

وزارة التعليم العالي و البحث العلمي

Ministère de l'enseignement supérieur et de la recherche scientifique

جامعة فرحات عباس سطيف

UNIVERSITE FERHAT ABBAS – SETIF

UFAS (ALGERIE)

Mémoire

Présenté au département d'Electrotechnique
Faculté de Technologie

Pour obtenir le diplôme

De Master

Commande des processus industriels

Par

CHERRAD Fateh

Thème

Identification des systèmes non linéaires

Soutenue le/...../2011 devant la commission d'examen composée de :

Dr. HARMAS. M.N.

M C à l'Université de Sétif

Président

Mr. AGGOUNE. L.

M A. A. à l'Université de Sétif

Encadreur

Mr. BOUROUBA. B.

M A. A. à l'Université de Sétif

Examineur

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Dédicaces

*Je dédie ce modeste travail à la personne qui
m'est la plus chère au monde ma mère celle
qui m'a élevée, éduquée et sacrifiée toutes les
belles années de sa vie pour me voir un jour
comme je suis.*

A mon cher père

A mes chers frères

A toute ma famille

A mes chères sœurs

A mes tantes saida et leila

A mon ami adel brahimi

A tous mes amis(es) et collègues

Remerciements

Nous remercions Dieu de nous avoir accorde des connaissances de la science et de nous avoir aidé à réaliser ce travail.

Nous tenons à remercier chaleureusement et respectivement tous ceux qu'ont contribués de près ou de loin à la réalisation de ce mémoire de fin d'étude

Nous tenons à remercier tous les enseignants qui nous ont suivies durant notre formation.

En fin nous remercions toute personne qui a aidé de près ou de loin à la finalisation de notre mémoire.

Tables des matières

Introduction générale	8
-----------------------------	---

CHAPITRE1 : Modélisation et identification des systèmes

1.1 Introduction	10
1.2 Notion de processus	10
1.3 Notion de modèle.....	11
1.3.1 Intérêts de Modèles.....	11
1.3.2 Principe de la modélisation.....	12
1.3.3 Les différents types de modèles	13
1.3.3.1 Modèles de connaissance.....	13
1.3.3.2 Modèles de représentation et modèles de conduite.....	13
1.4 Choix du modèle d'un processus.....	18
1.4.1 Modèle à temps continu.....	18
1.4.2 Modèles à temps discontinu ou discret.....	19
1.4.3 Modèle linéaire.....	19
1.4.4 Modèle non linéaire.....	20
1.5 L'identification.....	20
1.5.1 Procédure d'identification des systèmes.....	21
1.5.2 Types d'identification	23
1.5.2.1 Identification Paramétrique.....	23
1.5.2.2 Identification non paramétrique.....	23
1.5.3 Méthodes d'identification.....	23
1.5.3.1 Méthodes graphiques.....	24
1.5.3.2 Les méthodes récursives.....	24
1.5.3.3 Les méthodes non récursives.....	24
1.6 Adaptation et simplification des modèles.....	24
1-7 Conclusion.....	25

CHAPITRE2 : Identification des systèmes non linéaires

2.1 Introduction.....	27
2.2 Structures de modèles non linéaires.....	27
2.2.1 Modèle de Volterra-Wiener.....	29
2.2.2 Modèle de Hammerstein et de Wiener.....	30
2.2.3 Modèles flous.....	30
2.2.4 Modèles multi-experts.....	32
2.2.5 Modèle à base de réseaux de neurones.....	32
2.2.5.1 Perceptrons Multi-Couches.....	32
2.2.5.2 Réseaux de fonctions à base radiale.....	32

2.2.6 Machines à vecteurs supports.....	34
2.2.7 Multimodèles.....	34
2.3 Modèles dynamiques non linéaires.....	34
2.3.1 Modèle NFIR.....	35
2.3.2 Modèle NARX.....	35
2.3.3 Modèle NOE.....	36
2.3.4 Modèle NARMAX.....	37
2.4 Conclusion.....	38
CHAPITRE3 : Exemples d'application	
3.1 Introduction.....	40
3.2 Moindres carrés	40
THÉORÈME 3.1.....	41
3.3 Moindres carrés récursifs.....	42
THÉORÈME 3.2.....	43
3.4 Moindres carrés récursifs avec valeur d'oubli.....	44
THÉORÈME 3.3.....	44
3.5 Moindres carrés orthogonales.....	45
3.6 Exemples d'application	47
Système 1.....	47
Interprétation des résultats.	53
Système 2	53
Interprétation des résultats.....	57
3.7 Conclusion.....	57
Conclusion générale.....	58
Bibliographie.....	59
Résumé en français et en anglais.....	60

Liste des figures

Fig. 1.1 Représentation générale d'un processus.	10
Fig. 1.2 Méthodologie de la modélisation.	12
Fig. 1.3 Représentation par schéma-blocs.	16
Fig. 1.4 Régulation de processus.	16
Fig. 1.5 Graphe de fluence.	17
Fig. 1.6 Système à non linéarité séparable.	20
Fig. 1.7 Identification à partir du comportement entrée-sortie.	20
Fig. 1.8 Identification de type paramétrique.	21
Fig. 1.9 Procédure d'identification d'un système.	22
Fig. 2.1 modèle de Hammerstein.....	30
Fig. 2.2 modèle de Wiener.....	30
Fig. 2.3 Réseau de neurones RBF.....	33
Fig. 3.1 le signal d'entrée.....	49
Fig. 3.2 la sortie du système sans bruit.....	49
Fig. 3.3 bruit gaussienne.....	50
Fig. 3.4 la sortie du système avec bruit.....	50
Fig. 3.5 les paramètres réels/estimées pour $\lambda = 1.000$	51
Fig. 3.6 les paramètres réels/estimées pour $\lambda = 0.999$	51
Fig. 3.7 les paramètres réels/estimées pour $\lambda = 0.990$	52
Fig. 3.8 les paramètres réels/estimées pour $\lambda = 0.950$	52
Fig. 3.9 la sortie du système2 sans bruit.....	53
Fig. 3.10 bruit du système2.....	53
Fig. 3.11 la sortie du système2 avec bruit.....	54
Fig. 3.12 estimation du paramètre a.....	54
Fig. 3.13 estimation du paramètre b.....	55
Fig. 3.14 estimation du paramètre c.....	55
Fig. 3.15 estimation du paramètre d.....	56
Fig. 3.16 estimation du paramètre f.....	56

Liste des tableaux

Tableau 1.1 les systèmes continus, discrets.....	15
Tableau 3.1 les résultats de simulation pour les valeurs réelles et les valeurs estimées	53
Tableau 3.2 les résultats de simulation pour les valeurs réelles et les valeurs estimées	57

Nomenclature

1. Notations et Symboles

t, k	Indice de temps entier
$u(t)$	Signal d'entrée
$y(t)$	Signal de sortie
x	Vecteur d'état du système
\dot{x}	La dérivée totale dx/dt du vecteur état x
$y_s(\cdot)$	La sortie réelle du système
h	Le pas de prédiction
$\underline{u}(t)$	Un vecteur des éléments des entrées externes du système à l'instant courant t
$\underline{\tilde{y}}_s(t)$	Un vecteur dont les composantes sont liées à l'état du système à l'instant t
n_{u_k}	L'ordre de l'entrée u_k
n_i	Le nombre d'entrées
n_{y_s}	L'ordre de la sortie y_s
$e(t)$	Variable aléatoire de moyenne nulle et de variance σ^2 représentant le bruit.
$E(Y_s(t+h) t)$	L'espérance mathématique conditionnelle de $Y_s(t+h)$
$F(\cdot)$	Fonction non linéaire
$\varphi(t)$	Le vecteur de régression
θ	Vecteur de paramètres
$V(\theta, t)$	La fonction de perte des moindres carrés
$\varepsilon(t)$	L'erreur de prédiction du modèle
$h_i(\tau_1, \dots, \tau_i)$	La réponse impulsionnelle d'ordre i du système.
d_{u_k}	Le retard correspondant à l'entrée u_k
$\hat{\theta}$	Vecteur de paramètres estimé
N	Nombre de points de données
P	Variance de $\hat{\theta}$

$K(t)$	Facteur de pondération ou de gain
d_ε et n_ε	représentent le retard et la mémoire correspondant à l'erreur de prédiction ε .
$d_{\hat{y}}$ et $n_{\hat{y}}$	sont le retard et l'ordre associé à \hat{y} .

2. Lexique

SISO	Single Input Single Output (Single Input Single Output)
LS	Moindres Carrés (Least Squares)
RLS	Moindres Carrés récursifs (Recursive Least Squares)
MLP	Perceptrons Multi-Couches (Multi Layer Perceptron)
RBF	Les Réseaux de Fonctions à Base Radiale (Radial Basis Function)
SVM	Les machines à vecteurs supports (Support Vector Machines)
NFIR	Nonlinear Finite Impulse Response
NARX	Nonlinear AutoRegressive with eXogenous input
NOE	Nonlinear Output Error
NARMAX	Nonlinear AutoRegressive Moving Average with eXogenous input

Introduction générale

L'étude d'un système dynamique non linéaire, quelle que soit sa nature (industrielle, environnementale, financière, etc.) et quel que soit l'objectif visé (commande, optimisation du fonctionnement, prédiction, analyse du comportement, etc.), nécessite la mise en place d'une représentation capable de reproduire son comportement. Cette représentation, communément appelée modèle.

Parmi les différents types de modélisation qui existent, la modélisation mathématique connaît plus de succès grâce au progrès de l'informatique qui a permis une avancée significative des méthodes de calcul numérique. La modélisation mathématique d'un système est une représentation mathématique sous forme d'une relation liant les différentes variables régissant son fonctionnement. Cependant, la détermination de ces variables et de la structure de la relation qui les lie constitue souvent un problème majeur.

Pour certains systèmes, il est possible d'établir ces relations à partir de connaissances physiques, chimiques, biologiques ou autres : une telle représentation est appelée modèle de connaissance, modèle boîte blanche ou modèle théorique. Il est cependant très difficile, voire parfois impossible, d'établir de tels modèles pour des systèmes complexes.

Plusieurs approches ont été proposées pour la représentation des systèmes dynamiques non linéaires : méthode de linéarisation du système, représentation de Volterra-Wiener, représentations de Hammerstein et de Wiener, etc. Plus récemment, d'autres techniques ont vu le jour notamment celles basées sur les Réseaux de Neurones Artificiels (RNA), et les multimodèles.

Ce mémoire est composé de trois chapitres répartis comme suit :

Le premier chapitre présente les notions de base de la modélisation et l'identification des systèmes dynamiques.

La problématique de la représentation des systèmes dynamiques non linéaires est décrite dans le deuxième chapitre.

Dans le troisième chapitre on présente quelques exemples d'identification des systèmes non linéaires.

Enfin en va terminer par une conclusion générale.

CHAPITRE 1 : Modélisation et identification des systèmes

1.1 Introduction

L'automatique consiste en l'étude des systèmes réels des différentes disciplines scientifiques (Electronique, mécanique, thermique, chimie, écologie, biologie, économie, sociologie, physique, cosmologie...), en vue de l'analyse, de la prédiction, de la commande, de la surveillance et / ou de l'optimisation des systèmes. Généralement, la connaissance du modèle du système réel (on réalise une modélisation) est nécessaire dans l'étude [9].

Dans cette première partie, nous allons développer un aperçu détaillé sur la modélisation et l'identification des processus.

1.2 Notion de processus

Un processus est un système dynamique, c'est-à-dire un système évolutif pour lequel le temps joue un rôle fondamental.

Dans le cas général (fig. 1.1), un processus est un système traversé par des flux d'information, d'énergie et de matière tout en étant soumis à des perturbations ayant l'une des trois formes précitées [9].

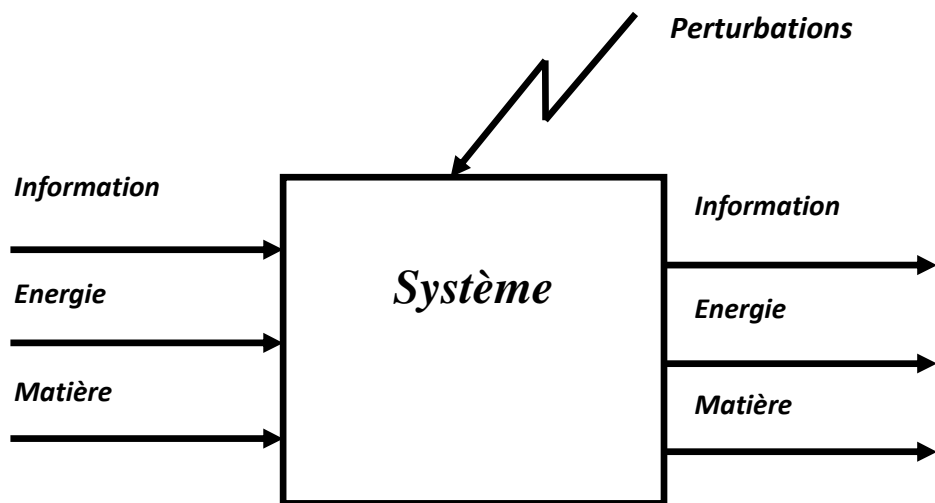


Fig. 1.1 Représentation générale d'un processus.

Du point de vue d'un observateur, un processus correspond à un système physique envisagé dans le cadre de l'évolution des échanges réalisés avec son environnement.

Diverses variables peuvent être mises en évidence sur un processus :

- des entrées de commande, qui permettent d'agir sur l'évolution du processus ;
- des entrées de perturbation, en général non contrôlables par l'utilisateur et qui agissent également sur le processus ;
- des sorties, variables mesurables ou au moins détectables, qui caractérisent l'action du processus sur son environnement ;
- des variables d'état, variables internes du système, dont l'action sur l'environnement n'est pas nécessairement perceptible mais dont l'évolution régit celle du processus.

L'étude et la commande d'un processus s'effectuent à partir d'un modèle de ce processus. Il existe plusieurs types de modèles, principalement les modèles de connaissance d'une part, et les modèles de représentation et de conduite d'autre part. Il est important de noter que dans tous les cas, le système existe indépendamment de tout modèle que l'on peut lui attribuer et que le modèle n'est le plus souvent qu'une simplification et une caractérisation de la réalité.

D'un point de vue plus épistémologique, on peut enfin remarquer que le modélisateur est, en fait, partie intégrante du modèle avec qu'il se trouve en interaction étroite :

-la nature du modèle dépend des objectifs de son concepteur ;

-le modèle adopté modifie la perception et la compréhension que le modélisateur a du processus.

1.3 Notion de modèle

Un modèle est un ensemble d'équations mathématiques qui représente au mieux les relations entre les entrées et les sorties du processus. Alors c'est une structure mathématique pouvant représenter le système étudié [13].

Un modèle mathématique est obtenu via deux approches :

- Par modélisation: modèle de connaissances/modèle boîte grise (Grey-box).
- Par identification: modèle de comportement entrées-sorties (commande)/modèle boîte noire (Black-box).

1.3.1 Intérêts de Modèles

Un modèle est utilisé :

- Dans la simulation; Exemple: simulateur d'avions.
- Dans la prédiction des sorties; Exemple: prévision d'évolution des actions dans le marché (boursier/prévision, météorologique/commande prédictive).
- Pour la synthèse des régulateurs; Exemple: contrôle de procédés industriels.

1.3.2 Principe de la modélisation

Cette partie traite de l'obtention pratique d'une représentation d'un système. L'organigramme de la figure 1.2 présente les différentes étapes permettant l'obtention d'un modèle représentatif du système étudié [11].

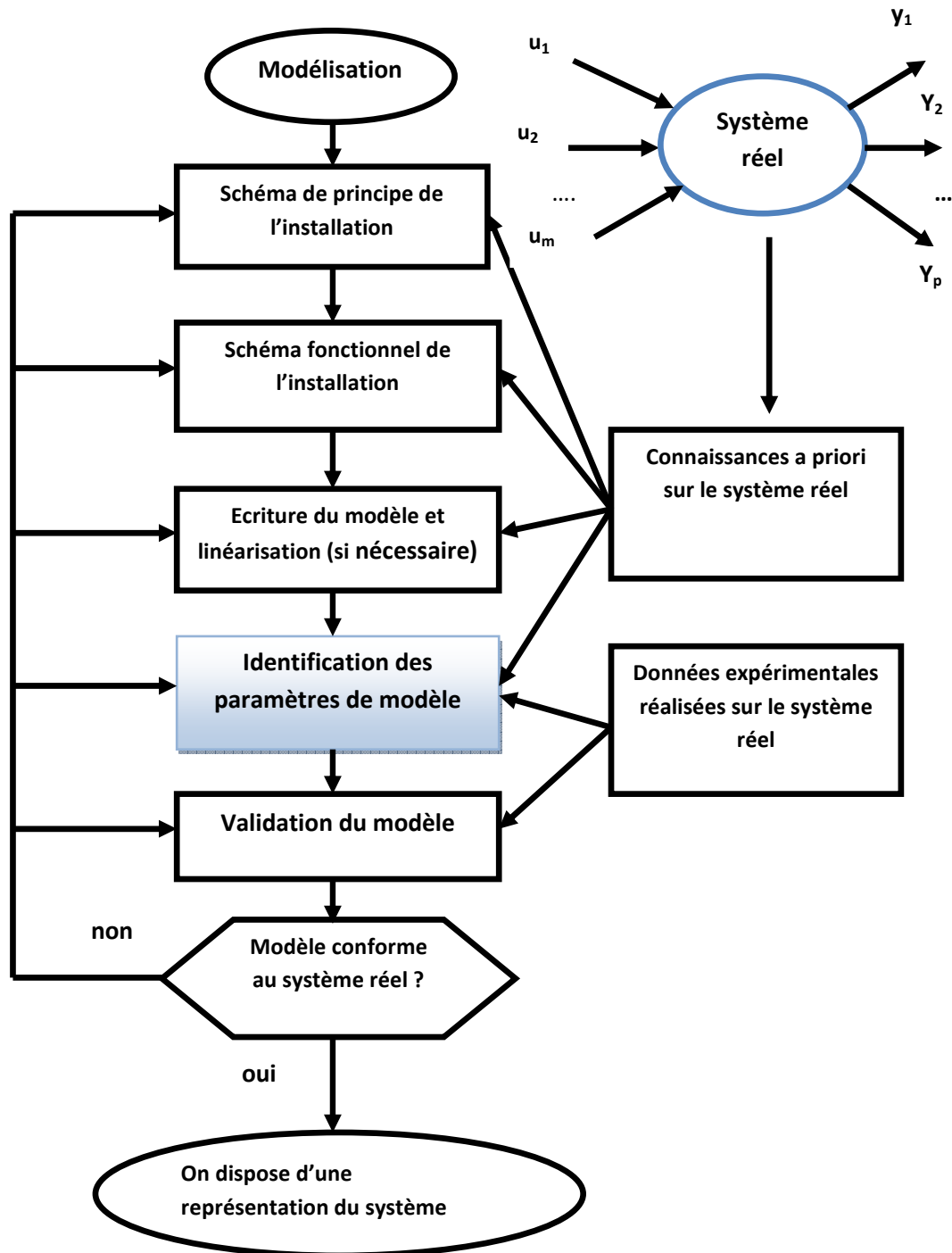


Fig. 1.2 Méthodologie de la modélisation.

