HAL
open science

# Validation de données à variance inconnue

Houcine Chafouk

## To cite this version:

Houcine Chafouk. Validation de données à variance inconnue. Automatique / Robotique. Université Henri Poincaré - Nancy 1, 1990. Français. NNT : 1990NAN10379 . tel-01747421

**HAL Id: tel-01747421**

**https://hal.univ-lorraine.fr/tel-01747421v1**

Submitted on 29 Mar 2018

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

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

UNIVERSITÉ DE LORRAINE

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10
http://www.cfcopies.com/V2/leg/leg_droi.php
http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm

Université de Nancy I

# THESE

Présentée pour l'obtention du

## DOCTORAT

**DE L'UNIVERSITE DE NANCY I**

**Spécialité Automatique**

par

**Houcine CHAFOUK**

# VALIDATION DE DONNEES A VARIANCE INCONNUE

soutenue publiquement le 27 Novembre 1990 devant la commission d'examen

| | | |
|---|---|---|
| Président | : | C. HUMBERT |
| Rapporteurs | : | E. OSTERTAG |
| | | M. DAROUACH |
| Examinateurs | : | J. RAGOT |
| | | D. MAQUIN |
| | | J. FAYOLLE |

*A mes parents*

# Avant propos

Le travail présenté dans ce manuscrit a été réalisé au Laboratoire d'Automatique et de Recherche Appliquée du Centre de Recherche en Automatique de Nancy et de l'Université de Nancy I.

Je tiens à remercier particulièrement Monsieur C. HUMBERT, professeur à l'Université de Nancy I et directeur du Laboratoire d'Automatique et de Recherche Appliquée, pour l'honneur qu'il me fait de présider ce jury. Qu'il trouve également exprimée dans ces quelques lignes ma gratitude pour ces précieux conseils, ses nombreux encouragements et sa constante disponibilité.

J'exprime également mes remerciements aux membres du jury qui m'ont honoré de leur présence lors de l'examen de ce travail.

Monsieur J. RAGOT, Professeur à l'Ecole Nationale Supérieure de Géologie, dont les interventions au cours de ma formation et l'examen critique de ce travail ont contribué à son élaboration.

Monsieur E. OSTERTAG, Professeur à l'Ecole Nationale Supérieure de Physique de Strasbourg pour l'intérêt qu'il a porté a ce travail.

Monsieur M. DAROUACH, Professeur à l'Institut Universitaire de Technologie de Longwy pour ses questions pertinentes et l'intérêt qu'il a bien voulu porter à ce travail.

Monsieur D. MAQUIN, Maître de Conférence à l'Institut Universitaire de Technologie de Longwy, dont l'examen approfondi de ce document et les conseils m'ont été d'un précieux apport.

Monsieur J. FAYOLLE, Directeur de la société Automatica, pour sa présence à mon jury et pour l'examen de ce document.

Ma reconnaissance s'adresse enfin à toutes les personnes qui m'ont prodigué leurs conseils et leur amitié dans l'élaboration de ce document. Je tiens enfin à remercier mes camarades et collègues du laboratoire pour les bons moments que j'ai passé en leur compagnie.

# TABLE DES MATIERES

TABLE DES SYMBOLES 4

INTRODUCTION 5

CHAPITRE I- CONCEPTS GENERAUX DE LA VALIDATION DE DONNEES

INTRODUCTION 8

I-1 GÉNÉRALITÉS 8

I-1-1 Caractéristiques d'une chaîne d'instrumentation 8
I-1-2 Les erreurs de mesure 9

I-1-2-1 Définition des erreurs de mesure 9
I-1-2-2 Loi de probabilité 10

I-1-3 Fidélité, justesse et précision d'un capteur 11
I-1-4 Etapes de la validation de données 12

I-2 NOTATION ET HYPOTHESES 14

I-2-1 Grandeurs vraies, mesurées et estimées 14
I-2-2 Hypothèses sur la répartition des erreurs 15
I-2-3 Equations de contrainte ou de bilan 16

I-3 BILAN AU SENS DU MAXIMUM DE VRAISEMBLANCE 17

I-3-1 Variance des erreurs de mesure connue 19
I-3-2 Variance des erreurs de mesure inconnue et générale 19
I-3-3 Variance des erreurs de mesure inconnue et diagonale 20

CONCLUSION 21

CHAPITRE II- ESTIMATION DE LA VARIANCE DES ERREURS DE MESURE

INTRODUCTION 24

II-1 GÉNÉRALITÉS 24

II-1-1 Codification et représentation d'un réseau de transport de matière 24
II-1-2 Forme canonique de la matrice d'incidence 26
II-1-3 Exemple d'application 26

II-2 ESTIMATION DE LA VARIANCE PAR MINIMISATION D'UN CRITERE QUADRATIQUE SOUS CONTRAINTES RÉSIDUELLES 27

II-2-1 Formulation du problème 27
II-2-2 Procédure de calcul 29

II-2-3 Essais de simulation 31

CONCLUSION 36

CHAPITRE III- RECONCILIATION DE DONNEES A VARIANCE PARTIELLEMENT INCONNUE

INTRODUCTION 38

**III-1 Estimation de variance des systèmes linéaires** 38

III-1-1 Formulation du problème 38
III-1-2 Algorithme de relaxation 39
III-1-3 Exemple d'application 40

**III-2 Estimation de variance des systèmes bilinéaires** 42

III-2-1 Position du problème 42
III-2-2 Résolution hiérarchisée 43
III-2-3 Exemples d'applications 44

CONCLUSION 48

CHAPITRE IV- RECONCILIATION DE DONNEES DES SYSTEMES MULTI-LINEAIRES A VARIANCE INCONNUE

INTRODUCTION 50

IV-1 RÉCONCILIATION DE DONNÉES DES SYSTEMES LINÉAIRES 50

IV-1-1 Formulation du problème 50
IV-1-2 Algorithme de relaxation 52
IV-1-3 Applications numériques 53
IV-1-4 Extension aux systèmes partiellement observés 59

IV-2 RÉCONCILIATION DE DONNÉES DES SYSTEMES BILINÉAIRES 62

IV-2-1 Position du problème 62
IV-2-2 Résolution dans le cas bilinéaire simple 64
IV-2-3 Résolution dans le cas bilinéaire étendu 73
IV-2-4 Extension aux systèmes bilinéaires partiellement observés 75

IV-2-4-1 Mise en évidence du réseau contenant la redondance de type linéaire-bilinéaire 76
IV-2-4-2 Estimation des variances des grandeurs bilinéaires 78

IV-3 RÉCONCILIATION DE DONNÉES DES SYSTEMES MULTI-LINÉAIRES 87

IV-3-1 Formulation du problème 87
IV-3-2 Résolution hiérarchisée 88

IV-3-2-1 Description de l'algorithme de résolution d'un système trilinéaire 89
IV-3-2-2 Application de la méthode à un système trilinéaire 91

CONCLUSION 97

CHAPITRE V- DETECTION DE CAPTEURS DEFAILLANTS PAR LA METHODE DU RAPPORT DE VRAISEMBLANCE GENERALISE

INTRODUCTION 100

V-1 DÉTECTION DE BIAIS DES SYSTEMES COMPLETEMENT MESURÉS 100

V-1-1 Présentation du problème 100
V-1-2 Propriétés statistiques des résidus des équations de bilan 101
V-1-3 Introduction des hypothèses de décision 102
V-1-4 Estimation de l'amplitude du biais 103
V-1-5 Détection de plusieurs défauts 105
V-1-6 Réconciliation de données 106
V-1-7 Essais en simulation 106

V-1-7-1 Application de la détection en mono-observation 106
V-1-7-2 Application de la détection en multi-observations 111

V-2 DÉTECTION DE DÉFAUTS DANS LE CAS DE DONNÉES MANQUANTES 113

V-2-1 Formulation du problème 113
V-2-2 Réconciliation de données 114
V-2-3 Exemple de simulation numérique 114

CONCLUSION 116

CONCLUSION 117

ANNEXES 119

REFERENCES 128

# TABLE DES SYMBOLES

| | |
|---|---|
| $A^T$ | Transposée de la matrice A |
| $A^{-1}$ | Inverse de la matrice A |
| $\lvert A \rvert$ | Déterminant de la matrice A |
| Tr(A) | Trace de la matrice A |
| Diag(A) | Matrice diagonale formée à l'aide de la diagonale de la matrice A |
| $\Lambda_X$ | Matrice diagonale telle que : $\Lambda_X(i,i) = X(i)$ où X est un vecteur |
| dim(X) | Dimension du vecteur X |
| $I_v$ ou I | Matrice identité de rang v |
| M | Matrice d'incidence |
| $M_R$ | Matrice des moments des carrés des résidus |
| $\varepsilon$ | Erreur de mesure |
| X | Mesure |
| $X^*$ | Grandeur vraie |
| $\hat{X}$ | Grandeur estimée |
| Esp($\varepsilon$) | Espérance mathématique de $\varepsilon$ |
| V ou Var($\varepsilon$) | Variance de $\varepsilon$ |
| $\sim N(m,V)$ | Distribué selon une loi normale de moyenne m et de variance V |
| $\sim \chi^2_n$ | Distribué selon une loi du khi 2 à n degrés de liberté |
| Log( ) | Fonction logarithme népérien |
| exp( ) | Fonction exponentielle |
| $\lVert X \rVert^2_p$ | Carré de la norme pondérée du vecteur X |
| $\Phi$ | Critère d'optimisation |
| L | Lagrangien |
| $\lambda$, $\mu$ | Multiplicateurs de Lagrange |

# INTRODUCTION

La commande et le suivi des installations industrielles nécessitent des données dont la cohérence a été éprouvée. L'une des techniques les plus utilisées actuellement est la validation de données par équilibrage de bilan-matière ou bilan-énergie. Cette méthode se décompose en trois phases :

- une décomposition préliminaire du système selon l'observabilité,

- une estimation des grandeurs réelle par équilibrage de bilan,

- un test de robustesse des résultats.

Les deux dernières phases sont étroitement liées au degré de connaissance de la matrice de variance-covariance des erreurs de mesure. Il est important de remarquer que presque toutes les techniques actuelles supposent que cette matrice est connue ou estimée par des méthodes classiques.

En ce sens, l'objectif de notre étude est de développer des algorithmes pour l'optimisation d'une fonction objectif sous contraintes égalités. Ces contraintes peuvent être linéaires ou multi-linéaires. En outre, nous appliquons ces algorithmes sur des systèmes en régime statique.

Le premier chapitre est consacré aux concepts généraux de la validation de données. Nous rappelons quelques définitions classiques à propos des erreurs de mesure. Ensuite, nous présentons la technique du maximum de vraisemblance et nous décrivons les différentes méthodes pour connaître la matrice de variance des erreurs de mesure.

Dans le second chapitre, nous estimons la matrice de variance des erreurs de mesure par minimisation d'un critère quadratique sous contraintes linéaires. Ensuite, nous étudions l'influence du nombre d'observations sur l'évolution de la matrice de variance.

En pratique, on peut disposer de certaines informations concernant la matrice de variance des erreurs de mesure. Ces informations sont, généralement, liées aux rapports de précision entre les différents capteurs. Cette idée a été exploitée et un algorithme de résolution est proposé dans le troisième chapitre. Nous avons appliqué cet algorithme aux systèmes régis par des équations linéaires et bilinéaires. Des exemples de simulation ont été utilisés pour illustrer cette méthode.

Le quatrième chapitre est consacré à la validation de données des systèmes multi-linéaires à variance inconnue. Un algorithme à itération directe est exposé pour estimer simultanément la matrice de variance des erreurs de mesure et les grandeurs réelles. Cette étude est étendue aux systèmes partiellement observés.

Les méthodes qui ont été présentées dans les chapitres précédents ont été appliquées à des données ne contenant pas de défauts. Le cinquième chapitre est consacré à la détection et la localisation des défauts ainsi que l'estimation de leurs

amplitudes. Nous présentons un algorithme de détection de défauts basé sur le rapport de vraisemblance généralisé (G.L.R).

Enfin cette étude est complétée par des annexes où nous avons rassemblé l'ensemble des "outils mathématiques" nécessaires à la compréhension et à l'utilisation des différents algorithmes qui ont été mis en œuvre.

# CHAPITRE I

## CONCEPTS GENERAUX DE LA VALIDATION DE DONNEES

# INTRODUCTION

La conduite des processus industriels recouvre trois aspects fondamentaux :

- l'acquisition des informations à partir de chaînes d'instrumentations,

- le traitement des informations,

- l'élaboration des lois de commande.

Les deux derniers aspects sont les domaines de prédilection de l'automaticien ; les nombreux travaux qui leur ont été consacré ont fait appel à un grand nombre de théories. Par contre, le problème de l'instrumentation est certainement celui dont les progrès sont les plus lents ; certains types de grandeur sont inaccessibles à la mesure et d'autres le sont très difficilement. En plus de ce problème, se pose celui de la validation de la mesure. Comment s'assurer de la qualité des informations collectées ?

Afin de vérifier la vraisemblance d'une mesure :

- la première solution consiste à multiplier les différents capteurs créant ainsi une redondance permettant, par comparaison, une évaluation plus fiable de la mesure en présence ou en l'absence d'un capteur défaillant. Cette méthode requiert l'emploi de différents capteurs entraînant, de ce fait, un surcoût de l'installation,

- la deuxième solution consiste à exploiter des modèles exacts liant les différentes grandeurs, par exemple, les équations de conservation de matière ou d'énergie.

Le principe de la validation de mesure consiste à exploiter les mesures (fournies par les capteurs) pour obtenir des grandeurs estimées qui se rapprochent "au mieux" des grandeurs vraies.

## I-1 GENERALITES

### I-1-1 CARACTÉRISTIQUES D'UNE CHAINE D'INSTRUMENTATION

Afin d'identifier les différents problèmes rencontrés lors de l'élaboration d'une information représentative d'une grandeur physique, nous avons représenté, figure I-1, le schéma d'une chaîne d'acquisition de mesure [**CERR 80**].

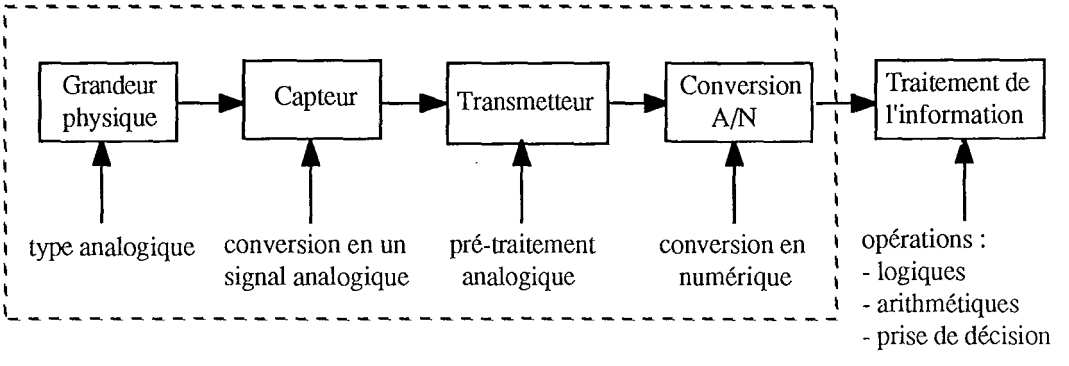



Figure I-1 : chaîne d'acquisition de mesure

A l'entrée de la chaîne, le capteur soumis à l'action du mesurande permet, directement, s'il est actif ou par le moyen de son conditionneur, s'il est passif, d'injecter dans la chaîne le signal électrique, support de l'information liée au mesurande.

A la sortie de la chaîne, ce signal électrique est converti sous une forme qui permet la lecture directe de la valeur cherchée du mesurande.

## I-1-2 LES ERREURS DE MESURE

Si l'on désigne par $x^*$ la valeur vraie d'une grandeur, par x la mesure de cette grandeur et par $\varepsilon$ l'erreur introduite dans cette mesure par la chaîne d'acquisition (provenant des capteurs, de la transmission mais aussi de l'opérateur), nous pouvons écrire la relation entre x, $x^*$ et $\varepsilon$ sous la forme :

$$x = x^* + \varepsilon \qquad \text{(I-1)}$$

Si l'on recommence un grand nombre de fois la mesure d'une grandeur, nous enregistrons, d'une opération à l'autre, une plus ou moins forte variation des résultats obtenus. Indépendamment de la méthode de mesure utilisée, on est amené à écrire :

- la première observation $x_1 = x^* + \varepsilon_1$
- la deuxième observation $x_2 = x^* + \varepsilon_2$
- .....
- la $m^{ème}$ observation $x_m = x^* + \varepsilon_m$

(I-2)

### I-1-2-1 Définition des erreurs de mesure

Les mesures collectées sur un procédé industriel sont souvent entachées d'erreurs, nous allons les définir par catégories :

- Les erreurs aléatoires : au cours du temps nécessaire à la mesure, certaines grandeurs (tension d'alimentation, température, champ électrique, ...) peuvent varier de façon aléatoire en produisant des effets de valeurs moyennes nulles. Les causes de ce type d'erreur sont généralement bien connues par l'opérateur, leur amplitude est faible et peut être bornée par étalonnage de la chaîne de mesure.

- Les erreurs systématiques : se distinguant par leur permanence et leur forte amplitude, elles sont liées, par exemple, au vieillissement ou au mauvais étalonnage d'un capteur. Elles peuvent se traduire par une dérive progressive ou une panne franche.

- Les erreurs accidentelles : caractérisées par leur forte amplitude, elles demeurent exceptionnelles. Elles ne se distinguent des erreurs systématiques que par leur aspect temporaire. Elles sont liées à une défaillance matérielle ou humaine momentanée. Par exemple, la grandeur à évaluer se trouve hors de la plage de fonctionnement normale de l'appareil de mesure, ou une brusque variation de la tension d'alimentation modifie les performances de la chaîne de mesure.

Nous allons rappeler quelques notions relatives aux erreurs aléatoires.

### I-1-2-2 Loi de probabilité

En pratique, les fonctions de distribution de la variable aléatoire x sont inconnues puisqu'elles correspondent à un cas limite : nombre infini de résultats et sensibilité très grande de l'appareil de mesure.

Cependant les mathématiciens utilisent différents modèles de distribution dont on peut admettre le plus souvent qu'ils décrivent correctement les situations expérimentales.

La distribution que nous utilisons par la suite est la loi de Gauss (ou loi normale). On dit qu'une variable x suit la loi normale lorsque sa densité de probabilité est une fonction définie par l'expression (I-3) et dont la représentation graphique est donnée par la figure I-2.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right) \qquad \text{(I-3)}$$

où $\sigma$ est l'écart-type de x et $\bar{x}$ sa moyenne.

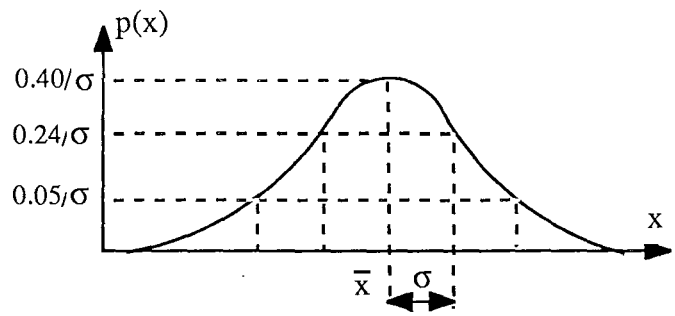



Figure I-2 : loi normale

Si la mesure d'une même valeur (inconnue) du mesurande a été répétée m fois, donnant les résultats $x_1, x_2, \ldots, x_m$, la valeur moyenne $\bar{x}$ est par définition :

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{(I-4)}$$

La dispersion de ces résultats est caractérisée par l'écart-type $\sigma$ défini par :

$$\sigma = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{x})^2} \qquad \text{(I-5)}$$

La courbe de distribution ne dépend que des deux paramètres $\sigma$ et $\bar{x}$. Nous allons représenter une loi normale de moyenne $\bar{x}$ et d'écart-type $\sigma$ par le symbole :

$$N(\bar{x},\sigma) \qquad \text{(I-6)}$$

De façon générale, la probabilité P d'avoir un résultat inférieur à x est donnée par l'intégrale :

$$P = \int_{-\infty}^{x} p(u)\, du \qquad \text{(I-7)}$$

Cette intégrale est appelée fonction de répartition, sa valeur dépend de $\bar{x}$ et de $\sigma$. On dit qu'un résultat de mesure x suit une loi normale si les conditions suivantes sont réalisées simultanément **[CEA 78]** :

- les causes d'erreur sont nombreuses,

- les erreurs sont du même ordre de grandeur,

- les fluctuations liées aux différentes causes d'erreur sont indépendantes.

Si la méthode de mesure a été correctement mise au point, ces conditions sont souvent vérifiées.

A présent nous allons rappeler quelques notions relatives aux erreurs accidentelles et systématiques.

### I-1-3 FIDÉLITÉ, JUSTESSE ET PRÉCISION D'UN CAPTEUR

Les erreurs accidentelles entraînent une dispersion des résultats lors de mesures répétées ; cependant leur traitement statistique permet de connaître la valeur la plus probable de la grandeur mesurée et de fixer les limites du bon fonctionnement d'un capteur.

La **fidélité** est la qualité d'un appareillage de mesure dont les erreurs accidentelles sont faibles : elle se traduit par des résultats de mesure groupés autour de leur valeur moyenne.

L'écart-type, dont l'importance reflète la dispersion des résultats, est souvent considéré comme l'erreur de fidélité : il permet ainsi une appréciation quantitative de la fidélité. La valeur du mesurande la plus probable, telle qu'elle résulte d'un ensemble de mesures, peut être connue avec une faible marge d'incertitude tout en étant éloignée de la valeur vraie du mesurande si des erreurs systématiques importantes se sont superposées aux erreurs accidentelles.

La **justesse** est la qualité d'un appareillage de mesure dont les erreurs systématiques sont réduites : la valeur la plus probable du mesurande qu'un tel appareillage permet de déterminer est alors très proche de la valeur vraie.

La **précision** qualifie l'aptitude d'un appareillage de mesure à donner des résultats qui sont proches de la valeur vraie du mesurande : un appareillage précis est donc à la fois fidèle et juste.

La figure I-3 illustre chacune de ces notions : sur cette figure p(x) désigne la fonction densité de probabilité de la variable x représentant la grandeur à mesurer.

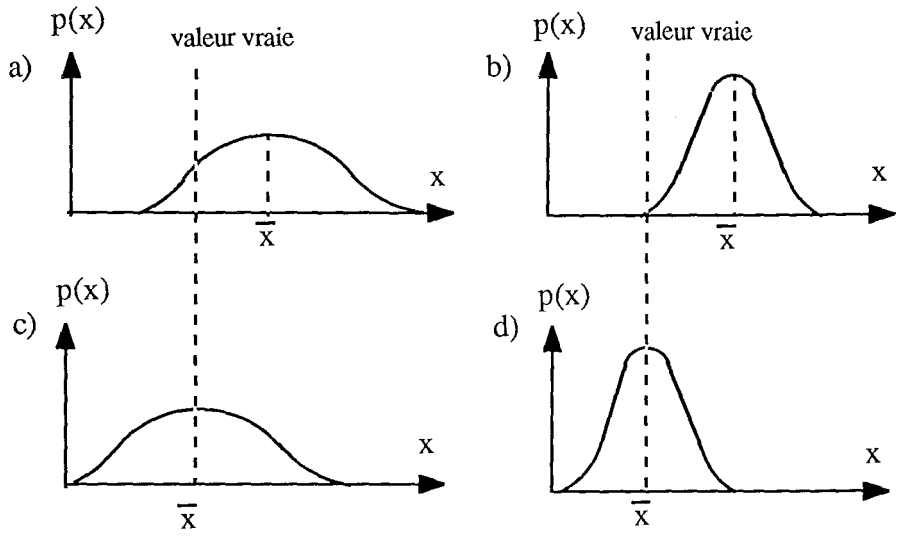



Figure I-3 : différents types de répartition des résultats de mesure

a) **appareillage ni juste, ni fidèle** : erreurs systématiques et accidentelles importantes.

b) **appareillage fidèle mais non juste** : erreurs systématiques importantes, erreurs accidentelles réduites.

c) **appareillage juste mais non fidèle** : erreurs systématiques faibles, erreurs accidentelles importantes.

d) **appareillage juste et fidèle donc précis** : erreurs systématiques et accidentelles faibles.

La précision peut être spécifiée numériquement par l'erreur de précision qui, compte tenu de toutes les causes d'erreur (systématiques et accidentelles) délimite l'intervalle autour de la valeur mesurée, à l'intérieur duquel on est assuré de trouver la valeur vraie. La mesure d'une grandeur physique avec une précision donnée exige :

- le choix d'une méthode de mesure,

- la sélection du capteur approprié,

- la conception et la réalisation de la chaîne de mesure associée.

Avant d'aborder le problème de l'équilibrage de bilan matière, il nous semble nécessaire de décrire les différentes étapes de la validation de données.

### I-1-4 ETAPES DE LA VALIDATION DE DONNÉES

En suivant l'organigramme de la figure I-4, la première préoccupation est celle du protocole des mesures. En effet, pour analyser l'état d'un procédé, pour en obtenir un diagnostic cohérent, il est nécessaire de disposer d'informations. Le protocole de mesures dépend :

- du choix des grandeurs à mesurer,

- du choix des capteurs aptes à mesurer ces grandeurs,

- du choix des signaux tests capables d'exciter convenablement le procédé.

La constitution de la base de données est un problème d'organisation informatique. En effet, devant la disparité des informations (qualitatives, quantitatives, logiques, ...) il est nécessaire de disposer de moyens de stockage obéissant à un protocole fiable, systématique, mais cependant flexible de façon à pouvoir évoluer en fonction de nouvelles structures de saisie d'informations **[FAYO 87a]**.

