjugué. L'équation (2.6) est discrétisée par un schéma implicite d'ordre 1 en temps. Le système d'optimalité est calculé à partir de l'équation (2.6) et des fonctions coût discrétisées.

Des tests ont été effectués pour différentes valeurs de r et différentes conditions initiales. Globalement, les résultats sont qualitativement les mêmes, c'est pourquoi nous ne présenterons que quelques exemples.

Dans le cas des sentinelles d'ordre 1 (indiquées par S1 dans les tableaux) les contrôles w sont obtenus après seulement quelques itérations de gradient

conjugué et sont de normes L^2 faibles. Dans le cas des sentinelles d'ordre 2 (indiquées par S_2) le nombre d'itérations est plus important et il est difficile d'obtenir une fonction coût inférieure à 10^{-6} ou 10^{-7} , alors que l'on atteint facilement le epsilon machine pour les sentinelles d'ordre 1. Les contrôles w obtenus sont de normes un peu plus importantes.

TEST 1 : Comparaison des sentinelles d'ordre 1 et d'ordre 2 lorsque les conditions initiales sont connues.

Les observations sont non bruitées et l'on suppose que la condition initiale est connue ($\tau_i = 0$). λ_1 varie en partant de valeurs petites vers des valeurs plus grandes.

$$C = (0, 0, 1)^T, (x_0, y_0, z_0) = (4,03, 3,64, 5,24), r = 23,45$$

$$\|w^{S1}\| = 1,4986 \quad \text{et} \quad \|w^{S2}\| = 6,7759 .$$

Tableau 2.1. — Comparaison sentinelle d'ordre 1 et d'ordre 2.

λ_1	10^{-4}	10^{-2}	0,1	0,5	1	3
$\lambda_1^{S1} - \lambda_1$	$4,01 \cdot 10^{-10}$	$4,01 \cdot 10^{-6}$	$3,93 \cdot 10^{-4}$	$8,99 \cdot 10^{-3}$	$3,14 \cdot 10^{-2}$	0,14
$\lambda_1^{S2} - \lambda_1$	$1,11 \cdot 10^{-14}$	$1,07 \cdot 10^{-8}$	$1,08 \cdot 10^{-5}$	$1,42 \cdot 10^{-3}$	$1,19 \cdot 10^{-2}$	$1,04 \cdot 10^{-2}$

$$C = (0, 0, 1)^T, (x_0, y_0, z_0) = (10, 10, 10), r = 10$$

$$\|w^{S1}\| = 1,0681 \quad \text{et} \quad \|w^{S2}\| = 2,8322 .$$

Tableau 2.2. — Comparaison sentinelle d'ordre 1 et d'ordre 2.

λ_1	10^{-4}	10^{-2}	0,1	0,5	1	3
$\lambda_1^{S1} - \lambda_1$	$1,62 \cdot 10^{-10}$	$1,57 \cdot 10^{-6}$	$1,11 \cdot 10^{-4}$	$-3,4 \cdot 10^{-3}$	$-5,9 \cdot 10^{-2}$	0,49
$\lambda_1^{S2} - \lambda_1$	$7,46 \cdot 10^{-15}$	$1,70 \cdot 10^{-9}$	$2,32 \cdot 10^{-6}$	$7,10 \cdot 10^{-4}$	$9,63 \cdot 10^{-3}$	$-5,5 \cdot 10^{-2}$

Des résultats semblables sont obtenus pour $C = (1, 0, 0)^T$ (avec $\|w^{S1}\| = 2,2661$ et $\|w^{S2}\| = 15,6239$).

TEST 2 : Influence de la perturbation de la condition initiale.

Les observations sont non bruitées, λ_1 est fixé et les τ_i varient (tous égaux) en partant de valeurs petites vers des valeurs plus grandes.