1.3.3 Les différents types de modèles

Il existe deux types :

1.3.3.1 Modèles de connaissance

Un modèle de connaissance est un modèle dont les caractéristiques et les équations ont été établies en faisant appel à des modèles plus généraux mettant en œuvre les lois de la physique, de la chimie, de la biologie, de l'économie,...les paramètres d'un tel modèle ont alors une interprétation physique directe : température, pression, courant, accélération, force... Ils sont beaucoup plus riches de signification que les modèles de représentation définis ci-dessous et contiennent toutes les informations utiles sur le processus étudié. Par contre, ils sont en général difficiles à déterminer et de mise en œuvre complexe [9].

Exemple: Contrôle de la vitesse d'un ventilateur entraîné par une MCC commandé par un Hacheur série.

1.3.3.2 Modèles de représentation et modèles de conduite

Ces modèles ne permettent pas, le plus souvent, d'interprétation physique des phénomènes étudiés. Ils sont constitués d'un ensemble de relations mathématiques qui vont relier dans un domaine d'évolution donné, les différentes variables du processus. Les paramètres de tels modèles peuvent n'avoir aucun sens physique particulier connu.

Exemple: Système de climatisation d'un grand centre commercial / conduite d'une bicyclette.

1.3.3.2.1 Modèles de règles

Ils correspondent à une description par règles de conduite, comme "si on actionne A deux fois, B avance" et sont issus de l'observation du système en fonctionnement. Souvent lourds à manier et limités du point de vue possibilités, ils se prêtent assez bien à une exploitation mettant en œuvre les techniques de l'intelligence artificielle.

1.3.3.2.2 Modèles fichiers

Ils sont constitués d'informations, sous forme de tableaux de données, reliant l'évolution des sorties du processus à l'évolution des entrées pour diverses classes d'entrées. Ces modèles constituent le plus souvent le point de départ en vue de l'élaboration de modèles plus évolués (en analyse structurale, ce niveau s'appelle le "système source").

1.3.3.2.3 Modèles entrées-sorties

Dans ce type de représentation, les entrées et sorties du processus sont liées par un ensemble de relations mathématiques (en analyse structurale c'est le niveau "système générateur"). Cet ensemble peut être composé de relations algébriques d'équations différentielles, et de relations récurrentes. A ces relations de type égalité, peuvent se rajouter des relations de type inégalité ou inclusion, permettant de tenir compte de contraintes ou d'éventuelles saturations. Dans le cas de processus linéaires stationnaires à état continu, les relations entrées-sorties peuvent être définies par des matrices de transfert (*en s ou en z*).

Les modèles entrées-sorties correspondent aux représentations du type "boîte noire".

1.3.3.2.4 Modèles d'état

Ils sont caractérisés par un ensemble de variables, en nombre minimum, regroupées dans un vecteur x de R^n appelé vecteur état, dont la connaissance à l'instant $t_0 \in R$ associée à la connaissance de l'évolution des entrées $u \in R^l$ sur l'intervalle $T = [t_0, \tau]$ permet, à partir du modèle, de prévoir l'évolution de $x(t)$ sur T . Dans l'hypothèse déterministe, il existe alors une fonction $\varphi(t, t_0, x(t_0), u[t_0, t])$ appelée fonction de transition d'état telle que :

$$x(t) = \varphi(t, t_0, x(t_0), u[t_0, t]),$$

Un tel modèle est appelé équation d'état sous forme explicite, par opposition aux modèles de types équations différentielles ou récurrentes qui sont des équations d'état sous forme implicite. Les systèmes à paramètres distribués peuvent s'interpréter comme des systèmes de vecteur état de dimension infinie.

Il est important de noter que les systèmes physiques sont non anticipatifs et ne peuvent pas dépendre d'événements futurs.

Dans l'ensemble R^n , certains états peuvent ne pas exister, ce qui conduit à définir l'ensemble des états admissibles $X \subset R^n$ comme l'ensemble des états x possibles. Par exemple, pour un registre de 4 bits, 16 états seulement sont possibles.

Pour un système simple de charge d'une capacité C à travers une résistance R , l'état (tension aux bornes de la capacité) peut être limité entre -250 et +250 Volts.

Si cet ensemble X est compact, le système est dit à état continu. Si X est dénombrable (donc non compact), on a un système à état discret.

Dans les deux exemples précédents, le registre est donc un système à états discrets, et le système RC est à états continus.

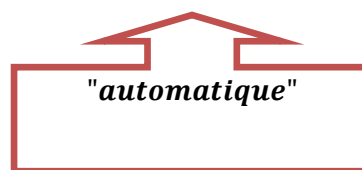
Si le nombre d'éléments de X est fini, le système est à états finis. C'est le cas en pratique pour tous les systèmes informatiques à états discrets, du fait des saturations de capacité mémoire.

Enfin, les systèmes échantillonnés correspondent à des systèmes à états continus observés ou commandés à des instants dénombrables. En pratique, ces instants sont très souvent générés par une (ou plusieurs) horloge(s) régulière(s), ce qui rend l'échantillonnage périodique.

Ceci peut se résumer par le tableau suivant :

Tableau 1.1 *les systèmes continus, discrets*

<div style="text-align: right;">x</div> <div style="text-align: left;">t</div>	<i>système à états continus</i>	<i>système à états discrets</i>
<i>temps continu</i>	<i>système continu exemple : équation différentielle</i>	<i>exemple : système asynchrone</i>
<i>temps discret</i>	<i>système discret exemple : équation de récurrence</i>	<i>exemple : système synchrone</i>



1.3.3.2.5 Modèles graphiques

Les modèles graphiques constituent un mode de représentation en général assez aisé à manipuler et possédant des propriétés remarquables susceptibles d'aider à la modélisation et parfois de permettre de la valider. Les principaux types de modèles graphiques sont :

A. Les schémas fonctionnels

Ce mode de description appelé également représentation par schéma-blocs est souvent utilisé de façon intuitive. Il correspond à une description directe des divers éléments du processus étudié, (fig. 1.3 et 1.4), faisant ou non intervenir les diverses relations mathématiques mises en œuvre.

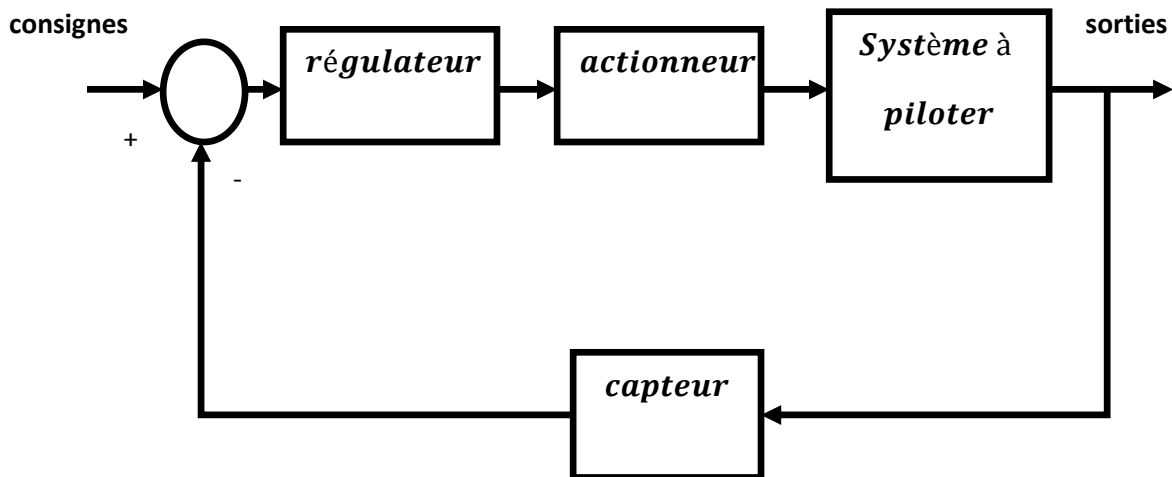


Fig. 1.3 Représentation par schéma-blocs.

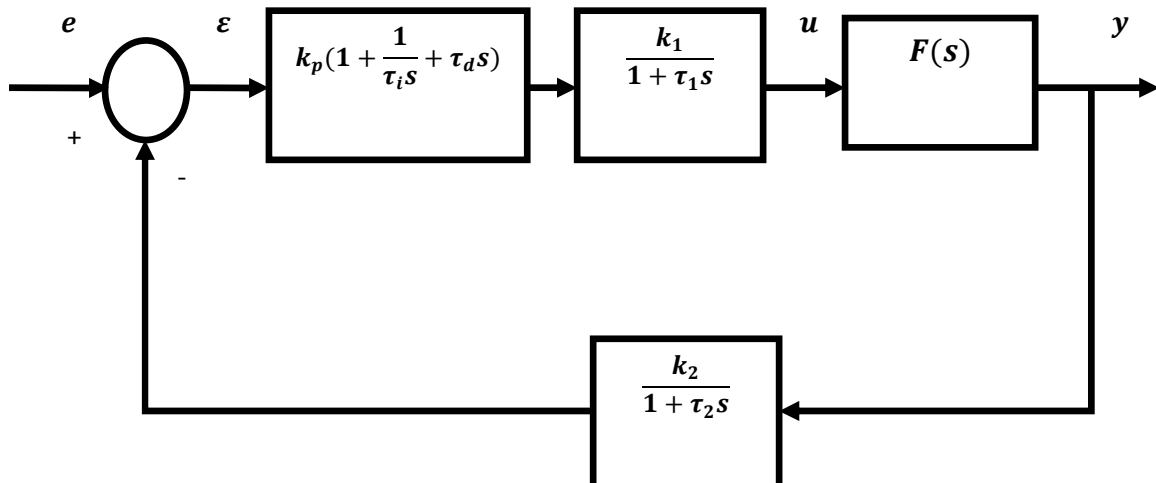


Fig. 1.4 Régulation de processus.

B. Les graphes de fluence

Ils correspondent à une représentation proche de celle des schémas fonctionnels (fig. 1.5) (ces deux représentations sont en effet duales l'une de l'autre).

L'intérêt des graphes de fluence apparaît principalement au niveau de la modélisation des systèmes linéaires, c'est-à-dire pour lesquels il est possible d'appliquer le théorème de superposition défini plus loin.

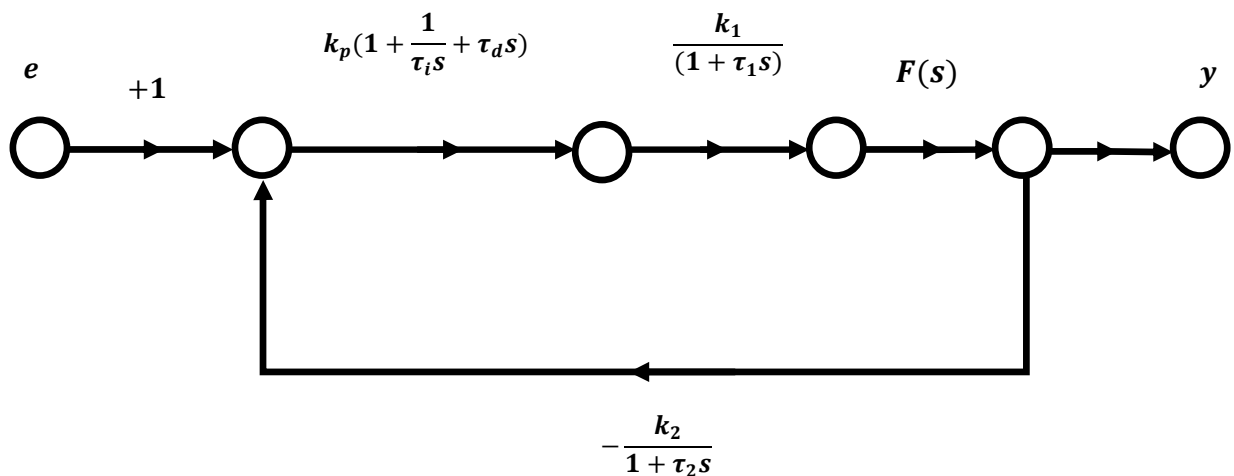


Fig. 1.5 Graphe de fluence.

C. Les réseaux de Pétri et les grafjets

Ces représentations sont particulièrement adaptées aux processus à événement discrets dont le nombre d'états est fini.

Très importants, principalement dans la caractérisation des systèmes logiques, des systèmes de fabrication automatisée et des ateliers flexibles, ces modes de représentation sont développés dans un volume spécifique de cette collection.

D. Les bond-graphs

Ils permettent une description des systèmes physiques (mécaniques, électriques, hydrauliques...) très bien adaptée à la modélisation des transferts de puissance, avec un langage unique quel que soit le domaine physique concerné.

1.4 Choix du modèle d'un processus

La détermination d'un modèle mathématique d'un processus nécessite en général diverse hypothèse simplificatrice afin de limiter sa complexité. Dans chaque application, il apparaît nécessaire de faire un compromis entre la finesse et la précision du modèle à mettre en œuvre d'une part, et la limite de complexité admissible, compte tenu des objectifs fixés, d'autre part [9].

Nous nous intéressons ici essentiellement aux processus pour lesquels les variables caractéristiques sont susceptibles de prendre un ensemble continu de valeurs appartenant à des intervalles fixés. De plus, nous ne traiterons que des modèles de processus à paramètres localisés, c'est-à-dire, décrits par des équations différentielles ordinaires ou des équations récurrentes. L'étude spécifique des systèmes à paramètres distribués, par exemple ceux dont l'évolution est régie par des équations aux dérivées partielles, est envisagée en détail dans un autre volume de cette collection [9].

Il est important de toujours se souvenir que le choix du modèle d'un processus dépend de l'utilisation prévue pour ce modèle.

Considérons à titre d'exemple un modèle d'automobile. Si le but du modèle est d'étudier la résistance au vent du véhicule, il suffit de prendre en compte la forme extérieure du véhicule et l'état des diverses surfaces qui le caractérisent. Si on s'intéresse à la résistance du véhicule en cas de choc, le modèle sera plus complexe et devra faire intervenir les caractéristiques géométriques et mécaniques de la carrosserie mais également du châssis et de l'ensemble de la structure interne de la voiture.

Lorsque l'objectif est de prévoir le comportement du véhicule en fonctionnement, le modèle pourra être supposée rigide mais par contre, on devra tenir compte des caractéristiques du moteur, des amortisseurs, etc....

On voit que les modèles utilisés les divers cas envisagés peuvent être très différents.

1.4.1 Modèle à temps continu

Dans ce type de modèle, le temps est une variable qui évolue continuellement sur un intervalle $T \subset \mathbb{R}$ prenant en croissant toutes les valeurs situées dans cet intervalle. C'est le cas, par exemple d'un modèle dans l'espace d'état de la forme :

$$\dot{x} = f(x, u, t, v),$$

$$y = h(x, u, t, v),$$

Modèle dans lequel \dot{x} représente la dérivée totale dx/dt du vecteur état $x \in R^n$ par rapport au temps, $u \in R^l$ le vecteur des commandes, t le temps, $v \in R^{n_v}$ un vecteur de perturbation, et $y \in R^m$ le vecteur des sorties.

1.4.2 Modèles à temps discontinu ou discret

Ce type de modèle est utilisé soit lorsque le processus étudié a naturellement un mode d'évolution séquentiel, soit lorsqu'on a adopté une représentation du processus qui correspond à une observation des variables d'état ou de sortie à des instants discrets t_k du temps avec $k \in Z$. On a dans ce cas une représentation dite échantillonnée, particulièrement bien adapté aux divers types de calculs à effectuer dans les problèmes de commande de processus par ordinateur.

En notant :

$$X_k = x(t_k),$$

$$X_{k+1} = f(x_k, u_k, k, v_k),$$

$$Y_k = h(x_k, u_k, k, v_k).$$

1.4.3 Modèle linéaire

Un tel modèle caractérise un processus susceptible d'être décrit par une équation différentielle ordinaire ou une équation récurrente à coefficients constantes (stationnaire) ou en fonction du temps (non stationnaire) ou par un ensemble d'équation de ce type.

La propriété fondamentale des systèmes linéaires s'exprime par le principe de superposition : si $y_1(t)$ et $y_2(t)$ représentent respectivement l'évolution des sorties du processus pour les entrées $u_1(t)$ et $u_2(t)$ sur un horizon et pour un état initial donnés alors à l'entrée $u(t)$:

$$u(t) = \alpha_1 u_1(t) + \alpha_2 u_2(t),$$

Où α_1 et α_2 sont réels, correspond la sortie :

$$y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t),$$

Sur le même horizon et pour les mêmes conditions initiales.

Les systèmes linéaires stationnaires admettent une représentation entrée-sortie par fonction ou matrice de transfert, en s ou en z selon qu'il s'agit d'un modèle continu ou discret.

1.4.4 Modèle non linéaire

Dans un modèle non linéaire, le théorème de superposition n'est plus valable. L'aspect non linéaire peut être intrinsèque et quasiment irréductible, comme dans la modélisation de la loi d'action de masse en chimie, ou correspondre à l'association d'éléments à caractéristiques non linéaires à des systèmes pour lesquels un modèle linéaire est utilisable comme c'est le cas des systèmes à non linéarités séparables (fig. 1.6).

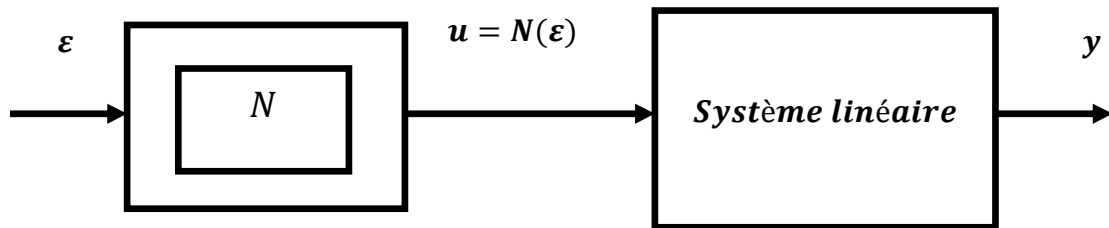


Fig. 1.6 Système à non linéarité séparable.