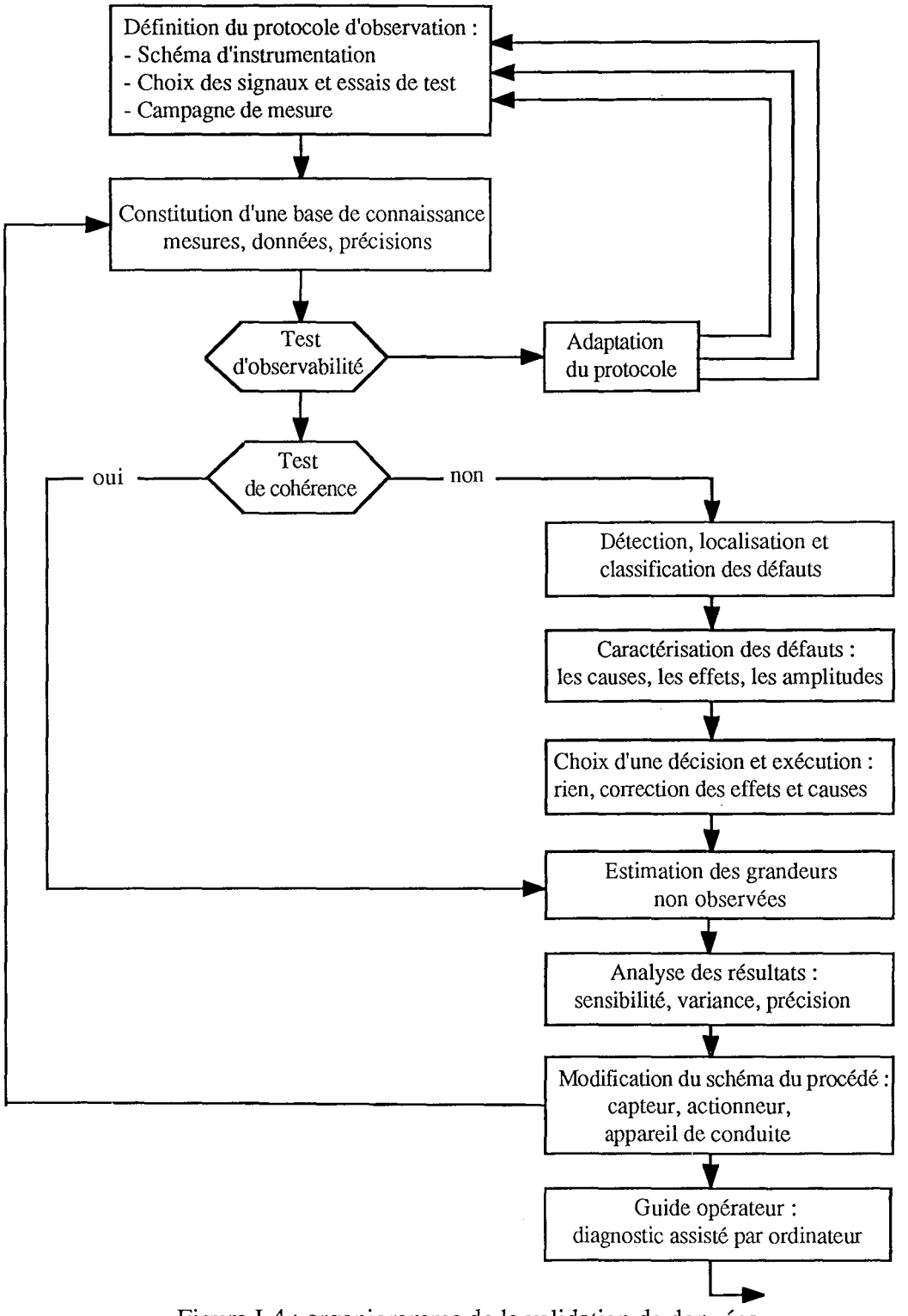



Figure I-4 : organigramme de la validation de données

Les tests d'observabilité requièrent une grande connaissance des objectifs et du procédé. Les tests doivent permettre de répondre aux questions suivantes : les informations disponibles sont-elles suffisantes pour décrire l'ensemble du comportement du procédé ? **[DARO 86 a]**.

Ces mêmes informations permettent-elles une redondance analytique ? Dispose-t-on des informations suffisantes pour détecter des défauts, des erreurs traduisant des mauvais fonctionnements du procédé, des capteurs, des dispositifs de commande ? Si ces tests permettent de conclure à l'inadéquation de la base de données il faudra en adapter le mode de constitution ; dans le cas contraire on admettra provisoirement que cette base de données est suffisante. Les tests de cohérence posent ensuite la question : les données sont-elles vraisemblablement "bonnes" ou au contraire "mauvaises". Des modèles de vraisemblance, utilisant des lois structurellement exactes et des lois empiriques provenant de l'apprentissage et d'hypothèses simplificatrices doivent être construites ; les conclusions issues de ces modèles devront être assorties de seuils de décision pour mettre en œuvre ou non une procédure de caractérisation des incohérences des informations **[COHE 82]**. Il est en effet important de pouvoir séparer dans la base de données ce qui, vraisemblablement, est correct de ce qui semble erroné ; la pratique montre également que les incohérences peuvent être de différentes nature : aléatoires, accidentelles, systématiques.

La caractérisation des incohérences à proprement parler concerne l'étude et la quantification des causes et des effets : l'estimation des amplitudes se fera de façon statistique si l'on dispose d'une redondance suffisante des informations **[HODO 79]**.

Comme conséquence de cette caractérisation une décision devra être prise :

- ne pas prolonger les investigations,

- corriger les effets des incohérences,

- ou encore corriger les incohérences.

Ces corrections se feront de façon statistique à partir de modèle de connaissance structurellement exacts et de modèles heuristiques ; la décision d'entreprendre cette phase de correction ou d'estimation d'incohérence sera consécutive à l'optimisation d'un critère de performance. A partir de données cohérentes ou rendues cohérentes, la phase suivante concerne l'estimation des grandeurs inaccessibles à la mesure ; cette estimation se fera par déduction à partir des modèles des procédés.

Aucun des résultats obtenus dans les phases précédentes n'a de valeur s'il n'est accompagné d'une étude de sensibilité, en fonction du nombre d'informations disponibles, de leur position sur le procédé et de leur précision. C'est à partir de l'estimation et de la précision de l'estimation des grandeurs que l'on pourra donner à l'opérateur une cartographie cohérente et complète de l'état du procédé.

Les informations ainsi fournies pourront servir de base à une prise de décision et de conduite du procédé.

## I-2 NOTATION ET HYPOTHESES

### I-2-1 GRANDEURS VRAIES, MESURÉES ET ESTIMÉES

Soit x la mesure délivrée par un appareil de mesure et x* le mesurande correspondant, on a :

$$x = x^* + \varepsilon + \delta \qquad \text{(I-8)}$$

où $\varepsilon$ est l'erreur de mesure aléatoire et $\delta$ l'erreur de mesure systématique.

Pour un appareil de mesure bien étalonné et fonctionnant correctement, on considère que l'erreur systématique est suffisamment faible pour être assimilée, au point de vue traitement, à une erreur aléatoire. Le problème qui se pose alors est de déterminer une estimation $\hat{x}$ du mesurande $x^*$ à partir de la mesure x.

Dans le cas général, un processus est caractérisé par **v** variables sur lesquelles on effectue **m** mesures que l'on rassemble dans la matrice X appelée matrice d'observations, de dimension (m x v).

$X^*$ représente la matrice des grandeurs vraies et X la matrice des grandeurs mesurées.

$$X = X^* + \varepsilon + \Delta \qquad \text{(I-9)}$$

$\varepsilon$ et $\Delta$ étant des matrices de dimension (m x v).

### I-2-2 HYPOTHESES SUR LA RÉPARTITION DES ERREURS

**Hypothèse 1**

L'erreur systématique $\Delta$ sur les variables est indépendante des observations et elle est constante.

La matrice des erreurs ou biais systématiques $\Delta$ de dimension m x v est ainsi constituée :

$$\Delta = \begin{bmatrix} \delta_1 & \delta_2 & \dots & \delta_v \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ \delta_1 & \delta_2 & \dots & \delta_v \end{bmatrix}$$

Ce qui s'écrit également :

$$\Delta = \mathbf{1} \,.\, \delta^T \qquad \text{(I-10)}$$

avec $\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ . \\ . \\ \delta_v \end{bmatrix}$

et $$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ . \\ . \\ 1 \end{bmatrix}$$ de dimension m x 1

**Hypothèse 2**

Les variables aléatoires $\varepsilon_{ij}$ (i = 1, ..., m et j = 1, ..., v) suivent une loi normale centrée et de variance $V_{jj}$ (termes diagonaux de la matrice V).

La matrice des erreurs de mesures, notée $\varepsilon$, de dimension (m x v) s'écrit :

$$\varepsilon = \begin{bmatrix} \varepsilon_{11} & \dots & \varepsilon_{1j} & \dots & \varepsilon_{1v} \\ \dots & \dots & \varepsilon_{ij} & \dots & \dots \\ \varepsilon_{m1} & \dots & \varepsilon_{mj} & \dots & \varepsilon_{mv} \end{bmatrix}$$

### I-2-3 EQUATIONS DE CONTRAINTE OU DE BILAN

Une équation de bilan traduit des lois de conservation de masse ou d'énergie ou, plus généralement, une relation entre les variables. Elle constitue en fait un modèle du système.

En général, les mesures ne vérifient pas ces équations de bilan. Equilibrer un bilan, consiste à estimer les grandeurs vraies qui doivent être cohérentes, c'est-à-dire vérifier les équations de bilan.

**Modèle d'un processus statique**

Le modèle d'un procédé est une représentation mathématique d'un système physique. Son élaboration conduit souvent à des hypothèses simplificatrices. D'une façon générale, les mesures prélevées sur un procédé ne vérifient que partiellement les équations de modèle. Ces écarts peuvent se justifier de la façon suivante :

- la structure du modèle est imparfaite,

- les paramètres du modèle sont inexacts,

- les mesures effectuées sont entachées d'erreurs.

Dans la suite de cet exposé, seules les erreurs de mesures seront prises en considération, les modèles seront supposés exacts. Ce cas particulier est celui des équations de bilan qui traduisent la conservation de la matière ou d'énergie.

Soit g la fonction vectorielle de dimension **m** correspondant aux **m** relations (ou équations de bilans) entre les variables. On a alors, pour les grandeurs vraies, le modèle du procédé :

$$g(X_i^*, \Theta^*) = 0 \qquad \text{pour } i=1, ..., m \qquad \text{(I-11)}$$

où $X_i^*$ est le vecteur des grandeurs vraies à l'instant i,

et $\Theta^*$ le vecteur des paramètres du modèle.

Au sens du maximum de vraisemblance, le meilleur estimateur $\hat{X}_i$ et $\hat{\Theta}$ des grandeurs vraies est celui qui maximise la densité de probabilité des grandeurs observées en respectant les équations de contraintes :

$$g(\hat{X}_i, \hat{\Theta}) = 0 \qquad \text{pour } i=1, ..., m \qquad \text{(I-12)}$$

## I-3 BILAN AU SENS DU MAXIMUM DE VRAISEMBLANCE

Les erreurs de mesure sont supposées statistiquement indépendantes des observations. La fonction densité de probabilité de $\varepsilon$ est égale au produit des fonctions densités de probabilité des vecteurs $\varepsilon_i$, qui sont formés des variables aléatoires $\varepsilon_{ij}$ pour $j = 1, ..., v$.

La fonction densité de probabilité $F(\varepsilon/V)$ s'écrit, en tenant compte de la deuxième hypothèse :

$$F(\varepsilon/V) = (2\pi)^{-\frac{vm}{2}} \prod_{i=1}^{m} |V|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{m} \| \varepsilon_i \|^2_{V^{-1}}\right) \qquad \text{(I-13)}$$

avec

$$\| \varepsilon_i \|^2_{V^{-1}} = \varepsilon_i^T V^{-1} \varepsilon_i \qquad \text{(I-14)}$$

Pour réaliser l'estimation des $\varepsilon_i$ nous utilisons la méthode basée sur le critère du maximum de vraisemblance.

Nous avons :

$$X_i = X_i^* + \varepsilon_i + \delta \qquad \text{pour } i = 1, ..., m \qquad \text{(I-15)}$$

N'ayant pas accès aux grandeurs vraies, posons :

$$X_i = \hat{X}_i + E_i(\hat{X}) + \delta \qquad \text{(I-16)}$$

où $E_i(\hat{X})$ représente le résidu $(X_i - \hat{X}_i - \delta)$ pour une estimation particulière $\hat{X}$ des grandeurs vraies obtenue après minimisation d'une certaine fonction coût.

Nous désignerons par fonction de vraisemblance, la densité de probabilité de $\hat{X}$ dans laquelle $\varepsilon_i$ est remplacé par les résidus $E_i(\hat{X})$.

$$F(\hat{X}/V) = (2\pi)^{-\frac{vm}{2}} \prod_{i=1}^{m} |V|^{-\frac{1}{2}} \exp(-\frac{1}{2} \sum_{i=1}^{m} \| X_i - \hat{X}_i - \delta \|^2_{V^{-1}}) \qquad \text{(I-17)}$$

avec

v : nombre de variables,
m : nombre d'observations,
V : matrice de covariance des erreurs de mesure de dimension v x v,
$X_i$ : vecteur colonne formé des éléments de la ligne i de la matrice X,
$\hat{X}_i$ : vecteur colonne formé des éléments de la ligne i de la matrice $\hat{X}$.

Le terme de l'exponentielle s'écrit :

$$\frac{1}{2} \| X_i - \hat{X}_i - \delta \|^2_{V^{-1}} = \frac{1}{2} \mathrm{tr} [ ( X_i - \hat{X}_i - \delta )^T V^{-1} ( X_i - \hat{X}_i - \delta ) ]$$

$$= \frac{1}{2} \mathrm{tr} [ V^{-1} ( X_i - \hat{X}_i - \delta )( X_i - \hat{X}_i - \delta )^T ]$$

d'où :

$$\frac{1}{2} \sum_{i=1}^{m} \| X_i - \hat{X}_i - \delta \|^2_{V^{-1}} = \frac{1}{2} \mathrm{tr} [ \sum_{i=1}^{m} V^{-1} ( X_i - \hat{X}_i - \delta )( X_i - \hat{X}_i - \delta )^T ] \qquad \text{(I-18)}$$

où le symbole **tr** désigne la trace d'une matrice, égale à la somme des termes diagonaux de cette matrice.

Enfin l'expression de la fonction de vraisemblance s'écrit :

$$F(\hat{X}/V) = (2\pi)^{-\frac{vm}{2}} \prod_{i=1}^{m} |V|^{-\frac{1}{2}} \exp ( -\frac{1}{2} \mathrm{tr}(V^{-1} M_R)) \qquad \text{(I-19)}$$

où $M_R$ est la matrice des moments des carrés des résidus définie par :

$$M_R = \sum_{i=1}^{m} [( X_i - \hat{X}_i - \delta)( X_i - \hat{X}_i - \delta)^T] \qquad \text{(I-20)}$$

Au sens du maximum de vraisemblance, les estimateurs $\hat{X}_i$ et $\hat{\Theta}$ maximiseront cette fonction de vraisemblance sous le respect des équations de contraintes :

$$g(\hat{X}_i, \hat{\Theta}) = 0 \qquad \text{pour } i = 1, ..., m \qquad \text{(I-21)}$$

Le problème d'équilibrage de bilan est donc traduit en termes d'optimisation sous contraintes, les grandeurs d'optimisation sont :

- les estimées $\hat{X}_i$,

- les paramètres $\hat{\Theta}$,

- et, dans certains cas, la matrice de variance des erreurs de mesure V.

Le problème d'estimation de ces grandeurs est étroitement lié à la nature des contraintes, à leurs dimensions, au nombre de variables du processus, à la présence éventuelle de biais systématique $\delta$ et à la structure de la matrice de covariance V.

De plus, ce problème est fonction du degré de connaissance de la matrice de covariance. Trois cas essentiels se présentent :

- matrice de variance-covariance V connue,

- matrice de variance-covariance V inconnue et générale,

- matrice de variance-covariance V inconnue et diagonale.

Dans les paragraphes suivants, nous développerons ces trois cas.

### I-3-1 VARIANCE DES ERREURS DE MESURE CONNUE

En général, la matrice de covariance V est connue, car élaborée à partir des précisions supposées connues des capteurs. Dans ce cas, nous pouvons formuler simplement le problème d'équilibrage de bilan par la minimisation du critère $\Phi$ par rapport à $\hat{X}_i$ et $\hat{\Theta}$ :

$$\Phi = tr(V^{-1}M_R) \qquad \text{(I-22)}$$

Sous les contraintes $g(\hat{X}_i,\hat{\Theta}) = 0$ pour $i = 1, ..., m$

La résolution du problème de minimisation du critère sous contraintes est effectuée à l'aide de la méthode des multiplicateurs de Lagrange. Nous conseillons aux lecteurs intéressés par cette procédure de consulter le livre de Y. Bard [**BARD 74**].

### I-3-2 VARIANCE DES ERREURS DE MESURE INCONNUE ET GÉNÉRALE

Dans le cas où la variance des erreurs de mesure est inconnue la fonction de vraisemblance s'écrit :

$$F(\hat{X}) = (2\pi)^{-\frac{vm}{2}} \prod_{i=1}^{m} |V|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} tr(V^{-1}M_R)\right) \qquad \text{(I-23)}$$

Après avoir dérivé F par rapport à la variance V [**BARD 74**], nous trouvons :

$$V = \frac{1}{m} M_R \qquad \text{(I-24)}$$

Nous en déduisons :

$$tr(V^{-1}M_R) = tr(V^{-1}mV) = m\, tr(I) = mv \qquad \text{(I-25)}$$

En substituant ce résultat dans l'expression de la fonction de vraisemblance, on trouve :

$$F(\hat{X}) = (2\pi)^{-\frac{vm}{2}} \prod_{i=1}^{m} |\frac{1}{m} M_R|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} mv\right) \qquad \text{(I-26)}$$

d'où :

$$\mathrm{Log}(F(\hat{X})) = -\frac{vm}{2}\mathrm{Log}(2\pi) - \frac{m}{2}\mathrm{Log}(\frac{1}{m}|M_R|) - \frac{1}{2}mv \qquad \text{(I-27)}$$

Maximiser $\mathrm{Log}(F(\hat{X}))$ revient à minimiser $\mathrm{Log}(|M_R|)$, dans le cas d'erreurs de mesure de variance-covariance inconnue et générale, le problème d'équilibrage de bilan matière se réduit à chercher :

$$\min \Phi = \mathrm{Log}(|M_R|) \qquad \text{par rapport à } \hat{X}_i \text{ et } \hat{\Theta} \qquad \text{(I-28)}$$

$$\text{Sous les contraintes } g(\hat{X}_i, \hat{\Theta}) = 0 \qquad \text{pour } i=1,\dots, m$$

### I-3-3 VARIANCE DES ERREURS DE MESURE INCONNUE ET DIAGONALE

Dans le cas où V est diagonale, l'expression (I-24) peut s'écrire :

$$V_{jj} = \frac{1}{m} M_{Rjj} \qquad j = 1,\dots, v \qquad \text{(I-29)}$$

où $V_{jj}$ et $M_{Rjj}$ sont les termes diagonaux des matrices V et $M_R$.

on en déduit :

$$|V| = \prod_{j=1}^{v} V_{jj} = \prod_{j=1}^{v} \frac{1}{m} M_{Rjj}$$

et

$$\mathrm{tr}(V^{-1}M_R) = \sum_{i=1}^{m} V_{jj}^{-1} M_{Rjj} = mv \qquad \text{(I-30)}$$

Le logarithme de la fonction de vraisemblance devient :

$$\mathrm{Log}(F(\hat{X})) = -\frac{vm}{2}\mathrm{Log}(2\pi) - \frac{m}{2}\sum_{j=1}^{v} \mathrm{Log}(\frac{1}{m} M_{Rjj}) - \frac{1}{2}mv \qquad \text{(I-31)}$$

En résumé, dans le cas de mesures non corrélées et de variance inconnue et diagonale, le problème d'équilibrage de bilan se réduit à chercher :

$$\min \Phi = \sum_{j=1}^{v} \mathrm{Log}(M_{Rjj}) = \mathrm{tr}(\mathrm{Log}(M_R)) \qquad \text{par rapport à } \hat{X}_i \text{ et } \hat{\Theta} \qquad \text{(I-32)}$$

$$\text{Sous les contraintes } g(\hat{X}_i, \hat{\Theta}) = 0 \qquad \text{pour } i = 1,\dots, m$$

Nous présentons la résolution de ce problème d'optimisation dans la suite de ce mémoire aux chapitres III et IV.

## CONCLUSION

Ce chapitre nous a permis de présenter les concepts généraux de la validation de données. Après avoir mis en évidence, analysé et classé les causes des différentes erreurs de mesure, la formulation de l'hypothèse de normalité de la distribution conduit à considérer la validation de données par équilibrage de bilan comme un problème d'estimation au sens du maximum de vraisemblance correspondant à la recherche d'extremum de fonctions.

Cette recherche d'extremum n'admet pas, en général, de solution analytique en raison de la structure des contraintes et aussi parce que la matrice de variance-covariance des erreurs de mesure n'est pas toujours connue. Y. Bard [**BARD 74**] a pu expliciter la formulation du problème en fonction du degré de connaissance de la matrice de variance.

Dans les chapitres suivants nous allons développer des techniques d'estimation de variance des erreurs de mesure ainsi que les grandeurs réelles.

# CHAPITRE II

## ESTIMATION DE LA VARIANCE DES ERREURS DE MESURE

## INTRODUCTION

Dans ce chapitre, nous développons une méthode d'estimation de la matrice de variance des erreurs de mesure par minimisation d'un critère quadratique basé sur la norme carrée des termes non diagonaux de cette matrice, en utilisant des contraintes issues de la matrice de covariance des résidus de bilan.

Avant de présenter cette méthode, nous allons rappeler quelques notions sur la représentation de réseaux de transport de matière ou d'énergie.

## II-1 GENERALITES

### II-1-1 CODIFICATION ET REPRÉSENTATION D'UN RÉSEAU DE TRANSPORT DE MATIERE

Les procédés auxquels nous nous intéressons, sont schématisés par un graphe orienté, où les nœuds représentent des unités élémentaires de production et les arcs, les voies de circulation de la matière. De plus, nous ferons l'hypothèse qu'il n'y a pas d'accumulation de matière au sein des nœuds ou, s'il y a accumulation, que celle-ci ne varie pas. Cette hypothèse n'est pas restrictive. En effet, une variation de stock dans un réservoir pourra se traduire par l'ajout d'une voie sur le nœud considéré.

Pour codifier les liaisons entre les différents nœuds, on définit la matrice d'incidence aux arcs M d'éléments $m_{ij}$ tels que :

| | |
|---|---|
| $m_{ij} = +1$ | lorsque le flux j arrive au nœud i, |
| $= -1$ | lorsque le flux j sort du nœud i, |
| $= 0$ | lorsque le flux j n'est pas connecté au nœud i. |

avec i variant de 1 à n (nombre de nœuds) et j variant de 1 à v (nombre de voies).

Un circuit industriel est fréquemment constitué de sous-ensembles plus ou moins interconnectés, chacun d'entre eux ayant une fonction particulière : dans un circuit de flottation par exemple, on distingue le sous-ensemble des "cellules laveuses", celui des "cellules épuiseuses". Chacun de ces sous-ensembles est, à son tour, constitué de nœuds élémentaires (une cellule dans le cas de la flottation) ; chaque nœud est connecté à un certain nombre de branches images des flux de matière transportés.

L'exemple de la figure II-1 représente un circuit élémentaire de transport de matière comportant quatre nœuds et huit voies :

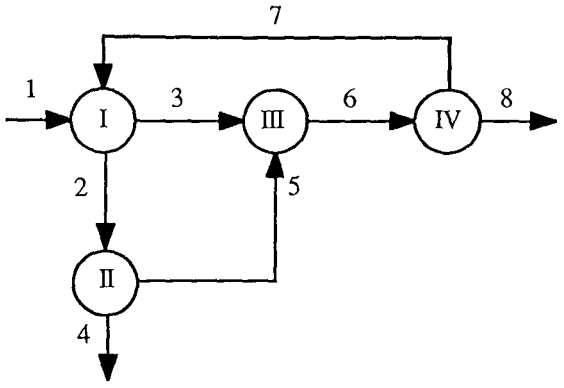



Figure II-1 : circuit élémentaire de transport de matière

Ainsi la matrice d'incidence associée à ce réseau est :

$$M = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}$$

Quelques règles doivent être respectées lors de la construction d'une matrice d'incidence :

- chaque colonne de la matrice comporte au moins un élément non nul et au plus deux éléments ; si une colonne contient deux éléments non nuls, ils sont de signes opposés. En effet, physiquement, une voie est liée au plus à deux nœuds.

- chaque ligne de la matrice d'incidence contient au minimum deux éléments non nuls, de signes opposés. Physiquement, chaque nœud doit comporter au moins une voie "d'entrée" et une voie "de sortie".

De part sa construction, une matrice d'incidence a un rang égal au nombre de nœuds du réseau qu'elle représente.

Les équations de bilan total du réseau de la figure II-1 s'écrivent :

- nœud I $\quad X_1^* - X_2^* - X_3^* + X_7^* = 0$

- nœud II $\quad X_2^* - X_4^* - X_5^* = 0$

- nœud III $\quad X_3^* + X_5^* - X_6^* = 0$

- nœud IV $\quad X_6^* - X_7^* - X_8^* = 0$

En posant $X^*$ le vecteur de dimension 8 formé de $X_i^*$ (i = 1, ..., 8), la représentation matricielle des équations de bilan total s'écrit :

$$MX^* = 0 \qquad \text{(II-1)}$$

où $X^*$ peut représenter le vecteur des débits vrais de dimension v,

et 0 le vecteur de dimension n dont toutes les composantes sont nulles.

### II-1-2 FORME CANONIQUE DE LA MATRICE D'INCIDENCE

La linéarité des équations de bilans rend explicite une partie $X_b$ des variables de débit en fonction des autres $X_h$. Cette partition s'obtient en isolant la partie régulière $M_b$ de M **[MAQU 87]**. Une nouvelle numérotation des voies permet d'écrire :

$$M = [\, M_b \;\mid\; M_h \,] \qquad \text{(II-2)}$$

où $M_b$ est une matrice de dimension (n x n) et $M_h$ une matrice de dimension (n x (v-n)).

$$X = [\, X_b \quad X_h \,]^T \qquad \text{(II-3)}$$

où $X_b$ est un vecteur de dimension (n x 1) et $X_h$ un vecteur de dimension ((v-n) x 1).