$C = (0, 0, 1)^T, (x_0, y_0, z_0) = (4,03, 3,64, 5,24), r = 23,45, \lambda_1 = 1,$
 λ_1^0 représente la valeur obtenue lorsque les $\tau_i = 0$.

Tableau 2.3. — Influence de la perturbation de la condition initiale.

τ_1	10^{-4}	10^{-2}	0,1	0,2	0,5	1,0
$\lambda_1^{s1} - \lambda_1$	$3,17 \cdot 10^{-2}$	$2,97 \cdot 10^{-2}$	$1,31 \cdot 10^{-3}$	$-5,0 \cdot 10^{-3}$	-0,31	-0,97
$\lambda_1^{s1} - \lambda_1^0$	$-1,9 \cdot 10^{-5}$	$3,03 \cdot 10^{-3}$	$3,04 \cdot 10^{-2}$	$-8,2 \cdot 10^{-2}$	-0,34	1,01
$\lambda_1^{s2} - \lambda_1$	$1,19 \cdot 10^{-2}$	$1,37 \cdot 10^{-2}$	$3,38 \cdot 10^{-2}$	$6,01 \cdot 10^{-2}$	0,15	0,20
$\lambda_1^{s2} - \lambda_1^0$	$1,61 \cdot 10^{-5}$	$1,72 \cdot 10^{-3}$	$2,01 \cdot 10^{-2}$	$4,82 \cdot 10^{-2}$	0,14	0,19

TEST 3 : Comparaisons sentinelles et moindres carrés.

λ_1 et les τ_i sont fixés. Les observations sont bruitées avec un bruit multiplicatif suivant une loi uniforme centrée sur 0. Les valeurs du bruit portées dans le tableau correspondent à l'amplitude maximum du bruit en %.

$C = (0, 0, 1)^T$, $(x_0, y_0, z_0) = (4,03, 3,64, 5,24)$, $r = 23,45$, $\lambda_1 = 1$, $\tau_i = 0,2$.

λ_1^0 représente la valeur obtenue pour une observation non bruitée et λ_1^{mc} la valeur du paramètre obtenue par la méthode des moindres carrés.

Tableau 2.4. — Comparaison sentinelles et moindres carrés.

bruit	$5 \cdot 10^{-3}$	$5 \cdot 10^{-2}$	0,1	0,25	0,5	1,0
$\lambda_1^{s1} - \lambda_1$	$-5,0 \cdot 10^{-2}$	$-4,8 \cdot 10^{-2}$	$-4,6 \cdot 10^{-3}$	$-3,9 \cdot 10^{-2}$	$-2,8 \cdot 10^{-2}$	$-6,6 \cdot 10^{-3}$
$\lambda_1^{s1} - \lambda_1^0$	$-2,1 \cdot 10^{-4}$	$2,2 \cdot 10^{-3}$	$-4,3 \cdot 10^{-2}$	$-1,1 \cdot 10^{-2}$	$-2,2 \cdot 10^{-2}$	$-4,3 \cdot 10^{-2}$
$\lambda_1^{s2} - \lambda_1$	$5,89 \cdot 10^{-2}$	$4,77 \cdot 10^{-2}$	$-1,7 \cdot 10^{-3}$	$3,53 \cdot 10^{-2}$	$-6,3 \cdot 10^{-2}$	-0,18
$\lambda_1^{s2} - \lambda_1^0$	$1,23 \cdot 10^{-3}$	$1,23 \cdot 10^{-2}$	$6,18 \cdot 10^{-2}$	$2,47 \cdot 10^{-2}$	0,12	0,24
$\lambda_1^{mc} - \lambda_1$	$1,6 \cdot 10^{-3}$	$1,6 \cdot 10^{-2}$	$3,2 \cdot 10^{-2}$	$8,8 \cdot 10^{-2}$	0,21*	0,41*

On obtient des résultats analogues avec un bruit additif.