1.5 L'identification

L'identification constitue une phase importante dans la définition du modèle, c'est par elle que le choix de la classe de modèles à adopter puis les valeurs des paramètres qui le caractérisent vont se préciser.

Identifier un processus (système), c'est chercher un modèle (dynamique) mathématique, appartenant à une classe de modèles connue, et qui, soumis à des signaux tests (en entrée), donne une réponse (dynamique et statique en sortie), la plus proche possible du système réel.

Le plus souvent l'identification s'effectue en optimisant un critère de qualité qui caractérise l'écart entre le comportement du processus (repéré par un ensemble de mesures), et celui de son modèle (étudié par simulation) pour un ensemble de sollicitations données.

De nombreuses méthodes d'identification sont précisées dans la suite correspondant le plus souvent à l'un des schémas des figures 1.7 et 1.8.

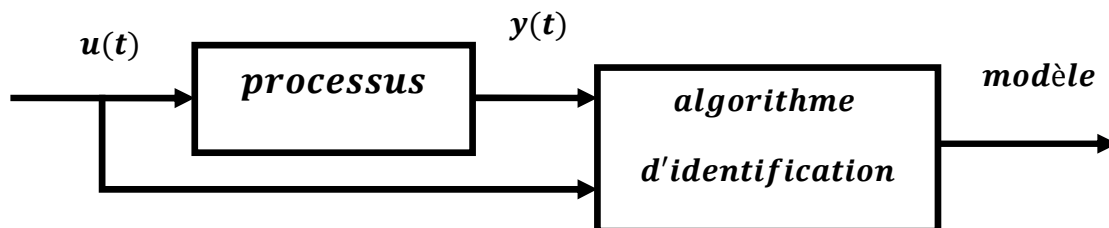


Fig. 1.7 Identification à partir du comportement entrée-sortie.

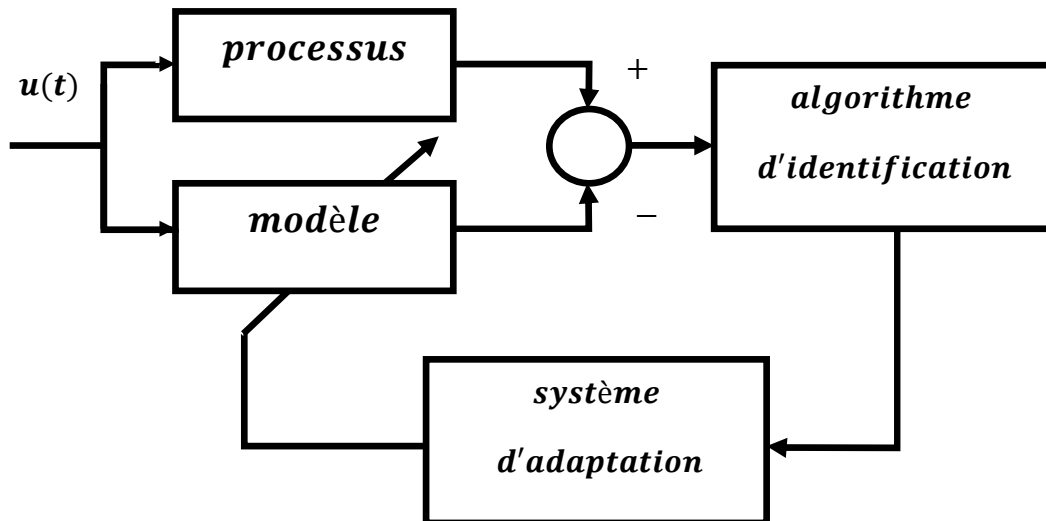


Fig. 1.8 Identification de type paramétrique.

1.5.1 Procédure d'identification des systèmes

Dans notre travail, nous utilisons la modélisation expérimentale ([5], [6]). L'établissement de ce type de modèle est une procédure itérative comportant cinq phases (voir figure 1.9) :

- 1. Extraction de données** : durant cette phase, des mesures sont effectuées sur les variables sensées caractériser le système. Ces variables peuvent être des variables externes qui agissent sur le système (entrées de commande ou perturbations mesurables), des variables internes qui traduisent l'état du système (variables d'état), ou la réponse du système (variable de sortie). Il existe souvent des perturbations non mesurables qui agissent sur le système (en entrée ou en sortie) rendant plus difficile sa modélisation.
- 2. Choix de la structure du modèle** : il s'agit de définir d'une façon formelle la relation expliquant le fonctionnement du système. Cette relation correspond à une famille de fonctions mathématiques dont une seule correspond au modèle recherché.
- 3. Choix du critère d'estimation paramétrique** : c'est le choix de la fonction objectif (fonction coût) dont l'optimisation (minimisation) permet de déterminer la structure du modèle de façon unique. Ce critère est fonction de l'écart entre la sortie du système et celle du modèle. Le critère quadratique est généralement choisi.
- 4. Estimation paramétrique** : de la famille de fonctions obtenues du choix de la structure du modèle, il faut en déterminer une qui représente convenablement le système. Il s'agit alors de trouver la valeur des paramètres permettant la satisfaction d'un critère de performance donné (optimisation de la fonction objectif).
- 5. Validation du modèle** : c'est une procédure qui permet d'évaluer l'exactitude (ou la fidélité) du modèle. Pendant cette phase, le modèle est testé avec des données non utilisées pendant la phase d'identification.

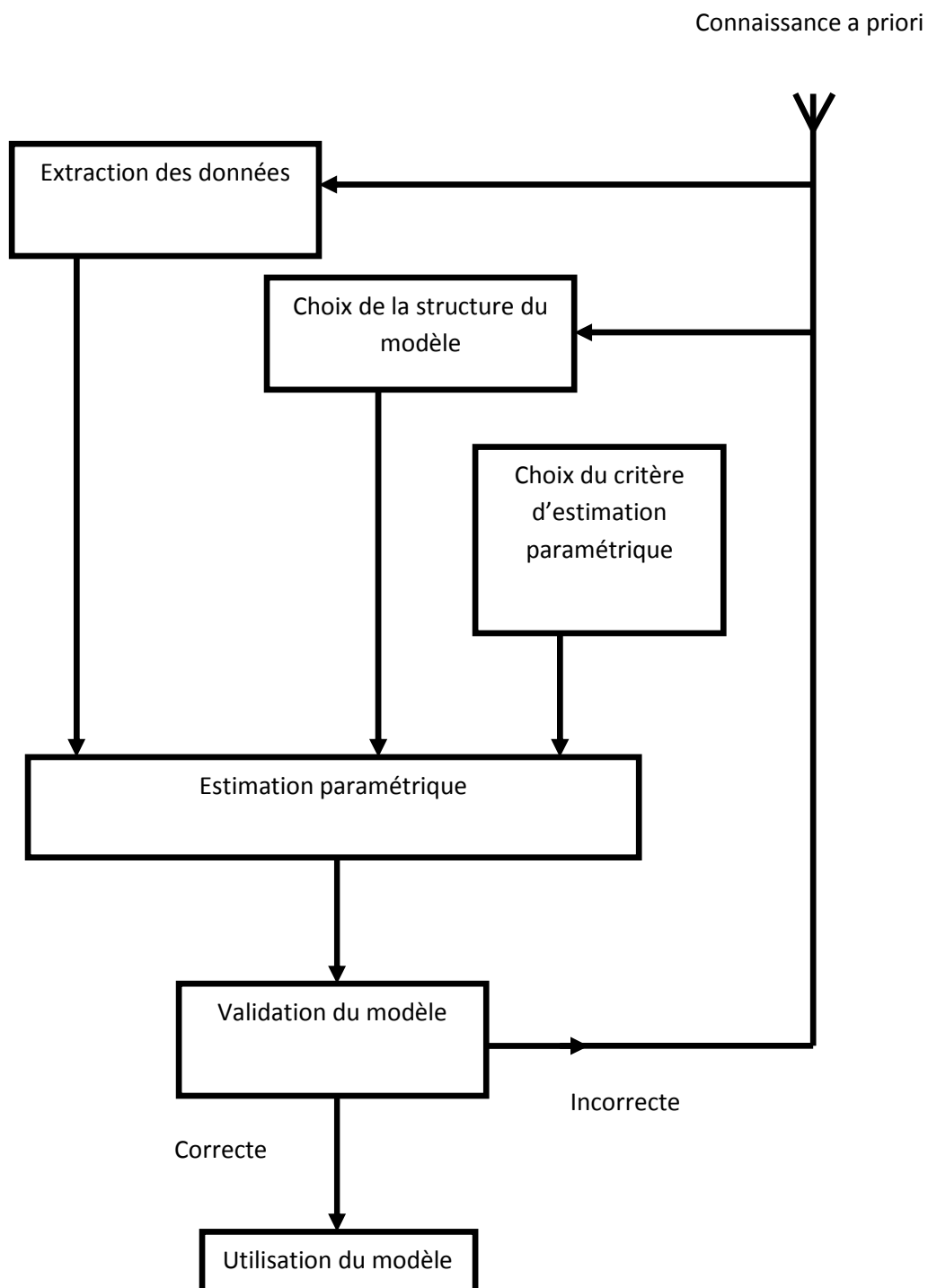


Fig. 1.9 *Procédure d'identification d'un système.*

1.5.2 Types d'identification

Le modèle à identifié peut être :

- ✚ Paramétrique: modèle caractérisé par un ensemble fini de paramètres
 - Fonction de transfert $G(p) = B(p)/A(p)$
 - Représentation d'état $dx/dt = Ax + Bu$, $y = Cx + Du$
 - Equations aux différences: $\sum a_i y(t-i) = \sum b_i u(t-i)$
- ✚ Non Paramétrique: Réponse indicielle/impulsionnelle ($g(t)$) / fréquentielle ($G(jw)$).
Selon sa on distingue deux types d'identification [13] :

1.5.2.1 Identification Paramétrique

Procédure en ligne (temps réel), se prête mieux à une commande adaptative; Peut s'effectuer alors que le système est en exploitation.

L'identification paramétrique se base sur l'estimation d'un certain nombre des paramètres d'un modèle choisi au préalable.

Les connaissances a priori sur le système à identifier peuvent guider ce choix. Cette dernière est la plus utilisée.

1.5.2.2 Identification non paramétrique

Souvent procédure hors ligne, nécessite des essais particuliers, le système doit être hors exploitation. Cette méthode est la moins utilisée.

1.5.3 Méthodes d'identification

Après avoir choisi la structure du modèle, on passe à la partie quantitative, c'est-à-dire, effectuer des valeurs numériques aux coefficients ou aux paramètres intervenant dans les équations décrivant le modèle, soit d'une manière définitive si le système est stationnaire, soit en permanence (temps réel) si le procédé évolue en fonction du temps. Ces paramètres sont reliés aux retards, gains, pôles et zéros de la fonction de transfert de système ([7], [8]).

Les méthodes à utiliser dépendent du type de modèle désiré (paramétrique ou non), des informations disponibles concernant le processus.

Généralement les méthodes sont classées en trois groupes :

- ❖ Les méthodes graphiques.
- ❖ Les méthodes récursives
- ❖ Les méthodes non récursives.

1.5.3.1 Méthodes graphiques

Souvent appelées méthodes classiques, dans ces méthodes le processus est attaqué par une entrée déterministe, c'est-à-dire, une forme d'excitation telle que (sinusoïde, échelon, impulsion de Dirac). Après avoir récolté la réponse du processus, on identifie graphiquement les paramètres.

Des exemples de ces méthodes sont :

- ✚ La méthode de Strejc : consiste à l'injection d'un échelon à l'entrée et l'obtention de la réponse indicielle du système.
- ✚ La méthode réponse fréquentielle : consiste à l'excitation du système par un signal sinusoïdale et l'enregistrements de la réponse fréquentielle.
- ✚ La méthode de la réponse Impulsionnelle : caractérise théoriquement la réponse du processus à une impulsion de Dirac.

1.5.3.2 Les méthodes récursives

Ces méthodes traitent en bloc les fichiers des données (entrées/sorties). Elles sont sensibles au bruit et perturbation.

Parmi ces méthodes on trouve

- La méthode des moindres carrés.
- La méthode de maximum de vraisemblance récursif.

1.5.3.3 Les méthodes non récursives

La plupart de ces méthodes permettent d'obtenir des modèles non paramétriques de type réponse temporelle ou fréquentielle avec précision bien meilleure que d'une approche directe.

1.6 Adaptation et simplification des modèles

Le choix d'un modèle "optimum" diffère considérablement en fonction de l'objectif. Nous avons vu par exemple que pour un processus non linéaire, le modèle utilisable dans un fonctionnement en régulateur autour d'un point de consigne peut parfois être linéaire, par contre si on s'intéresse au comportement du processus pendant la phase d'évolution qui a permis d'atteindre ce régime permanent, le choix du modèle précédent pourrait conduire à des résultats aberrants [9].

De même pour un processus possédant des variables de dynamiques très séparées, par exemple, une variable lente et une variable rapide, il peut être intéressant et même parfois nécessaire d'utiliser différents modèles suivant qu'on s'intéresse aux variations lentes ou rapides de processus. Dans un problème de régulation, il n'est pas toujours nécessaire de conserver tous les modes d'un modèle linéaire d'un processus dont la présence peut parfois

complicquer le calcul sans pour autant permettre, compte tenu de la précision de l'identification et de celle du comportement souhaité, d'accroître les performances du processus réel.

1-7 Conclusion

Elaborer un modèle de comportement d'un processus à partir des entrées/sorties nécessite à la fois de moyens logiciels pour effectuer les calculs et un savoir faire pour spécifier correctement les mesures à faire et ensuite interpréter correctement les résultats du calcul effectué et de la valider d'une manière censée un modèle significatif. Susceptible de représenter au mieux le système et qui soit le plus proche possible au système réel.

Notamment le logiciel MATLAB est un moyen très performant au calcul matriciel, à l'identification et au traitement des signaux.

Enfin, il est rare qu'on puisse trouver un modèle satisfaisant tous les critères en passant qu'une seule fois par les différents étapes d'identification, on est obligé de revoir soit la structure du modèle, ou le protocole d'identification qui doit être modifié.

CHAPITRE 2 : Identification des systèmes non-linéaires

2.1 Introduction

Dans plusieurs domaines de la science, il est souvent nécessaire de disposer des représentations de systèmes afin de maîtriser ou d'optimiser leur fonctionnement, voire de mieux comprendre leur évolution. Ces systèmes peuvent être artificiels (systèmes industriels), naturels (système écologique, biologique, etc.), sociaux ou financiers. La modélisation est une représentation du système dans le but de satisfaire un besoin préalablement défini. La représentation qu'on va faire du système, donc le modèle à établir, dépend principalement des objectifs visés. Dans les sciences pour l'ingénieur, on utilise principalement le modèle physique ou le modèle mathématique. Le travail qui fait l'objet de ce mémoire est basé sur la modélisation mathématique. L'étude se limite aux systèmes dynamiques non linéaires à temps discret de type SISO (mono-entrées, mono-sortie).

Nous abordons dans ce chapitre différentes structures de modèles de systèmes dynamiques non linéaires, les types les plus courants de modèles non linéaires, les critères d'estimation paramétrique et de sélection de modèles ainsi que les méthodes d'estimation paramétrique des modèles non linéaires.

2.2 Structures de modèles non linéaires

La construction d'un modèle « boîte noire » d'un système repose sur l'hypothèse selon laquelle il existe une relation déterministe liant les entrées du système à sa sortie [5]. D'une manière générale, le modèle prédictif de comportement d'un système dynamique non linéaire peut s'écrire sous la forme:

$$y_s(t+h) = F_s(\underline{u}(t), \underline{\tilde{y}}_s(t)) + e(t+h) \quad (2.1)$$

où :

$y_s(t+h)$ est la sortie du système à l'instant $t+h$, h étant le pas de prédiction ;

$F_s(\cdot)$ est une fonction non linéaire déterministe inconnue, appelée prédicteur théorique ;

$\underline{u}(t)$ est un vecteur dont les composantes sont des éléments des entrées externes du système à l'instant courant t et/ou aux instants antérieurs

$$\underline{u}(t) = \left[u_1(t), u_1(t-1), \dots, u_1(t-n_{u_1}+1), \dots, u_k(t), u_k(t-1), \dots, \right. \\ \left. u_k(t-n_{u_k}+1), \dots, u_{n_i}(t), u_{n_i}(t-1), \dots, u_{n_i}(t-n_{u_{n_i}}+1) \right]^T$$

avec n_{u_k} l'ordre de l'entrée u_k , $k = 1, \dots, n_i$, n_i étant le nombre d'entrées ;

$\underline{\tilde{y}}_s(t)$ est un vecteur dont les composantes sont liées à l'état du système à l'instant courant t et/ou aux instants antérieurs ; on peut avoir par exemple :

$$\underline{\tilde{y}}_s(t) = [y_s(t), y_s(t-1), \dots, y_s(t-n_{y_s}+1)]^T$$

Avec

- n_{y_s} l'ordre de la sortie y_s ;
- $e(t+h)$ est une variable aléatoire de moyenne nulle et de variance σ^2 représentant le bruit.

La représentation formelle donnée par la relation (2.1) intègre les connaissances a priori du système ainsi que des hypothèses concernant son comportement : connaissance des variables descriptives, caractère statique ou dynamique, linéaire ou non linéaire, présence ou absence de perturbations.

Le modèle est statique si la sortie à l'instant t ne dépend que des entrées à l'instant t , il est dynamique si la sortie dépend aussi des entrées et/ou sorties antérieures. Le modèle est linéaire si $y_s(t+h)$ est une combinaison linéaire de $\underline{u}(t)$ et de $\underline{\tilde{y}}_s(t)$. Dans le cas contraire, il est non linéaire. Le modèle est récurrent si $\underline{\tilde{y}}_s(t)$ comporte des variables d'état estimés à l'instant courant t et/ou aux instants antérieurs. A l'inverse, si la sortie du modèle ne dépend que des entrées et/ou des sorties mesurées, le modèle est non récurrent.

L'estimation paramétrique des modèles non linéaires récurrents est un problème majeur en identification. L'un des principaux apports de ce travail est la mise en place d'une méthodologie d'estimation paramétrique de modèles non linéaires récurrents présentée dans le chapitre 3.

En considérant la sortie du système comme une variable aléatoire $Y_s(t)$ dont une réalisation est $y_s(t)$, on peut écrire :

$$Y_s(t+h) = E(Y_s(t+h)|t) + e(t+h) \quad (2.2)$$

où $E(Y_s(t+h)|t)$ est l'espérance mathématique conditionnelle de $Y_s(t+h)$, lorsqu'on dispose de toutes les informations disponibles jusqu'à l'instant courant t , et $e(t+h)$ correspond à la partie non prédictible de $Y_s(t+h)$ à l'instant t (erreur de modélisation).