On déduit de l'équation de bilan :

$$X_b = - M_b^{-1} M_h X_h \qquad \text{(II-4)}$$

Le vecteur X peut s'exprimer uniquement en fonction de $X_h$ :

$$X = \begin{bmatrix} X_b \\ \\ X_h \end{bmatrix} = \begin{bmatrix} - M_b^{-1} M_h \\ \\ I \end{bmatrix} X_h \qquad \text{(II-5)}$$

Compte tenu de cette décomposition on peut écrire la matrice d'incidence sous la forme équivalente :

$$\bar{M} = [\, I \quad M_b^{-1} M_h \,] \qquad \text{(II-6)}$$

### II-1-3 EXEMPLE D'APPLICATION

Nous décomposons la matrice d'incidence en deux blocs, dont le premier $M_b$ contient les colonnes 1, 2, 3 et 6 et le second bloc $M_h$ contient les autres colonnes :

$$M_b = \begin{bmatrix} 1 & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

et $$M_h = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

Nous obtenons d'après la relation (II-6), l'expression suivante :

$$\bar{M} = \begin{bmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & -1 \end{bmatrix}$$

## II-2 ESTIMATION DE LA VARIANCE PAR MINIMISATION D'UN CRITERE QUADRATIQUE SOUS CONTRAINTES RÉSIDUELLES

Durant les dernières années, plusieurs publications sont apparues sur le problème d'estimation et de validation de données **[MAH 82]**, **[CROW 83]** et **[DARO 86]**. La plupart des techniques utilisées sont basées sur l'hypothèse que les erreurs de mesure sont des variables aléatoires de caractéristiques statistiques connues. Presque toutes les techniques supposent que la matrice de variance des erreurs de mesure V est connue. Très peu d'auteurs ont montré comment ils ont obtenu cette matrice de variance.

Dans ce paragraphe, nous étudions une méthode d'estimation de la variance V qui utilise la matrice de variance-covariance des résidus des équations de bilan obtenue sur plusieurs campagnes de mesure.

Disposant de la matrice de variance-covariance des résidus, comment pourra-t-on extraire au mieux la matrice V ?

Une réponse à cette question a été apportée par Almasy **[ALMA 84]**. Il a présenté une procédure de calcul et des conditions nécessaires pour la résolution du problème que nous avons développé dans ce chapitre.

### II-2-1 FORMULATION DU PROBLEME

Soit $X_i$ le $i^{ème}$ élément d'un vecteur contenant v variables mesurées, alors la variance empirique de m observations de $X_i$ est donnée par l'expression suivante :

$$Var(X_i) = \frac{1}{m-1} \sum_{k=1}^{m} ( X_{ik} - \bar{X}_i )^2 \qquad \text{(II-7)}$$

et la covariance de $X_i$ et $X_j$ est donnée par :

$$Cov(X_i,X_j) = \frac{1}{m-1} \sum_{k=1}^{m} ( X_{ik} - \bar{X}_i )( X_{jk} - \bar{X}_j ) \qquad \text{(II-8)}$$

où $\bar{X}_i$ et $\bar{X}_j$ sont les moyennes respectives de $X_i$ et $X_j$ (i = 1, ..., v et j = 1, ..., v).

Comme nous l'avons indiqué dans le paragraphe précédent, les équations de contrainte qui sont nécessaires pour la validation de données s'écrivent :

$$MX^* = 0 \qquad \text{(II-9)}$$

où M est la matrice d'incidence de dimension n x v.

Dans ce chapitre, nous supposons que les mesures ne sont pas biaisées. Le vecteur des erreurs de mesure est défini par :

$$\varepsilon = X - X^* \qquad \text{(II-10)}$$

où $\varepsilon$ est le vecteur des erreurs de mesure de dimension v.

A partir des expressions (II-9) et (II-10) nous pouvons écrire le vecteur des résidus de dimension n :

$$r = MX = M\varepsilon \qquad \text{(II-11)}$$

Si les mesures ne sont pas biaisées, les moments d'ordre 1 et 2 de r s'écrivent :

$$Esp(r) = MEsp(\varepsilon) = 0 \qquad \text{(II-12)}$$

et

$$Var(r) = MEsp(\varepsilon\varepsilon^T)\, M^T$$

$$= MVM^T$$

$$Var(r) = H \qquad \text{(II-13)}$$

Après avoir calculé les résidus de bilan, nous pouvons exprimer la matrice de variance-covariance de ces résidus. A partir de cette matrice, nous ne pouvons pas déduire directement la matrice V.

Dans le cas où les mesures sont délivrées par des capteurs différents, on suppose que ces mesures sont non corrélées. On notera cependant que cette hypothèse peut ne pas être vérifiée dans le cas où les capteurs comportent des éléments communs tels que l'alimentation. Dans ce cas les mesures peuvent être faiblement corrélées.

En conséquence, l'estimation de V peut être basée sur la minimisation, par rapport aux termes diagonaux (d) et non diagonaux (t) de V, du critère défini par :

$$\begin{cases} \min\limits_{d,t} \Phi = \dfrac{1}{2}\, t^T t \\ \text{Sous la contrainte } H - MVM^T = 0 \end{cases} \qquad \text{(II-14)}$$

### II-2-2 PROCÉDURE DE CALCUL

Pour exposer plus précisément le problème et la procédure de résolution, nous devons utiliser la forme vectorielle des matrices qui sont données à l'annexe A. En utilisant cette notation, l'équation (II-13) peut être écrite :

$$\mathrm{vec}(H) = (M \otimes M)\mathrm{vec}(V) \qquad \text{(II-15)}$$

Puisque H et V sont deux matrices symétriques, nous n'utilisons que la partie triangulaire supérieure ou inférieure (diagonale comprise) pour nos calculs.

Soit **d** le vecteur colonne des éléments diagonaux de V, et **t** le vecteur des éléments non diagonaux, la relation liant les dimensions des deux vecteurs est :

$$\dim(t) = \dim(d)[\dim(d) - 1]/2 \qquad \text{(II-16)}$$

où $\dim(d) = v$

et $\dim(t) = \dfrac{v(v-1)}{2}$

Les vecteurs d et t sont obtenus à partir de vec(V) en multipliant par des matrices appropriées D et T :

$$d = D\ \mathrm{vec}(V) \qquad \text{(II-17)}$$

$$t = T\ \mathrm{vec}(V) \qquad \text{(II-18)}$$

la matrice D permet de sélectionner les termes diagonaux de la matrice de variance et la matrice T permet de sélectionner les termes non diagonaux.

Réciproquement, il est possible d'exprimer V en fonction de d et t par l'expression suivante :

$$\mathrm{vec}(V) = D^{+}d + T^{+}t \qquad \text{(II-19)}$$

avec des matrices $D^{+}$ et $T^{+}$ qui permettent de reconstituer la matrice V à partir des vecteur d et t. Les matrices D, T, $D^{+}$ et $T^{+}$ contiennent des 0 et des 1.

De la même manière, nous pouvons définir un vecteur **p** associé à H tel que :

$$p = P\ \mathrm{vec}(H) \qquad \text{(II-20)}$$

La dimension de p vérifie la relation suivante :

$$\dim(p) = \frac{1}{2}\dim(r)[\dim(r) - 1]$$

où $\dim(p) = \frac{1}{2}n(n-1)$ (II-21)

Par substitution des équations (II-15) et (II-19) dans l'équation (II-20), on a :

$$p = P(M \otimes M)( D^{+}d + T^{+}t )$$

$$p = Ad + Bt \qquad \text{(II-22)}$$

Notons que, pour effectuer le calcul, nous n'avons pas besoin des matrices T, D, $T^{+}$ et $D^{+}$, mais la notation matricielle permet une présentation claire et concise de cette méthode.

En remplaçant l'équation de contrainte (II-13) par l'expression (II-22), le problème d'estimation de la matrice de variance-covariance des erreurs de mesure défini en (II-14) devient :

$$\begin{cases} \min\limits_{d,t} \Phi = \frac{1}{2} t^{T} t \\ \text{Sous la contrainte } p = Ad + Bt \end{cases} \qquad \text{(II-23)}$$

Pour résoudre ce problème d'optimisation, nous appliquons la méthode des multiplicateurs de Lagrange pour minimiser le critère $\Phi$ sous le respect de la contrainte.

Le lagrangien associé à ce problème s'écrit :

$$L = \Phi + \lambda^{T} ( p - Ad - Bt ) \qquad \text{(II-24)}$$

Il est stationnaire quand :

$$\frac{\partial L}{\partial t} = t - B^{T}\lambda = 0 \qquad \text{(II-25-1)}$$

$$\frac{\partial L}{\partial d} = - A^{T}\lambda = 0 \qquad \text{(II-25-2)}$$

$$\frac{\partial L}{\partial \lambda} = p - Ad - Bt = 0 \qquad \text{(II-25-3)}$$

A l'aide des équations (II-25-1) et (II-25-3), on trouve :

$$\lambda = (BB^{T})^{-1}(p - Ad) \qquad \text{(II-26)}$$

On remplace $\lambda$ par son expression dans l'équation (II-25-2) :

$$A^{T}(BB^{T})^{-1}(p - Ad) = 0 \qquad \text{(II-27)}$$

En posant :

$$R = (BB^{T})^{-1} \qquad \text{(II-28)}$$

et

$$S = (A^T R A)^{-1} \tag{II-29}$$

La solution de ce problème est alors donnée par :

$$d = S A^T R p \tag{II-30}$$

et

$$t = B^T (R - R A S A^T R) p \tag{II-31}$$

Les conditions nécessaires pour l'existence de l'inverse de R et S sont :

$$\dim(t) \geq \dim(p) \geq \dim(d) \tag{II-32}$$

Notons que ces conditions par elles-mêmes ne garantissent pas l'existence de R et S. Pour chaque réseau nous avons besoin de calculer les coefficients des matrices $SA^TR$ et $B^T(R - RASA^TR)$ une seule fois et d'appliquer les résultats à l'équation (II-30) et (II-31) à chaque fois qu'un p est obtenu.

Comme vérification de la robustesse de la procédure, nous devons nous attendre à ce qu'elle donne exactement V dans le cas particulier où elle est diagonale. En réalité ce cas peut être montré aisément à partir de l'équation (II-22) appliquée à (II-30) et (II-31).

### II-2-3 ESSAIS DE SIMULATION

La performance de la méthode proposée a été évaluée par simulation sur plusieurs exemples. A partir d'un vecteur de grandeurs réelles vérifiant les équations de bilan, nous lui avons associé un vecteur des erreurs de mesure normalement distribuées avec une matrice de variance spécifiée.

200 points de mesures ont été utilisés pour calculer la matrice de covariance des résidus H qui est utilisée pour estimer la matrice de variance des erreurs de mesure. Cette matrice sera appelée matrice de variance des erreurs de mesure estimées par la méthode indirecte, ou tout simplement l'estimée de V.

#### Premier essai

Le premier essai utilise le réseau de la figure II-1 comportant quatre nœuds et huit voies. Nous avons simulé des mesures de débit massique sur toutes les voies du réseau. Les écarts-type réels des erreurs de mesure $\sigma^*$ sont :

$$\sigma^* = [\ 2.236 \quad 1.115 \quad 2.107 \quad 0.528 \quad 0.606 \quad 2.384 \quad 0.709 \quad 1.627\ ]$$

L'évolution chronologique des mesures pour les huit voies est donnée à la figure II-2.

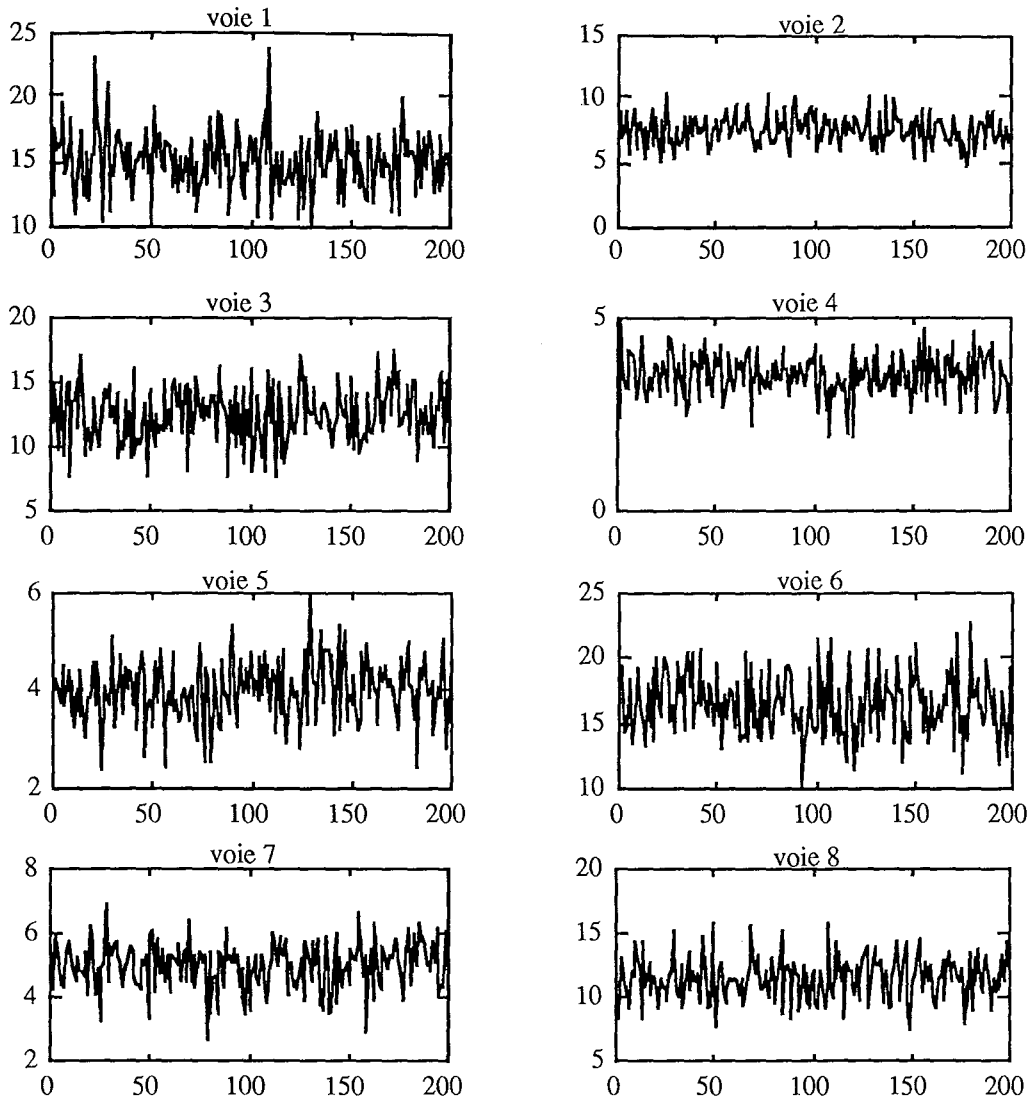



Figure II-2 : évolution temporelle des mesures

Le tableau II-1 représente l'évolution des termes diagonaux de la matrice de variance des erreurs de mesure en fonction du nombre d'observations.

| m | voie 1 | voie 2 | voie 3 | voie 4 | voie 5 | voie 6 | voie 7 | voie 8 |
|---|---|---|---|---|---|---|---|---|
| 50 | 2.944 | 1.272 | 1.272 | 0.974 | 1.182 | 2.138 | 0.461 | 1.969 |
| 100 | 2.598 | 1.182 | 1.870 | 0.693 | 0.926 | 2.127 | 0.548 | 1.621 |
| 150 | 2.593 | 1.126 | 2.081 | 0.507 | 0.789 | 2.382 | 0.590 | 1.783 |
| 200 | 2.392 | 1.171 | 2.029 | 0.573 | 0.815 | 2.300 | 0.506 | 1.819 |

Tableau II-1 : écarts-type des erreurs de mesure

Nous trouvons finalement, pour les 200 observations, les écarts-type estimés suivants :

$\sigma = [\ 2.392 \quad 1.171 \quad 2.029 \quad 0.573 \quad 0.815 \quad 2.300 \quad 0.506 \quad 1.819\ ]$

Pour chaque voie, on calcule un taux de correction défini par :

$$\tau(\%) = \frac{|\text{ valeur réelle - valeur estimée }|}{\text{valeur réelle}} \times 100 \qquad \text{(II-33)}$$

Le vecteur des taux de correction $\tau$ est :

$\tau = [\ 6.97 \quad 5.00 \quad 3.71 \quad 8.51 \quad 34.39 \quad 3.52 \quad 28.64 \quad 11.84\ ]$

La matrice de variance-covariance $H_{mesurée}$ des résidus de bilan calculée directement à l'aide de la matrice des résidus à partir de l'expression (II-13) s'écrit :

$$H_{mesurée} = \begin{bmatrix} 11.01 & -1.23 & -4.28 & 0.39 \\ -1.23 & 1.79 & -0.75 & -0.06 \\ -4.28 & -0.75 & 10.03 & -5.37 \\ 0.39 & -0.06 & -5.37 & 8.43 \end{bmatrix}$$

Après avoir estimé la matrice de variance des erreurs de mesures, nous pouvons déduire la matrice $H_{estimée}$ à l'aide de l'expression (II-13) :

$$H_{estimée} = \begin{bmatrix} 11.47 & -1.37 & -4.12 & -0.26 \\ -1.37 & 2.36 & -0.66 & 0 \\ -4.12 & -0.66 & 10.07 & -5.29 \\ -0.26 & 0 & -5.29 & 8.86 \end{bmatrix}$$

On constate que le taux de correction calculé pour la voie 5 est plus élevé. Ceci peut être expliqué par la faible amplitude de cette mesure et de sa variance.

### Deuxième essai

Pour illustrer l'application de cette méthode sur des réseaux de moyenne taille, nous simulons des données sur le réseau d'une installation chimique représenté par la figure II-3, constitué de onze nœuds et vingt huit voies [**SERT 86**].

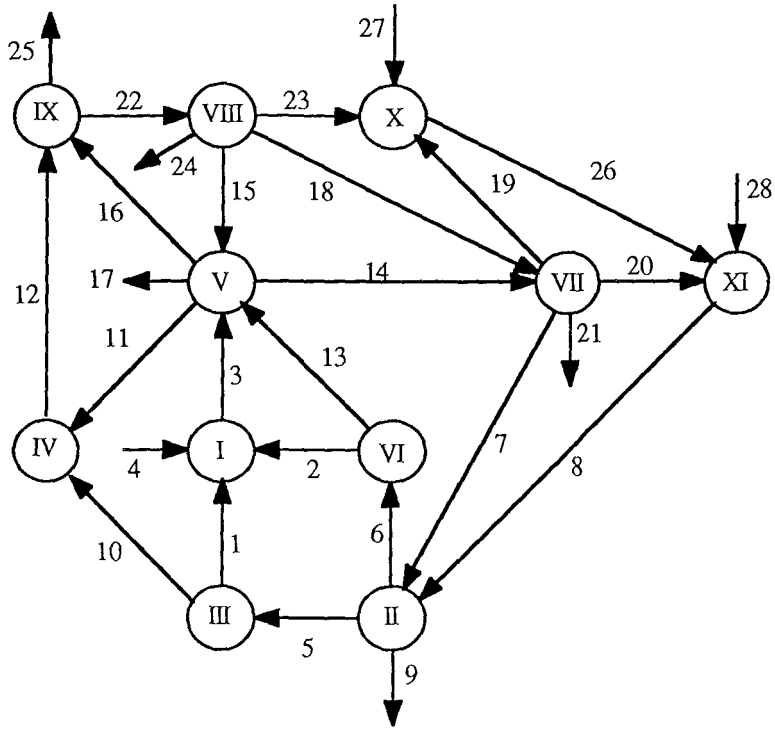



Figure II-3 : réseau de transport de matière

Les évolutions temporelles de quelques mesures sont données aux figures II-4 et II-5.

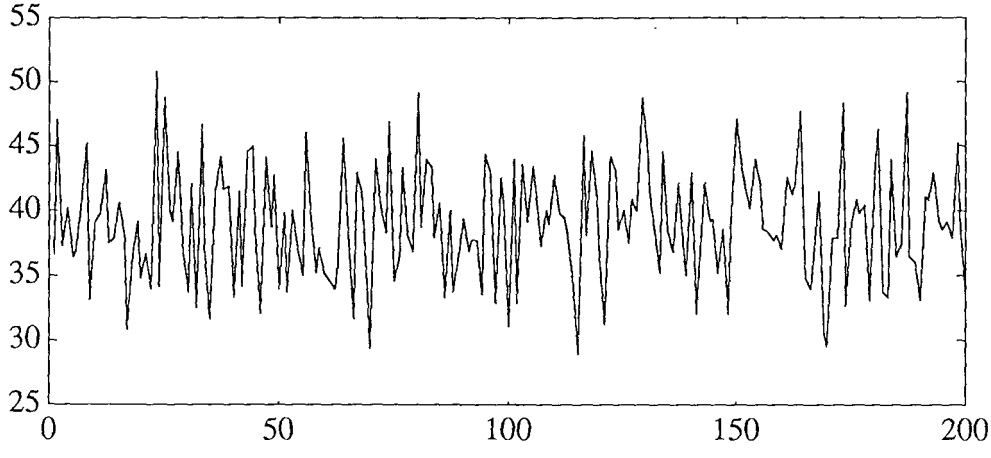



Figure II-4 : évolution temporelle des mesures de la voie 2

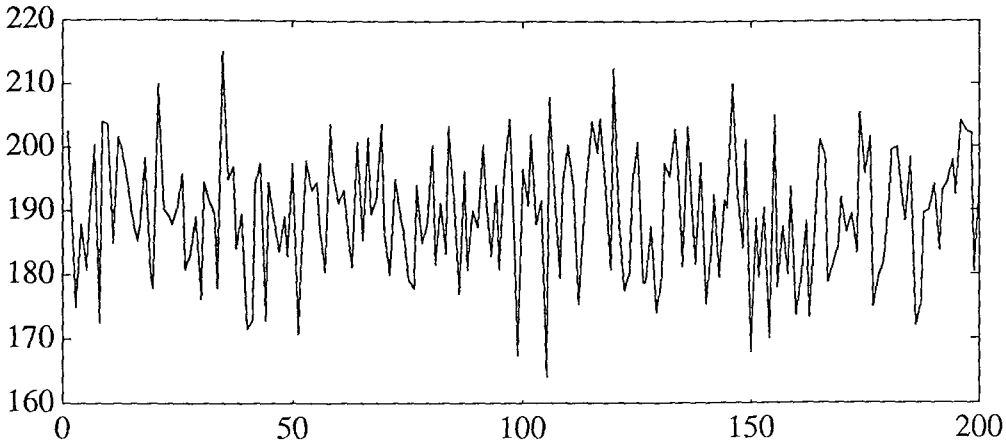



Figure II-5 : évolution temporelle des mesures de la voie 8

Le tableau II-2 représente les écarts-type réels et estimés des erreurs de mesure.

| voie | écarts-type réels | écarts-type estimés | taux de correction |
|---|---|---|---|
| 1 | 4.963 | 5.224 | 5.26 |
| 2 | 4.422 | 3.391 | 9.73 |
| 3 | 8.896 | 8.919 | 0.25 |
| 4 | 5.913 | 5.177 | 12.45 |
| 5 | 6.514 | 5.449 | 16.33 |
| 6 | 5.935 | 6.269 | 5.64 |
| 7 | 3.528 | 4.178 | 18.43 |
| 8 | 9.738 | 9.888 | 1.54 |
| 9 | 5.443 | 5.905 | 8.49 |
| 10 | 4.219 | 5.244 | 24.29 |
| 11 | 5.176 | 3.919 | 24.28 |
| 12 | 6.678 | 6.725 | 0.70 |
| 13 | 3.958 | 4.145 | 4.72 |
| 14 | 4.144 | 4.320 | 4.26 |
| 15 | 5.717 | 5.478 | 4.18 |
| 16 | 3.290 | 3.394 | 3.14 |
| 17 | 2.708 | 3.007 | 11.03 |
| 18 | 3.552 | 3.925 | 10.49 |
| 19 | 1.854 | 2.398 | 29.33 |
| 20 | 2.462 | 2.719 | 10.42 |
| 21 | 2.799 | 4.461 | 59.34 |
| 22 | 3.872 | 4.636 | 17.74 |
| 23 | 2.347 | 2.497 | 6.37 |
| 24 | 5.436 | 5.708 | 5.00 |
| 25 | 6.358 | 6.538 | 2.83 |
| 26 | 6.809 | 6.488 | 4.72 |
| 27 | 6.117 | 6.152 | 0.57 |
| 28 | 6.511 | 6.707 | 3.00 |

Tableau II-2 : grandeurs réelles et estimées

Comme nous l'avons signalé dans le premier essai, nous constatons que les mesures qui ont les plus faibles amplitudes (voies 10, 11, 19 et 21) ont les plus grands taux de correction.

## CONCLUSION

Au travers de nombreux essais de simulation nous avons constaté que lorsque le nombre de mesures est relativement grand, l'estimée de V tend vers la valeur vraie de V. Mais il est important de remarquer que, si nous travaillons avec un nombre d'observations inférieur à 100, les valeurs estimées de V sont, généralement, différentes des valeurs réelles. Par exemple, si la matrice de variance réelle est diagonale, alors la matrice de variance estimée peut avoir des éléments non diagonaux non nuls.

En outre, la méthode proposée ne permet pas d'estimer simultanément les grandeurs réelles et les variances des erreurs de mesure. L'application de cette méthode est étroitement liée à la dimension des réseaux. Ceci nous amène à développer une méthode qui permet de résoudre ces problèmes, qui fait l'objet du chapitre suivant.

# CHAPITRE III

## RECONCILIATION DE DONNEES A VARIANCE PARTIELLEMENT INCONNUE

## INTRODUCTION

Dans ce chapitre, nous appliquerons la technique du maximum de vraisemblance pour estimer la matrice de variance des erreurs de mesure. Pour effectuer cette estimation, nous disposons d'une information partielle sur la matrice de variance des erreurs de mesure ; cette information concerne les rapports de précision entre les différents capteurs.

Nous présentons la réconciliation de données appliquée aux systèmes régis par des équations linéaires et bilinéaires en régime statique.

Concernant la présentation de nos résultats, nous nous intéresserons uniquement aux variances estimées. Dans le cas des systèmes linéaires, nous ferons varier le nombre d'observations afin de voir son influence sur les estimations. Des exemples de simulation seront traités dans le cas linéaire et bilinéaire.

## III-1 ESTIMATION DE VARIANCE DES SYSTEMES LINÉAIRES

### III-1-1 FORMULATION DU PROBLEME

Le problème d'estimation des grandeurs vraies et de la matrice de variance des erreurs de mesure revient à maximiser le logarithme de la fonction de vraisemblance par rapport $\hat{X}$ et V. Ceci revient à minimiser le critère $\Phi$ par rapport à $\hat{X}$ et V :

$$\Phi = \frac{m}{2} \mathrm{Log}|V| + \frac{1}{2} \mathrm{tr}\,(V^{-1} M_R) \qquad \text{(III-1)}$$

sous les contraintes $M\hat{X} = 0$

où $$M_R = \sum_{i=1}^{m} (\hat{X} - X_i)(\hat{X} - X_i)^T \qquad \text{(III-2)}$$

avec

$M_R$ : matrice des moments des carrés des résidus de dimension (v x v),
$X_i$ : vecteur de mesure de dimension v pour l'observation i,
$\hat{X}$ : vecteur des estimées de dimension v,
m : nombre d'observations.