Lorsque l'observation n'est pas bruitée, la méthode des moindres carrés permet d'identifier λ_1 et les τ_i de façon quasiment exacte (mais au prix d'un coût calcul relativement plus important que dans le cas des sentinelles). Dans le cas de données bruitées, la résolution du problème de moindres carrés devient difficile et peut même échouer après de nombreuses itérations. Le temps de calcul entre les sentinelles et les moindres carrés est de l'ordre de 15 à 20 en faveur des sentinelles.

Les sentinelles se comportent bien face aux perturbations avec des résultats meilleurs pour les sentinelles d'ordre 1 que d'ordre 2, ceci peut s'expliquer par le fait que la norme de w est plus grande dans le deuxième

(*) Difficulté pour minimiser la fonction coût, nombre important d'itérations, existence probable de minima locaux.

cas, ce qui a tendance à amplifier les perturbations. Et d'une façon générale, on remarque que la stabilité de l'identification est d'autant meilleure que la norme de w est faible.

3. IDENTIFICATION DE TERMES DE POLLUTION DANS LE CAS D'UN SYSTÈME DE RÉACTION-DIFFUSION NON LINÉAIRE

3.1. Présentation du problème

On cherche à identifier les amplitudes des pollutions λ_i déversées dans une rivière Ω . La forme générale du modèle (A. Okubo [11], L. C. Wrobel [13]) est le suivant :

$$\begin{cases} \frac{\partial y}{\partial t} - \mathcal{A}y + F(y) = \sum_{i=1}^n \lambda_i f_i(t) \delta(x - x_i) \\ y(x, 0) = \xi \\ y = 0 \quad \text{sur} \quad \Gamma \times]0, T[\end{cases} \quad (3.1)$$

Où \mathcal{A} est un opérateur différentiel elliptique du deuxième ordre et F est une fonction non linéaire dérivable et bornée. Les $f_i(t)$ représentent la manière dont la pollution est déversée aux points x_i de la rivière. Et ξ représente la donnée initiale, totalement inconnue, qui ne nous intéresse nullement. On observe y sur un ouvert ω de la rivière Ω .

3.2. Méthode des moindres carrés

La fonction coût J s'écrit dans ce cas :

$$J(\xi, \lambda) = \int_{\omega \times]0, T[} (y - y_{\text{obs}})^2 dx dt. \quad (3.2)$$

Et on cherche

$$\inf_{\xi, \lambda} J(\xi, \lambda).$$

THÉORÈME 3.1 : *On suppose que Ω est un ouvert de \mathbb{R}^n , $n \geq 2$. Alors il existe $(\xi^*, \lambda^*) \in G \times \mathbb{R}^n$ vérifiant*

$$J(\xi^*, \lambda^*) \leq J(\xi, \lambda) + \xi \quad \forall (\xi, \lambda) \in G \times \mathbb{R}^n \quad \text{et} \quad \forall \xi > 0.$$

Où G est un espace défini par complétion de $L^2(\Omega)$ (J.-L. Lions [9]).

Pour démontrer ce théorème on aura besoin du lemme suivant :

LEMME 3.1 : Soit le problème suivant

$$\begin{cases} \frac{\partial y}{\partial t} - \mathcal{A}y = h(x, t) + \sum_{i=1}^n \lambda_i f_i(t) \delta(x - x_i) \\ y(x, 0) = \xi \\ y = 0 \text{ sur } \Gamma \times]0, T[\\ \|\xi\|_{H^\beta} \leq R, |\lambda|_{\mathbb{R}^n} \leq R \end{cases} \quad (3.3)$$

où h est une fonction appartenant à un borné B de $L^2(\Omega \times]0, T[)$ et β est un réel positif petit. Soit $(\bar{\xi}, \bar{\lambda})$ le couple optimal

$$J(\bar{\xi}, \bar{\lambda}) \leq J(\xi, \lambda) \quad \forall \|\xi\|_{H^\beta} \leq R \quad \text{et} \quad \forall |\lambda|_{\mathbb{R}^n} \leq R.$$

Alors $\forall \varepsilon > 0, \exists R > 0$ indépendant de h tel que

$$\left| J(\bar{\xi}, \bar{\lambda}) - \inf_{\xi, \lambda} J(\xi, \lambda) \right| \leq \varepsilon.$$

Remarques :

— Le couple optimal existe car le couple (ξ, λ) est dans une boule de rayon R .