La sortie y du système, donnée par le modèle $F_s(\cdot)$, est :

$$y(t+h) = E(Y_s(t+h)|t) = F_s(\underline{u}(t), \underline{\tilde{y}}_s(t)) \quad (2.3)$$

La fonction $F_s(\cdot)$ est approchée par la fonction $F(\cdot)$ de structure connue paramétrée par un vecteur $\underline{\theta}$. La sortie \hat{y} estimée par ce prédicteur réel est :

$$\hat{y}(t+h) = F(\underline{u}(t), \underline{\tilde{y}}_s(t), \underline{\theta}) = F(\underline{\varphi}(t), \underline{\theta}) \quad (2.4)$$

où $\underline{\varphi}(t) = [\underline{u}(t)^T, \underline{\hat{y}}_s(t)^T]^T$ est le vecteur de régression ou d'information constitué de l'ensemble des variables explicatives du système dans la zone de validité du modèle. Il est obtenu par la concaténation des éléments des vecteurs $\underline{u}(t)$ et $\underline{\hat{y}}_s(t)$. Chaque élément de $\underline{\varphi}(t)$ est appelée régresseur.

L'erreur de prédiction du modèle est obtenue par :

$$\varepsilon(t) = y_s(t) - \hat{y}(t) \quad (2.5)$$

La détermination des régresseurs est un problème majeur rencontré en modélisation expérimentale. L'idéal est de ne sélectionner que les variables caractéristiques des non linéarités du système. Ces variables étant inconnues, plusieurs démarches sont proposées pour la sélection des entrées. L'une d'elle, comme expliquée dans, consiste à choisir un ensemble de variables d'entrée le plus grand possible permettant d'obtenir le « modèle complet ». Les performances de ce modèle sont comparées à celles des modèles dont les variables d'entrée constituent des sous-ensembles des variables du « modèle complet ».

Le modèle qui présente les meilleures performances est choisi. Avec cette démarche le nombre de modèle croît de façon exponentielle avec le nombre de variables, ce qui complique la mise en œuvre. Des stratégies sous-optimales mais plus simples à mettre en œuvre sont souvent utilisées en pratique [5].

Dans ce travail, nous ne traitons pas le problème de la sélection des entrées et considérons que les variables permettant de décrire les systèmes à étudier sont connues.

La description du système donnée par l'équation (2.4) conduit à deux questions fondamentales : l'identification structurelle de la fonction $F(\cdot)$ et l'estimation paramétrique.

L'identification structurelle consiste à choisir une structure de $F(\cdot)$ adéquate pour la description du système. L'estimation paramétrique de $F(\cdot)$ consiste à estimer la valeur $\underline{\hat{\theta}}$ de $\underline{\theta}$ qui minimise une norme de l'écart entre la sortie réelle du système y_s et celle du modèle \hat{y} pour l'ensemble des observations :

$$\underline{\hat{\theta}} = \arg \min_{\underline{\theta}} \left(\left\| y_s(t+h) - F(\underline{\varphi}(t), \underline{\theta}) \right\| \right) \quad (2.6)$$

La valeur $\underline{\hat{\theta}}$ détermine la structure choisie pour $F(\cdot)$ de façon unique.

La variété des modèles de représentation des systèmes dynamiques non linéaires repose sur le choix de la structure de la fonction $F(\cdot)$.

2.2.1 Modèle de Volterra-Wiener

Le modèle de Volterra-Wiener, l'un des plus anciens, est basé sur une décomposition en série, soit pour un système mono-entrée :

$$y(t) = \sum_{i=1}^{\infty} y_i(t) \quad (2.7)$$

$$y_i(t) = \sum_{\tau_1=0}^{t-1} \dots \sum_{\tau_i=0}^{t-1} h_i(\tau_1, \dots, \tau_i) u(t - \tau_1) u(t - \tau_i) \quad (2.8)$$

où $h_i(\tau_1, \dots, \tau_i)$ est la réponse impulsionnelle d'ordre i du système.

Cette représentation n'a pas connu beaucoup de succès en identification, vu le nombre très élevé de paramètres qu'elle engendre.

2.2.2 Modèle de Hammerstein et de Wiener

Les modèles de Hammerstein et de Wiener (voire figures 2.1 et 2.2) sont constitués de blocs distincts de modèles dynamiques linéaires et de modèles statiques non linéaires, interconnectés en série et/ou en parallèle ([1], [2]).

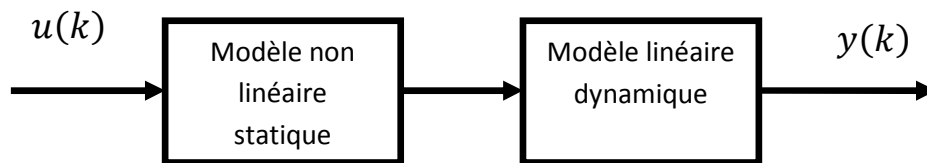


Fig. 2.1 *modèle de Hammerstein.*

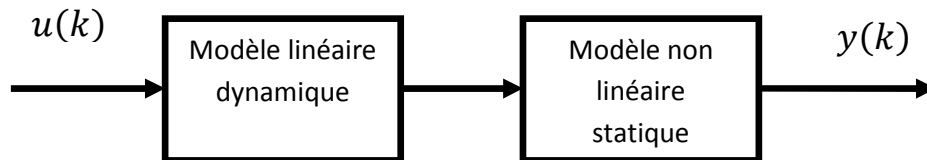


Fig. 2.2 *modèle de Wiener.*

2.2.3 Modèles flous

Les modèles flous ont connu un développement important ces dernières années. Ils permettent une représentation des systèmes non linéaires en intégrant les imprécisions du raisonnement humain. Un modèle flou est constitué d'un ensemble de règles comportant des variables linguistiques. Les règles sont de la forme :

si prémisse alors conséquence

Par exemple :

si la vitesse est faible alors appuyer sur l'accélérateur

La variable vitesse est la variable de prémisse. Afin d'évaluer son degré de véracité, on utilise des ensembles flous correspondant à des qualificatifs de la variable. Pour l'exemple

considéré, les qualificatifs de la variable de prémisses vitesse peuvent être : faible, moyenne ou élevée. Le degré d'appartenance d'une variable de prémisses à un ensemble flou est donné par une fonction d'appartenance. Pour l'exemple considéré, ce serait une fonction indiquant le degré d'appartenance de la variable de prémisses vitesse à l'ensemble flou faible (ou aux autres ensembles flous). Les fonctions d'appartenance peuvent être de différentes formes : triangulaire, trapézoïdale, gaussienne, sigmoïde, etc.

Un modèle flou est constitué principalement de quatre parties :

- la fuzzification : transforme les valeurs numériques de l'entrée en un ensemble flou ;
- la base de connaissance : comprend une base de règles permettant de décrire le système et une base de données comportant les paramètres des fonctions d'appartenance et ceux des parties conséquences des règles ;
- le moteur d'inférence : permet de déterminer, pour une entrée donnée, l'ensemble flou de sortie à l'aide de la base de connaissance ;
- la défuzzification : transforme l'ensemble flou de sortie en une valeur numérique.

Il existe principalement trois types de modèles flous: modèles flous linguistiques proposés par Zadeh en 1973 et Mamdani en 1977, le modèle à relation floue de Pedrycz (1941) et le modèle de Takagi-Sugeno.

Dans le modèle de Takagi-Sugeno, qui a connu beaucoup de succès notamment dans l'industrie, la partie conséquence de chaque règle (r_i) est une fonction affine des variables de prémisses :

$$r_i: \text{si } (x_{i1} \text{ est } A_{i1}, x_{i2} \text{ est } A_{i2}, \dots, x_{in} \text{ est } A_{in}) \text{ alors } y_i = \underline{\varphi}^T \theta_{i1} + \theta_{i0}$$

où:

A_{ij} est l'ensemble flou correspondant à la variable de prémisses x_{ij} , $j = 1, \dots, n$,

n étant le nombre de variables ;

$\underline{\varphi}^T = [x_{i1}, x_{i2}, \dots, x_{in}]^T$ vecteur composé des variables de prémisses ;

θ_{i1} vecteur de paramètres et θ_{i0} constante scalaire.

La sortie globale du modèle est obtenue par l'expression :

$$y = \frac{\sum_{i=1}^M v_i(\underline{\varphi}) y_i}{\sum_{j=1}^M v_j(\underline{\varphi})} \quad (2.9)$$

où v_i est le degré de véracité de la règle r_i .

L'expression (2.9) peut se mettre sous la forme :

$$y = \sum_{i=1}^M w_i y_i \quad (2.10)$$

$w_i = v_i(\underline{\varphi}) / \sum_{j=1}^M v_j(\underline{\varphi})$ étant le degré de véracité normalisé de la règle r_i .

2.2.4 Modèles multi-experts

Un modèle multi-experts (ou mélange d'expert - « mixture of experts » en anglais) est composé de plusieurs modèles ou réseaux d'experts et d'un réseau de déclenchement.

Ce dernier permet d'évaluer la sortie du modèle par une combinaison linéaire des sorties des réseaux experts. La sortie de chaque réseau expert est pondérée par la probabilité que la sortie du modèle soit égale à celle de l'expert. Ces modèles appartiennent à la classe des multi modèles.

2.2.5 Modèle à base de réseaux de neurones

Initialement étudiés en vue de modéliser le comportement du cerveau humain, les modèles à base de réseaux de neurones sont aujourd'hui des outils de calculs mathématiques sophistiqués utilisés dans des domaines très divers. Un modèle neuronal est constitué de plusieurs unités de calcul élémentaires (les neurones artificiels), fonctionnant en parallèle.

Chaque neurone reçoit des informations (qui peuvent être les entrées du modèle ou les sorties d'autres neurones), les traite, et envoie le résultat du traitement vers d'autres neurones. Le ou les neurones de sorties permet (tent) de reproduire le comportement du système à modéliser. Ces types de réseaux sont capables de représenter des systèmes très complexes. Il existe différents types de réseaux de neurones artificiels dont les Perceptrons Multicouches.

2.2.5.1 Perceptrons Multi-Couches

Les réseaux de neurones de type Perceptrons Multi-Couches (Multi Layer Perceptron - MLP) sont des réseaux à propagation avant, composés d'une ou plusieurs couches cachées et d'une couche de sortie. Chaque couche du réseau est composée de neurones artificiels.

La première couche cachée reçoit l'information provenant des entrées. L'information est traitée et transmise vers les couches suivantes jusqu'à la dernière. Les MLP sont connus comme étant des approximateurs universels et sont très utilisés dans des problèmes de régression non linéaire.

2.2.5.2 Réseaux de fonctions à base radiale

Les Réseaux de Fonctions à Base Radiale (Radial Basis Function - RBF) ont une architecture semblable à celle des MLP, mais ne comporte généralement que deux couches : une couche cachée et une couche de sortie. L'architecture d'un réseau RBF est représentée sur la figure 2.3.

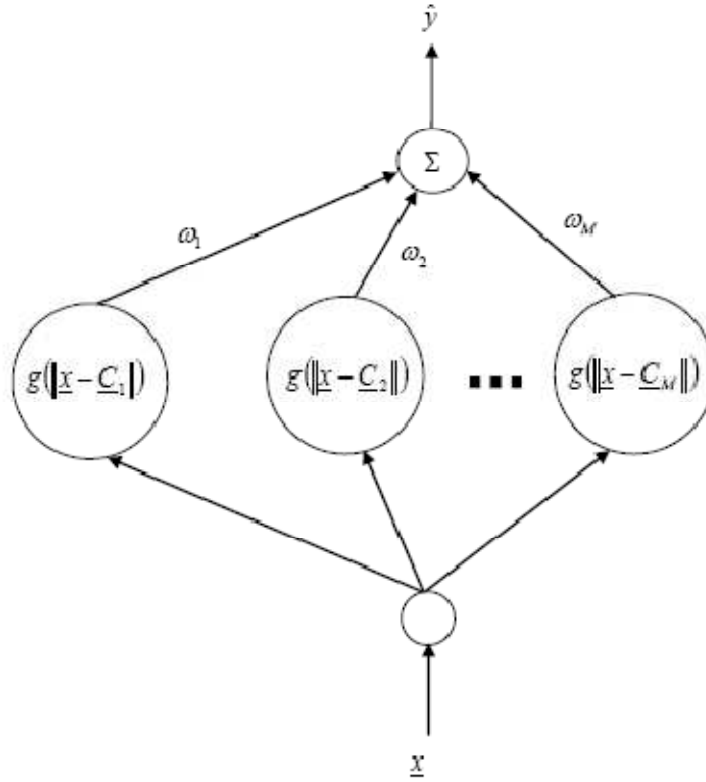


Fig. 2.3 Réseau de neurones RBF.

Un RBF permet de réaliser une décomposition sur une base de fonctions radiales :

$$\hat{y} = \sum_{i=1}^M w_i g \left(\left\| \underline{x} - \underline{C}_i \right\|_{\Sigma_i} \right) \quad (2.11)$$

où

\hat{y} est la sortie du réseau ;

g est la fonction d'activation à base radiale ;

w_i représente le poids de la connexion entre le neurone i ($i = 1, \dots, M$) et le neurone de sortie ;

\underline{x} est le vecteur d'entrée ;

\underline{C}_i est un vecteur désignant le centre de la fonction d'activation du neurone i ;

Σ_i est une matrice qui détermine l'étendue et l'orientation de la fonction d'activation dans l'espace des variables de régression ;

$\left\| \underline{x} - \underline{C}_i \right\|_{\Sigma_i}$ désigne la distance entre le point courant x et le centre C_i :

$$\left\| \underline{x} - \underline{C}_i \right\|_{\Sigma_i} = (\underline{x} - \underline{C}_i)^T \Sigma_i^{-1} (\underline{x} - \underline{C}_i)$$

Il existe plusieurs types de fonction à base radiale, mais la fonction gaussienne est la plus utilisée pour les RBF. Pour un neurone i de la couche cachée du réseau, la fonction d'activation pour une entrée \underline{x} est :

$$g(\|\underline{x} - \underline{C}_i\|_{\Sigma_i}) = g_i(\underline{x}) = \exp [(\underline{x} - \underline{C}_i)^T \sum_i^{-1} (\underline{x} - \underline{C}_i)]$$

Les réseaux RBF sont, comme les MLP, des approximateurs universels, et sont utilisés dans des applications telles que la classification, la reconnaissance de forme ou la prédiction. Ils sont faciles à estimer mais généralisent moins bien que les MLP et leur taille croît vite avec la dimension de l'entrée.

2.2.6 Machines à vecteurs supports

Les machines à vecteurs supports (Support Vector Machines - SVM) sont des outils d'apprentissage statistiques introduits par Vapnik en 1995. Les SVM sont généralement utilisées dans des problèmes de classification. L'approche permet de définir des surfaces complexes dans des espaces de dimensions importantes, avec des représentations très concises. Si les méthodes classiques d'apprentissage se basent sur la minimisation de l'erreur d'apprentissage (risque empirique), le principal avantage des SVM est la possibilité de déterminer une marge d'erreur (le risque structurel) valable pour la validation. Basés sur le principe des SVM, les Support Vector Regression - SVR permettent de traiter des problèmes de régression (linéaire ou non linéaire). Récemment plusieurs travaux ont été consacrés à l'utilisation des SVR pour l'approximation de fonctions et pour la prédiction. Cependant un des inconvénients de cette approche est le coût de calcul souvent élevé.

2.2.7 Multimodèles

La modélisation par multimodèles consiste à décomposer un système complexe en plusieurs sous-systèmes pour lesquels on peut élaborer des modèles locaux simples. Chaque modèle local tente alors de représenter le système dans un domaine de fonctionnement bien défini. Le modèle global du système est obtenu par une combinaison des modèles locaux. Les multimodèles apparaissent ainsi comme une généralisation de plusieurs types de modèles tels que les modèles flous, les modèles multi-experts, etc.

2.3 Modèles dynamiques non linéaires

Les connaissances a priori du système à modéliser, prises en compte dans la représentation (2.4), intègrent le choix des variables explicatives et la manière dont le bruit agit sur le système [5].

Le bruit peut agir sur la sortie, sur l'état ou sur la sortie et l'état à la fois. Ces informations sont spécifiées lors de la constitution du vecteur de régression qui permet ainsi de déterminer la classe du modèle non linéaire. On distingue différentes classes de modèles non linéaires. Le choix d'un de ces modèles est indépendant du choix de la structure de la fonction $F(\cdot)$. Ainsi,

un même modèle non linéaire peut par exemple être représenté avec une structure neuronale ou multimodèle. Le choix de la structure doit permettre d'obtenir un modèle parcimonieux (faible erreur avec un minimum de paramètres). Tout au long de ce travail, les performances des structures MLP et multimodèles sont étudiées à travers l'implantation de quelques modèles non linéaires qui sont présentons ici.

2.3.1 Modèle NFIR

Le modèle NFIR (Nonlinear Finite Impulse Response) permet de représenter des systèmes non linéaires dont la sortie ne dépend que des entrées à l'instant courant et/ou aux instants antérieurs.

$$y_s(t+1) = F_s(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k}+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2)) + e(t+1) \quad (2.12)$$

où :

$d_{u_k} \geq 1$ est le retard correspondant à l'entrée u_k ;

$n_{u_k} \geq 1$ est la mémoire sur l'entrée u_k , avec $k = 1, \dots, n_i$, n_i nombre d'entrées.

Le pas de prédiction h défini dans (2.1) peut être déterminé par :

$$h = \min(d_{u_k}), \quad k = 1, \dots, n_i$$

Le prédicteur réel du modèle NFIR est la fonction non linéaire $F(\cdot)$ de structure connue, paramétrée par un vecteur $\underline{\theta}$ qui, convenablement choisi, permettrait d'obtenir une erreur de prédiction ayant les mêmes caractéristiques statistiques que le bruit e :

$$\hat{y}(t+1) = F(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k}+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), \underline{\theta}) \quad (2.13)$$

Puisque le vecteur de régression du modèle NFIR ne dépend que des entrées externes, le modèle est non récurrent.

2.3.2 Modèle NARX

Le modèle NARX (Nonlinear AutoRegressive with eXogenous input) permet de représenter des systèmes dynamiques non linéaires dont la sortie dépend des entrées passées

et des sorties mesurées passées. Il permet également la représentation de systèmes dynamiques non linéaires avec « bruit d'état » ([1], [5]) :

$$x(t+1) = F_s(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k}+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), \dots, x(t-d_x+1), x(t-d_x), \dots, x(t-d_x-n_x+2)) + e(t+1) \quad (2.14)$$

$$y_s(t) = x(t)$$

où :

x désigne l'état du système et y_s sa sortie ;

$d_x \geq 1$ et $n_x \geq 1$ représentent le retard et l'ordre de la variable x respectivement.