En pratique, la matrice de variance des erreurs de mesure peut s'exprimer en fonction d'un facteur d'échelle et d'une matrice $V_0$ qui donne une information sur le rapport de précision entre les différents capteurs.

$$V = s^2 V_0 \qquad \text{(III-3)}$$

Le déterminant de V s'écrit :

$$|V| = s^{2v} |V_0| \qquad \text{(III-4)}$$

En remplaçant (III-4) dans (III-1), le critère suivant sera minimisé par rapport à s et $\hat{X}$ :

$$\Phi = mv\ \mathrm{Log}(s) + \frac{1}{2s^2}(\mathrm{tr}\ (V_0^{-1}\ M_R)) \qquad \text{(III-5)}$$

Pour résoudre ce problème d'optimisation, nous utilisons la technique des multiplicateurs de Lagrange :

$$L = \Phi + \lambda^T M\hat{X} \qquad \text{(III-6)}$$

où $\lambda$ est le vecteur des multiplicateurs de Lagrange de dimension n.

L est stationnaire quand :

$$\frac{\partial L}{\partial \hat{X}} = \frac{1}{s^2} V_0^{-1} \sum_{i=1}^{m} (\hat{X} - X_i) + M^T\lambda = 0 \qquad \text{(III-7-1)}$$

$$\frac{\partial L}{\partial s} = mv - \frac{1}{s^2}\mathrm{tr}(V_0^{-1}M_R) = 0 \qquad \text{(III-7-2)}$$

$$\frac{\partial L}{\partial \lambda} = M\hat{X} = 0 \qquad \text{(III-7-3)}$$

En général, ce problème n'admet pas de solution analytique. Nous proposons un calcul hiérarchisé basé sur une méthode de relaxation.

### III-1-2 ALGORITHME DE RELAXATION

A partir de l'équation (III-7-2) nous trouvons :

$$s^2 = \frac{1}{mv}\mathrm{tr}\ (V_0^{-1}M_R) \qquad \text{(III-8)}$$

En résolvant le sous-système contenant les équations (III-7-1) et (III-7-2), nous calculons $\lambda$ et $\hat{X}$ en fonction de V :

$$\lambda = m\ (MVM^T)^{-1}M\bar{X} \qquad \text{(III-9)}$$

$$\hat{X} = P\bar{X} \qquad \text{(III-10)}$$

avec $P = (I - VM^T(MVM^T)^{-1}M)$ (III-11)

et $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$

La résolution s'effectue selon une structure hiérarchisée à deux niveaux selon le schéma de la figure III-1.

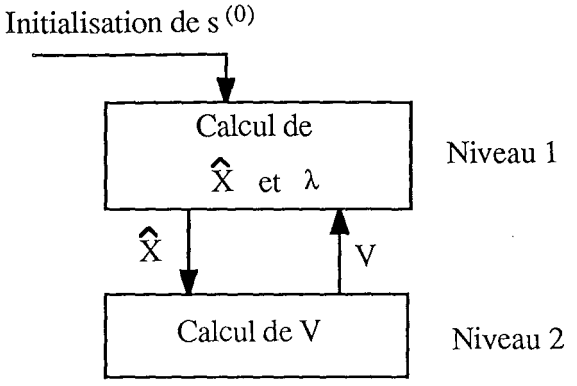



Figure III-1 : schéma du calcul hiérarchisé

L'arrêt de ce calcul itératif s'effectue par un test sur les dérivées premières du lagrangien à partir d'un seuil fixé par l'utilisateur.

### III-1-3 EXEMPLE D'APPLICATION

La figure III-2 représente le schéma simplifié d'un réseau de transport de matière constitué de quatre nœuds et huit voies. Sur chaque voie, 100 observations ont été simulées avec des écarts-type réels regroupés dans le vecteur $\sigma^*$ :

$$\sigma^* = [\ 2.739 \quad 1.937 \quad 2.500 \quad 1.323 \quad 1.414 \quad 2.872 \quad 1.581 \quad 2.398\ ]$$

La valeur réelle du facteur d'échelle est :

$$s^* = 1.7321$$

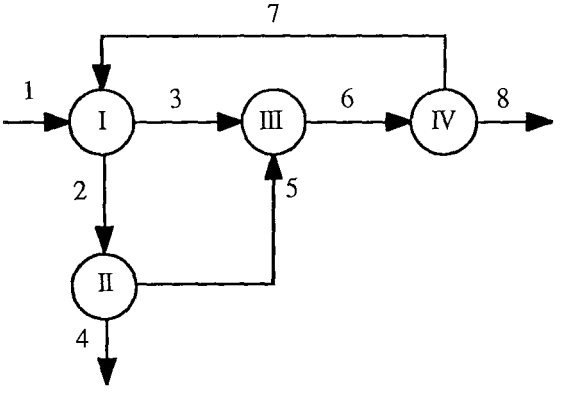



Figure III-2 : réseau élémentaire de transport de matière

A titre indicatif, les évolutions chronologiques des débits des voies 3 et 7 sont présentés aux figures III-3 et III-4.

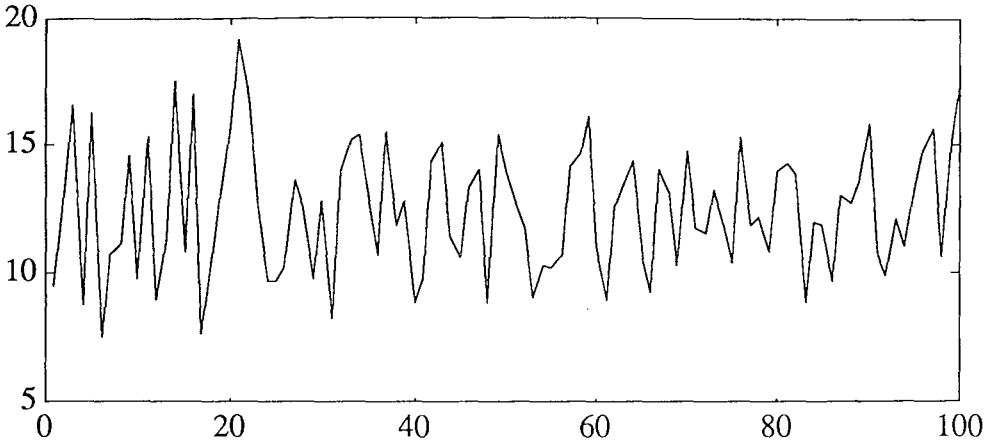



Figure III-3 : évolution chronologique des mesures de la voie 3

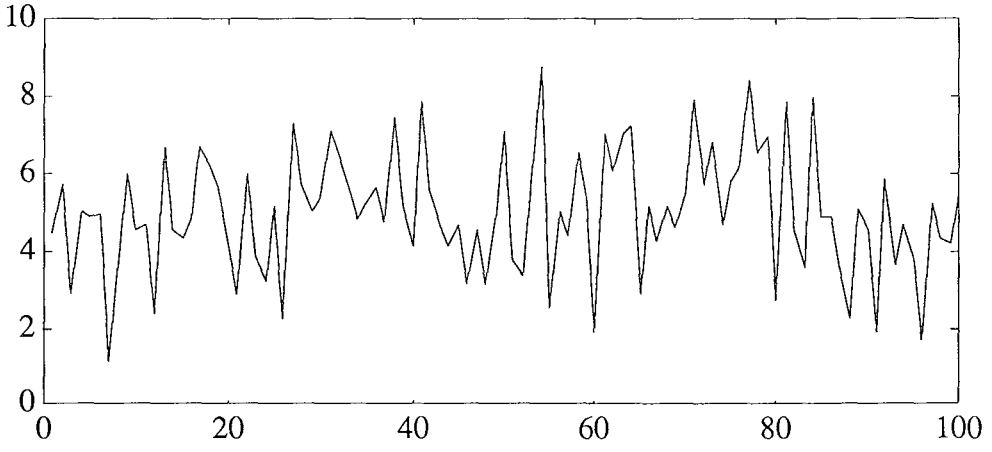



Figure III-4 : évolution chronologique des mesures de la voie 7

Le tableau III-1 montre l'évolution des écarts-type en fonction du nombre d'observations.

| m | voie 1 | voie 2 | voie 3 | voie 4 | voie 5 | voie 6 | voie 7 | voie 8 |
|---|---|---|---|---|---|---|---|---|
| 10 | 2.621 | 1.853 | 2.344 | 1.172 | 1.387 | 2.749 | 1.483 | 2.285 |
| 20 | 2.810 | 1.987 | 2.514 | 1.254 | 1.487 | 2.947 | 1.589 | 2.450 |
| 30 | 2.707 | 1.914 | 2.421 | 1.211 | 1.433 | 2.839 | 1.532 | 2.360 |
| 40 | 2.669 | 1.888 | 2.388 | 1.194 | 1.413 | 2.800 | 1.510 | 2.327 |
| 50 | 2.682 | 1.896 | 2.399 | 1.199 | 1.419 | 2.813 | 1.517 | 2.338 |
| 60 | 2.769 | 1.958 | 2.477 | 1.238 | 1.465 | 2.904 | 1.566 | 2.414 |
| 70 | 2.765 | 1.955 | 2.473 | 1.236 | 1.463 | 2.900 | 1.564 | 2.411 |

Tableau III-1 : écarts-types estimés

Pour 100 observations le vecteur des écarts-type estimés est :

$\sigma = [\ 2.761 \quad 1.952 \quad 2.469 \quad 1.235 \quad 1.461 \quad 2.895 \quad 1.562 \quad 2.407]$

La valeur estimée du facteur d'échelle est :

$s = 1.7460$

Nous constatons que, lorsque le nombre d'observations est relativement grand, les écarts-type des erreurs de mesure s'approchent des valeurs réelles.

## III-2 ESTIMATION DE VARIANCE DES SYSTEMES BILINÉAIRES

### III-2-1 POSITION DU PROBLEME

Dans le cas des systèmes bilinéaires, nous désignerons par $X_i$ les débits massiques et $Y_i$ les teneurs ($i = 1, ..., m$) :

$$X_i = X^* + \varepsilon_{xi} \qquad \text{pour } i = 1, ..., m$$

$$Y_i = Y^* + \varepsilon_{yi} \qquad \text{pour } i = 1, ..., m$$

avec

- $m$ : nombre d'observations,
- $X_i$ : vecteur de mesures des débits à l'observation i,
- $Y_i$ : vecteur de mesures des teneurs à l'observation i,
- $Y^*$ : vecteur des grandeurs réelles de dimension v,
- $X^*$ : vecteur des grandeurs réelles de dimension v,
- $\varepsilon_{xi}$ : vecteur des erreurs de mesure de débits à l'observation i,
- $\varepsilon_{yi}$ : vecteur des erreurs de mesure de teneurs à l'observation i.

Comme dans le cas des systèmes linéaires nous posons :

$$V_x = s_x^2\, V_{x0} \tag{III-12}$$

$$V_y = s_y^2\, V_{y0} \tag{III-13}$$

avec

$V_{x0}$ et $V_{y0}$ matrices diagonales de dimension (v x v),
$s_x$ facteur d'échelle de la variance des débits,
$s_y$ facteur d'échelle de la variance des teneurs.

Le problème d'estimation de la matrice de variance des erreurs de mesure des systèmes bilinéaires revient à minimiser le critère $\Phi$ par rapport à $\hat{X}$, $\hat{Y}$, $s_x$ et $s_y$ :

$$\Phi = mv\ \mathrm{Log}\,(s_x) + mv\ \mathrm{Log}\,(s_y) + \frac{1}{2s_x^2}\mathrm{tr}(V_{x0}^{-1}M_{Rx}) + \frac{1}{2s_y^2}\mathrm{tr}(V_{y0}^{-1}M_{Ry})$$

sous les contraintes linéaires $M\hat{X} = 0$ (III-14)
et bilinéaires $M(\hat{X} * \hat{Y}) = 0$

où $M_{Rx} = \sum_{i=1}^{m} (\hat{X} - X_i)(\hat{X} - X_i)^T]$

et $M_{Ry} = \sum_{i=1}^{m} (\hat{Y} - Y_i)(\hat{Y} - Y_i)^T]$

Le lagrangien associé à ce problème s'écrit :

$$L = \Phi + \lambda^T M\hat{X} + \mu^T M(\hat{X} * \hat{Y}) \qquad \text{(III-15)}$$

il est stationnaire quand :

$$\frac{\partial L}{\partial \hat{X}} = \frac{1}{s_x^2} V_{x0}^{-1} \sum_{i=1}^{m} (\hat{X} - X_i) + M^T\lambda + (M * \hat{Y})^T\mu = 0 \qquad \text{(III-16-1)}$$

$$\frac{\partial L}{\partial \hat{Y}} = \frac{1}{s_y^2} V_{y0}^{-1} \sum_{i=1}^{m} (\hat{Y} - Y_i) + (M * \hat{X})^T\mu = 0 \qquad \text{(III-16-2)}$$

$$\frac{\partial L}{\partial s_x} = mv - \frac{1}{s_x^2} tr(V_{x0}^{-1} M_{Rx}) = 0 \qquad \text{(III-16-3)}$$

$$\frac{\partial L}{\partial s_y} = mv - \frac{1}{s_y^2} tr(V_{y0}^{-1} M_{Ry}) = 0 \qquad \text{(III-16-4)}$$

$$\frac{\partial L}{\partial \lambda} = M\hat{X} = 0 \qquad \text{(III-16-5)}$$

$$\frac{\partial L}{\partial \mu} = M(\hat{X} * \hat{Y}) = 0 \qquad \text{(III-16-6)}$$

$\lambda$ et $\mu$ sont les vecteurs des multiplicateurs de Lagrange de dimension n.

Pour la résolution de ce système non linéaire nous proposons un calcul hiérarchisé basé sur la méthode de relaxation.

### III-2-2 RÉSOLUTION HIÉRARCHISÉE

Comme dans le cas des systèmes linéaires, nous utilisons la méthode de résolution à itération directe.

Des équations (III-16-3) et (III-16-4), nous pouvons déduire les expressions de $s_x$ et $s_y$ en fonction de $\hat{X}$ et $\hat{Y}$.

$$s_x = \frac{1}{mv} tr(V_{x0}^{-1} M_{Rx}) \qquad \text{(III-17)}$$

et

$$s_y = \frac{1}{mv} tr(V_{y0}^{-1} M_{Ry}) \qquad \text{(III-18)}$$

A l'aide des équations (III-16-1) et (III-16-5), nous pouvons exprimer $\lambda$ et $\hat{X}$ :

$$\lambda = (MV_xM^T)^{-1}M[m\,\bar{X} - V_x(M * \hat{Y})^T\mu] \qquad \text{(III-19)}$$

$$\hat{X} = P_x[\bar{X} - \frac{1}{m}V_x(M * \hat{Y})^T\mu] \qquad \text{(III-20)}$$

avec

$$P_x = I - V_xM^T(MV_xM^T)^{-1}M \qquad \text{(III-21)}$$

et $$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i \qquad \text{(III-22)}$$

A l'aide des équations (III-16-2) et (III-16-6), nous pouvons exprimer $\mu$ et $\hat{Y}$ :

$$\mu = m\,[(M * \hat{X})V_y(M * \hat{X})^T]^{-1}(M * \hat{X})\,\bar{Y} \qquad \text{(III-23)}$$

$$\hat{Y} = P_y\bar{Y} \qquad \text{(III-24)}$$

avec

$$P_y = I - V_y(M * \hat{X})^T[(M * \hat{X})V_y(M * \hat{X})^T]^{-1}(M * \hat{X}) \qquad \text{(III-25)}$$

et $$\bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i \qquad \text{(III-26)}$$

$P_x$ et $P_y$ sont des matrices de projection de dimension (v x v).

La résolution du système non linéaire s'effectue alors selon une structure hiérarchisée à deux niveaux :

- le premier niveau calcule $\lambda$ et $\hat{X}$ à partir de $s_x$ puis détermine la valeur de $s_x$,

- le deuxième niveau calcule $\mu$ et $\hat{Y}$ à partir de $s_y$ et de l'estimation $\hat{X}$ transmise par le premier niveau puis détermine la valeur de $s_y$.

### III-2-3 EXEMPLES D'APPLICATIONS

#### Première application

Nous reprenons le réseau de la figure III-2 sur lequel nous simulons les mêmes mesures de débits et nous rajoutons les mesures de teneurs. Les écarts-type réels des débits et des teneurs sont :

$$\sigma_1^* = [\ 2.739 \quad 1.937 \quad 2.500 \quad 1.323 \quad 1.414 \quad 2.872 \quad 1.581 \quad 2.398\ ]$$

$$\sigma_2^* = [\ 0.500 \quad 0.424 \quad 0.458 \quad 0.289 \quad 0.515 \quad 0.473 \quad 0.212 \quad 0.548\ ]$$

Les valeurs réelles des facteurs d'échelles sont :

$$s_x^* = 1.7321$$

$$s_y^* = 1.2247$$

Les évolutions temporelles des teneurs des voies 3 et 7 sont données aux figures III-5 et III-6 (les évolutions temporelles des débits de ces voies ayant été données aux figures III-3 et III-4).

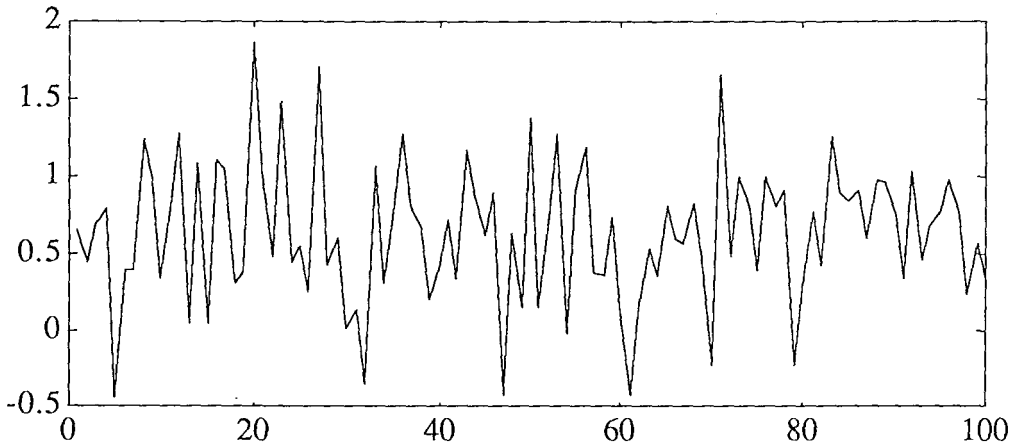



Figure III-5 : évolution temporelle des mesures de teneur de la voie 3

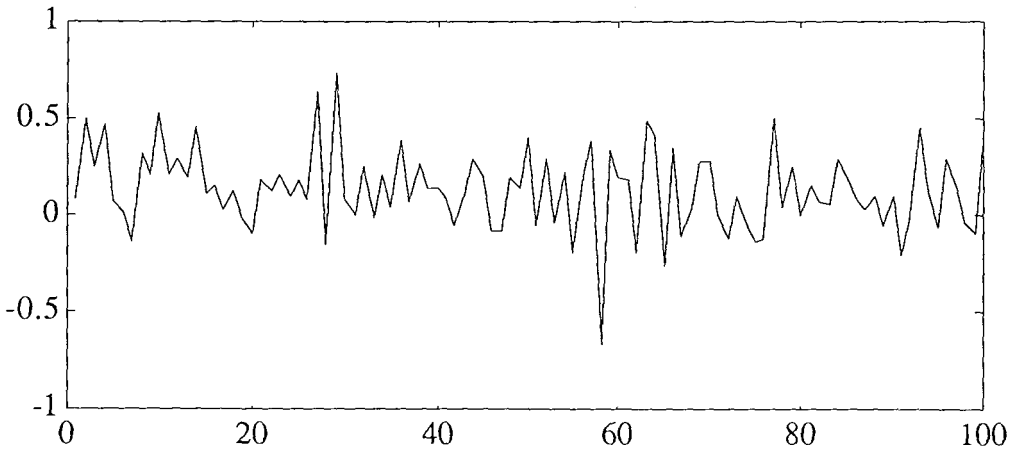



Figure III-6 : évolution temporelle des mesures de teneur de la voie 7

Les résultats obtenus pour une centaine d'observations sont :

$\sigma_1 = [\ 2.725 \quad 1.927 \quad 2.487 \quad 1.316 \quad 1.407 \quad 2.858 \quad 1.573 \quad 2.386\ ]$

$\sigma_2 = [\ 0.497 \quad 0.422 \quad 0.456 \quad 0.287 \quad 0.512 \quad 0.470 \quad 0.211 \quad 0.546\ ]$

Les valeurs estimées des facteurs d'échelles sont :

$s_x = 1.7234$

$s_y = 1.2186$

Nous constatons que les estimations obtenues par cette méthode sont très proches des grandeurs réelles.

## Deuxième application

Dans cette application, nous considérons un réseau de transport de matière de taille moyenne formé de sept nœuds et treize voies (figure III-7). Cent observations ont été simulées sur toutes les voies de ce réseau.

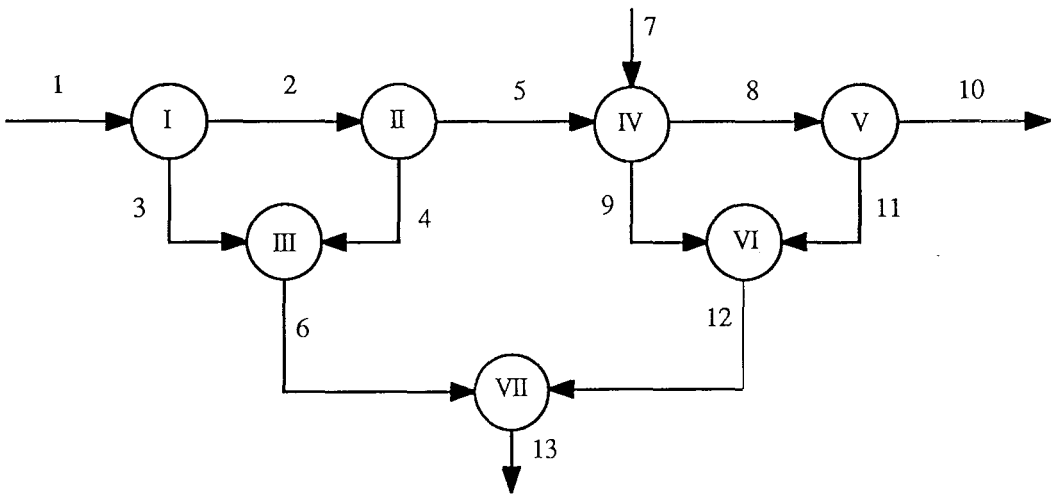



Figure III-7 : réseau de transport de matière

Les valeurs réelles des facteurs d'échelles sont :

$s_x^* = 1.5811$

$s_y^* = 1.1832$

A titre d'exemple, les évolutions chronologiques du débit et de la teneur de la voie 8 sont présentées respectivement aux figures III-8 et III-9.

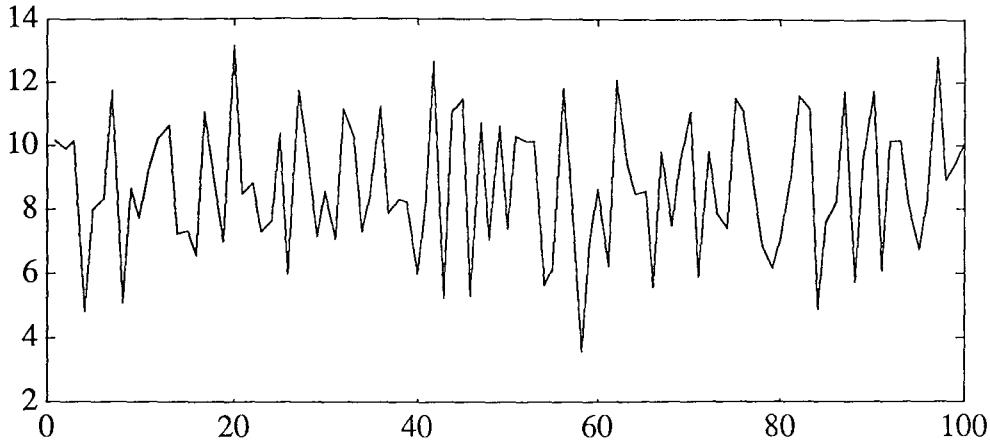



Figure III-8 : évolution chronologique des mesures de débit de la voie 8

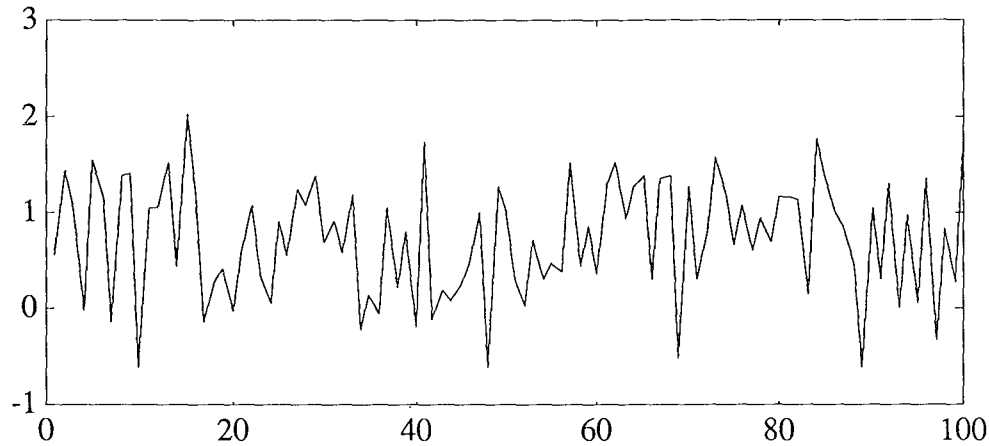



Figure III-9 : évolution chronologique des mesures de teneur de la voie 8

Les résultats de la simulation et de l'estimation des écarts-type sont donnés dans le tableau III-2 pour 100 observations.

| voie | écarts-type des débits | | écarts-type des teneurs | |
|---|---|---|---|---|
| | réels | estimés | réels | estimés |
| 1 | 3.571 | 3.553 | 0.659 | 0.656 |
| 2 | 2.806 | 2.792 | 0.648 | 0.645 |
| 3 | 2.208 | 2.197 | 0.678 | 0.674 |
| 4 | 1.696 | 1.687 | 0.706 | 0.702 |
| 5 | 2.236 | 2.225 | 0.612 | 0.609 |
| 6 | 2.784 | 2.769 | 0.688 | 0.684 |
| 7 | 1.275 | 1.268 | 0.510 | 0.507 |
| 8 | 2.092 | 2.081 | 0.607 | 0.604 |
| 9 | 1.500 | 1.493 | 0.552 | 0.549 |
| 10 | 1.458 | 1.450 | 0.524 | 0.522 |
| 11 | 1.500 | 1.493 | 0.676 | 0.672 |
| 12 | 2.121 | 2.111 | 0.617 | 0.614 |
| 13 | 3.500 | 3.483 | 0.663 | 0.659 |

Tableau III-2 : écarts-type réels et estimés

Les valeurs estimées des facteurs d'échelles sont :

$s_x = 1.5732$

et $s_y = 1.1773$

En comparant les valeurs réelles des facteurs d'échelles des débits et des teneurs respectivement avec leurs estimées, nous remarquons que les écarts relatifs entre ces valeurs sont faibles.