— $\inf_{\xi \in G, \lambda \in \mathbb{R}^n} J(\xi, \lambda)$ existe (J.-L. Lions [9], B.-E. Aïnseba [1]).

Preuve du lemme : Soit h un élément de $H^{-2\gamma, -\gamma}(\Omega \times]0, T[)$ pour $0 < \gamma < \frac{1}{4}$. Soit G_R l'application :

$$\begin{aligned} G_R : H^{-2\gamma, -\gamma}(\Omega \times]0, T[) &\rightarrow L^2(\omega \times]0, T[) \\ h &\rightarrow y_{\omega \times]0, T[}(h, \bar{\xi}, \bar{\lambda}). \end{aligned}$$

La condition d'optimalité pour $(\bar{\xi}, \bar{\lambda})$ s'écrit :

$$\int_{\omega \times]0, T[} (y(h, \bar{\xi}, \bar{\lambda}) - y_{\text{obs}})(y(0, \xi, \lambda) - y(0, \bar{\xi}, \bar{\lambda})) \, dx \, dt \geq 0$$

$$\forall \|\xi\|_{H^\beta} \leq R \quad \text{et} \quad \forall |\lambda|_{\mathbb{R}^n} \leq R.$$

En écrivant la condition d'optimalité pour deux valeurs h_1, h_2 de h , en choisissant (ξ, λ) égal à $(\bar{\xi}^2, \bar{\lambda}^2)$ puis à $(\bar{\xi}^1, \bar{\lambda}^1)$ et en ajoutant on obtient :

$$\int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2)) \times$$

$$\times (y(0, \bar{\xi}^1 - \bar{\xi}^2, \bar{\lambda}^1 - \bar{\lambda}^2)) \, dx \, dt \leq 0.$$

Ceci étant montrons que les G_R sont équicontinues en R .

$$\begin{aligned} \int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2))^2 dx dt &= \\ &= \int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2)) \\ &\quad (y(0, \bar{\xi}^1 - \bar{\xi}^2, \bar{\lambda}^1 - \bar{\lambda}^2) + y(h_1 - h_2, 0, 0)) dx dt \\ &\leq \int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2)) y(h_1 - h_2, 0, 0) dx dt \\ &\leq \left(\int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2))^2 dx dt \right)^{1/2} \\ &\quad \left(\int_{\omega \times]0, T[} y(h_1 - h_2, 0, 0)^2 dx dt \right)^{1/2}. \end{aligned}$$

On en déduit donc :

$$\begin{aligned} \int_{\omega \times]0, T[} (y(h_1, \bar{\xi}^1, \bar{\lambda}^1) - y(h_2, \bar{\xi}^2, \bar{\lambda}^2))^2 dx dt &\leq \\ &\leq \int_{\omega \times]0, T[} y(h_1 - h_2, 0, 0)^2 dx dt \\ &\leq C \|h_1 - h_2\|_{L^2(\Omega \times]0, T[)}^2. \end{aligned}$$

Où C est une constante indépendante de R .