Le prédicteur réel $F(\cdot)$ du model NARX est :

$$\hat{y}(t+1) = F(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k}+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), \dots, y_s(t-d_{y_s}+1), (t-d_{y_s}), \dots, y_s(t-d_{y_s}-n_{y_s}+2), \underline{\theta}) \quad (2.15)$$

où :

$d_{y_s} \geq 1$ est le retard correspondant à la sortie mesurée y_s ;

$n_{y_s} \geq 1$ est l'ordre de la partie autorégressive de y_s .

Le pas de prédiction h défini dans (2.1) peut être déterminé par :

$$h = \min(d_{u_k}, d_{y_s})$$

Le modèle NARX (appelé aussi modèle à Erreur d'Equation ou « Equation Error » en anglais) est un modèle non récurrent, tous les régresseurs étant des mesures faites sur le système (le vecteur de régression est constitué des entrées passées u_k et des sorties mesurées passées de y_s).

2.3.3 Modèle NOE

Le modèle NOE (Nonlinear Output Error) permet de représenter des systèmes dynamiques non linéaires par estimation de la sortie à un instant t à partir des entrées antérieures et des sorties antérieures estimées [5]. Il permet également la représentation de systèmes dynamiques non linéaires avec « bruit de sortie » :

$$\begin{aligned}
x(t+1) = F_s(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k} \\
+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), \\
u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), x(t-d_x+1), \dots, x(t \\
-d_x), \dots, x(t-d_x-n_x+2)) \quad (2.16)
\end{aligned}$$

$$y_s(t) = x(t) + e(t)$$

Le prédicteur réel $F(\cdot)$ du model NOE est :

$$\begin{aligned}
\hat{y}(t+1) = F(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k} \\
+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), \\
u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), \\
\hat{y}(t-d_{\hat{y}}+1), \hat{y}(t-d_{\hat{y}}), \dots, \hat{y}(t-d_{\hat{y}}-n_{\hat{y}}+2), \underline{\theta}) \quad (2.17)
\end{aligned}$$

où $d_{\hat{y}}$ et $n_{\hat{y}}$ sont le retard et l'ordre associé à \hat{y} .

Le pas de prédiction est :

$$h = \min(d_{u_k}, d_{\hat{y}})$$

Le vecteur de régression est constitué des entrées passées u_k et des sorties passées estimées \hat{y} , ce qui fait du modèle NOE un modèle récurrent. C'est un modèle qui convient bien à la simulation puisqu'il n'utilise pas les sorties réelles du système.

2.3.4 Modèle NARMAX

NARMAX désigne Nonlinear AutoRegressive Moving Average with eXogenous input.

Ce modèle permet de représenter un système dynamique non linéaire par estimation de la sortie à un instant t à partir des entrées passées, des sorties mesurées passées et des erreurs de prédiction passées. Le modèle NARMAX est aussi utilisé pour représenter un système dynamique non linéaire en présence de « bruit d'état et bruit de sortie » ([5], [1]):

$$\begin{aligned}
y_s(t+1) = F_s(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k} \\
+1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_i}(t-d_{u_{n_i}}+1), \\
u_{n_i}(t-d_{u_{n_i}}), \dots, u_{n_i}(t-d_{u_{n_i}}-n_{u_{n_i}}+2), \dots, y_s(t-d_{y_s}+1), y_s(t \\
-d_{y_s}), \dots, y_s(t-d_{y_s}-n_{y_s}+2), \dots, e(t-d_e+1), e(t-d_e), \dots, e(t-d_e \\
-n_e+2)) + e(t+1) \quad (2.18)
\end{aligned}$$

$d_e \geq 1$ et $n_e \geq 1$ sont le retard et la mémoire correspondant à la variable aléatoire e .

Les valeurs $e(t-d_e+1), \dots, e(t-d_e), \dots, e(t-d_e-n_e+2)$ n'étant pas mesurables, ce prédicteur est irréalisable. Cependant en supposant que ces valeurs peuvent être estimées par celles de l'erreur de prédiction $\varepsilon(t) = y_s(t) - \hat{y}(t)$, on peut obtenir le prédicteur réel

du modèle NARMAX :

$$\begin{aligned} \hat{y}(t+1) = F_s(u_1(t-d_{u_1}+1), u_1(t-d_{u_1}), \dots, u_1(t-d_{u_1}-n_{u_1}+2), \dots, u_k(t-d_{u_k} \\ +1), u_k(t-d_{u_k}), \dots, u_k(t-d_{u_k}-n_{u_k}+2), \dots, u_{n_k}(t-d_{u_{n_k}}+1), u_{n_k}(t \\ -d_{u_{n_k}}), \dots, u_{n_k}(t-d_{u_{n_k}}-n_{u_{n_k}}+2), \dots, y_s(t-d_{y_s}+1), y_s(t \\ -d_{y_s}), \dots, y_s(t-d_{y_s}-n_{y_s}+2), \dots, \varepsilon(t-d_\varepsilon+1), \varepsilon(t-d_\varepsilon), \dots, \varepsilon(t-d_\varepsilon \\ -n_\varepsilon+2), \underline{\theta}) \end{aligned} \quad (2.19)$$

où $d_\varepsilon \geq 1$ et $n_\varepsilon \geq 1$ représentent le retard et la mémoire correspondant à l'erreur de prédiction ε .

Le pas de prédiction est :

$$h = \min(d_{u_k}, d_{y_s}, d_\varepsilon)$$

Le vecteur de régression du modèle NARMAX est constitué des entrées passées uk , des sorties mesurées passées de y_s et des erreurs de prédiction ε . C'est donc un modèle récurrent.

2.4 Conclusion

Dans ce chapitre, nous avons présenté les différents aspects liés à la représentation des systèmes dynamiques non linéaires. Différents modèles non linéaires ainsi que différentes structures de représentations permettant leur implantation ont également été présentés.

On a représenté uniquement quelques structures des modèles paramétriques les plus fréquemment utilisées.

Une présentation compacte de la méthode des moindres carrés est donnée dans le chapitre suivant.

CHAPITRE 3 : Exemples d'application

3.1 Introduction

La détermination en ligne des paramètres d'un processus est un élément clé dans tous les domaines de l'automatique. Ce chapitre présente quelques exemples d'estimation des paramètres en temps réel. Les principaux éléments d'identification d'un système sont la structure du modèle, conception de l'expérience, l'estimation des paramètres et la validation. Puisque l'identification du système est exécutée automatiquement dans les systèmes d'adaptation, il est essentiel d'avoir une bonne compréhension de tous les aspects du problème. La sélection de la structure du modèle et le paramétrage sont des questions fondamentales. Les problèmes d'identification sont considérablement facilités si les modèles sont linéaires par rapport aux paramètres. La conception de l'expérience est cruciale pour l'identification du système avec succès. Dans les problèmes de contrôle cela revient à la sélection du signal d'entrée. Choisir un signal d'entrée requiert une certaine connaissance du processus et l'utilisation prévue du modèle. Dans les systèmes adaptatifs il y a une complication supplémentaire, car le signal d'entrée est généré par la rétroaction. Dans certains cas, cela ne permet pas de déterminer les paramètres de façon unique, une situation qui a de lourdes conséquences.

3.2 Moindres carrés

Karl Friedrich Gauss a formulé le principe des moindres carrés à la fin du XVIII^e siècle et l'a utilisé pour déterminer les orbites des planètes et des astéroïdes. Gauss a déclaré que, selon ce principe ([3], [4]), les paramètres inconnus d'un modèle mathématique doit être choisis de manière à ce que la somme des carrés des différences entre les valeurs effectivement observées et les valeurs calculées, multiplié par le nombre qui mesurent le degré de précision, est un minimum.

La méthode des moindres carrés peut être appliquée à une grande variété de problèmes. Elle est particulièrement simple pour un modèle mathématique qui peut être décrit sous la forme :

$$y(i) = \varphi_1(i)\theta_1^0 + \varphi_2(i)\theta_2^0 + \dots + \varphi_n(i)\theta_n^0 = \varphi^T(i)\theta^0 \quad (3.1)$$

où y est la variable observée, $\theta_1^0, \theta_2^0, \dots, \theta_n^0$ sont les paramètres du modèle, $\varphi_1, \varphi_2, \dots, \varphi_n$ sont des fonctions connues qui peuvent dépendre d'autres variables connues. Les vecteurs :

$$\begin{aligned} \varphi^T(i) &= (\varphi_1(i) \ \varphi_2(i) \ \dots \ \varphi_n(i)) \\ \theta &= (\theta_1^0 \ \theta_2^0 \ \dots \ \theta_n^0)^T \end{aligned}$$

ont également été introduites. Le modèle est indexé par la variable i , qui représente souvent le temps. On supposera d'abord que l'ensemble des indices est un ensemble discret. Les variables φ_i sont appelées les *variables de régression* ou les *régresseurs*, et le modèle de l'équation. (3.1) est aussi appelé un *modèle de régression*. Les paires d'observations et régresseurs $\{(y(i), \varphi_n(i)), i = 1, 2, \dots, t\}$ sont obtenus à partir d'une expérience. Le problème est de déterminer les paramètres de telle sorte que les résultats calculés à partir du modèle de l'équation. (3.1) conviennent aussi étroitement que possible avec les variables mesurées $y(i)$

dans le sens des moindres carrés. Autrement dit, le paramètre θ doit être choisi pour minimiser la fonction de perte des moindres carrés [3].

$$V(\theta, t) = \frac{1}{2} \sum_{i=1}^t (y(i) - \varphi^T(i) \theta)^2 \quad (3.2)$$

Puisque la variable mesurée y est linéaire dans les paramètres θ^0 , et le critère des moindres carrés est quadratique, le problème (3.2) admet une solution analytique. Introduire les notations

$$\begin{aligned} Y(t) &= (y(1) \ y(2) \ \dots \ y(t))^T \\ E(t) &= (\varepsilon(1) \ \varepsilon(2) \ \dots \ \varepsilon(t))^T \end{aligned}$$

$$\Phi(t) = \begin{pmatrix} \varphi^T(1) \\ \vdots \\ \varphi^T(t) \end{pmatrix}$$

$$P(t) = (\Phi^T(t) \Phi(t))^{-1} = \left(\sum_{i=1}^t \varphi(i) \varphi^T(i) \right)^{-1} \quad (3.3)$$

où les résidus $\varepsilon(i)$ sont définis par

$$\varepsilon(i) = y(i) - \hat{y}(i) = y(i) - \varphi^T(i) \theta$$

Avec ces notations, la fonction de perte (3.2) peut être écrite comme

$$V(\theta, t) = \frac{1}{2} \sum_{i=1}^t \varepsilon^2(i) = \frac{1}{2} E^T E = \frac{1}{2} \|E\|^2$$

où E peut être écrit comme

$$E = Y - \hat{Y} = Y - \Phi \theta \quad (3.4)$$

La solution au problème des moindres carrés est donnée par le théorème suivant.

THÉORÈME 3.1 « L'estimation de moindres carrés »

La fonction de l'équation. (3.2) est minimal pour les paramètres $\hat{\theta}$ tels que

$$\Phi^T \Phi \hat{\theta} = \Phi^T Y \quad (3.5)$$

Si la matrice $\Phi^T \Phi$ est non-singulière, le minimum est unique et donnée par

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y \quad (3.6)$$

Preuve : La fonction de perte de l'équation. (3.2) peut être écrite comme

$$\begin{aligned}
2V(\theta, t) &= E^T E = (Y - \Phi\theta)^T (Y - \Phi\theta) \\
&= Y^T Y - Y^T \Phi\theta - \theta^T \Phi^T Y + \theta^T \Phi^T \Phi\theta
\end{aligned} \tag{3.7}$$

Puisque la matrice $\Phi^T \Phi$ est toujours définie positive ou nulle, la fonction V a un minimum. La fonction de perte est quadratique en θ . Le minimum se trouve à plusieurs façons. Une de ces façons est de déterminer le gradient de l'équation (3.7) par rapport à θ . Le gradient est nul lorsque l'équation (3.5) est satisfaite. Une autre façon de trouver le minimum est de compléter le carré. Nous obtenons

$$\begin{aligned}
2V(\theta, t) &= Y^T Y - Y^T \Phi\theta - \theta^T \Phi^T Y + \theta^T \Phi^T \Phi\theta + Y^T \Phi (\Phi^T \Phi)^{-1} \Phi^T Y - Y^T \Phi (\Phi^T \Phi)^{-1} \Phi^T Y \\
&= Y^T (I - \Phi (\Phi^T \Phi)^{-1} \Phi^T) Y \\
&\quad + (\theta - (\Phi^T \Phi)^{-1} \Phi^T Y)^T \Phi^T \Phi (\theta - (\Phi^T \Phi)^{-1} \Phi^T Y)
\end{aligned} \tag{3.8}$$

Le premier terme du côté droit est indépendant de θ . Le second terme est toujours positif. Le minimum est obtenu pour

$$\theta = \hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y$$

et le théorème est démontré.

Remarque 1. L'équation (3.5) est appelée *l'équation normale*. L'équation (3.6) peut être écrite comme

$$\hat{\theta}(t) = \left(\sum_{i=1}^t \varphi(i) \varphi^T(i) \right)^{-1} \left(\sum_{i=1}^t \varphi(i) y(i) \right) = P(t) \left(\sum_{i=1}^t \varphi(i) y(i) \right) \tag{3.9}$$

Remarque 2. La condition pour que la matrice $\Phi^T \Phi$ est inversible est appelé la condition *d'excitation*.

3.3 Moindres carrés récurrents

Dans beaucoup des domaines de contrôle les observations sont obtenues de façon séquentielle en temps réel. Il est alors souhaitable de faire les calculs de manière récursive pour gagner du temps de calcul. Le calcul de l'estimation des moindres carrés peut être organisé de telle sorte que les résultats obtenus au temps $t - 1$ peuvent être utilisé pour obtenir les estimations à l'instant t . Dans ce cas la solution de l'équation (3.6) au problème des moindres carrés sera réécrite sous une forme récursive. ([3], [4]).

Soit $\hat{\theta}(t - 1)$ désigner les estimations de moindres carrés basée sur les mesures $t - 1$. Supposons que la matrice $(\Phi^T \Phi)^{-1}$ est non singulière pour tout instant t . Il résulte de la définition de $P(t)$ dans l'équation (3.3) que :

$$\begin{aligned}
P^{-1}(t) &= \Phi^T(t)\Phi(t) = \sum_{i=1}^t \varphi(i)\varphi^T(i) \\
&= \sum_{i=1}^{t-1} \varphi(i)\varphi^T(i) + \varphi(t)\varphi^T(t) \\
&= P^{-1}(t-1) + \varphi(t)\varphi^T(t)
\end{aligned} \tag{3.10}$$

L'estimation des moindres carrés $\hat{\theta}(t)$ est donnée par l'équation (3.9) :

$$\hat{\theta}(t) = P(t) \left(\sum_{i=1}^t \varphi(i)y(i) \right) = P(t) \left(\sum_{i=1}^{t-1} \varphi(i)y(i) + \varphi(t)y(t) \right)$$

De les Equations (3.9) et (3.10) Il résulte que

$$\sum_{i=1}^{t-1} \varphi(i)y(i) = P^{-1}(t-1)\hat{\theta}(t-1) = P^{-1}(t)\hat{\theta}(t-1) - \varphi(t)\varphi^T(t)\hat{\theta}(t-1)$$

L'estimation à l'instant t peut être maintenant écrite comme

$$\begin{aligned}
\hat{\theta}(t) &= \hat{\theta}(t-1) - P(t)\varphi(t)\varphi^T(t)\hat{\theta}(t-1) + P(t)\varphi(t)y(t) \\
&= \hat{\theta}(t-1) - P(t)\varphi(t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \\
&= \hat{\theta}(t-1) + K(t)\varepsilon(t)
\end{aligned}$$

où

$$\begin{aligned}
K(t) &= P(t)\varphi(t) \\
\varepsilon(t) &= y(t) - \varphi^T(t)\hat{\theta}(t-1)
\end{aligned}$$

Le résiduelle $e(t)$ peut être interprété comme l'erreur dans la prévision du signal $y(t)$ une longueur d'avance sur la base de l'estimation $\hat{\theta}(t-1)$.

THÉORÈME 3.2 « estimation des moindres carrés récursif (RLS) »

Supposons que la matrice $\Phi(t)$ est de rang plein, c'est-à-dire $\Phi(t)\Phi^T(t)$ est non-singulière, pour tout $t \geq t_0$. On donne $\hat{\theta}(t_0)$ et $P(t_0) = (\Phi(t_0)\Phi^T(t_0))^{-1}$, l'estimation de moindres carrés $\hat{\theta}$ vérifie alors les équations récursives suivantes :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(y(t) - \varphi^T(t)\hat{\theta}(t-1)) \tag{3.11}$$

$$K(t) = P(t)\varphi(t) = P(t-1)\varphi(t)(I + \varphi^T(t)P(t-1)\varphi(t))^{-1} \tag{3.12}$$

$$\begin{aligned}
P(t) &= P(t-1) - P(t-1)\varphi(t)(I + \varphi^T(t)P(t-1)\varphi(t))^{-1}\varphi^T(t)P(t-1) \\
&= (I - K(t)\varphi^T(t))P(t-1)
\end{aligned} \tag{3.13}$$

3.4 Moindres carrés récurrents avec facteur d'oubli

Dans le modèle des moindres carrés (3.1) les paramètres θ_i^0 sont supposés constants. Dans plusieurs problèmes d'adaptation, il est intéressant d'examiner la situation dans laquelle les paramètres sont variables dans le temps [4]. Deux cas peuvent être couverts d'extensions simples de la méthode des moindres carrés. Dans un tel cas les paramètres sont supposés changer brusquement, mais rarement, dans l'autre cas paramètres changent continuellement, mais lentement. Le cas de variations brutales des paramètres peuvent être couverts par la réinitialisation. La matrice P dans l'algorithme des moindres carrés (théorème 3.3) est ensuite périodiquement remis à un αI , où α est un grand nombre. Cela implique que le gain $K(t)$ dans l'estimateur devient important et l'estimation peut être mise à jour avec une plus grande étape. Une version plus sophistiquée consiste à exécuter n estimateurs en parallèle, qui sont remis à zéro de manière séquentielle. L'estimation est alors choisie en utilisant une logique de décision. Une approche pragmatique est tout simplement de remplacer le critère des moindres carrés de l'équation (3.2) avec

$$V(\theta, t) = \frac{1}{2} \sum_{i=1}^t \lambda^{t-i} (y(i) - \varphi^T(i)\theta)^2 \tag{3.14}$$

où λ est un paramètre tel que $0 < \lambda \leq 1$. Le paramètre λ est appelé le facteur d'oubli.