## CONCLUSION

Les résultats que nous avons obtenu montrent que la méthode se révèle efficace pour l'estimation des variances. D'autres essais que nous n'avons pas présentés ont montré qu'une légère variation autour des valeurs de la matrice (± 10% des valeurs initiales) représentant le rapport de précision entre les différents capteurs, modifie peu les résultats des estimations, tandis qu'une grande variation des valeurs initiales de cette matrice altère les résultats des estimations.

Pour une bonne estimation des variances, la connaissance de la matrice $V_0$ est fondamentale. Dans le cas où celle-ci est inconnue, cette méthode ne peut s'appliquer, on aura alors recours à d'autres techniques telle que celle que nous présentons dans le chapitre suivant.

# CHAPITRE IV

## RECONCILIATION DE DONNEES DES SYSTEMES MULTI-LINEAIRES A VARIANCE INCONNUE

## INTRODUCTION

Dans ce chapitre, nous appliquerons la technique du maximum de vraisemblance pour estimer la matrice de variance des erreurs de mesure et les estimées des grandeurs réelles. Pour effectuer cette estimation, nous supposerons que les erreurs de mesure suivent une loi normale de moyenne nulle et de variance constante.

L'estimation de ces grandeurs ne se fera que si nous disposons d'une série chronologique de mesures (plusieurs campagnes).

Dans un premier temps, nous présentons la réconciliation de données appliquée aux systèmes linéaires en régime statique dans le cas où l'ensemble des variables est complètement observé. Une extension de cette méthode aux cas des systèmes partiellement observés sera ensuite formulée.

Dans un deuxième temps, nous étendons cette méthode aux systèmes bilinéaires complètement ou partiellement mesurés. Nous considérons le cas d'un seul constituant et de plusieurs constituants.

Enfin, nous généralisons la méthode aux systèmes multi-linéaires complètement mesurés. La complexité de la formulation de cette méthode est telle que nous formaliserons le problème pour un système trilinéaire puis, nous généraliserons les équations obtenues aux systèmes multi-linéaires.

Pour prouver l'efficacité de ces méthodes, des exemples numériques sont exposés. Nous simulerons des mesures évoluant autour d'un ou plusieurs points de fonctionnement.

## IV-1 RECONCILIATION DE DONNEES DES SYSTEMES LINEAIRES

### IV-1-1 FORMULATION DU PROBLEME

Après avoir réalisé plusieurs campagnes de mesure, nous obtenons une matrice d'observation X contenant les mesures brutes représentant l'état du système pendant la durée des campagnes de mesure.

A l'instant i, le vecteur des mesures $X_i$ de dimension v s'exprime en fonction du vecteur des grandeurs réelles X* et du vecteur des erreurs de mesure $\varepsilon_i$ :

$$X_i = X^* + \varepsilon_i \qquad \text{pour } i = 1, ..., m \qquad \text{(IV-1)}$$

où m est le nombre d'observations.

D'après les hypothèses faites sur les erreurs de mesure (cf. chapitre I), la fonction de vraisemblance de ces erreurs est définie par :

$$F(\varepsilon_i) = (2\pi)^{-\frac{vm}{2}} |V|^{-\frac{m}{2}} \exp(-\frac{1}{2} tr(V^{-1} M_R)) \qquad \text{(IV-2)}$$

où $tr(V^{-1} M_R)$ représente la trace du produit matriciel $(V^{-1} M_R)$.

La matrice d'observations est constituée de plusieurs sous-matrices correspondant aux différents régimes statiques. Par conséquent, la matrice des moments des résidus, de dimension v x v, s'écrit :

$$M_R = \sum_{j=1}^{p} \sum_{i=1}^{m_j} [(\hat{X}_j - X_{ij}) (\hat{X}_j - X_{ij})^T] \qquad \text{(IV-3)}$$

où $\sum_{j=1}^{p} m_j = m$

avec

$m_j$ : nombre d'observations de la $j^{ème}$ zone de fonctionnement statique,
$p$ : nombre total des zones de fonctionnement statique,
$\hat{X}_j$ : vecteur des estimées qui correspond à la $j^{ème}$ zone de fonctionnement,
$X_{ij}$ : vecteur de mesures de dimension v correspondant à la $j^{ème}$ zone et à l'observation i.

Dans la suite de notre étude, nous noterons V la variance estimée et non $\hat{V}$.

Le problème d'estimation des grandeurs vraies et de la matrice de variance des erreurs de mesure revient à maximiser le logarithme de la fonction de vraisemblance par rapport à $\hat{X}_j$ et V. Ceci revient à minimiser le critère $\Phi$ par rapport à $\hat{X}_j$ et V :

$$\Phi = \frac{m}{2} \mathrm{Log}|V| + \frac{1}{2} \sum_{j=1}^{p} \sum_{i=1}^{m_j} \|\hat{X}_j - X_{ij}\|^2_{V^{-1}} \qquad \text{(IV-4)}$$

Sous les contraintes $\qquad M\hat{X}_j = 0 \qquad$ pour j = 1, ..., p

Pour résoudre ce problème d'optimisation, nous utilisons la technique des multiplicateurs de Lagrange. Définissons le lagrangien :

$$L = \Phi + \sum_{j=1}^{p} \lambda_j^T M\hat{X}_j \qquad \text{(IV-5)}$$

où $\lambda_j$ est le vecteur des multiplicateurs de Lagrange de dimension n correspondant à la $j^{ème}$ zone de fonctionnement statique.

Les conditions de stationnarité du lagrangien pour j = 1, ..., p sont :

$$\frac{\partial L}{\partial \hat{X}_j} = V^{-1} \sum_{i=1}^{m_j} (\hat{X}_j - X_{ij}) + M^T \lambda_j = 0 \qquad \text{(IV-6-1)}$$

$$\frac{\partial L}{\partial V} = \frac{m}{2} V^{-1} - \frac{1}{2} V^{-1} \mathrm{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (\hat{X}_j - X_{ij})(\hat{X}_j - X_{ij})^T] V^{-1} = 0 \qquad \text{(IV-6-2)}$$

$$\frac{\partial L}{\partial \lambda_j} = M\hat{X}_j \qquad = 0 \qquad \text{(IV-6-3)}$$

En général, ce problème n'admet pas de solution analytique. Nous proposons un calcul hiérarchisé basé sur une méthode de relaxation pour résoudre le système non linéaire précédent.

### IV-1-2 ALGORITHME DE RELAXATION

Parmi les algorithmes de résolution utilisant le principe de la relaxation, nous utilisons celui à itération directe **[RAGO 90a]**.

De l'équation (IV-6-2), nous pouvons déduire l'expression de V en fonction de $\hat{X}_j$ :

$$V = \frac{1}{m}\mathrm{diag}[\sum_{j=1}^{p}\sum_{i=1}^{m_j}(\hat{X}_j - X_{ij})(\hat{X}_j - X_{ij})^T] \qquad \text{(IV-7)}$$

A partir de l'équation (IV-6-1), nous obtenons :

$$\hat{X}_j = \frac{1}{m_j}[\sum_{i=1}^{m_j} X_{ij} - VM^T\lambda_j] \qquad \text{pour } j = 1,\dots, p \qquad \text{(IV-8)}$$

En remplaçant $\hat{X}_j$ par son expression dans l'équation (IV-6-3), nous obtenons :

$$\lambda_j = (MVM^T)^{-1}M\sum_{i=1}^{m_j} X_{ij} \qquad \text{pour } j = 1,\dots, p \qquad \text{(IV-9)}$$

Nous trouvons finalement en reportant IV-9 dans IV-8 :

$$\hat{X}_j = P\bar{X}_j \qquad \text{pour } j = 1,\dots, p \qquad \text{(IV-10)}$$

avec

$$P = I_v - VM^T(MVM^T)^{-1}M \qquad \text{(IV-11)}$$

et

$$\bar{X}_j = \frac{1}{m_j}\sum_{i=1}^{m_j} X_{ij} \qquad \text{pour } j = 1,\dots, p \qquad \text{(IV-12)}$$

où

P est la matrice de projection de dimension v x v,

$I_v$ la matrice identité de dimension v,

$\bar{X}_j$ la moyenne des mesures pour les p zones de fonctionnement statique.

La résolution du système non linéaire s'effectue alors selon une structure hiérarchisée à deux niveaux :

- le premier niveau calcule $\lambda_j$ et $\hat{X}_j$ à partir de V initialisée à la matrice identité,

- le deuxième niveau calcule V à l'aide de l'estimation $\hat{X}_j$ transmise par le premier niveau.

Le schéma du calcul est donné par la figure IV-1.

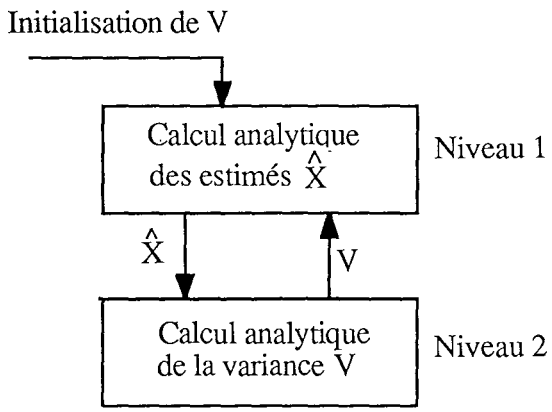



Figure IV-1 : algorithme de relaxation

Les deux niveaux de calcul sont repris jusqu'à convergence de la solution ; le test d'arrêt est effectué sur les dérivées première du lagrangien à partir d'un seuil fixé par l'utilisateur (choisi dans nos exemple à $10^{-5}$).

*Remarque*

Dans le cas de systèmes de grande dimension, l'application directe de l'équation (IV-10) pour le calcul de l'estimation peut présenter quelques difficultés numériques dues à la dimension de la matrice de projection et à son conditionnement. La forme particulièrement creuse de la matrice M peut être utilisée avec profit pour résoudre ce type de difficulté. Une autre technique consiste à utiliser un calcul récursif pour calculer l'estimation $\hat{X}_j$ par ajout progressif de contraintes **[RAGO 89]**.

## IV-1-3 APPLICATIONS NUMÉRIQUES

### Application à un réseau de petite dimension

Considérons le réseau de transport de matière représenté par la figure IV-2 comportant quatre nœuds et huit voies :

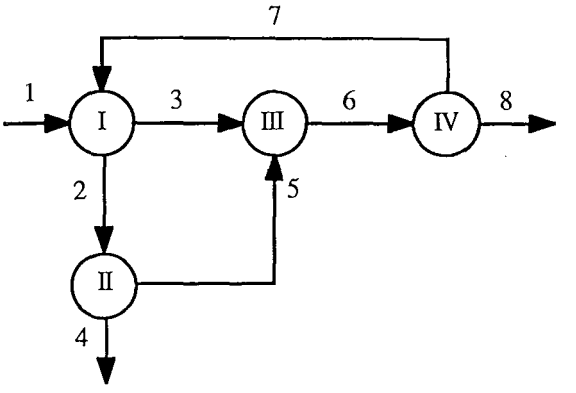



Figure IV-2 : réseau de transport de matière

Dans une première application, nous simulerons des mesures de débit massique sur toutes les voies du réseau de la figure IV-2. Ces mesures sont générées à partir de grandeurs vérifiant les équations de bilan et de bruits gaussiens de moyennes nulles et de variances constantes. La figure IV-3 indique, pour les 8 voies et au cours du temps, les variations des différentes mesures.

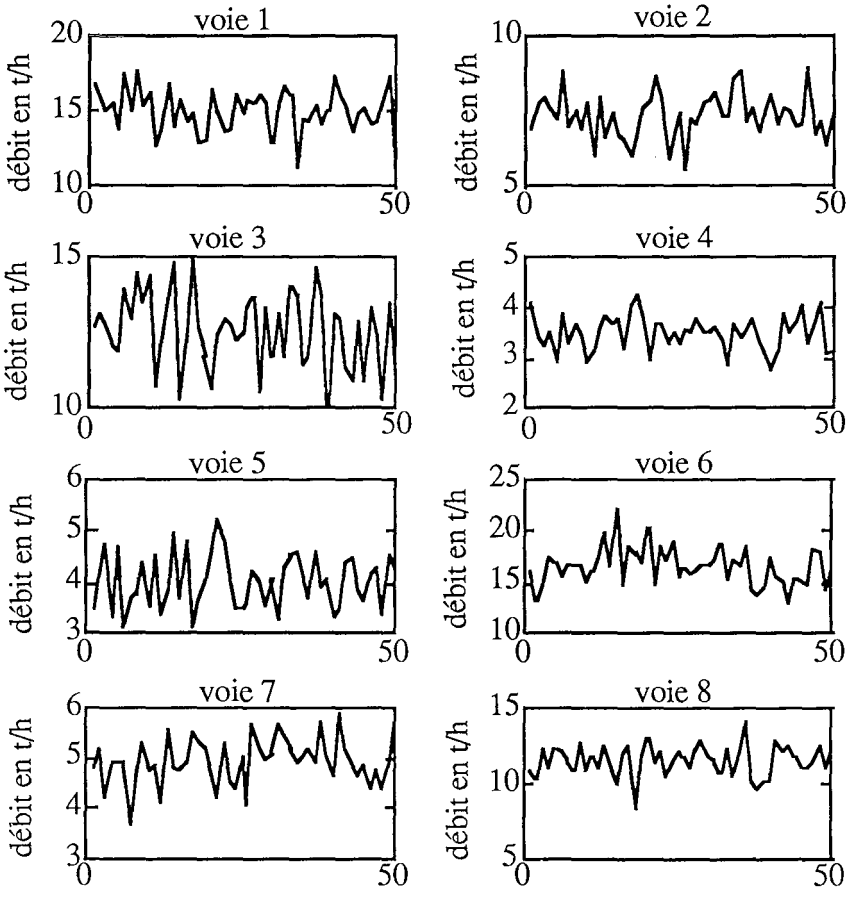



Figure IV-3 : évolution chronologique des débits massiques

Les résultats de l'estimation des grandeurs réelles et des écarts-type des erreurs de mesure sont présentés dans le tableau IV-1.

| voie | grandeur réelle | grandeur estimée | écart-type réel | écart-type estimé |
|---|---|---|---|---|
| 1 | 15.00 | 15.04 | 1.436 | 1.422 |
| 2 | 7.50 | 7.44 | 0.748 | 0.748 |
| 3 | 12.50 | 12.54 | 1.245 | 1.233 |
| 4 | 3.50 | 3.49 | 0.347 | 0.345 |
| 5 | 4.00 | 3.95 | 0.526 | 0.525 |
| 6 | 16.50 | 16.49 | 1.836 | 1.818 |
| 7 | 5.00 | 4.95 | 0.452 | 0.448 |
| 8 | 11.50 | 11.54 | 1.002 | 0.993 |

Tableau IV-1 : grandeurs réelles et estimées

Pour la deuxième application, relative au même réseau, nous avons simulé quatre zones de fonctionnement. Sur la figure IV-4, la première zone est comprise entre l'observation 1 et 10, la deuxième entre 11 et 25, la troisième entre 26 et 40 et la quatrième entre 41 et 50.

Les écarts-type réels des erreurs de mesure sont :

$$\sigma^* = [\ 2.94 \quad 0.91 \quad 2.35 \quad 0.18 \quad 0.45 \quad 5.69 \quad 0.29 \quad 1.60\ ]$$

La matrice des grandeurs réelles (formée de 8 lignes représentant les variables du réseau et de 4 colonnes représentant les 4 points de fonctionnement) qui a servi à générer les mesures est :

$$X_{\text{réelle}} = \begin{bmatrix} 15.00 & 22.50 & 17.25 & 15.00 \\ 7.50 & 11.25 & 8.62 & 7.50 \\ 12.50 & 18.75 & 14.38 & 12.50 \\ 3.50 & 5.25 & 4.03 & 3.50 \\ 4.00 & 6.00 & 4.60 & 4.00 \\ 16.50 & 24.75 & 18.98 & 16.50 \\ 5.00 & 7.50 & 5.75 & 5.00 \\ 11.50 & 17.25 & 13.23 & 11.50 \end{bmatrix}$$

L'évolution temporelle des mesures est représentée par la figure IV-4.

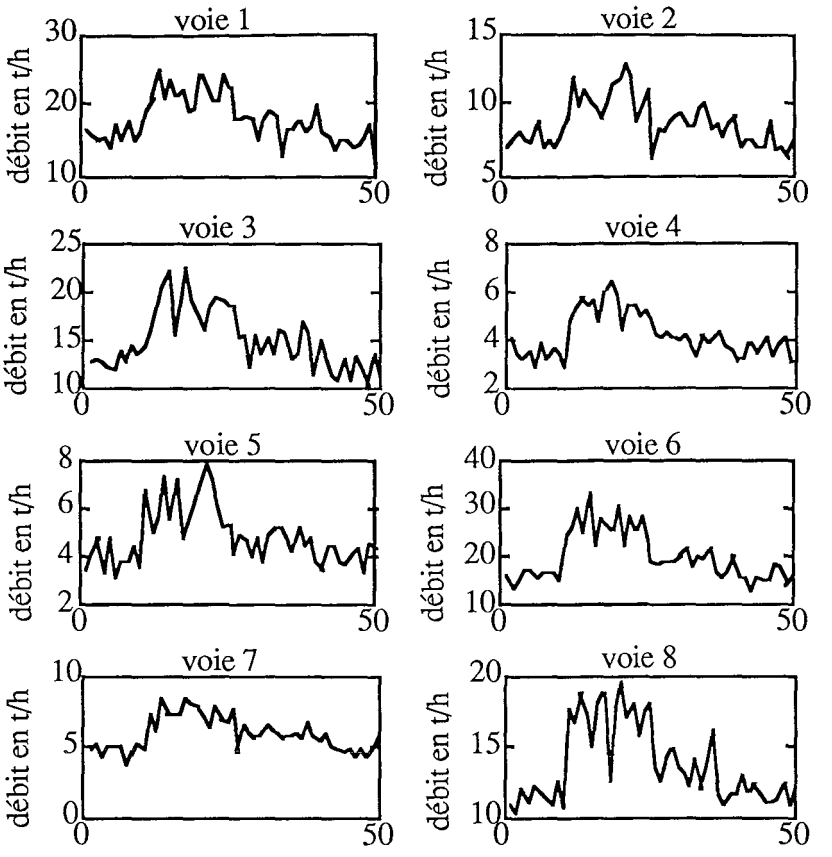



Figure IV-4 : évolution chronologique des débits massiques

Au bout de 5 itérations, nous obtenons les estimées $\hat{X}$ et les écarts-type des erreurs de mesure des quatre zones de fonctionnement. Les écarts-type estimés des erreurs de mesure sont :

$$\sigma = [\ 2.79 \quad 0.91 \quad 2.21 \quad 0.17 \quad 0.45 \quad 5.62 \quad 0.27 \quad 1.62\ ]$$

et la matrice des grandeurs estimées :

$$\hat{X} = \begin{bmatrix} 15.38 & 22.53 & 17.17 & 14.89 \\ 7.39 & 11.15 & 8.57 & 7.51 \\ 12.76 & 18.80 & 14.49 & 12.27 \\ 3.51 & 5.24 & 3.97 & 3.53 \\ 3.88 & 5.90 & 4.59 & 3.97 \\ 16.64 & 24.70 & 19.08 & 16.25 \\ 4.77 & 7.41 & 5.88 & 4.89 \\ 11.87 & 17.29 & 13.20 & 11.36 \end{bmatrix}$$

Nous constatons des écarts très faibles entre les valeurs réelles des grandeurs et leurs estimées ce qui montre l'efficacité de la méthode proposée.

**Application à un réseau de dimension moyenne [SERT 86]**

Nous considérons le réseau de la figure IV-5 comportant onze nœuds et vingt huit voies sur lesquelles nous avons simulé cinquante observations évoluant autour de trois points de fonctionnement. La première plage de fonctionnement est comprise entre l'observation 1 et 20, la deuxième entre 21 et 35 et la troisième entre 36 et 50.

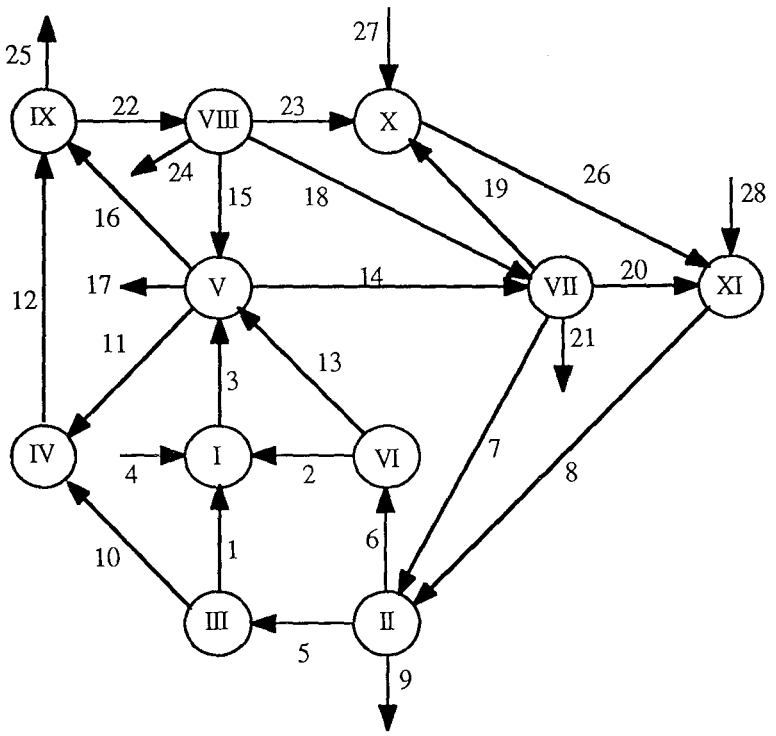



Figure IV-5 : réseau d'une installation chimique

Les résultats de la simulation des débits massiques ainsi que les écarts-type des erreurs de mesure sont présentés dans le tableau IV-2.

| | débit massique réel | | | débit massique estimé | | | écart-type | |
|---|---|---|---|---|---|---|---|---|
| voie | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 | réel | estimé |
| 1 | 49.26 | 73.88 | 56.64 | 49.32 | 73.67 | 56.67 | 1.41 | 1.39 |
| 2 | 39.10 | 58.65 | 44.96 | 39.01 | 58.79 | 44.98 | 1.25 | 1.24 |
| 3 | 158.29 | 237.43 | 182.03 | 158.34 | 237.28 | 182.06 | 5.07 | 5.06 |
| 4 | 69.93 | 104.89 | 80.42 | 70.01 | 104.82 | 80.36 | 2.13 | 2.12 |
| 5 | 84.86 | 127.29 | 97.59 | 84.88 | 127.25 | 97.57 | 2.43 | 2.42 |
| 6 | 70.43 | 105.65 | 81.00 | 70.31 | 105.85 | 81.03 | 2.13 | 2.11 |
| 7 | 24.89 | 37.34 | 28.63 | 24.89 | 37.33 | 28.64 | 0.83 | 0.82 |
| 8 | 189.65 | 284.47 | 218.09 | 189.63 | 284.54 | 218.07 | 5.58 | 5.55 |
| 9 | 59.25 | 88.87 | 68.13 | 59.33 | 88.77 | 68.09 | 1.96 | 1.96 |
| 10 | 35.60 | 53.40 | 40.94 | 35.56 | 53.58 | 40.86 | 1.17 | 1.16 |
| 11 | 53.59 | 80.38 | 61.62 | 53.74 | 79.92 | 61.75 | 1.66 | 1.63 |
| 12 | 89.19 | 133.79 | 102.57 | 89.30 | 133.50 | 102.61 | 2.90 | 2.91 |
| 13 | 31.33 | 47.01 | 36.03 | 31.30 | 47.06 | 36.05 | 0.98 | 0.97 |
| 14 | 34.33 | 51.50 | 39.49 | 34.35 | 51.58 | 39.42 | 1.09 | 1.08 |
| 15 | 65.37 | 98.06 | 75.18 | 65.19 | 98.31 | 75.26 | 1.92 | 1.90 |
| 16 | 21.65 | 32.48 | 24.90 | 21.64 | 32.56 | 24.85 | 0.69 | 0.68 |
| 17 | 14.67 | 22.00 | 16.87 | 14.71 | 21.97 | 16.82 | 0.42 | 0.42 |
| 18 | 25.23 | 37.85 | 29.01 | 25.20 | 37.76 | 29.12 | 0.72 | 0.71 |
| 19 | 6.87 | 10.31 | 7.90 | 6.83 | 10.31 | 7.96 | 0.21 | 0.20 |
| 20 | 12.13 | 18.19 | 13.94 | 12.16 | 18.22 | 13.85 | 0.34 | 0.34 |
| 21 | 15.68 | 23.51 | 18.03 | 15.65 | 23.48 | 18.08 | 0.54 | 0.53 |
| 22 | 29.98 | 44.97 | 34.47 | 29.84 | 44.97 | 34.68 | 1.00 | 0.99 |
| 23 | 11.02 | 16.53 | 12.67 | 10.99 | 16.55 | 12.70 | 0.31 | 0.30 |
| 24 | 59.10 | 88.65 | 67.96 | 58.84 | 88.97 | 68.11 | 1.88 | 1.85 |
| 25 | 80.86 | 121.29 | 92.99 | 81.10 | 121.08 | 92.78 | 2.40 | 2.38 |
| 26 | 92.72 | 139.09 | 106.63 | 92.89 | 139.02 | 106.42 | 2.85 | 2.82 |
| 27 | 75.82 | 112.24 | 86.05 | 75.06 | 112.15 | 85.76 | 2.01 | 1.97 |
| 28 | 84.79 | 127.20 | 97.52 | 84.57 | 127.30 | 97.78 | 2.58 | 2.57 |

Tableau IV-2 : grandeurs réelles et estimées

L'analyse des résultats présentés au tableau IV-2 permet de constater que, pour chaque plage de fonctionnement statique, les estimées des écarts-type et des débits massiques sont très proches des valeurs réelles. La convergence est obtenue au bout d'une dizaine d'itérations ce qui est faible compte tenu de la dimension du réseau.