Dans le cas linéaire ou affine le minimum est atteint pour $(\xi, \lambda) \in G \times \mathbb{R}^n$; mais H^β est dense dans L^2 qui lui même est dense dans G , donc :

$$\lim_{R \rightarrow \infty} J(\bar{\xi}, \bar{\lambda}) = \inf_{\xi \in H^\beta(\Omega), \lambda \in \mathbb{R}^n} J(\xi, \lambda) = \inf_{\xi \in G, \lambda \in \mathbb{R}^n} J(\xi, \lambda).$$

On en déduit alors que $G_R(h)$ est convergente dans $L^2(\Omega)$, pour tout h , quand $R \rightarrow \infty$. Le théorème d'Ascoli nous permet alors d'affirmer que les G_R convergent uniformément sur tout compact de $H^{-2\gamma, -\gamma}(\Omega \times]0, T[)$. Mais comme l'injection de $L^2(\Omega \times]0, T[)$ dans $H^{-2\gamma, -\gamma}(\Omega \times]0, T[)$ est compacte, le lemme est démontré pour tout F borné dans $L^2(\Omega \times]0, T[)$.

Démonstration du théorème 3.1 : Considérons le problème linéaire suivant :

$$\begin{cases} \frac{\partial y}{\partial t} - \mathcal{A}y + F(z) = \sum_{i=1}^n \lambda_i f_i(t) \delta(x - x_i) \\ y(x, 0) = \xi \\ y = 0 \quad \text{sur} \quad \Gamma \times]0, T[. \end{cases} \tag{3.4}$$

Soit $R \geq 0$ donné et $\beta > 0$ petit. Considérons la multi-application \mathcal{V}_ε de $L^2(\Omega \times]0, T[)$ dans lui-même, définie par :

$$\mathcal{V}_\varepsilon(z) = \{y \text{ solution de (3.4)} / \exists \xi_\varepsilon, \lambda_\varepsilon$$

$$\text{avec } \|\xi_\varepsilon\|_{H^\beta} \leq R, \quad \|\lambda_\varepsilon\| \leq R$$

$$\text{et } \left| J(\xi_\varepsilon, \lambda_\varepsilon) - \inf_{\xi \in G, \lambda \in \mathbb{R}^n} J(\xi, \lambda) \right| < \varepsilon \}.$$

Démontrer le théorème 3.1 revient alors à prouver l'existence d'un point fixe de la multi-application \mathcal{V}_ε . Ceci est fait en utilisant le théorème du point fixe de Kakutani [3] ; il suffit alors de prouver que \mathcal{V}_ε est une correspondance d'un ensemble convexe compact dans lui-même, héli-continue supérieurement, à valeurs convexes fermées non vides.

Il est clair que $\mathcal{V}_\varepsilon(z)$ est convexe, compact car pour toute suite y_n de $\mathcal{V}_\varepsilon(z)$, il existe (ξ_n, λ_n) , bornée dans $H^\beta \times \mathbb{R}^n$, dont on peut extraire une sous suite convergente dans $L^2 \times \mathbb{R}^n$. On construit donc une sous suite y_{n_k} de y_n convergente.

$\mathcal{V}_\varepsilon(z)$ est non vide pour tout z de $L^2(\Omega \times]0, T[)$; en effet F étant borné, on peut utiliser le lemme précédent et choisir R suffisamment grand pour que le minimum soit atteint à ε près.

Pour montrer que \mathcal{V}_ε est semi continue supérieurement, on va raisonner par l'absurde : pour tout voisinage \mathcal{H} de $\mathcal{V}_\varepsilon(z)$ dans $L^2(\Omega \times]0, T[)$, on pourrait trouver une suite z_n qui converge vers z , dont on peut extraire une sous suite toujours notée z_n qui converge vers z presque partout et une suite $y_n \in \mathcal{V}_\varepsilon(z_n)$, mais $y_n \notin \mathcal{H}$.

F étant continue et bornée, $F(z_n)$ converge vers $F(z)$ dans $L^2(\Omega \times]0, T[)$ fort. Notons par $\xi_n = y_n(0, x)$, on a $\|\xi_n\|_{H^\beta} \leq R$ et $\|\lambda_n\| \leq R$. On peut donc extraire une sous suite $\{\xi_n\}$ convergente vers ξ dans $H^\beta(\Omega)$ faible et dans $L^2(\Omega)$ fort. De même on extrait une sous suite $\{\lambda_n\}$ convergente vers λ . La boule fermée, de rayon R , de $H^\beta(\Omega)$ est convexe, donc faiblement fermée, alors $\|\xi\|_{H^\beta} \leq R$. De même on a $\|\lambda\| \leq R$.