La fonction de perte de l'équation (3.14) implique qu'une pondération variable dans le temps des données est introduite. Les données les plus récentes est donnée unité de poids, mais les données qui est n unités de temps anciens est pondérée par λ^n . La méthode est donc appelé oubli exponentiel ou actualisation exponentielle. En répétant les calculs effectués pour le théorème 3.2 pour la fonction de perte de l'équation (3.14), on obtient les résultats suivants :

THÉORÈME 3.3 : « moindres carrés récurrents avec facteur d'oubli »

Supposons que la matrice $\Phi(t)$ est de rang plein pour $t \geq t_0$. Le vecteur des paramètres θ , qui minimise l'équation (3.14), est donné de manière récursive

$$\begin{aligned}
\hat{\theta}(t) &= \hat{\theta}(t-1) - K(t) \left(y(t) - \varphi^T(t)\hat{\theta}(t-1) \right) \\
K(t) &= P(t)\varphi(t) = P(t-1)\varphi(t)(\lambda + \varphi^T(t)P(t-1)\varphi(t))^{-1} \\
P(t) &= (I - K(t)\varphi^T(t))P(t-1)/\lambda
\end{aligned} \tag{3.15}$$

Un inconvénient moindres carrés récurrents avec facteur d'oubli exponentiel est que les données sont réduits même si $P(t)\varphi(t) = 0$. Cette condition implique que $y(t)$ ne contient

pas de nouvelles informations sur le paramètre θ . Dans ce cas, il résulte de l'Eq. (3.15) que la matrice augmente de façon exponentielle avec λ .

3.5 Moindres carrés orthogonaux:

Soit le modèle:

$$y(t) = \sum_{i=1}^M \varphi_i(t) \theta_i + \varepsilon(t) \quad (3.16)$$

où $t = 1, \dots, N$ et N est le nombre de données, M est le nombre de paramètres estimés, les $\varphi_i(t)$ sont les variables explicatives, $\varepsilon(t)$ est une erreur de modélisation qui n'est pas corrélée avec les régresseurs et les θ_i sont les paramètres à estimer [10].

L'objectif donc est d'estimer les paramètres θ_i pour $i = 1, \dots, M$

L'algorithme des moindres carrés orthogonaux transforme l'ensemble des régresseurs φ_i dans des vecteurs de base orthogonaux

En MCO, l'équation (3.16) peut être transformée comme suite :

$$y(t) = \sum_{i=1}^M w_i(t) g_i + \varepsilon(t) \quad (3.17)$$

où les g_i sont des coefficients constants et $w_i(t)$ sont les données orthogonales construit comme :

$$\sum_{i=1}^N w_i(t) w_j(t) = 0 \quad (3.18)$$

où $i \neq j$.

Les polynômes orthogonaux peuvent être calculés en appliquant les procédures de Gram-Schmidt comme décrit par Chen et al. Toutefois, Korenberg et al ont défini une procédure plus simple de construire une famille de polynômes orthogonaux.

on à

$$w_0(t) = \varphi_0(t) = 1$$

$$w_i(t) = \varphi_i(t) - \sum_{j=1}^{i-1} \alpha_{ji} w_j(t) \quad (3.19)$$

où $i = 1, \dots, M$ et

$$\alpha_{ji} = \frac{\sum_{t=1}^N \varphi_i(t) w_j(t)}{\sum_{t=1}^N w_j^2(t)} \quad (3.20)$$

où $j = 1, \dots, i - 1$.

En multipliant l'équation (3.19) par $w_j(t)$ et en prenant la valeur attendue pour les deux côtés de l'équation on aura :

$$E[w_j(t)w_i(t)] = E[w_j(t)\varphi_i(t)] - \alpha_{ji}E[w_j^2(t)] \quad (3.21)$$

où $E[.]$ est l'opérateur prévu.

En appliquant les propriétés orthogonales comme définie dans les équations (3.18), (3.21) on trouve :

$$0 = E[w_j(t)\varphi_i(t)] - \alpha_{ji}E[w_j^2(t)] \quad (3.22)$$

Le coefficient constant de α_{ji} peut alors être calculée comme :

$$\alpha_{ji} = \frac{E[w_j(t)\varphi_i(t)]}{E[w_j^2(t)]} \quad (3.23)$$

La prochaine étape est d'obtenir le coefficient constant pour les paramètres estimés θ_i . L'équation (3.17) est multipliée par $w_j(t)$, ce qui donne l'équation suivante :

$$w_j(t)y(t) = \sum_{i=1}^M w_j(t)g_i(t)w_i(t) + w_j(t)\varepsilon(t) \quad (3.24)$$

En utilisant la propriété orthogonale [équation (3.18)] à l'équation (3.24) et en prenant la valeur attendue, on aura :

$$E[w_j(t)y(t)] - g_j E[w_j^2(t)] \quad (3.25)$$

Les coefficients d'estimations des paramètres peuvent être calculés comme :

$$g_i = \frac{E[y(t)w_i(t)]}{E[w_i^2(t)]} \quad (3.26)$$

où $i = 0, \dots, M$.

Les paramètres estimés peuvent être calculée en utilisant les équations suivantes :

$$\theta_M = g_M \quad (3.27)$$

et

$$\theta_i = g_i - \sum_{j=i+1}^M \alpha_{ij} \theta_j \quad (3.28)$$

où $i = M - 1, M - 2, \dots, 1$.

La dérivation simple du l'ERR (error-reduction ratio) est résumée dans [7] et est calculée à partir de l'équation suivante:

$$ERR_i = \frac{g_i^2 E[w_i^2(t)]}{E[y^2(t)]} \quad (3.29)$$

Pour inclure tous les termes dans le vecteur d'entrée, on utilise l'équation suivante :

$$1 - \sum_{i=1}^{n_r} ERR_i < \rho_t \quad (3.30)$$

où ρ_t est la tolérance souhaitée.

L'algorithme de moindres carrés orthogonaux peut être résumé comme suit:

- (i) Définir les ordres de l'entrée et de la sortie (nu et ny).
- (ii) Définir la valeur ρ_t , le critère pour arrêter la régression.
- (iii) former les régresseurs w_i comme il est indiqué aux équations (3.19) et (3.20).
- (iv) calculer le coefficient g_i par l'équation (3.27).
- (v) l'ERR est calculée à partir de l'équation (3.29).
- (vi) Le critère pour choisir le nombre de termes à inclure dans le modèle final est calculé à partir de l'équation (3.30). Pour $1 - \sum_{i=1}^{n_r} ERR_i < \rho_t$ les relations (3.18) à (3.21) sont répétées avec les termes qui ont été sélectionnés.
- (vii) Les estimateurs des paramètres du modèle peut alors être calculée par la substitution en arrière comme indiqué dans l'équation (3.27) et (3.28).

3.6 Exemples d'application :

Dans cette section on va appliquer la méthode de moindre carré récursif pour l'identification des paramètres de quelques systèmes non linéaires.

Système 1 : $y(t) + ay(t - 1) = b * \cos(u(t - 1)) + e(t)$

où

$$\hat{\theta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$$

$$\Phi = (-y(t-1) \cos(u(t-1)))$$

$$a = -0.8, \quad b = 0.5 \quad e(t) \text{ est un bruit blanc à 0 moyenne}$$

$[u(t), y(t)]_{k=1}^N$. où $N = 1000$ est le nombre maximal de points de simulation.

Les conditions initiales sont :

$$P(0) = 100 \times I$$

$$\hat{\theta}(0) = 0$$

Les figures ci dessous montrent les résultats de simulation.