Pour pouvoir comparer la méthode exposée dans le second chapitre (méthode 1) avec celle que nous venons de développer (méthode 2), le même exemple de la figure II-1, les mêmes mesures et les mêmes écarts-type des erreurs de mesure ont été employés. Nous avons regroupé les résultats obtenus par ces deux méthodes dans le tableau IV-3.

| voie | écarts-type réels | écarts-type estimés méthode 1 | méthode 2 |
|---|---|---|---|
| 1 | 2.236 | 2.392 | 2.231 |
| 2 | 1.115 | 1.171 | 1.112 |
| 3 | 2.107 | 2.029 | 2.102 |
| 4 | 0.528 | 0.573 | 0.526 |
| 5 | 0.606 | 0.815 | 0.605 |
| 6 | 2.384 | 2.300 | 2.378 |
| 7 | 0.709 | 0.506 | 0.707 |
| 8 | 1.627 | 1.819 | 1.623 |

Tableau IV-3 : résultats des deux méthodes

L'analyse de ce tableau montre clairement que certains résultats obtenus par la méthode 1 sont plus écartés des valeurs réelles que ceux obtenus par la méthode 2.

### IV-1-4 EXTENSION AUX SYSTEMES PARTIELLEMENT OBSERVÉS

Nous pouvons étendre la méthode précédente pour estimer la variance et les grandeurs réelles des systèmes globalement observables et partiellement observés. L'introduction d'un modèle de mesure qui, dans notre cas particulier, se présente sous la forme d'une matrice de sélection des voies mesurées autorise une approche globale du problème.

On considère toujours un réseau à n nœuds et v voies (dont seulement k sont mesurées), décrit par sa matrice d'incidence M. Le vecteur des mesures s'exprime en fonction du vecteur des grandeurs vraies :

$$X = HX^* + \varepsilon \qquad \text{(IV-13)}$$

avec

X : vecteur de mesures de dimension k,
X*: vecteur des grandeurs vraies de dimension v,
ε : vecteur des erreurs de mesure de dimension k,
H : matrice de sélection de dimension k x v.

Le processus de mesure étant supposé direct, la matrice H traduit l'absence ou la présence d'une mesure sur chaque voie du réseau [**DARO 86**].

Le problème d'estimation des variances et des grandeurs vraies consiste à chercher le minimum par rapport à $\hat{X}_j$ et V de :

$$\Phi = \frac{m}{2}\mathrm{Log}|V| + \frac{1}{2}\sum_{j=1}^{p}\sum_{i=1}^{m_j} \|H\hat{X}_j - X_{ij}\|^2_{V^{-1}} \qquad \text{(IV-14)}$$

Sous les contraintes $M\hat{X}_j = 0$ pour j = 1, ..., p

Le lagrangien associé à ce problème :

$$L = \Phi + \sum_{j=1}^{p} \lambda_j^T M\hat{X}_j \qquad \text{(IV-15)}$$

est stationnaire pour $j = 1, ..., p$ quand :

$$\frac{\partial L}{\partial \hat{X}_j} = H^T V^{-1} \sum_{i=1}^{m_j} (H\hat{X}_j - X_{ij}) + M^T \lambda_j = 0 \qquad \text{(IV-16-1)}$$

$$\frac{\partial L}{\partial V} = \frac{m}{2} V^{-1} - \frac{1}{2} V^{-1} \operatorname{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (H\hat{X}_j - X_{ij})(H\hat{X}_j - X_{ij})^T] V^{-1} = 0 \qquad \text{(IV-16-2)}$$

$$\frac{\partial L}{\partial \lambda_j} = M\hat{X}_j = 0 \qquad \text{(IV-16-3)}$$

Comme précédemment, nous proposons un calcul hiérarchisé basé sur la méthode de relaxation pour résoudre ce système non linéaire.

Le système d'équations (IV-16) admet une solution unique si et seulement si la matrice $(H^T V^{-1} H + M^T M)$ est régulière. Cette condition d'observabilité globale du système se résume par **[DARO 86]** :

$$\operatorname{rang}\begin{bmatrix} H \\ M \end{bmatrix} = v \qquad \text{(IV-17)}$$

Si cette condition n'est pas satisfaite, il faut procéder à l'extraction de la partie observable du système.

De l'équation (IV-16-2), nous pouvons déduire l'expression de V en fonction de $\hat{X}_j$.

$$V = \frac{1}{m} \operatorname{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (H\hat{X}_j - X_{ij})(H\hat{X}_j - X_{ij})^T] \qquad \text{(IV-18)}$$

En posant :

$$G = m_j H^T V^{-1} H + M^T M \qquad \text{(IV-19)}$$

on déduit de (IV-16-1) et (IV-16-2) :

$$\hat{X}_j = G^{-1} [H^T V^{-1} \sum_{i=1}^{m_j} X_{ij} - M^T \lambda_j] \qquad \text{pour } j = 1, .., p \qquad \text{(IV-20)}$$

En remplaçant $\hat{X}_j$ par son expression dans l'équation (IV-16-3), nous obtenons :

$$\lambda_j = (MG^{-1}M^T)^{-1}MG^{-1}H^TV^{-1}\sum_{i=1}^{m_j} X_{ij} \qquad \text{pour } j = 1, .., p \qquad \text{(IV-21)}$$

En substituant $\lambda_j$ dans $\hat{X}_j$, nous trouvons :

$$\hat{X}_j = m_j\, P\, G^{-1}\, H^T\, V^{-1}\bar{X}_j \qquad \text{pour } j = 1, .., p \qquad \text{(IV-22)}$$

avec

$$P = I_v - G^{-1}M^T(MG^{-1}M^T)^{-1}M \qquad \text{(IV-23)}$$

et

$$\bar{X}_j = \frac{1}{m_j}\sum_{i=1}^{m_j} X_{ij} \qquad \text{pour } j = 1, .., p \qquad \text{(IV-24)}$$

Pour résoudre ce système non linéaire, nous appliquons l'algorithme de relaxation utilisé précédemment en prenant pour matrice de variance initiale la matrice identité.

## Application au réseau de la figure IV-2

Dans un premier temps, considérons les mesures de débit dont l'évolution chronologique est donnée par la figure IV-3. Pour notre étude, nous ne tenons pas compte des mesures effectuées sur la voie 8 ; celle-ci est considérée comme non mesurée. Les résultats de l'estimation des grandeurs réelles et les écarts-type des erreurs de mesures sont présentés dans le tableau IV-4. On obtient des résultats tout à fait comparables à ceux obtenus précédemment.

| voie | grandeur estimée | écart-type estimé |
|---|---|---|
| 1 | 15.05 | 1.422 |
| 2 | 7.44 | 0.748 |
| 3 | 12.56 | 1.233 |
| 4 | 3.49 | 0.345 |
| 5 | 3.95 | 0.525 |
| 6 | 16.51 | 1.818 |
| 7 | 4.95 | 0.448 |
| 8 | 11.56 | 0.997 |

Tableau IV-4 : grandeurs estimées

Dans un deuxième temps, nous ne disposons pas des mesures de débit sur les voies 7 et 8. Les résultats de l'estimation des grandeurs réelles et des écarts-type des erreurs de mesure sont présentés dans le tableau IV-5.

| voie | grandeur estimée | écart-type estimé |
|---|---|---|
| 1 | 15.00 | 1.422 |
| 2 | 7.44 | 0.748 |
| 3 | 12.58 | 1.233 |
| 4 | 3.49 | 0.345 |
| 5 | 3.95 | 0.524 |
| 6 | 16.54 | 1.818 |
| 7 | 5.02 | 0.452 |
| 8 | 11.51 | 1.015 |

Tableau IV-5 : grandeurs estimées

On constate donc que la méthode s'applique bien au cas des systèmes partiellement observés mais observables et que l'estimation de la variance est peu sensible, pour cet exemple, à l'absence de certaines mesures.

## IV-2 RECONCILIATION DE DONNEES DES SYSTEMES BILINEAIRES

Pour des systèmes de transport de matière ou d'énergie, on utilise souvent des débits massiques, des débits volumiques, des teneurs, des densités... Certaines de ces grandeurs sont liées entre elles par des équations linéaires (débits massiques), bilinéaires (débits et teneurs) et généralement multi-linéaires.

### IV-2-1 POSITION DU PROBLEME

Dans le cas des systèmes bilinéaires, nous désignerons par $X_i$ les débits massiques ($i = 1, ..., m$) et $Y_i$ les teneurs :

$$X_i = X^* + \varepsilon_{xi} \qquad \text{pour } i = 1, ..., m \qquad \text{(IV-25)}$$

$$Y_i = Y^* + \varepsilon_{yi} \qquad \text{pour } i = 1, ..., m \qquad \text{(IV-26)}$$

avec

- $m$ : nombre d'observations,
- $X_i$ : vecteur de mesure des débits à l'observation i,
- $Y_i$ : vecteur de mesure des teneurs à l'observation i,
- $Y^*$ : vecteur des grandeurs réelles de dimension v,
- $X^*$ : vecteur des grandeurs réelles de dimension v,
- $\varepsilon_{xi}$ : vecteur des erreurs de mesure de débits à l'observation i,
- $\varepsilon_{yi}$ : vecteur des erreurs de mesure de teneurs à l'observation i.

D'après les hypothèses citées précédemment, la fonction de densité de probabilité des erreurs de mesure est définie par :

$$F(\varepsilon_{xi},\varepsilon_{yi}) = (2\pi)^{-\frac{vm}{2}} |V_x|^{-\frac{m}{2}} |V_y|^{-\frac{m}{2}} \exp(-\frac{1}{2}\sum_{i=1}^{m} \| \varepsilon_{xi} \|^2_{V_x^{-1}} + \sum_{i=1}^{m} \| \varepsilon_{yi} \|^2_{V_y^{-1}})) \tag{IV-27}$$

où $V_x$ et $V_y$ sont respectivement les matrices de variance des débits et des teneurs et $|V_x|$ et $|V_y|$ les déterminants de $V_x$ et $V_y$.

Le problème d'estimation de la matrice de variance des erreurs de mesure des systèmes bilinéaires revient à minimiser le critère $\Phi$ par rapport à $\hat{X}_i$, $\hat{Y}_i$ , $V_x$ et $V_y$ :

$$\Phi = \frac{m}{2}\mathrm{Log}|V_x| + \frac{m}{2}\mathrm{Log}|V_y| + \frac{1}{2}\sum_{j=1}^{p}\sum_{i=1}^{m_j} \|\hat{X}_j - X_{ij}\|^2_{V_x^{-1}} + \sum_{j=1}^{p}\sum_{i=1}^{m_j} \|\hat{Y}_j - Y_{ij}\|^2_{V_y^{-1}}$$

sous les contraintes linéaires $M\hat{X}_j = 0$ (IV-28)

et bilinéaires $M(\hat{X}_j * \hat{Y}_j) = 0$ pour $j = 1, ..., p$

avec

$$\sum_{j=1}^{p} m_j = m$$

où $m_j$ est le nombre d'observations de la $j^{ème}$ plage de fonctionnement et p, le nombre total de ces plages.

Le lagrangien associé à ce problème s'écrit :

$$L = \Phi + \sum_{j=1}^{p} \lambda_j^T M\hat{X}_j + \sum_{j=1}^{p} \mu_j^T M(\hat{X}_j * \hat{Y}_j) \tag{IV-29}$$

il est stationnaire pour $j = 1, ..., p$ quand :

$$\frac{\partial L}{\partial \hat{X}_j} = V_x^{-1} \sum_{i=1}^{m_j}(\hat{X}_j - X_{ij}) + M^T\lambda_j + (M * \hat{Y}_j)^T\mu_j = 0 \tag{IV-30-1}$$

$$\frac{\partial L}{\partial \hat{Y}_j} = V_y^{-1} \sum_{i=1}^{m_j}(\hat{Y}_j - Y_{ij}) + (M * \hat{X}_j)^T\mu_j = 0 \tag{IV-30-2}$$

$$\frac{\partial L}{\partial V_x} = \frac{m}{2}V_x^{-1} - \frac{1}{2}V_x^{-1}\mathrm{diag}[\sum_{j=1}^{p}\sum_{i=1}^{m_j}(\hat{X}_j - X_{ij})(\hat{X}_j - X_{ij})^T]\, V_x^{-1} = 0 \tag{IV-30-3}$$

$$\frac{\partial L}{\partial V_y} = \frac{m}{2} V_y^{-1} - \frac{1}{2} V_y^{-1} \mathrm{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (\hat{Y}_j - Y_{ij})(\hat{Y}_j - Y_{ij})^T] \, V_y^{-1} = 0 \qquad \text{(IV-30-4)}$$

$$\frac{\partial L}{\partial \lambda_j} = M\hat{X}_j = 0 \qquad \text{(IV-30-5)}$$

$$\frac{\partial L}{\partial \mu_j} = M(\hat{X}_j * \hat{Y}_j) = 0 \qquad \text{(IV-30-6)}$$

$\lambda_j$ et $\mu_j$ sont les vecteurs des multiplicateurs de Lagrange de dimension n correspondant à la $j^{ème}$ plage de fonctionnement.

En général, ce problème n'admet pas de solution analytique. Pour la résolution du système non linéaire précédent, nous proposons un calcul hiérarchisé basé sur la méthode de relaxation.

## IV-2-2 RÉSOLUTION DANS LE CAS BILINÉAIRE SIMPLE

Comme dans le cas des systèmes linéaires, nous utilisons la méthode de résolution à itération directe.

Des équations (IV-30-3) et (IV-30-4), nous pouvons déduire les expressions de $V_x$ et $V_y$ en fonction de $\hat{X}_j$ et $\hat{Y}_j$.

$$V_x = \frac{1}{m} \mathrm{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (\hat{X}_j - X_{ij})(\hat{X}_j - X_{ij})^T] \qquad \text{(IV-31)}$$

et

$$V_y = \frac{1}{m} \mathrm{diag}[\sum_{j=1}^{p} \sum_{i=1}^{m_j} (\hat{Y}_j - Y_{ij})(\hat{Y}_j - Y_{ij})^T] \qquad \text{(IV-32)}$$

A l'aide des équations (IV-30-1) et (IV-30-5), nous pouvons exprimer $\lambda_j$ et $\hat{X}_j$ :

$$\lambda_j = (MV_xM^T)^{-1}M[m_j \bar{X}_j - V_x(M * \hat{Y}_j)^T \mu_j] \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-33)}$$

$$\hat{X}_j = P_x[\bar{X}_j - \frac{1}{m_j} V_x(M * \hat{Y}_j)^T \mu_j] \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-34)}$$

avec

$$P_x = I - V_xM^T(MV_xM^T)^{-1}M \qquad \text{(IV-35)}$$

et

$$\bar{X}_j = \frac{1}{m_j} \sum_{i=1}^{m_j} X_{ij} \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-36)}$$

où $\bar{X}_j$ est la moyenne de $X_{ij}$ correspondant au $j^{ème}$ régime statique.

A l'aide des équations (IV-30-2) et (IV-30-6), nous pouvons exprimer $\mu_j$ et $\hat{Y}_j$ :

$$\mu_j = m_j [(M * \hat{X}_j)V_y(M * \hat{X}_j)^T]^{-1}(M * \hat{X}_j)\, \bar{Y}_j \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-37)}$$

$$\hat{Y}_j = P_y \bar{Y}_j \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-38)}$$

avec

$$P_y = I - V_y(M * \hat{X}_j)^T[(M * \hat{X}_j)V_y(M * \hat{X}_j)^T]^{-1}(M * \hat{X}_j) \qquad \text{(IV-39)}$$

et $$\bar{Y}_j = \frac{1}{m_j}\sum_{i=1}^{m_j} Y_{ij} \qquad \text{pour } j = 1, ..., p \qquad \text{(IV-40)}$$

$P_x$ et $P_y$ sont des matrices de projection de dimension (v x v),

et $\bar{X}_j$ et $\bar{Y}_j$ sont les moyennes respectives des mesures $X_j$ et $Y_j$ pour les p points de fonctionnement.

La résolution du système non linéaire s'effectue alors selon une structure hiérarchisée à deux niveaux :

- le premier niveau calcule $\lambda_j$ et $\hat{X}_j$ à partir de $V_x$ (initialisée à la matrice identité),

- le deuxième niveau calcule $\mu_j$ et $\hat{Y}_j$ à partir de $V_y$ et de l'estimation $\hat{X}_j$ transmise par le premier niveau.

A la fin de chaque niveau, nous calculons les matrices de variance $V_x$ et $V_y$.

La méthode de relaxation que nous proposons pour résoudre ce problème d'optimisation utilise un algorithme à itération directe dont le schéma est donné par la figure IV-6.

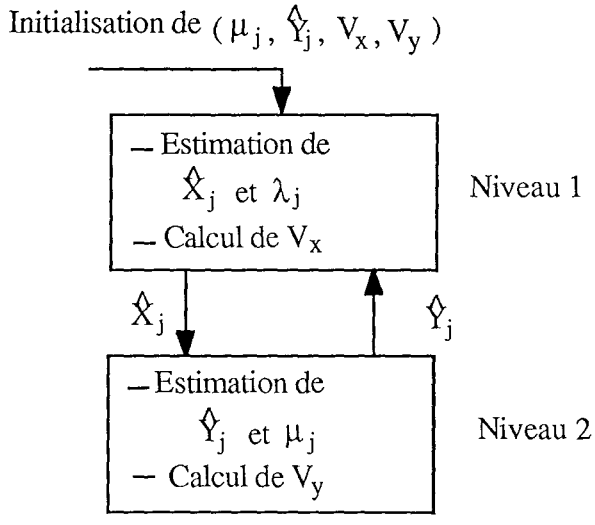



Figure IV-6 : algorithme de relaxation

## Applications numériques

Dans le premier essai, nous présentons deux applications ; pour la première, les mesures évoluent sur une seule plage de fonctionnement statique et pour la deuxième, les mesures sont réparties sur trois plages de fonctionnement statique.

La méthode a ensuite été appliquée à un réseau de plus grande dimension que celui utilisé lors du premier essai. Les mesures varient autour d'un seul point de fonctionnement statique

## Premier essai

Considérons le réseau de transport de matière représenté par la figure IV-7, comportant quatre nœuds et huit voies :

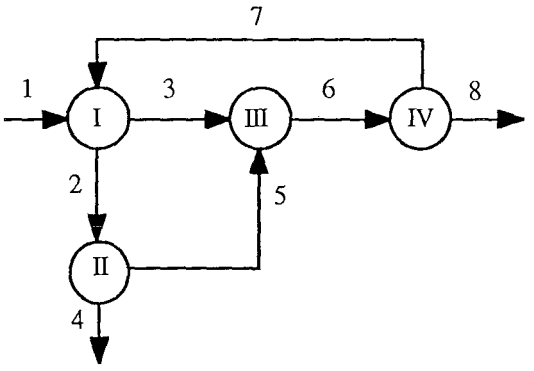



Figure IV-7 : réseau de transport de matière

Pour illustrer l'utilisation de la méthode précédente, nous simulerons des mesures des débits massiques et des teneurs sur toutes les voies autour d'un seul point de fonctionnement. L'évolution de quelque mesures de débits massiques et de teneurs est présentée par les figures IV-8 et IV-9.

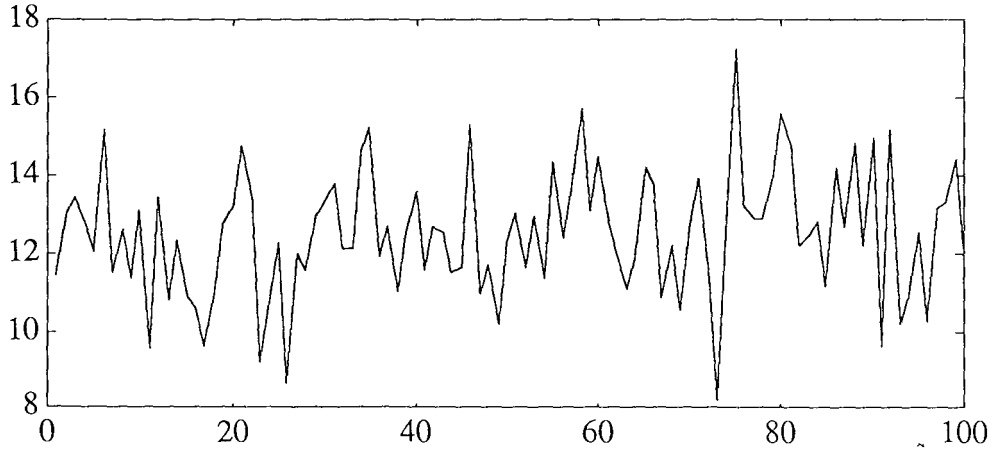



Figure IV-8 : évolution chronologique des débits de la voie 3

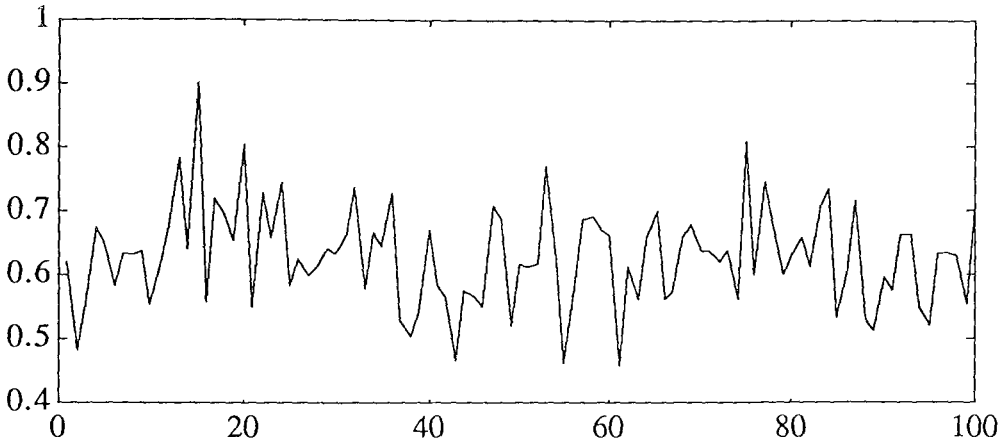



Figure IV-9 : évolution chronologique des teneurs de la voie 3

Les résultats de l'estimation des grandeurs réelles et des écarts-type des erreurs de mesures sont présentés dans les tableaux IV-6 et IV-7.

| voie | débits réels | teneurs réelles | écart-type des débits | écart-type des teneurs |
|---|---|---|---|---|
| 1 | 15.00 | 0.75 | 1.708 | 0.102 |
| 2 | 7.50 | 0.54 | 0.951 | 0.070 |
| 3 | 12.50 | 0.63 | 1.621 | 0.077 |
| 4 | 3.50 | 0.25 | 0.414 | 0.031 |
| 5 | 4.00 | 0.79 | 0.458 | 0.096 |
| 6 | 16.50 | 0.67 | 2.020 | 0.079 |
| 7 | 5.00 | 0.13 | 0.635 | 0.016 |
| 8 | 11.50 | 0.90 | 1.291 | 0.112 |

Tableau IV-6 : grandeurs réelles

| voie | débits estimés | teneurs estimées | écart-type des débits | écart-type des teneurs |
|---|---|---|---|---|
| 1 | 15.01 | 0.74 | 1.702 | 0.105 |
| 2 | 7.49 | 0.53 | 0.955 | 0.071 |
| 3 | 12.51 | 0.64 | 1.623 | 0.075 |
| 4 | 3.50 | 0.26 | 0.414 | 0.032 |
| 5 | 4.02 | 0.79 | 0.458 | 0.095 |
| 6 | 16.51 | 0.68 | 2.020 | 0.079 |
| 7 | 5.00 | 0.13 | 0.635 | 0.016 |
| 8 | 11.52 | 0.90 | 1.290 | 0.114 |

Tableau IV-7 : grandeurs estimées

Nous appliquerons la même procédure au réseau de la figure IV-8 sur lequel nous avons simulé trois régimes de fonctionnement, le premier entre les observations 1 et 20, le deuxième entre 20 et 40 et le troisième entre 40 et 50. Les évolutions chronologiques des débits et des teneurs sont représentées par les graphiques des figures IV-10 et IV-11 :

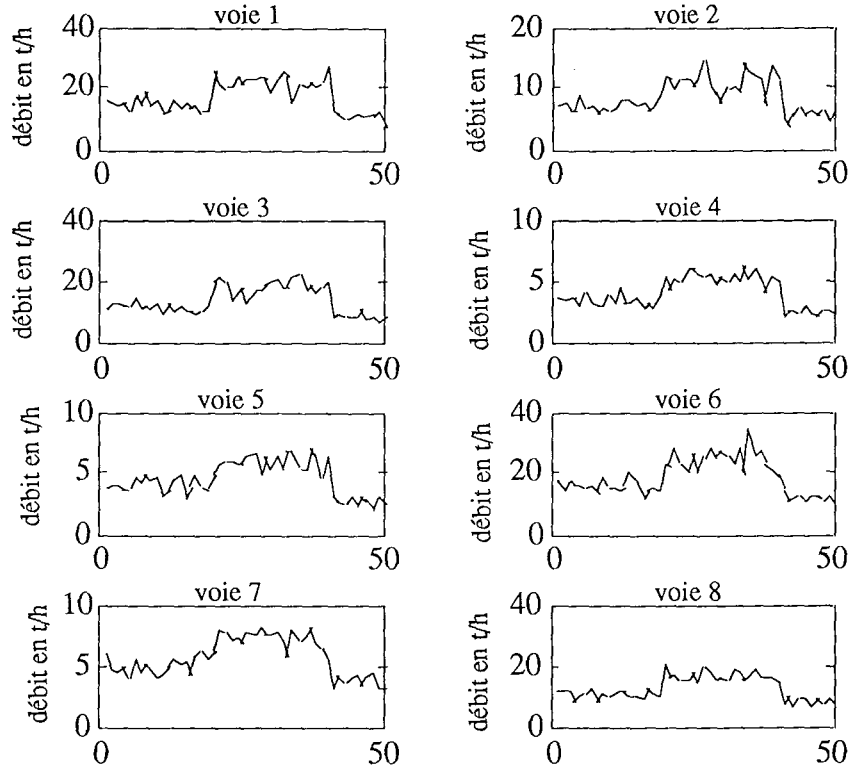



Figure IV-10 : évolution chronologique des débits

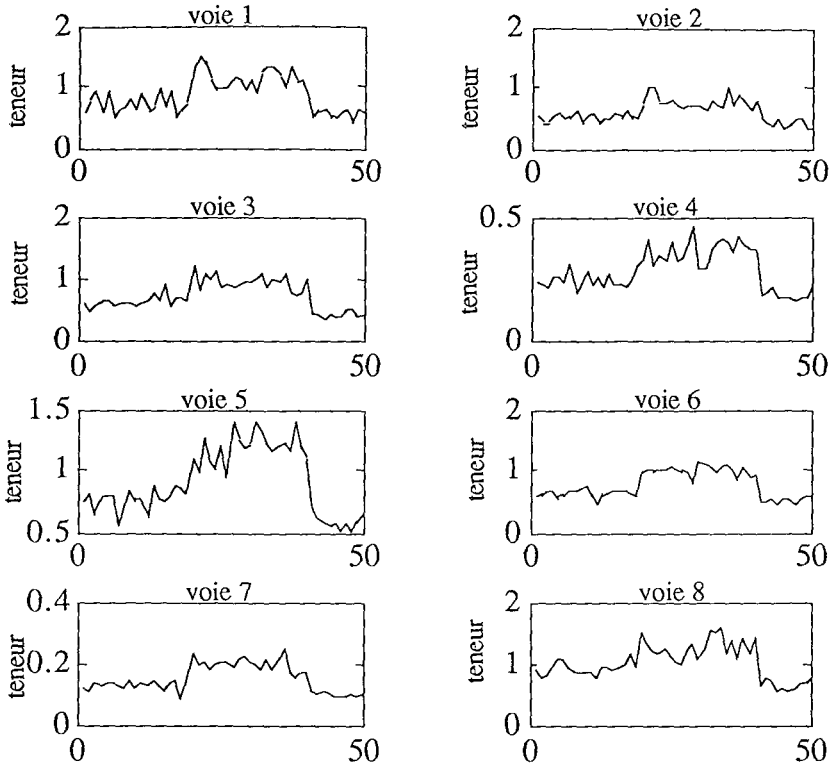



Figure IV-11 : évolution chronologique des teneurs

Les résultats de la simulation des grandeurs réelles et des écarts-type des erreurs de mesure sont présentés dans le tableau IV-8.