Soit $y(z, \xi, \lambda)$ la solution de (3.4), correspondant à z, ξ, λ . Alors y_n converge vers y dans $L^2(\Omega \times]0, T[)$ et donc dans $L^2(\omega \times]0, T[)$. On en déduit alors que $y(z, \xi, \lambda) \in \mathcal{V}_\varepsilon(z)$ ce qui contredit $y_n \notin \mathcal{H}$.

Finalement lorsque z parcourt $L^2(\Omega \times]0, T[)$, $F(z)$ reste borné dans $L^2(\Omega \times]0, T[)$ et donc $y(z, \xi, \lambda)$ est borné dans $L^2(\Omega \times]0, T[)$. L'image de \mathcal{V}_ε est un compact de $L^2(\Omega \times]0, T[)$ et est incluse dans une boule convexe de $L^2(\Omega \times]0, T[)$. L'enveloppe convexe de l'image de \mathcal{V}_ε est donc compacte. Si l'on restreint \mathcal{V}_ε à $\text{conv}(\text{Im } \mathcal{V}_\varepsilon)$, on peut appliquer le théorème de Kakutani [3]. Il existe alors un point fixe de la multi-application \mathcal{V}_ε . L'infimum est donc atteint.

3.3. Résultats numériques

L'identification des paramètres λ_i dans ce problème non linéaire par la méthode des sentinelles se ramène à un problème de contrôle linéaire, comme nous l'avons remarqué au paragraphe (1.1). Le calcul de w s'effectue en minimisant la fonction coût suivante :

$$J(w) = \frac{1}{2} \left((P(0))^2 + \left(\int_0^T \beta P(t) dt - (\alpha_1, 0, \dots, 0) \right)^2 \right) \quad (3.5)$$

où P est donné par (1.2) et $\beta_i = f_i(t) \delta(x - x_i)$.

Que ce soit dans le cas des sentinelles ou des moindres carrés, la minimisation des fonctions coût est obtenue en utilisant un algorithme de gradient conjugué sur les problèmes discrétisés (fonctions coût et équations d'état discrétisées). La discrétisation en espace des équations (1.2) et (3.1) est effectuée par éléments finis de type $P1$, la discrétisation en temps est effectuée par un schéma implicite d'ordre un. Dans le cas des moindres carrés, à chaque pas de temps, le système d'équations non linéaires obtenu est résolu par une méthode de Newton.

Comme exemple nous avons considéré un tronçon de rivière dont la longueur est de 4 620 mètres et la largeur de 60 mètres (fig. 3.1).

La variable y de l'équation (3.1) représente la D.B.O. (demande en oxygène dissout) (A. Okubo [11], L. C. Wrobel [13]). L'opérateur \mathcal{A} est donné par :

$$\mathcal{A} = V \cdot \nabla y - D_1 \frac{\partial^2 y}{\partial x_1^2} - D_2 \frac{\partial^2 y}{\partial x_2^2} + Ky.$$

V est le vecteur vitesse de transport et D_1 , D_2 sont les coefficients de dispersion. K représente le coefficient de réaction de la partie linéaire, la partie non linéaire du terme de réaction (J.-P. Kernevez [5]) est donnée par

$$F(y) = \frac{y}{1 + |y|}.$$

On suppose que le fluide est idéal et irrotationnel, la vitesse sera donc exprimée à partir du gradient de potentiel de vitesse.

$$\begin{cases} \Delta_D \Phi = 0 \\ \Phi = \Phi_d \text{ sur } S_1 \text{ et } V = \nabla \phi. \\ \frac{\partial \Phi}{\partial \nu_D} = \Phi_n \text{ sur } S_2 \end{cases}$$

Φ_d et Φ_n sont données. La valeur maximale de la vitesse est 7 Km/j. Le coefficient de dispersion longitudinale D_1 est 8 m²/s. Nous supposons que la dispersion transversale D_2 est donnée par la loi d'Okubo ([11]).

$$\frac{D_2}{D_1} = \left(\frac{L_2}{L_1} \right)^{4/3}$$

où L_2 et L_1 représentent la largeur et la longueur de la rivière.