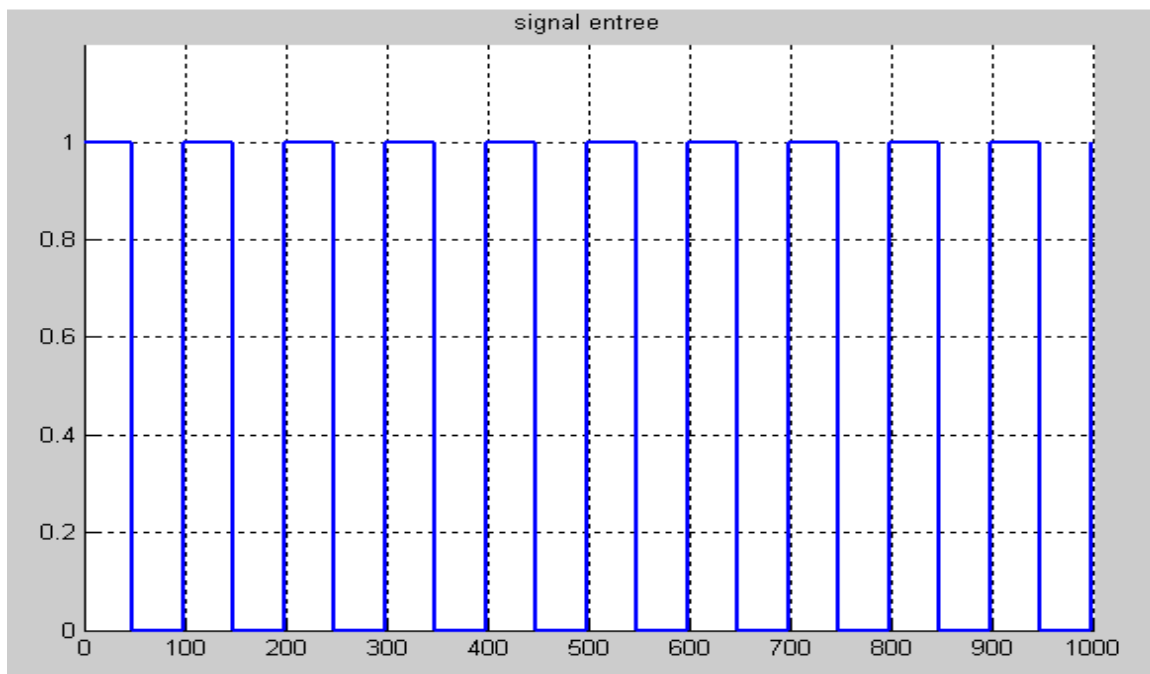


Fig. 3.1 le signal d'entrée

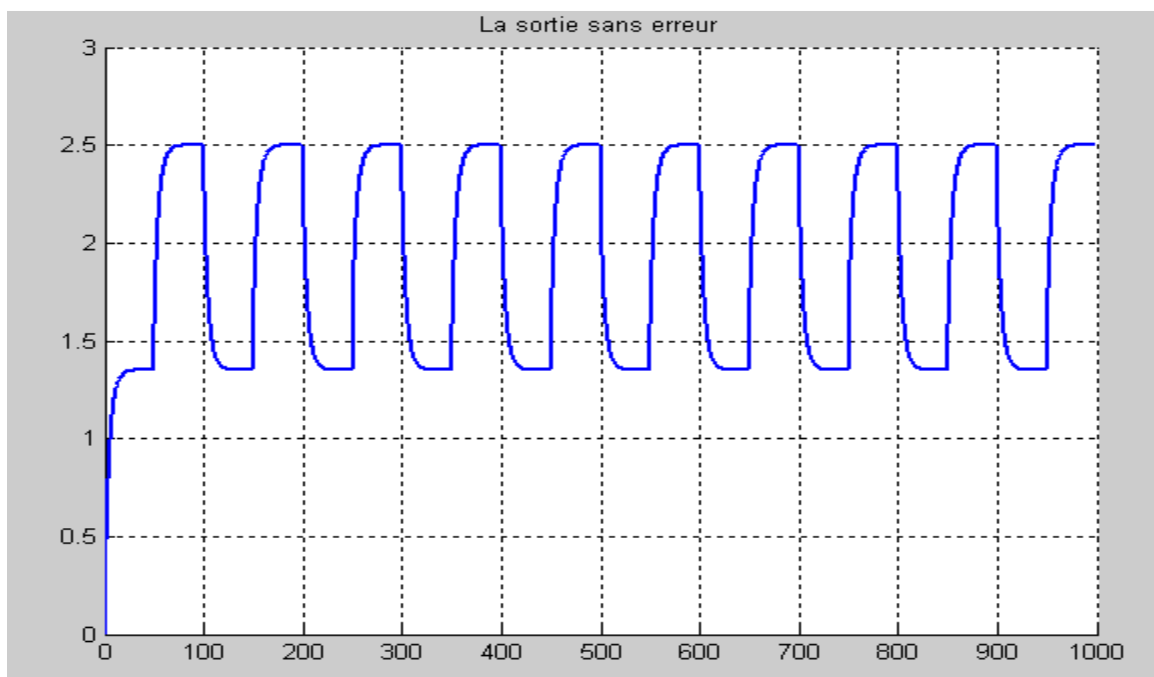


Fig. 3.2 la sortie du système1 sans bruit

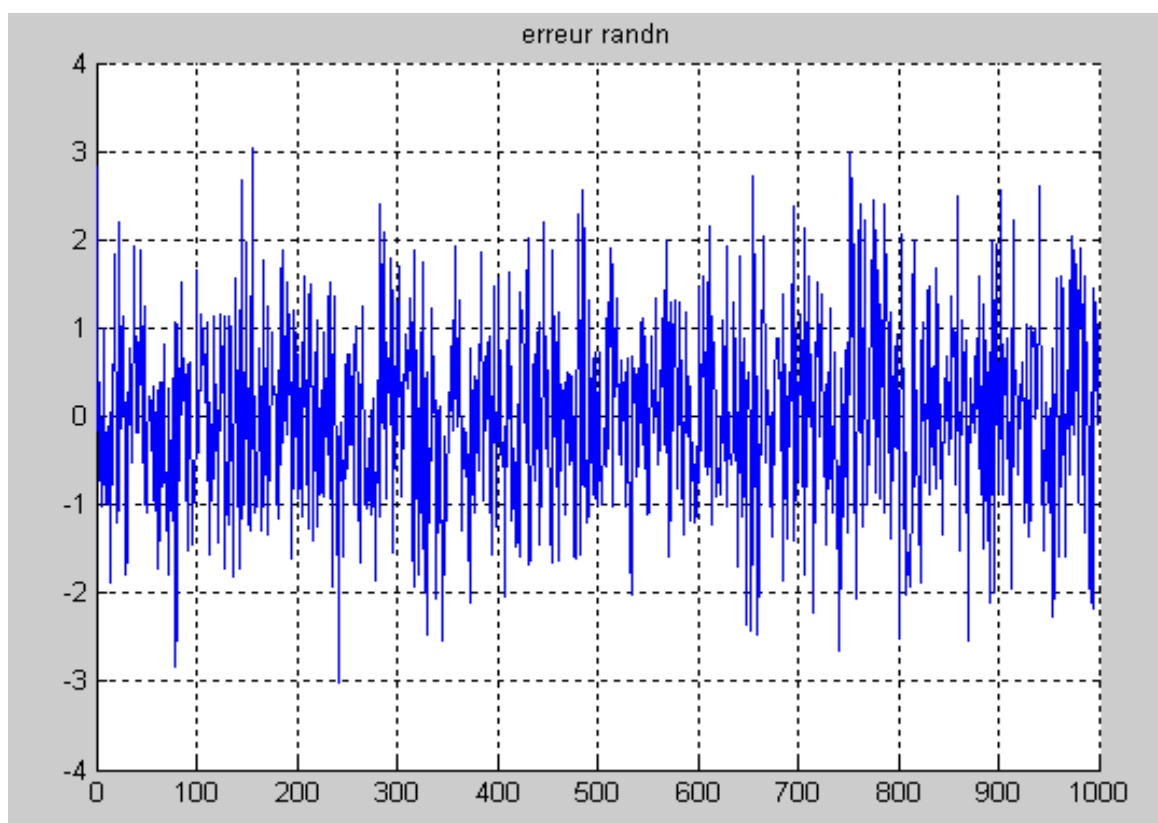


Fig. 3.3 *bruit gaussienne*

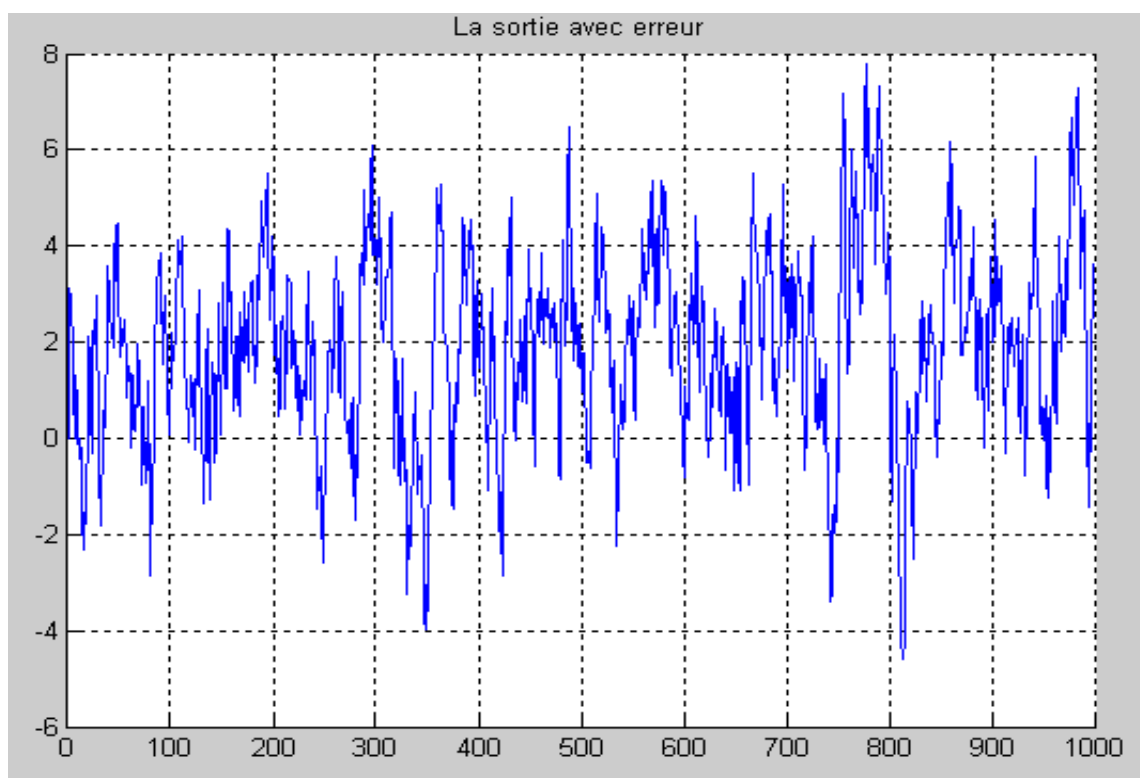


Fig. 3.4 *la sortie du système 1 avec bruit*

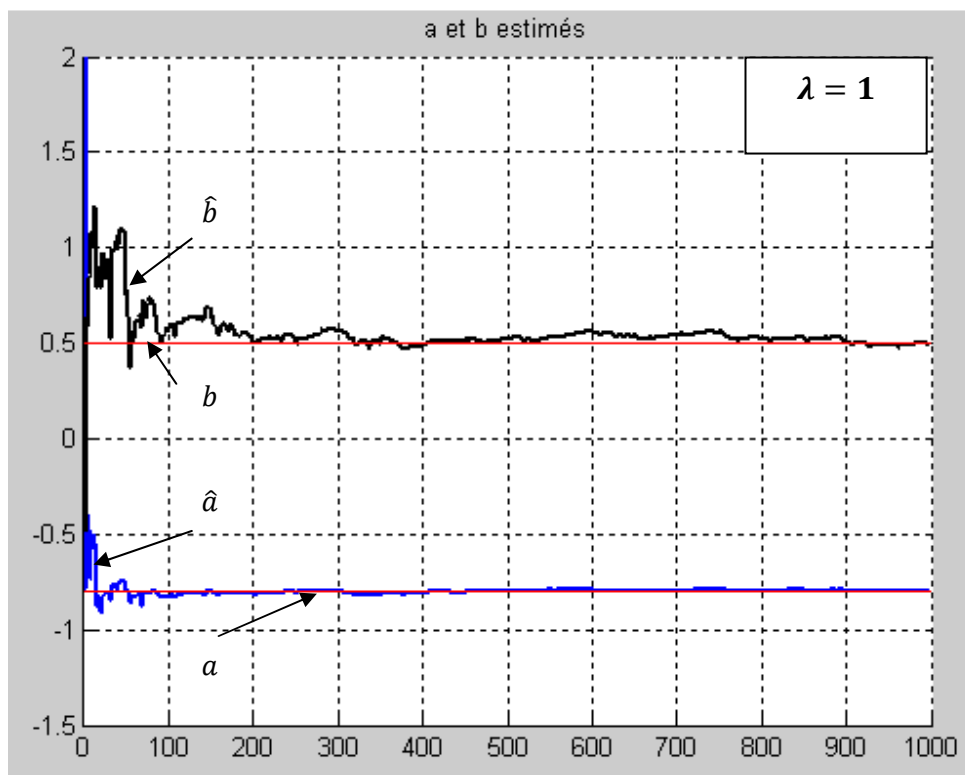


Fig. 3.5 les paramètres réels/estimés pour $\lambda = 1$

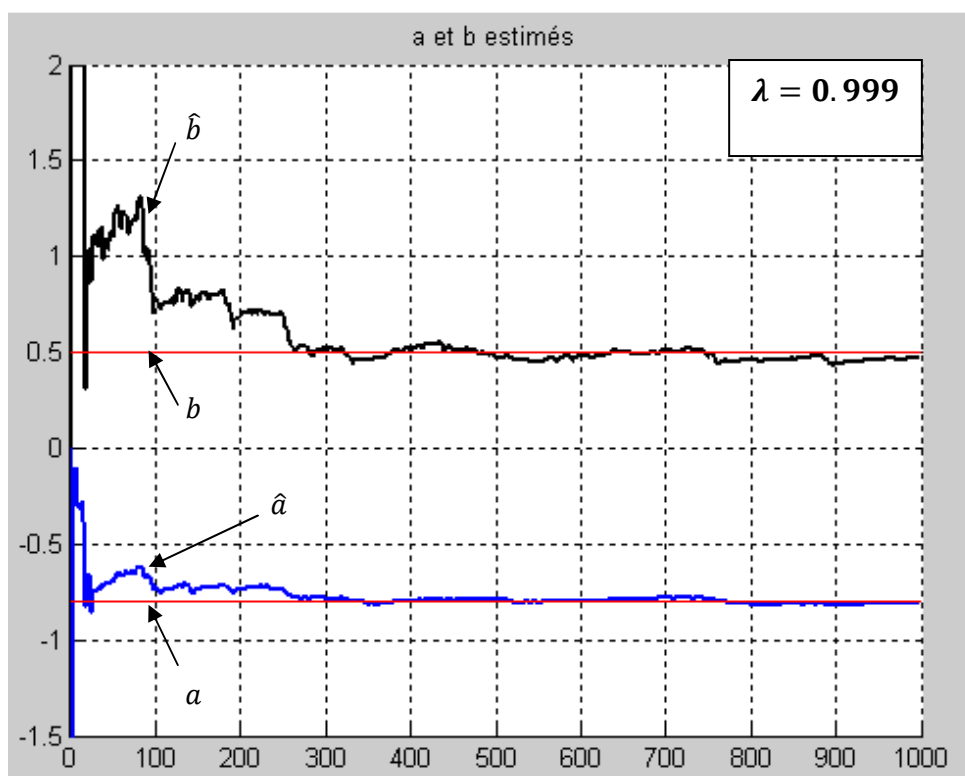


Fig. 3.6 les paramètres réels/estimés pour $\lambda = 0.999$

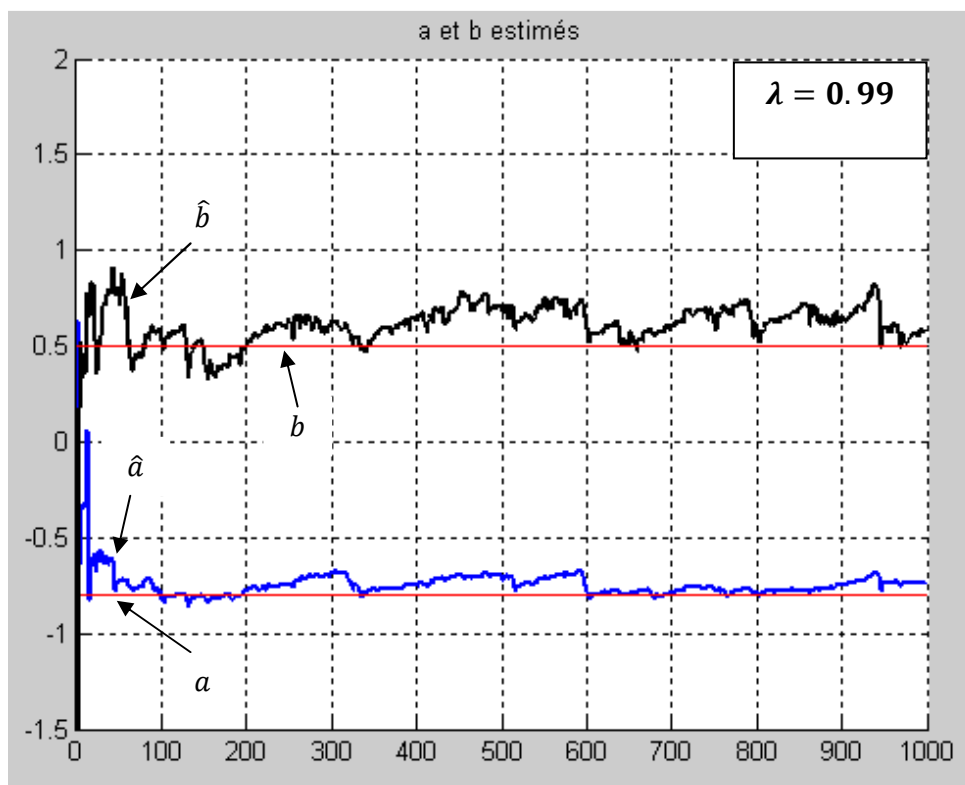


Fig. 3.7 les paramètres réels/estimées pour $\lambda = 0.99$

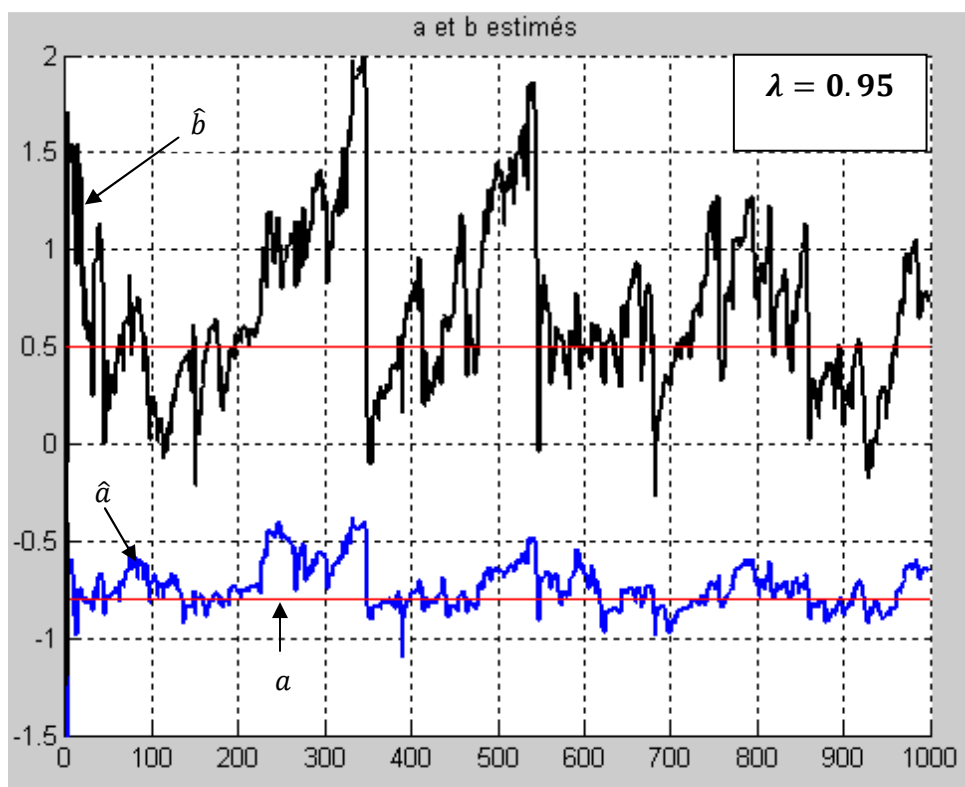


Fig. 3.8 les paramètres réels/estimées pour $\lambda = 0.95$

Interprétation des résultats:

Dans cet exemple nous avons utilisé la méthode de Moindres carrés récursifs (RLS) avec le facteur d'oubli, (voir section 3.4).

Le tableau 3.1 présente les résultats obtenus pour différentes valeurs λ . Les Fig. 3.5, 3.6, 3.7, 3.8 donnent les trajectoires d'évolution des paramètres réels et des paramètres estimés pour les différentes valeurs de facteur d'oubli.

Tableau 3.1 : les résultats de simulation pour les valeurs réelles et les valeurs estimées

paramètres	a	b
réels	-0.8	0.5
Estimées ($\lambda = 1.000$)	-0.8154	0.4741
Estimées ($\lambda = 0.999$)	-0.8110	0.4728
Estimées ($\lambda = 0.990$)	-0.7254	0.5883
Estimées ($\lambda = 0.950$)	-0.6506	0.7267

Système 2 : système à 5 paramètres(NARMAX)

$$y(t) = 0.5y(t-1) + u(t-2) + 0.5e(t-1) + 0.1u^2(t-1) + 0.2u(t-1) * e(t-2) + e(t)$$

$$a = 0.5;$$

$$b = 1;$$

$$c = 0.5;$$

$$d = 0.1;$$

$$f = 0.2;$$

$$\theta^T = [a \ b \ c \ d \ f], \hat{\theta}^T = [\hat{a} \ \hat{b} \ \hat{c} \ \hat{d} \ \hat{f}]$$

Les conditions initiales sont :