| | débit massique réel | | | teneur réelle | | | écarts-type réels | |
|---|---|---|---|---|---|---|---|---|
| voie | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 | écart-type débit | écart-type teneur |
| 1 | 15.00 | 22.50 | 11.25 | 0.75 | 1.12 | 0.56 | 2.01 | 0.14 |
| 2 | 7.84 | 11.20 | 5.58 | 0.54 | 0.81 | 0.41 | 1.35 | 0.08 |
| 3 | 12.47 | 18.72 | 9.12 | 0.67 | 0.93 | 0.46 | 1.86 | 0.10 |
| 4 | 3.50 | 5.25 | 2.63 | 0.25 | 0.38 | 0.19 | 0.46 | 0.04 |
| 5 | 4.00 | 6.00 | 3.00 | 0.79 | 1.19 | 0.60 | 0.56 | 0.10 |
| 6 | 16.50 | 24.75 | 12.37 | 0.67 | 1.00 | 0.50 | 2.51 | 0.08 |
| 7 | 5.00 | 7.50 | 3.75 | 0.14 | 0.20 | 0.10 | 0.67 | 0.02 |
| 8 | 11.50 | 17.25 | 8.63 | 0.90 | 1.35 | 0.68 | 1.46 | 0.14 |

Tableau IV-8 : grandeurs réelles

Les résultats de l'estimation des grandeurs réelles et des écarts-type des erreurs de mesure sont présentés dans le tableau IV-9.

| | débit massique estimé | | | teneur estimée | | | écarts-type estimés | |
|---|---|---|---|---|---|---|---|---|
| voie | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 | écart-type débit | écart-type teneur |
| 1 | 15.40 | 22.42 | 11.05 | 0.78 | 1.11 | 0.56 | 2.38 | 0.16 |
| 2 | 7.84 | 11.20 | 5.58 | 0.54 | 0.81 | 0.42 | 1.48 | 0.09 |
| 3 | 12.47 | 18.72 | 9.12 | 0.67 | 0.93 | 0.47 | 2.12 | 0.12 |
| 4 | 3.66 | 5.22 | 2.58 | 0.25 | 0.37 | 0.20 | 0.53 | 0.04 |
| 5 | 4.18 | 5.99 | 3.00 | 0.79 | 1.20 | 0.61 | 0.54 | 0.11 |
| 6 | 16.65 | 24.71 | 12.12 | 0.70 | 1.00 | 0.50 | 2.70 | 0.09 |
| 7 | 4.90 | 7.50 | 3.65 | 0.13 | 0.22 | 0.10 | 0.68 | 0.03 |
| 8 | 11.75 | 17.21 | 8.47 | 0.94 | 1.34 | 0.68 | 1.90 | 0.15 |

Tableau IV-9 : grandeurs estimées

**Deuxième essai**

Considérons le réseau de transport de matière représenté par la figure IV-12, comportant sept nœuds et treize voies :

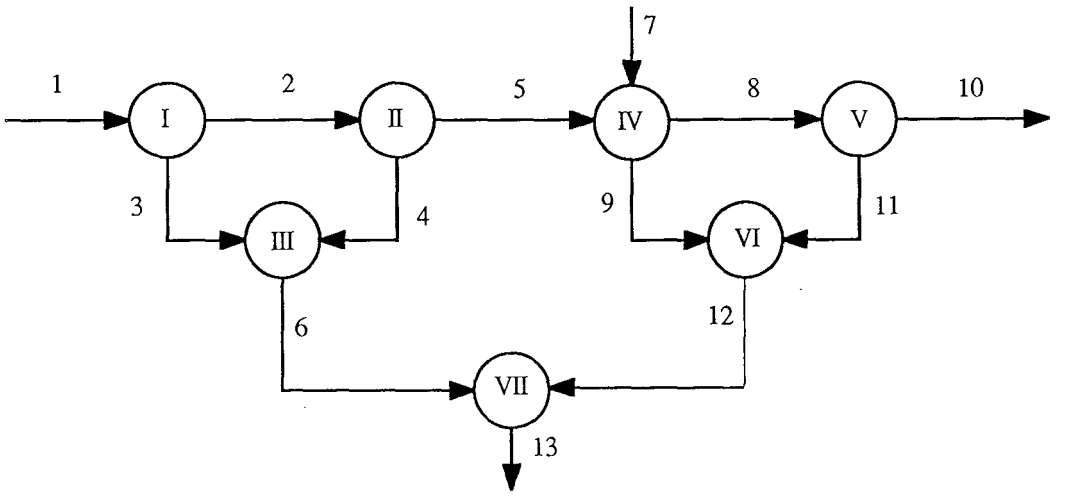



Figure IV-12 : réseau de transport de matière

Sa matrice d'incidence est :

| voie | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 0 | 0 | 1 | 1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -1 | -1 | 0 | 0 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | -1 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | -1 |

Sur ce réseau, nous avons simulé les données des débits et des teneurs de toutes les voies pour un seul régime statique. Les résultats de la simulation des débits et des teneurs avec leurs écarts-type sont représentés dans le tableau IV-10 :

| voie | débits réels | teneurs réelles | écart-type des débits | écart-type des teneurs |
|---|---|---|---|---|
| 1 | 25.50 | 0.870 | 3.724 | 0.104 |
| 2 | 15.75 | 0.840 | 2.433 | 0.091 |
| 3 | 9.75 | 0.918 | 1.313 | 0.116 |
| 4 | 5.75 | 0.996 | 0.886 | 0.104 |
| 5 | 10.00 | 0.750 | 1.402 | 0.094 |
| 6 | 15.50 | 0.947 | 2.224 | 0.115 |
| 7 | 3.25 | 0.520 | 0.461 | 0.057 |
| 8 | 8.75 | 0.737 | 1.236 | 0.088 |
| 9 | 4.50 | 0.610 | 0.665 | 0.070 |
| 10 | 4.25 | 0.550 | 0.557 | 0.070 |
| 11 | 4.50 | 0.913 | 0.620 | 0.108 |
| 12 | 9.00 | 0.761 | 1.209 | 0.091 |
| 13 | 24.50 | 0.879 | 3.543 | 0.092 |

Tableau IV-10 : grandeurs réelles

A titre indicatif, les évolutions chronologiques des mesures du débit et de la teneur de la voie 5 sont données par les figures IV-13 et IV-14.

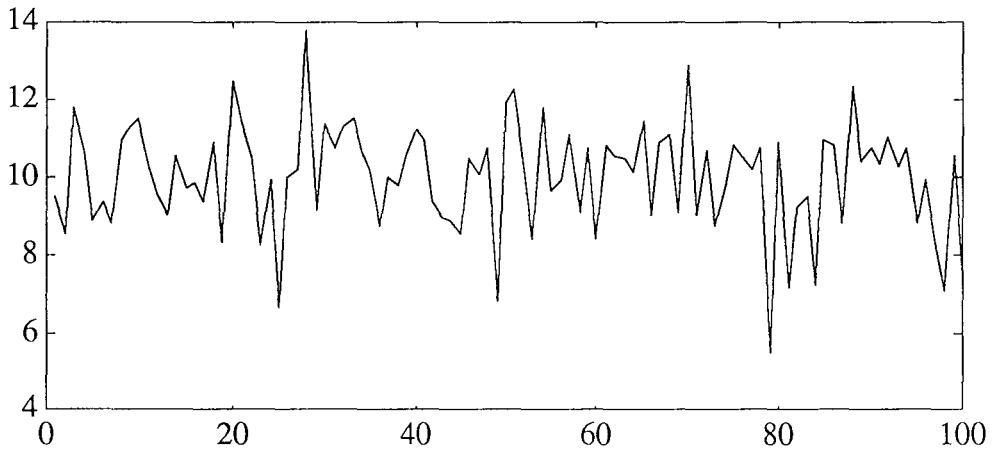



Figure IV-13 : évolution temporelle des mesures du débit

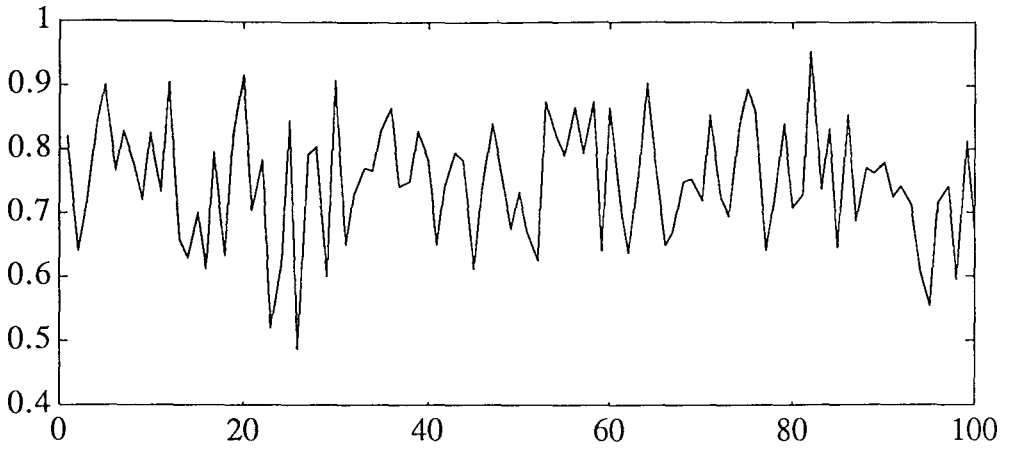



Figure IV-14 : évolution temporelle des mesures de teneur

L'estimation des débits et des teneurs ainsi que les écarts-types des erreurs de mesures sont présentés dans le tableau IV-11 :

| voie | débits estimés | teneurs estimées | écart-type des débits | écart-type des teneurs |
|---|---|---|---|---|
| 1 | 25.51 | 0.871 | 3.705 | 0.103 |
| 2 | 15.75 | 0.842 | 2.421 | 0.091 |
| 3 | 9.74 | 0.917 | 1.307 | 0.115 |
| 4 | 5.75 | 0.995 | 0.881 | 0.103 |
| 5 | 10.02 | 0.750 | 1.395 | 0.093 |
| 6 | 15.51 | 0.947 | 2.213 | 0.114 |
| 7 | 3.24 | 0.521 | 0.459 | 0.057 |
| 8 | 8.74 | 0.740 | 1.230 | 0.087 |
| 9 | 4.52 | 0.610 | 0.661 | 0.069 |
| 10 | 4.23 | 0.556 | 0.554 | 0.070 |
| 11 | 4.51 | 0.911 | 0.617 | 0.107 |
| 12 | 9.01 | 0.761 | 1.203 | 0.090 |
| 13 | 24.50 | 0.878 | 3.525 | 0.091 |

Tableau IV-11 : grandeurs estimées

Les tableaux IV-12 et IV-13 représentent respectivement l'évolution des écarts-type estimés des débits et des teneurs en fonction du nombre d'observations.

| m | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2.11 | 2.58 | 3.37 | 3.27 | 3.44 | 3.62 | 3.43 | 3.72 | 3.77 | 3.70 |
| 2 | 1.49 | 2.01 | 2.50 | 2.51 | 2.62 | 2.55 | 2.41 | 2.36 | 2.38 | 2.42 |
| 3 | 1.35 | 1.34 | 1.31 | 1.29 | 1.26 | 1.21 | 1.26 | 1.26 | 1.24 | 1.30 |
| 4 | 1.07 | 1.04 | 0.96 | 0.92 | 0.92 | 0.89 | 0.87 | 0.91 | 0.88 | 0.88 |
| 5 | 1.18 | 1.14 | 1.41 | 1.29 | 1.33 | 1.32 | 1.30 | 1.34 | 1.38 | 1.39 |
| 6 | 1.57 | 1.78 | 1.78 | 2.13 | 2.14 | 2.28 | 2.25 | 2.22 | 2.22 | 2.21 |
| 7 | 0.48 | 0.47 | 0.47 | 0.45 | 0.48 | 0.47 | 0.47 | 0.45 | 0.46 | 0.46 |
| 8 | 1.48 | 1.32 | 1.27 | 1.24 | 1.32 | 1.27 | 1.26 | 1.25 | 1.23 | 1.23 |
| 9 | 0.51 | 0.53 | 0.52 | 0.53 | 0.52 | 0.60 | 0.61 | 0.64 | 0.65 | 0.66 |
| 10 | 0.63 | 0.61 | 0.58 | 0.56 | 0.58 | 0.54 | 0.54 | 0.55 | 0.55 | 0.55 |
| 11 | 0.49 | 0.62 | 0.63 | 0.61 | 0.62 | 0.60 | 0.60 | 0.64 | 0.63 | 0.61 |
| 12 | 1.29 | 1.18 | 1.17 | 1.17 | 1.14 | 1.16 | 1.17 | 1.17 | 1.14 | 1.20 |
| 13 | 3.47 | 3.83 | 3.75 | 3.61 | 3.58 | 3.83 | 3.65 | 3.73 | 3.60 | 3.52 |

Tableau IV-12 : écarts-type estimés des débits

Au travers de ce tableau nous constatons que l'évolution des écarts-type des débits est relativement constante à partir de l'observation 30, elle est aussi constante pour les écarts-type des teneurs comme le montre le tableau IV-13.

| m | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.06 | 0.07 | 0.08 | 0.08 | 0.10 | 0.09 | 0.10 | 0.10 | 0.10 | 0.10 |
| 2 | 0.10 | 0.12 | 0.11 | 0.10 | 0.09 | 0.10 | 0.09 | 0.09 | 0.09 | 0.09 |
| 3 | 0.11 | 0.10 | 0.11 | 0.11 | 0.11 | 0.11 | 0.11 | 0.12 | 0.12 | 0.12 |
| 4 | 0.08 | 0.09 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.10 | 0.10 | 0.10 |
| 5 | 0.07 | 0.09 | 0.11 | 0.10 | 0.10 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| 6 | 0.11 | 0.13 | 0.13 | 0.12 | 0.13 | 0.12 | 0.12 | 0.12 | 0.12 | 0.11 |
| 7 | 0.06 | 0.05 | 0.05 | 0.06 | 0.05 | 0.06 | 0.05 | 0.06 | 0.06 | 0.06 |
| 8 | 0.06 | 0.08 | 0.08 | 0.08 | 0.08 | 0.09 | 0.09 | 0.09 | 0.09 | 0.09 |
| 9 | 0.06 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| 10 | 0.07 | 0.06 | 0.07 | 0.06 | 0.06 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| 11 | 0.12 | 0.12 | 0.12 | 0.11 | 0.11 | 0.11 | 0.10 | 0.10 | 0.11 | 0.11 |
| 12 | 0.08 | 0.08 | 0.09 | 0.08 | 0.08 | 0.08 | 0.09 | 0.09 | 0.09 | 0.09 |
| 13 | 0.10 | 0.09 | 0.09 | 0.08 | 0.08 | 0.09 | 0.08 | 0.08 | 0.09 | 0.09 |

Tableau IV-13 : écarts-type estimés des teneurs

### IV-2-3 RÉSOLUTION DANS LE CAS BILINÉAIRE ÉTENDU

L'extension de la méthode au cas multi-constituants s'effectue de manière fort simple. Le deuxième niveau sera constitué de **nc** (nombre de constituants) sous-niveaux, réalisant chacun l'estimation d'un vecteur de paramètres de Lagrange et celle du vecteur de la teneur considérée. Le premier niveau utilisera l'ensemble des informations provenant de ces nc sous-niveaux pour calculer l'estimation $\hat{X}$.

Pour ne pas alourdir la représentation matricielle, nous travaillerons autour d'un seul point de fonctionnement. La généralisation pour plusieurs régimes statiques se fera sans difficulté en utilisant le même principe que dans la procédure mono-constituant.

Pour résoudre ce problème, on minimise le critère $\Phi$ par rapport à $\hat{X}$, $V_x$, $\hat{Y}_k$ et $V_{y_k}$ (k = 1, ..., nc) :

$$\Phi = \frac{m}{2}\mathrm{Log}|V_x| + \frac{m}{2}\sum_{k=1}^{nc}\mathrm{Log}|V_{y_k}| + \frac{1}{2}\sum_{i=1}^{m}\|\hat{X} - X_i\|^2_{V_x^{-1}} + \sum_{k=1}^{nc}\sum_{i=1}^{m}\|\hat{Y}_k - Y_{ki}\|^2_{V_{y_k}^{-1}}$$

$$\text{sous les contraintes} \quad \begin{aligned} M\hat{X} &= 0 \\ M(\hat{X} * \hat{Y}_k) &= 0 \qquad \text{pour } k = 1, ..., nc \end{aligned} \qquad \text{(IV-41)}$$

Le lagrangien associé à ce problème s'écrit :

$$L = \Phi + \lambda^T M\hat{X} + \sum_{k=1}^{nc}\mu_k^T M(\hat{X} * \hat{Y}_k) \qquad \text{(IV-42)}$$

En annulant les dérivées premières du lagrangien par rapport à $\hat{X}$, $\hat{Y}_k$, $V_x$, $V_{yk}$, $\lambda$ et $\mu_k$, nous pouvons déduire, comme dans le cas bilinéaire simple, les expressions suivantes :

$$V_x = \frac{1}{m}\mathrm{diag}[\sum_{i=1}^{m}(\hat{X} - X_i)(\hat{X} - X_i)^T] \qquad \text{(IV-43)}$$

$$V_{yk} = \frac{1}{m}\mathrm{diag}[\sum_{i=1}^{m}(\hat{Y}_k - Y_{ki})(\hat{Y}_k - Y_{ki})^T] \qquad \text{(IV-44)}$$

$$\lambda = (MV_xM^T)^{-1}M[m\,\bar{X} - V_x\sum_{k=1}^{nc}N_{y_k}^T\mu_k] \qquad \text{(IV-45)}$$

$$\hat{X} = P_x[\bar{X} - \frac{1}{m}V_x\sum_{k=1}^{nc}N_{y_k}^T\mu_k] \qquad \text{(IV-46)}$$

$$\mu_k = m\,(N_xV_{yk}N_x^T)^{-1}N_x\,\bar{Y}_k \qquad \text{(IV-47)}$$

$$\hat{Y}_k = P_{yk}\bar{Y}_k \qquad \text{(IV-48)}$$

avec

$$P_x = I - V_xM^T(MV_xM^T)^{-1}M \qquad \text{(IV-49)}$$

$$P_{yk} = I - V_{yk}N_x^T(N_xV_{yk}N_x^T)^{-1}N_x \qquad \text{(IV-50)}$$

et $$\bar{X} = \frac{1}{m} \sum_{i=1}^{m} X_i \qquad \text{(IV-51)}$$

$$\bar{Y}_k = \frac{1}{m} \sum_{j=1}^{m} Y_{kj} \qquad \text{(IV-52)}$$

En posant :

$$N_x = M * \hat{X} \qquad \text{(IV-53)}$$

$$N_{y_k} = M * \hat{Y}_k \qquad \text{(IV-54)}$$

L'algorithme de relaxation utilisé pour résoudre ce problème est constitué de deux niveaux, le premier est similaire à celui du problème mono-constituant, par contre le second niveau est constitué de **nc** sous-niveaux (figure IV-15).

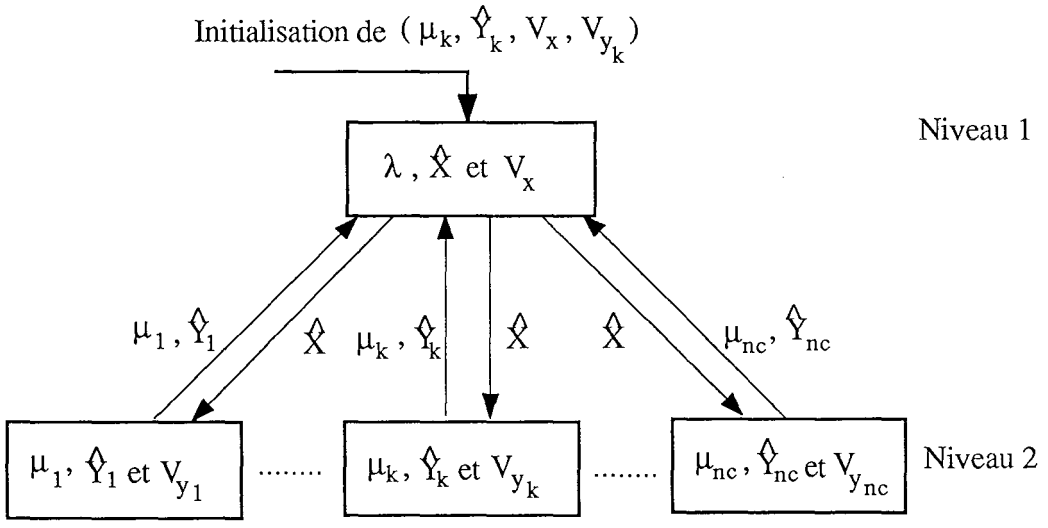



Figure IV-15 : algorithme de relaxation en multi-constituants

## IV-2-4 EXTENSION AUX SYSTEMES PARTIELLEMENT OBSERVÉS

La méthode présentée au paragraphe précédent n'est applicable que dans le cas où l'ensemble des grandeurs est mesuré, or la décomposition issue de l'observabilité **[FAYO 87]** conduit à résoudre un système dans lequel certains débits peuvent être inconnus. Il convient donc d'élaborer une méthode de résolution répondant à cette situation.

Nous ne formaliserons pas le problème de l'observabilité des systèmes bilinéaires. Nous nous contenterons de mettre en évidence la structure particulière des équations de redondance et nous proposerons une méthode d'estimation adaptée à cette structure.

Considérons le réseau représenté par la figure IV-16 formé de sept nœuds et treize voies. Les positions des mesures de débit sont représentées par des x et les positions des mesures de teneurs par des ..

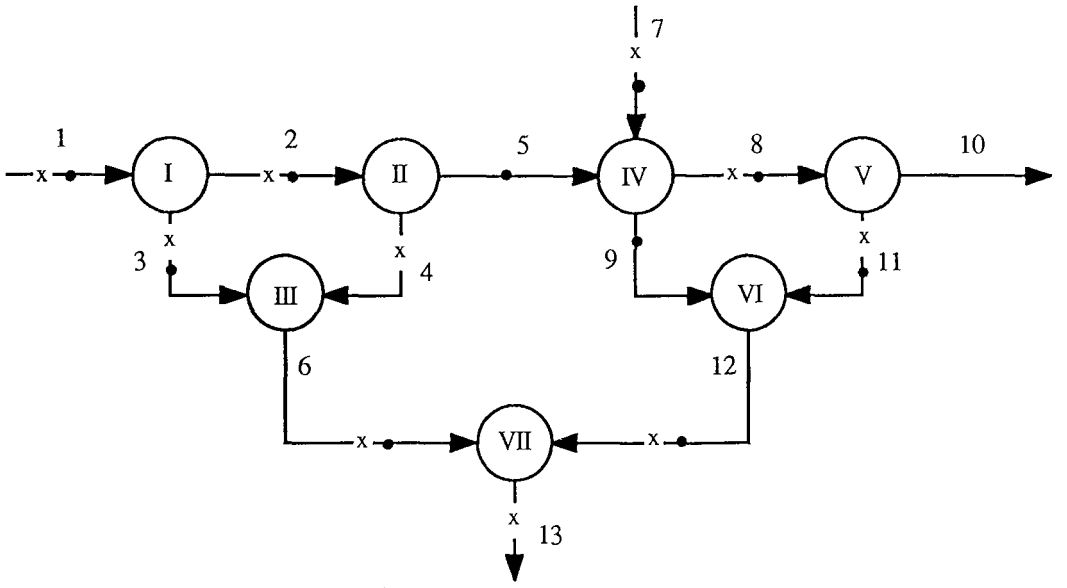



Figure IV-16 : placement des capteurs de mesure

### IV-2-4-1 Mise en évidence du réseau contenant la redondance de type linéaire-bilinéaire

L'application de la procédure d'observabilité développée par J. Fayolle **[FAYO 87]** au réseau de la figure IV-16 permet d'obtenir les équations de redondance suivantes :

- linéaires :

$$\hat{X}_1 - \hat{X}_2 - \hat{X}_3 = 0$$

$$\hat{X}_2 + \hat{X}_3 - \hat{X}_5 - \hat{X}_6 = 0$$

$$\hat{X}_5 + \hat{X}_7 - \hat{X}_8 - \hat{X}_9 = 0$$

$$\hat{X}_9 + \hat{X}_{11} - \hat{X}_{12} = 0$$

- bilinéaires :

$$\hat{X}_1\hat{Y}_1 - \hat{X}_2\hat{Y}_2 - \hat{X}_3\hat{Y}_3 = 0$$

$$\hat{X}_2\hat{Y}_2 + \hat{X}_3\hat{Y}_3 - \hat{X}_5\hat{Y}_5 - \hat{X}_6\hat{Y}_6 = 0$$

$$\hat{X}_5\hat{Y}_5 + \hat{X}_7\hat{Y}_7 - \hat{X}_8\hat{Y}_8 - \hat{X}_9\hat{Y}_9 = 0$$

$$\hat{X}_9\hat{Y}_9 + \hat{X}_{11}\hat{Y}_{11} - \hat{X}_{12}\hat{Y}_{12} = 0$$

- et trois équations linéaires supplémentaires :

$$\hat{X}_3 + \hat{X}_4 - \hat{X}_6 = 0$$

$$\hat{X}_2 + \hat{X}_7 + \hat{X}_{11} - \hat{X}_4 - \hat{X}_8 - \hat{X}_{12} = 0$$

$$\hat{X}_6 + \hat{X}_{12} - \hat{X}_{13} = 0$$

Les équations linéaires s'expriment matriciellement :

$$A\hat{X} = 0 \qquad \text{(IV-55)}$$

où A est une matrice d'incidence de dimension $n_a \times v_a$,
$n_a$ le nombre d'équations de type linéaire,
$v_a$ le nombre de débits intervenant dans les équations.