On cherche à estimer les λ_j exprimés en Kg/j.

Nous avons comparé les résultats obtenus avec les sentinelles et les moindres carrés lorsque l'observation est bruitée. Les tests suivants synthétisent les résultats obtenus.

Le nombre de nœuds de maillage inclus dans l'observatoire est de 3. On prend $\Phi_n = 0$, $\Phi_d = 7\ 000$ en aval et 3 500 en amont.

TEST 1 :

Les pollutions et l'observatoire sont placés comme indiqué sur la figure 3.1. La valeur du paramètre λ à identifier est de 0,1. Le w minimisant la fonction coût est de norme $\|w\| = 65,63$.

Bruit en %	0	1	5	10	20	50
λ^S	0,124	0,117	0,12	0,128	0,13	$-7,4 \cdot 10^{-2}$
λ^{mdr}	0,10027	0,1005	0,1108	0,2099*	$7,21 \cdot 10^{-2*}$	$-5,9 \cdot 10^{-2*}$

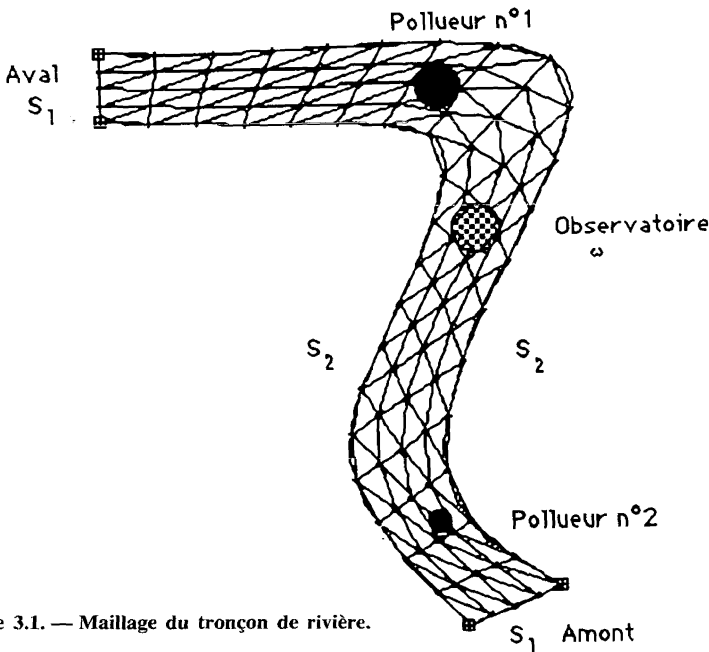


Figure 3.1. — Maillage du tronçon de rivière.

(*) Échec de la procédure d'optimisation après de nombreuses itérations.

TEST 2 :

On se place dans le même cadre que dans le test précédent mais avec deux pollutions placés en amont de l'observatoire. Les valeurs des deux paramètres à identifier sont de 0,1.

Pour l'identification de λ_1 le w minimisant la fonction coût est de norme $\|w\| = 134,12$.

Bruit en %	0	1	2	3	4	5	10
λ_1^S	0,11	0,124	0,155	0,146	$-7,3 \cdot 10^{-2}$	$5,5 \cdot 10^{-2}$	0,151
λ_1^{ndr}	0,1001	$9,810^{-2}$	0,1008	$9,6 \cdot 10^{-2}$	$4,6 \cdot 10^{-2}$	0,16	0,22

Pour l'identification de λ_2 le w minimisant la fonction coût est de norme $\|w\| = 882,7$.