$$P(0) = 100 \times I$$

$$\hat{\theta}(0) = 0$$

Les figures ci dessous montrent les résultats de simulation

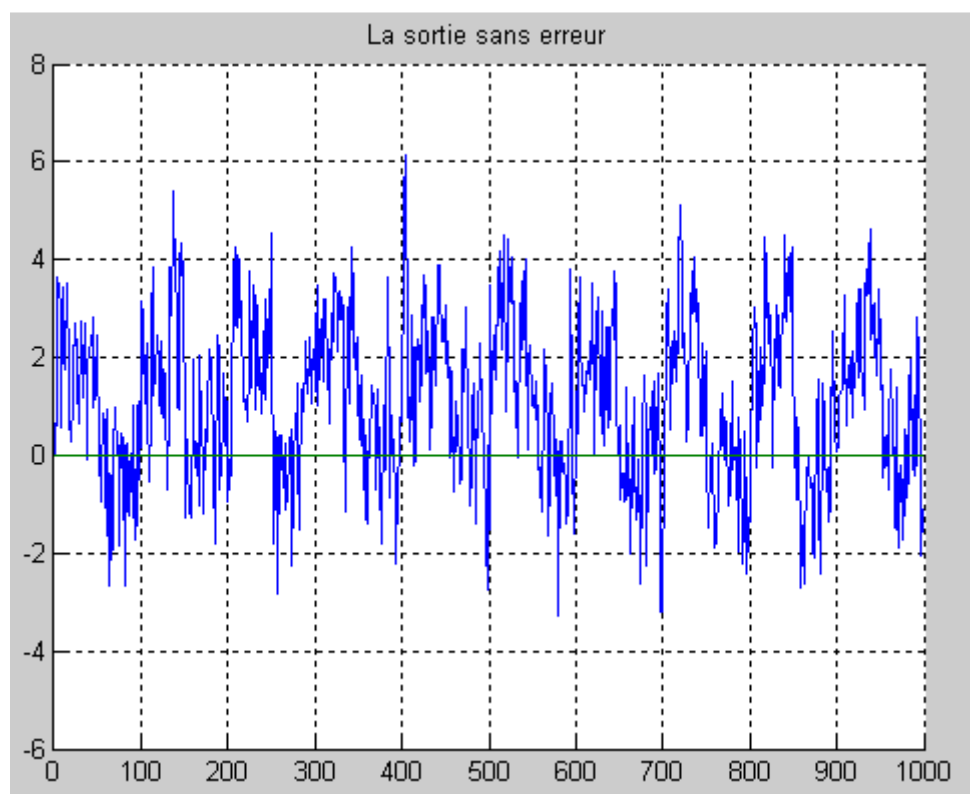


Fig. 3.9 la sortie du système2 sans bruit

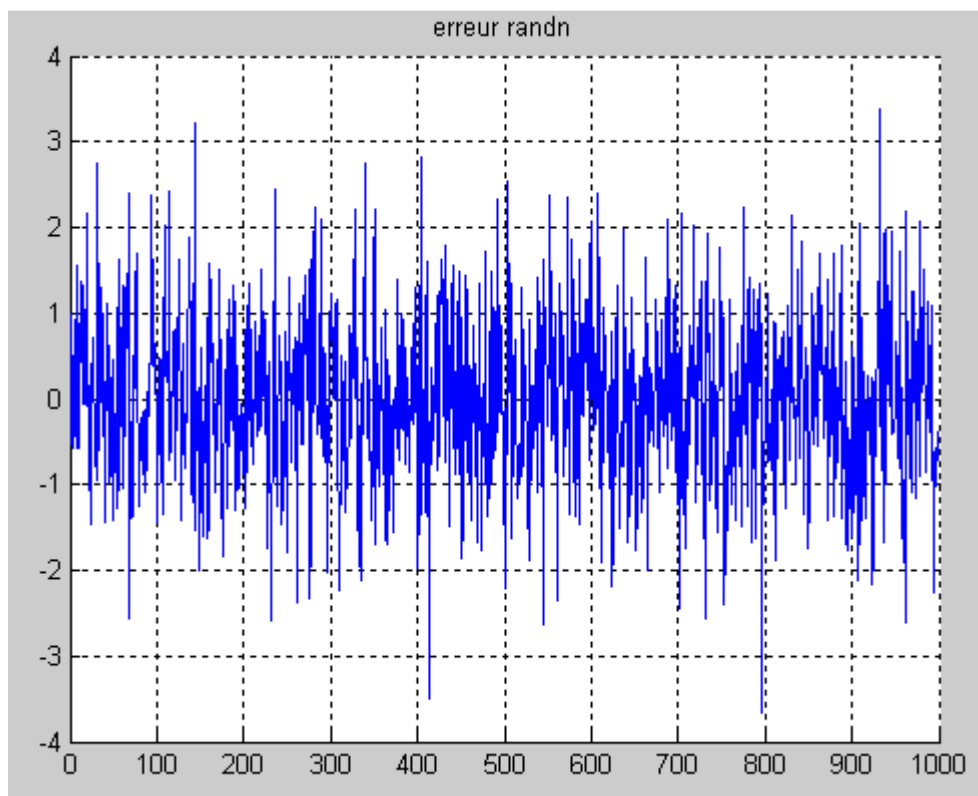


Fig. 3.10 bruit du système2

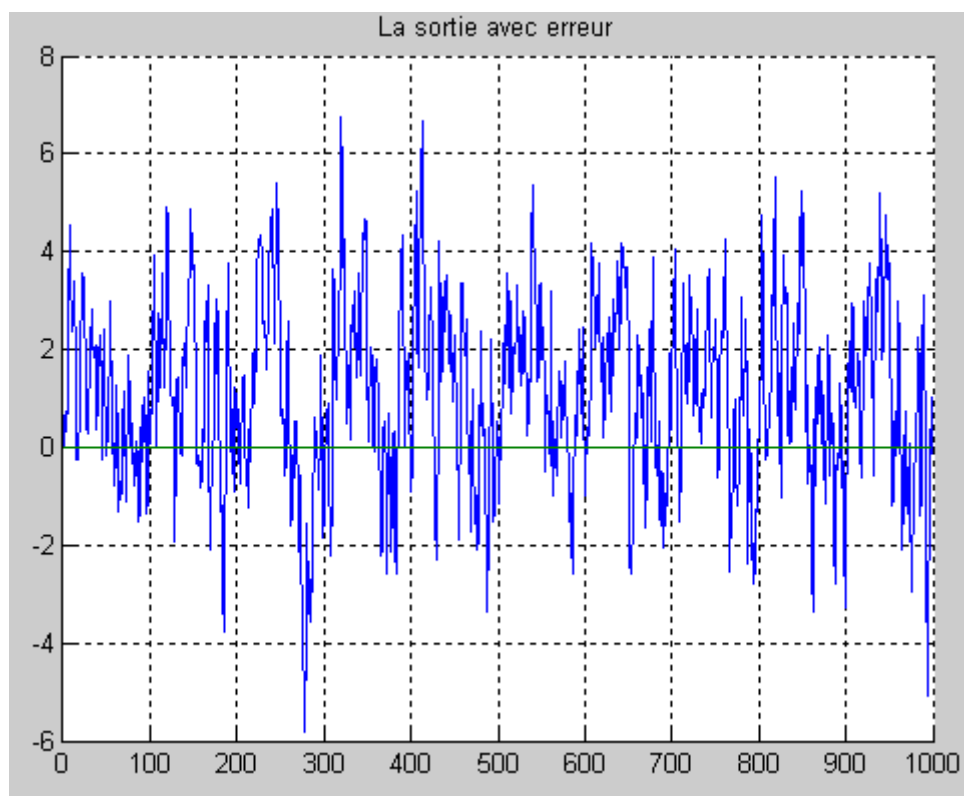


Fig. 3.11 la sortie du système2 avec bruit

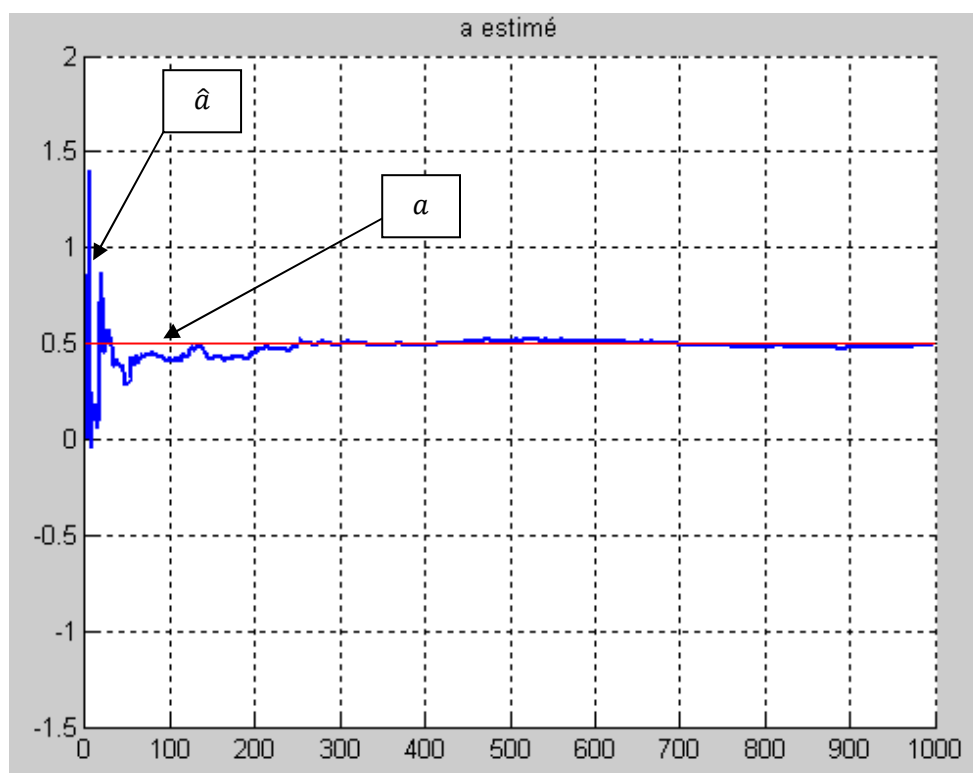
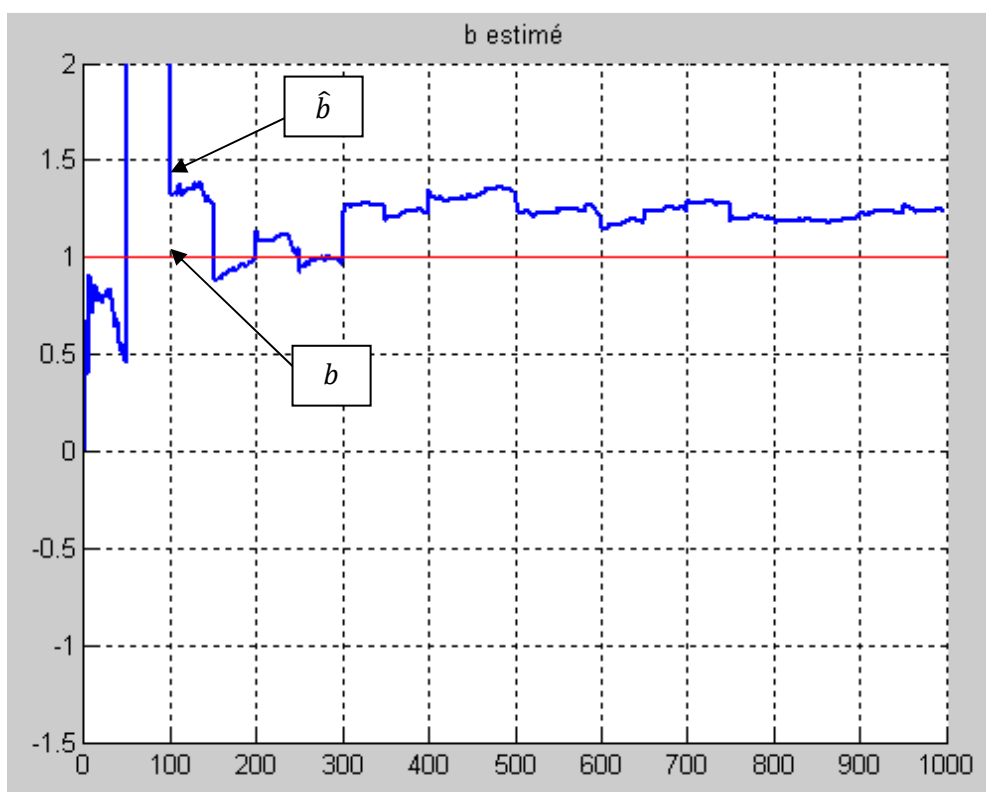
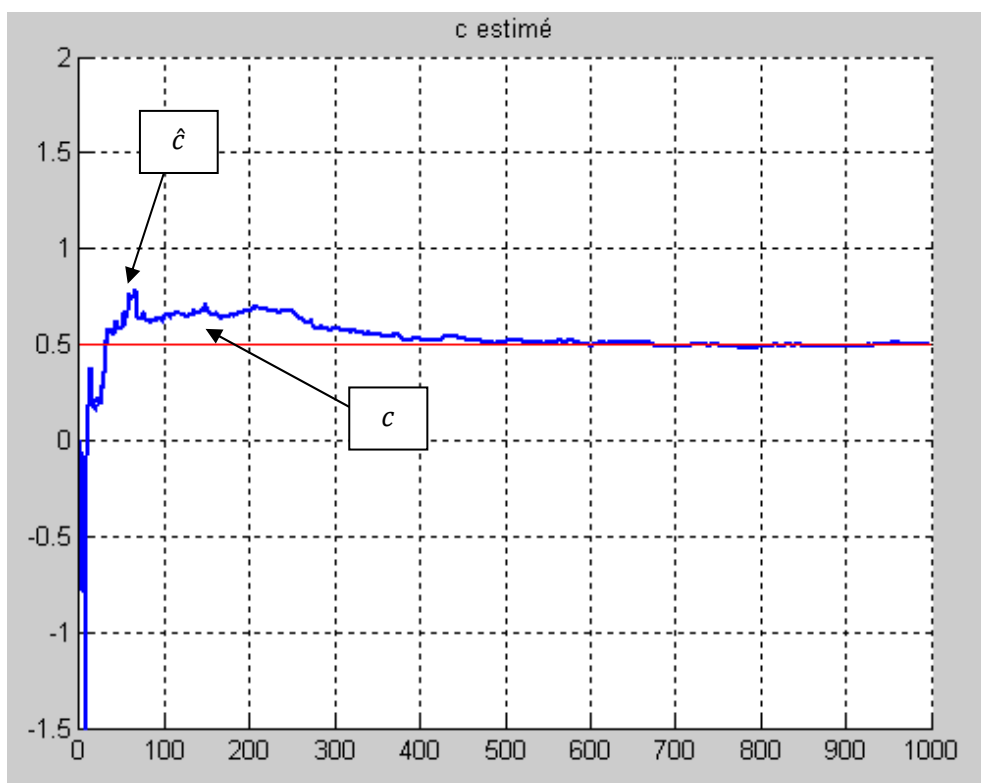


Fig. 3.12 estimation du paramètre a

**Fig. 3.13** estimation du paramètre b **Fig. 3.14** estimation du paramètre c

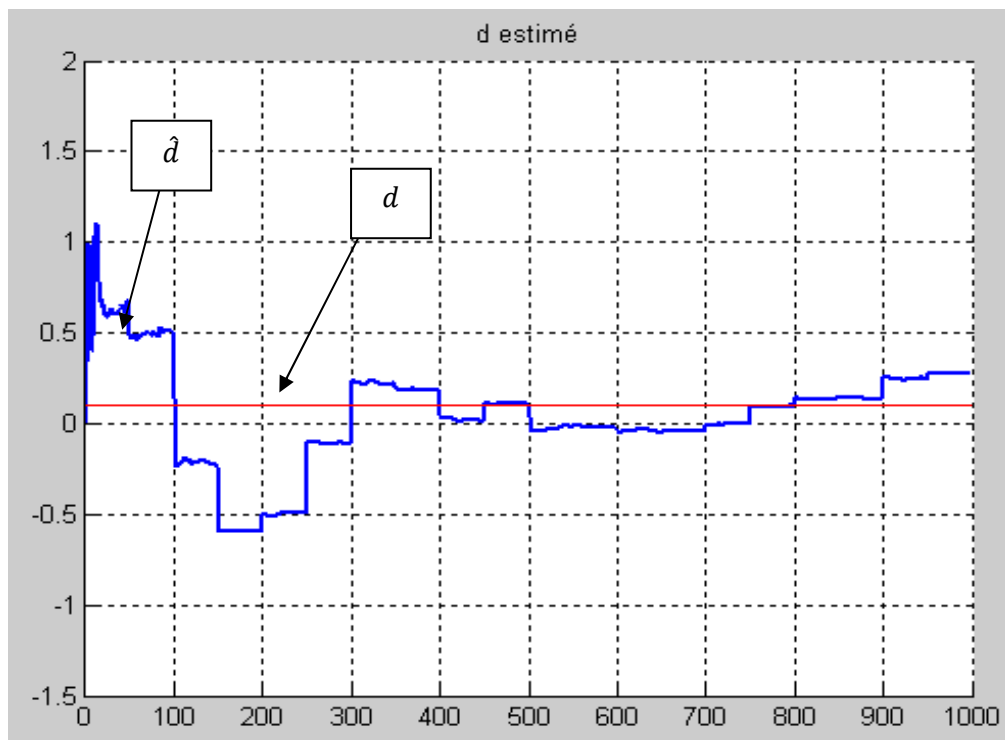


Fig. 3.15 estimation du paramètre d

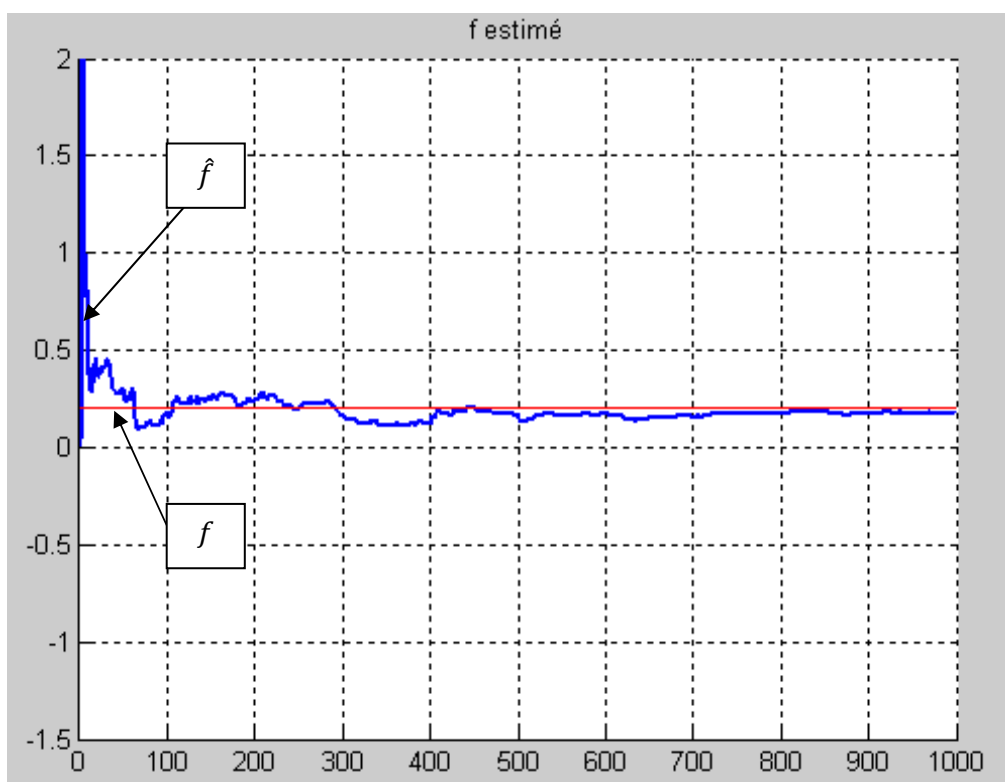


Fig. 3.16 estimation du paramètre f

Interprétation des résultats.

Dans cet exemple il y a 5 paramètres à estimer a, b, c, d, f , ces paramètres sont estimés pour l'ensemble des données $[u(t), y(t)]_{k=1}^N$. où $N = 1000$ est le nombre maximal de points de données obtenues.

Le tableau 3.2 présente les résultats de simulation pour les valeurs réelles et les valeurs estimées pour $N = 1000$. Les Fig. 3.12, 3.13, 3.14, 3.15, 3.16, donnent les trajectoires de la convergence des paramètres réels et les paramètres estimés.

Tableau 3.2 : les résultats de simulation pour les valeurs réelles et les valeurs estimées

paramètres	a	b	c	d	f
réels	0.5000	1	0.5000	0.1	0.2
estimées	0.5019	0.6870	0.5367	0.2114	0.1765

Les deux paramètres b, d ne convergent pas, donc on peut remarquer que l'estimateur des moindres carrés récursifs ne permet pas d'estimer les paramètres d'un modèle NARMAX, pour résoudre ce problème il faut utiliser une autre méthode. La méthode des moindres carrés orthogonaux est bien adaptée pour ce type de modèle.

3.7 Conclusion

Une présentation compacte de l'estimateur des méthodes de moindres carrés a été donnée dans ce chapitre.

La méthode des moindres carrés est particulièrement simple si le modèle a la propriété d'être linéaire dans les paramètres. Dans ce cas, l'estimation des paramètres peut être calculée analytiquement.

On a appliqué la méthode des moindres carrés récursifs pour estimer les paramètres de deux systèmes non linéaires.

En déduire que La méthode des moindres carrés est une technique de base pour l'estimation des paramètres.

Conclusion générale

La modélisation mathématique est un outil indispensable à la représentation de systèmes pour maîtriser ou optimiser leur fonctionnement, ou comprendre leur évolution.

Cette technique de modélisation est constituée de la modélisation théorique et de la modélisation expérimentale. La première nécessite la connaissance des lois régissant le fonctionnement du système, alors que la deuxième ne nécessite que des mesures sur le système à modéliser.

Différentes structures de modèles non linéaires ont été présentées : modèles NFIR, NARX, NOE et NARMAX...etc.

Ce travail est fait l'objet d'estimation des paramètres des systèmes non linéaires par la méthode des moindres carrés récurrents, des exemples d'identification des systèmes non linéaires ont été présentés. Il y a d'autres méthodes qui ne sont pas mentionnées dans ce travail tels que : la méthode du gradient, la méthode de maximum de vraisemblance qu'ils peuvent être utilisés pour l'estimation.

On peut dire que le modèle exact d'un procédé n'existe pas. Il est rare qu'on puisse trouver un modèle satisfaisant tous les critères en passant qu'une seule fois par les différentes étapes d'identification, on est obligé de revoir soit la structure du modèle, ou le protocole d'identification qui doit être modifié.

Comme perspective nous souhaitons étendre les méthodes d'estimation présentées au cas des systèmes non linéaires MIMO et MISO.

Bibliographie

- [1] A. Janczak, «Identification of Nonlinear Systems», Springer-Verlag Berlin Heidelberg 2005.
- [2] Er-Wei Bai, «An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems» *Automatica*, (3):333–338, 1998.
- [3] Ioan D. Landau, Alina Besançon-Voda « Identification des systèmes», HERMES Science Europe Ltd, 2001.
- [4] Karl Johan Åström, Björn Wittenmark, «Adaptive Control», Addison-Wesley Publishing Company, Inc 1995.
- [5] Lamine THIAW «Identification de systèmes dynamiques non-linéaires par réseaux de neurones et multimodèles». Thèse Doctorat, LiSSi - Université Paris12 Val-de-Marne 61, avenue du Général de Gaulle 94010, Créteil cedex, janvier 2008.
- [6] Lennart Ljung « System Identification: Theory for the User 2nd Edition » Prentice Hall, 1999.
- [7] Mahrougui Mabrouk, Redouane Elhadi « Identification paramétrique et non paramétrique avec une application sur un processus industriel », Ingénieur d'état en automatique, 1996.
- [8] Mjelloud Rebiha, Aoune Khadidja « Identification par logique floue des systèmes dynamiques non linéaires », Ingénieur d'état en automatique, 1999.
- [9] Pierre Borne, G. Dauphin – Tanguy, J.P Richard, F. Rotella, I. Zambettakis «Modélisation et identification des processus (T1,T2) », Technip, 1992.
- [10] Robiah Ahmad & Hishamuddin Jamaluddin «orthogonal least square algorithm and its application for modelling suspension system » *Universiti Teknologi Malaysia Journal Teknologi*, 34(A) Jun 2001: 71–8.
- [11] Rosario Toscano « Commande et diagnostic des systèmes dynamiques », Ellipses, 2005.
- [12] Syed Omer Farooq, « Identification and Control of a Class of Nonlinear Systems», master of science in systems engineering, King Fahd University OF Petroleum & Minerals, Dhahran, Saudi Arabia, June 2003.
- [13] Y. Rochdi «Identification des systèmes dynamiques» cours FST Settat-Master ATSII, 2010.

Résumé en français et en anglais

Résumé

Il existe de nombreux systèmes (biologiques, des systèmes de contrôle et systèmes de communication) qui peuvent être représentés par des modèles non linéaires. L'identification des systèmes non linéaires est un problème fondamental dans les domaines des systèmes de contrôle et de communication. De nombreuses études sur les systèmes non linéaires ont été réalisées du point de vue d'identification des paramètres. L'objectif de ce travail est l'identification paramétrique des systèmes non linéaires SISO par la méthode des moindres carrés et ces variantes.

Abstract

There are many systems in the field of biological systems, control systems and communication systems that can be represented by non-linear models. Identification of non-linear systems is a fundamental problem in the fields of control and communication systems. Many investigations of non-linear systems have been carried out from the viewpoint of parameter identification. The objective of this work is to identify a parametric nonlinear SISO systems by the method of least squares and these variants.