Bruit en %	0	1	5	10
λ_2^S	0,11	$1,58 \cdot 10^{-2}$	$-6,01 \cdot 10^{-2}$	-2,31
λ_2^{ndr}	0,101	0,14	0,69	2,69

Remarque :

— La nature du bruit influe sur les résultats. Par exemple l'identification se dégrade plus rapidement avec un bruit décentré.

— La stabilité de l'identification est d'autant meilleure que la norme de w est faible.

Les sentinelles et les moindres carrés donnent des résultats qualitativement identiques. Cependant les sentinelles offrent l'avantage de nécessiter un coût calcul 20 à 30 fois inférieur à celui des moindres carrés.

CONCLUSION

Alors que dans le cas des problèmes linéaires, on montre que les valeurs du paramètre à identifier obtenues par la méthode des sentinelles et des moindres carrés sont les mêmes ([2]), il n'en est pas de même dans le cas des problèmes non linéaires. Lorsque l'observation n'est pas bruitée les moindres carrés donnent généralement une meilleure estimation que les sentinelles principalement quand le paramètre à identifier devient grand. Quand les données sont bruitées la méthode des moindres carrés devient très coûteuse et peut même échouer après de nombreuses itérations, alors que les sentinelles sont relativement robustes face aux perturbations de l'observation. De plus le calcul de w ne dépend pas de l'observation ce qui rend les sentinelles particulièrement avantageuses lorsque l'on dispose de plusieurs essais de mesures.

Remerciements. Les auteurs remercient le Professeur J.-L. Lions pour les commentaires et aides qu'il leur a apportés.

REFERENCES

- [1] B. E. AINSEBA, 1992, *Contrôlabilité exacte, Identifiabilité, Sentinelles*, Thèse, Université de Technologie de Compiègne.
- [2] B. E. AINSEBA, J. P. KERNEVEZ, R. LUCE, *Application des sentinelles à l'identification des pollutions dans une rivière*, à paraître.
- [3] J.-P. AUBIN, 1984, *L'analyse non linéaire et ses motivations économiques*, Masson, 1984.
- [4] J. HENRY, 1978, *Contrôle d'un réacteur enzymatique à l'aide de modèles à paramètres distribués. Quelques problèmes de contrôlabilité de systèmes paraboliques*, Thèse d'état, Paris 6.
- [5] J.-P. KERNEVEZ, 1980, *Enzyme mathematics*, North-Holland.
- [6] J.-L. LIONS, 1992, *Sentinelles pour les systèmes distribués*, Masson.
- [7] J.-L. LIONS, 1988, *Sur les sentinelles des systèmes distribués : Cas des conditions initiales incomplètes (P 819-823). Conditions frontières, termes sources, coefficients incomplètement connus (P 865-870)*, C.R.A.S. 307, 819-823, 865-870.
- [8] J.-L. LIONS, 1988, *Contrôlabilité exacte, perturbations et stabilisation des systèmes distribués*, vol. 1 et 2, Masson.
- [9] J.-L. LIONS, *communication personnelle*.
- [10] R. LUCE, 1991, *Problème inverse. Contrôlabilité exacte des systèmes régis par des équations aux dérivées partielles*, Contrôle des rouleaux de Bénard, Thèse, Université de Technologie de Compiègne.
- [11] A. OKUBO, 1980, *Diffusion and ecological problems ; Mathematical models*, Springer Verlag.
- [12] C. SPARROW, *The Lorenz Equations : Bifurcations, Chaos and Strange Attractors*, Applied Mathematical Sciences, vol. 41, Springer-Verlag.
- [13] L. C. WROBEL, C. A. BREBBIA, 1991, *Water pollution : Modelling, measuring and prediction*, Elsevier.