Les équations bilinéaires s'écrivent

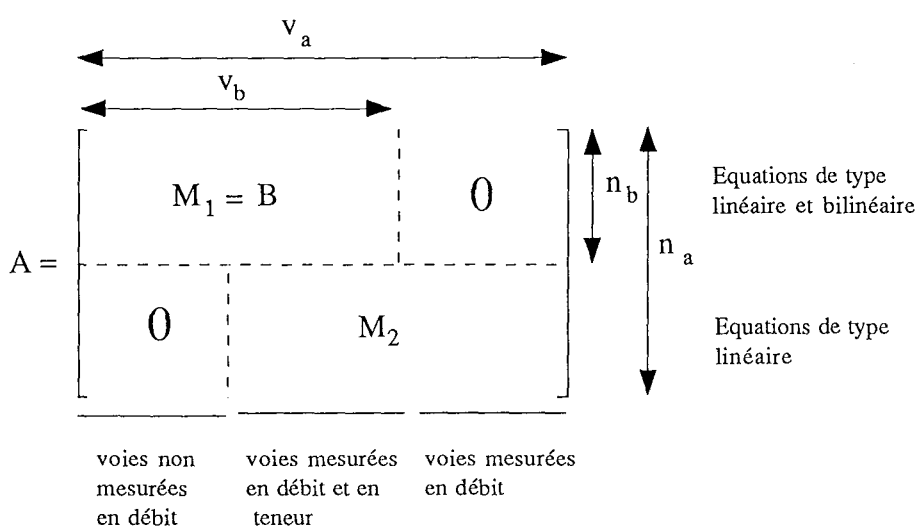
$$(B * \hat{Y})S\hat{X} = 0 \qquad \text{(IV-56)}$$

où B est une matrice d'incidence de dimension $n_b \times v_b$,
S une matrice de sélection des débits intervenant dans des équations de type bilinéaire, de dimension $v_b \times v_a$,
$\hat{X}$ est un vecteur de dimension $v_a$ et $\hat{Y}$ un vecteur de dimension $v_b$.

Après répartition des voies en trois groupes,

- voies mesurées en teneurs et non mesurées en débit,
- voies mesurées en teneurs et en débit,
- voies mesurées en débit,

la matrice A prend la forme générale suivante :




La matrice S s'écrit après classement des voies selon l'ordre de la forme ci-dessus :

$$S = [\, I_{v_b} \quad 0 \,] \qquad \text{(IV-57)}$$

Pour l'exemple précédent, la matrice A s'écrit :

| voies | 5 | 9 | 1 | 2 | 3 | 6 | 7 | 8 | 11 | 12 | 4 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Equations linéaires et bilinéaires | 0<br>-1<br>1<br>0 | 0<br>0<br>-1<br>1 | 1<br>0<br>0<br>0 | -1<br>1<br>0<br>0 | -1<br>1<br>0<br>0 | 0<br>-1<br>0<br>0 | 0<br>0<br>1<br>0 | 0<br>0<br>-1<br>0 | 0<br>0<br>0<br>1 | 0<br>0<br>0<br>-1 | 0<br>0<br>0<br>0 | 0<br>0<br>0<br>0 |
| Equations linéaires | 0<br>0<br>0 | 0<br>0<br>0 | 0<br>0<br>0 | 0<br>1<br>0 | 1<br>0<br>0 | -1<br>0<br>1 | 0<br>1<br>0 | 0<br>-1<br>0 | 0<br>1<br>0 | 0<br>-1<br>1 | 1<br>-1<br>0 | 0<br>0<br>-1 |
| | voies mesurées en teneurs et non mesurées en débit | | | | voies mesurées en teneurs et en débit | | | | | | voies mesurées en débit | |

### IV-2-4-2 Estimation des variances des grandeurs bilinéaires

Après décomposition selon l'observabilité [**DARO 86**], le problème d'estimation des grandeurs bilinéaires, dans le cas d'un seul régime de fonctionnement statique, s'écrit :

Chercher le minimum par rapport à $\hat{X}$, $\hat{Y}$, $V_x$ et $V_y$ du critère :

$$\Phi = \frac{m}{2}\mathrm{Log}|V_x| + \frac{m}{2}\mathrm{Log}|V_y| + \frac{1}{2}\sum_{i=1}^{m} \|H\hat{X} - X_i\|^2_{V_x^{-1}} + \sum_{i=1}^{m} \|\hat{Y} - Y_i\|^2_{V_y^{-1}}$$

sous les contraintes $A\hat{X} = 0$ (IV-58)

$(B * \hat{Y})S\hat{X} = 0$

où H est la matrice d'observation des débits mesurés. Nous nous placerons dans le cas de mesures directes et nous aurons alors $H = [0 \quad I_k]$,

$X_i$ est le vecteur des débits mesurés à l'observation i de dimension k,

$\hat{X}$ est le vecteur des débits estimés de dimension $v_a$,

$Y_i$ est le vecteur des teneurs à l'observation i de dimension $v_b$,

$\hat{Y}$ est le vecteur des teneurs estimées de dimension $v_b$,

S est la matrice de sélection des débits intervenant dans les équations bilinéaires de dimension $v_b \times v_a$ avec $S = [I_{v_b} \quad 0]$,

$V_x$ et $V_y$ sont les matrices de variance des mesures de débit et de teneur de dimensions respectives $k \times k$ et $v_b \times v_b$.

Le lagrangien associé à ce problème s'écrit :

$$L = \Phi + \lambda^T A\hat{X} + \mu^T (B * \hat{Y})S\hat{X} \qquad \text{(IV-59)}$$

il est stationnaire quand :

$$\frac{\partial L}{\partial \hat{X}} = H^T V_x^{-1} \sum_{i=1}^{m} (H\hat{X} - X_i) + A^T\lambda + S^T (B * \hat{Y})^T\mu \quad = 0 \qquad \text{(IV-60-1)}$$

$$\frac{\partial L}{\partial \hat{Y}} = V_y^{-1} \sum_{i=1}^{m} (\hat{Y} - Y_i) + (B * S\hat{X})^T\mu \quad = 0 \qquad \text{(IV-60-2)}$$

$$\frac{\partial L}{\partial V_x} = \frac{m}{2} V_x^{-1} - \frac{1}{2} V_x^{-1} \operatorname{diag}[\sum_{i=1}^{m} (H\hat{X} - X_i)(H\hat{X} - X_i)^T] V_x^{-1} \quad = 0 \qquad \text{(IV-60-3)}$$

$$\frac{\partial L}{\partial V_y} = \frac{m}{2} V_y^{-1} - \frac{1}{2} V_y^{-1} \operatorname{diag}[\sum_{i=1}^{m} (\hat{Y} - Y_i)(\hat{Y} - Y_i)^T] V_y^{-1} \quad = 0 \qquad \text{(IV-60-4)}$$

$$\frac{\partial L}{\partial \lambda} = A\hat{X} \quad = 0 \qquad \text{(IV-60-5)}$$

$$\frac{\partial L}{\partial \mu} = (B * \hat{Y})\, S\hat{X} \quad = 0 \qquad \text{(IV-60-6)}$$

En général, ce problème n'admet pas de solution analytique. Nous proposons un calcul hiérarchisé basé sur la méthode de relaxation pour résoudre le système non linéaire précédent.

A partir des équations (IV-60-3) et (IV-60-4), nous pouvons déduire les expressions de $V_x$ et $V_y$ en fonction de $\hat{X}_i$ et $\hat{Y}_i$.

$$V_x = \frac{1}{m} \operatorname{diag}[\sum_{i=1}^{m} (H\hat{X} - X_i)(H\hat{X} - X_i)^T] \qquad \text{(IV-61)}$$

et

$$V_y = \frac{1}{m} \operatorname{diag}[\sum_{i=1}^{m} (\hat{Y}_i - Y_i)(\hat{Y} - Y_i)^T] \qquad \text{(IV-62)}$$

Posons :

$$N_x = B * S\hat{X} \qquad \text{(IV-63)}$$

$$N_y = (B * \hat{Y})S \qquad \text{(IV-64)}$$

La démarche suivante permet alors d'exprimer l'estimateur $\hat{X}$ ; prémultiplions l'équation (IV-60-5) par $A^T$ :

$$A^T A\hat{X} = 0 \qquad \text{(IV-65)}$$

puis (IV-60-6) par $N_y^T$ :

$$N_y^T N_y \hat{X} = 0 \qquad \text{(IV-66)}$$

Additionnons les relations (IV-65), (IV-66) et (IV-60-1), nous obtenons :

$$G\hat{X} = H^T V_x^{-1} \sum_{i=1}^{m} X_i - A^T\lambda - N_y^T \mu \qquad \text{(IV-67)}$$

avec $G = mH^T V_x^{-1} H + A^T A + N_y^T N_y$ (IV-68)

A l'aide des équations (IV-60-1) et (IV-60-5), nous pouvons exprimer $\lambda$ et $\hat{X}$ :

$$\lambda = (AG^{-1}A^T)^{-1} AG^{-1}(H^T V_x^{-1} \sum_{i=1}^{m} X_i - N_y^T \mu) \qquad \text{(IV-69)}$$

$$\hat{X} = G^{-1} P_x (H^T V_x^{-1} \sum_{i=1}^{m} X_i - N_y^T \mu) \qquad \text{(IV-70)}$$

avec

$$P_x = I - A^T(AG^{-1}A^T)^{-1} A \qquad \text{(IV-71)}$$

Les relations (IV-60-2) et (IV-60-6), permettent d'exprimer $\mu$ et $\hat{Y}$ :

$$\mu = (N_x V_y N_x^T)^{-1} N_x \sum_{i=1}^{m} Y_i \qquad \text{(IV-72)}$$

$$\hat{Y} = P_y \bar{Y} \qquad \text{(IV-73)}$$

avec

$$P_y = I - V_y N_x^T (N_x V_y N_x^T)^{-1} N_x \qquad \text{(IV-74)}$$

et $\bar{Y} = \frac{1}{m} \sum_{i=1}^{m} Y_i$ (IV-75)

Afin de faciliter la mise en œuvre numérique de ces calculs, nous pouvons faire apparaître les matrices diagonales $\Lambda_{S\hat{X}}$ et $\Lambda_{\hat{Y}}$ formées respectivement des vecteurs $S\hat{X}$ et $\hat{Y}$.

D'après (IV-60-2) nous avons :

$$\Lambda_{S\hat{X}} B^T \mu = V_y^{-1} \sum_{i=1}^{m} (Y_i - \hat{Y}) \qquad \text{(IV-76)}$$

d'où

$$B^T\mu = \Lambda_{S\hat{X}}^{-1} V_y^{-1} \sum_{i=1}^{m}(Y_i - \hat{Y}) \qquad \text{(IV-77)}$$

Le vecteur $\hat{X}$ peut alors s'exprimer uniquement en fonction de X, Y et $\hat{Y}$ :

$$\hat{X} = P_x\, G\, [\, H^T V_x \sum_{i=1}^{m} X_i - S^T \Lambda_{\hat{Y}} \Lambda_{S\hat{X}}^{-1} V_y^{-1} \sum_{i=1}^{m}(Y_i - \hat{Y})] \qquad \text{(IV-78)}$$

de même, l'expression de l'estimation $\hat{Y}$ devient :

$$\hat{Y} = \Lambda_{S\hat{X}}^{-1} [\, \Gamma - \Gamma B^T (B\Gamma B^T)^{-1} B\Gamma \,]\, \Lambda_{S\hat{X}}^{-1} V_y^{-1} \sum_{i=1}^{m} Y_i \qquad \text{(IV-79)}$$

avec

$$\Gamma = \Lambda_{S\hat{X}}\, V_y\, \Lambda_{S\hat{X}} \qquad \text{(IV-80)}$$

**Description de l'algorithme de relaxation**

La résolution du système non linéaire s'effectue alors selon une structure hiérarchisée à deux niveaux :

- le premier niveau calcule $\lambda$ et $\hat{X}$ à partir de $V_x$ (initialisée à la matrice identité) et de $\hat{Y}$ initialisé à $\bar{Y}$,

- le second niveau calcule $\hat{Y}$ et $\mu$ après avoir calculé $\Gamma$ en fonction de l'estimation $\hat{X}$ transmise par le niveau supérieur.

A la fin de chaque niveau, nous calculons les matrices de variance $V_x$ et $V_y$.

La méthode de relaxation que nous proposons pour résoudre le problème d'optimisation (IV-58) est décrite par la figure IV-17 :

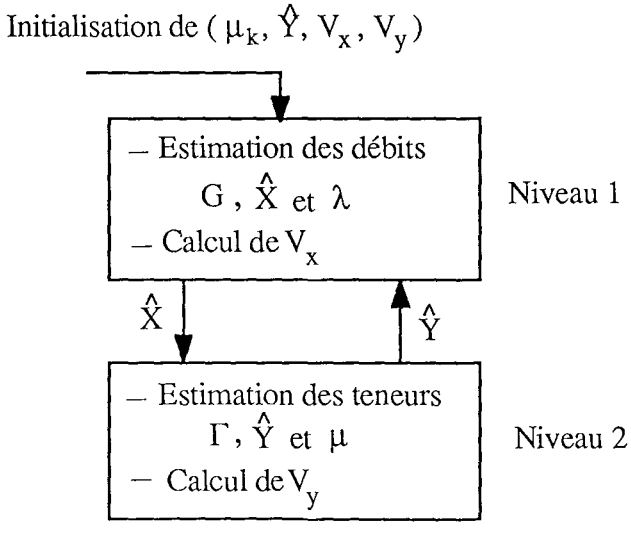



Figure IV-17 : algorithme de relaxation

L'arrêt de ce calcul itératif s'effectue par un test sur les dérivées du critère $\Phi$.

L'extension de cette méthode au cas multi-constituants est immédiate. Si l'on dispose de la mesure de nc constituants, le second niveau sera décomposé en nc sous-niveaux indépendants réalisant chacun l'estimation d'une teneur. La matrice G intervenant au premier niveau prendra alors la forme suivante :

$$G = mH^T V_x^{-1} H + A^T A + \sum_{j=1}^{nc} N_{y_j}^T N_{y_j} \qquad \text{(IV-81)}$$

avec

$$N_{y_j} = (B * \hat{Y}_j)S \qquad \text{pour } j = 1, ..., nc \qquad \text{(IV-82)}$$

La procédure d'estimation se déroulera selon le schéma de la figure IV-10. A chaque itération, il sera cependant nécessaire de s'assurer que la matrice G est régulière.

**Application de la méthode au réseau de la figure IV-16 (cas d'un seul constituant)**

Sur le réseau de la figure IV-16, nous effectuons dix mesures de débit et dix mesures de teneurs en un seul constituant. La matrice A contenant les équations linéaires-bilinéaires et les équations linéaires s'écrit sous la forme :

| voie | 5 | 9 | 1 | 2 | 3 | 6 | 7 | 8 | 11 | 12 | 4 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | -1 | 0 | 0 | 1 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 1 | -1 | 0 | 0 | -1 | 0 | 1 | -1 | 0 | 0 | 0 | 0 |
| 4 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | -1 | 0 | 0 |
| 5 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | -1 | 1 | -1 | -1 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | -1 |

La matrice B contenant les équations linéaires-bilinéaires s'écrit :

| voie | 5 | 9 | 1 | 2 | 3 | 6 | 7 | 8 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 |
| 2 | -1 | 0 | 0 | 1 | 1 | -1 | 0 | 0 | 0 | 0 |
| 3 | 1 | -1 | 0 | 0 | -1 | 0 | 1 | -1 | 0 | 0 |
| 4 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | -1 |

A titre indicatif, les évolutions chronologiques du débit de la voie 1 et de la teneur de la voie 5 sont données respectivement par les figures IV-18 et IV-19.

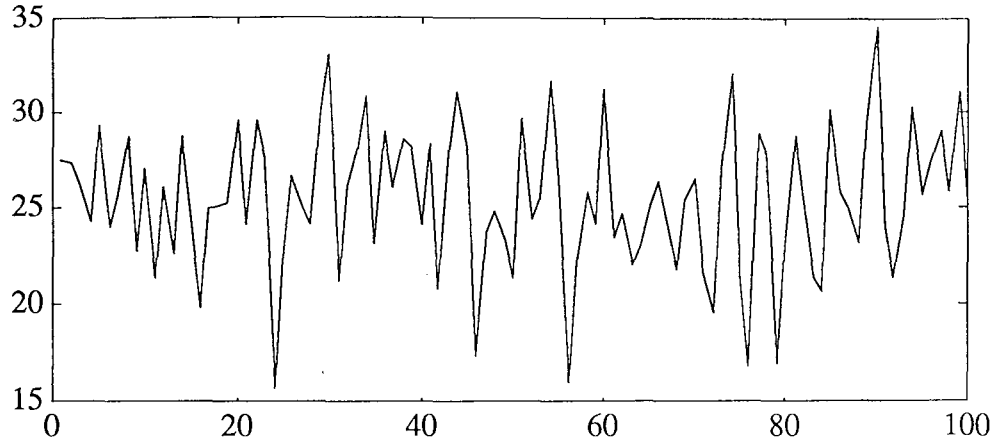



Figure IV-18 : évolution chronologique du débit de la voie 1

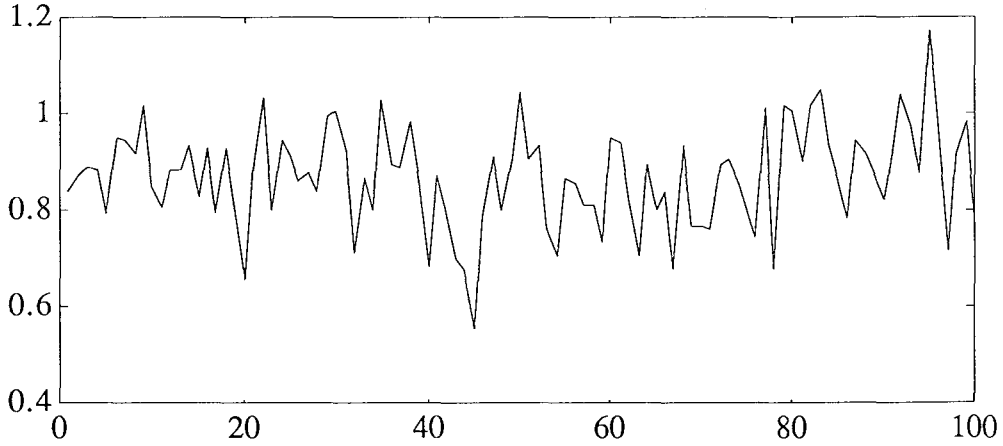



Figure IV-19 : évolution chronologique de la teneur de la voie 1

Les résultats de l'estimation sont rassemblés au tableau IV-14 :

| | débit massique | | écart-type des débits | | teneur | | écart-type des teneurs | |
|---|---|---|---|---|---|---|---|---|
| voie | réel | estimé | réel | estimé | réelle | estimée | réel | estimé |
| 1 | 25.50 | 25.48 | 1.30 | 1.31 | 0.87 | 0.86 | 0.10 | 0.11 |
| 2 | 15.75 | 15.74 | 2.21 | 2.22 | 0.84 | 0.83 | 0.09 | 0.09 |
| 3 | 9.75 | 9.71 | 0.45 | 0.45 | 0.92 | 0.93 | 0.11 | 0.12 |
| 4 | 5.75 | 5.72 | 4.33 | 4.34 | 0.99 | 1.00 | 0.44 | 0.44 |
| 5 | 10.00 | 10.13 | 3.70 | 3.70 | 0.75 | 0.76 | 0.09 | 0.09 |
| 6 | 15.50 | 15.51 | 1.23 | 1.23 | 0.95 | 0.95 | 0.11 | 0.11 |
| 7 | 3.25 | 3.24 | 0.61 | 0.61 | 0.52 | 0.51 | 0.06 | 0.06 |
| 8 | 8.75 | 8.75 | 1.20 | 1.21 | 0.74 | 0.73 | 0.08 | 0.09 |
| 9 | 4.50 | 4.52 | 2.42 | 2.41 | 0.61 | 0.61 | 0.07 | 0.07 |
| 10 | 4.25 | 4.24 | 1.54 | 1.55 | 0.55 | 0.55 | 0.09 | 0.08 |
| 11 | 4.50 | 4.47 | 0.88 | 0.87 | 0.91 | 0.90 | 0.10 | 0.11 |
| 12 | 9.00 | 8.98 | 3.52 | 3.50 | 0.76 | 0.76 | 0.09 | 0.09 |
| 13 | 24.50 | 24.51 | 3.77 | 3.76 | 0.88 | 0.87 | 0.07 | 0.08 |

Tableau IV-14 : grandeurs réelles et estimées des débits et des teneurs

**Application à un réseau en présence de plusieurs constituants**

La figure IV-20 représente le schéma d'un réseau de transport de matière constitué de six nœuds et douze voies. Sur cette figure, les x représentent les voies mesurées en débits massiques tandis que les • représentent les voies mesurées en teneurs.

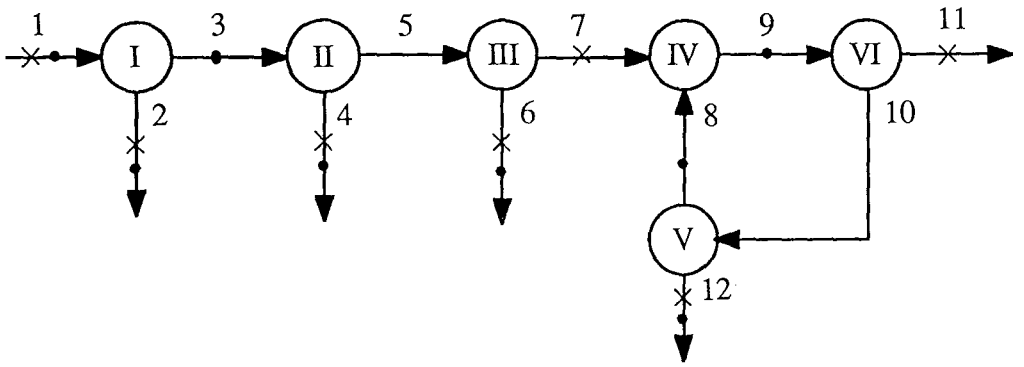



Figure IV-20 : réseau de transport de matière

On effectue sur ce procédé sept mesures de débit et neuf mesures de teneur en 3 constituants. Compte tenu de la localisation des points de mesures, l'étude de l'observabilité [**MAQU 87**] permet de constater que le réseau est complètement observable à la condition que les deux teneurs $Y_8$ et $Y_9$ soient distinctes.

La matrice A contenant les équations linéaires-bilinéaires et les équations linéaires s'écrit sous la forme :

| voie | 3 | 8 | 9 | 1 | 4 | 11 | 2 | 6 | 12 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | -1 | -1 | 0 | -1 | -1 | 0 |
| 2 | 0 | 0 | 0 | 1 | -1 | -1 | -1 | -1 | -1 | 0 |
| 3 | 0 | 1 | -1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4 | 0 | 0 | 0 | 1 | -1 | 0 | -1 | -1 | 0 | -1 |

La matrice B contenant les équations linéaires-bilinéaires s'écrit :

| voie | 3 | 8 | 9 | 1 | 4 | 11 | 2 | 6 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 0 | -1 | -1 | 0 | -1 | -1 |
| 2 | 0 | 0 | 0 | 1 | -1 | -1 | -1 | -1 | -1 |
| 3 | 0 | 1 | -1 | 0 | 0 | 1 | 0 | 0 | 0 |

Nous avons simulé des mesures sur toutes les voies puis nous avons éliminé les mesures de débits et de teneurs des voies supposées non mesurées pour pouvoir estimer les variances dans le cas des systèmes bilinéaires partiellement observés. Les

résultats de la simulation et de l'estimation des débits et des teneurs ainsi que leurs écarts-type sont rassemblés dans les tableaux IV-15 et IV-16.

| | débit massique | | teneur 1 | | teneur 2 | | teneur 3 | |
|---|---|---|---|---|---|---|---|---|
| voie | réel | écart type | réelle | écart type | réelle | écart type | réelle | écart type |
| 1 | 2106 | 400.5 | 0.65 | 0.11 | 2.14 | 0.30 | 29.75 | 4.40 |
| 2 | 220 | 41.2 | 0.54 | 0.10 | 2.16 | 0.35 | 23.29 | 3.92 |
| 3 | 1885 | 435.3 | 0.66 | 0.09 | 2.14 | 0.30 | 30.51 | 4.19 |
| 4 | 600 | 112.1 | 1.66 | 0.23 | 4.56 | 0.66 | 37.44 | 5.38 |
| 5 | 1286 | 254.2 | 0.19 | 0.05 | 1.01 | 0.20 | 27.27 | 5.30 |
| 6 | 171 | 38.0 | 0.61 | 0.07 | 3.92 | 0.56 | 34.97 | 4.95 |
| 7 | 1115 | 206.5 | 0.13 | 0.02 | 0.57 | 0.09 | 26.09 | 4.22 |
| 8 | 250 | 49.3 | 0.20 | 0.03 | 0.72 | 0.12 | 41.42 | 6.75 |
| 9 | 1364 | 271.1 | 0.14 | 0.03 | 0.60 | 0.10 | 28.89 | 4.75 |
| 10 | 759 | 139.2 | 0.19 | 0.04 | 0.86 | 0.12 | 47.51 | 6.80 |
| 11 | 605 | 118.5 | 0.08 | 0.01 | 0.28 | 0.04 | 5.53 | 0.79 |
| 12 | 510 | 106.6 | 0.18 | 0.02 | 0.92 | 0.15 | 50.49 | 8.55 |

Tableau IV-15 : grandeurs réelles des débits et des teneurs

A titre indicatif, les évolutions chronologiques du débit de la voie 6 et des teneurs des voies 2, 4 et 6 sont données respectivement par les figures IV-21, IV-22, IV-23 et IV-24.

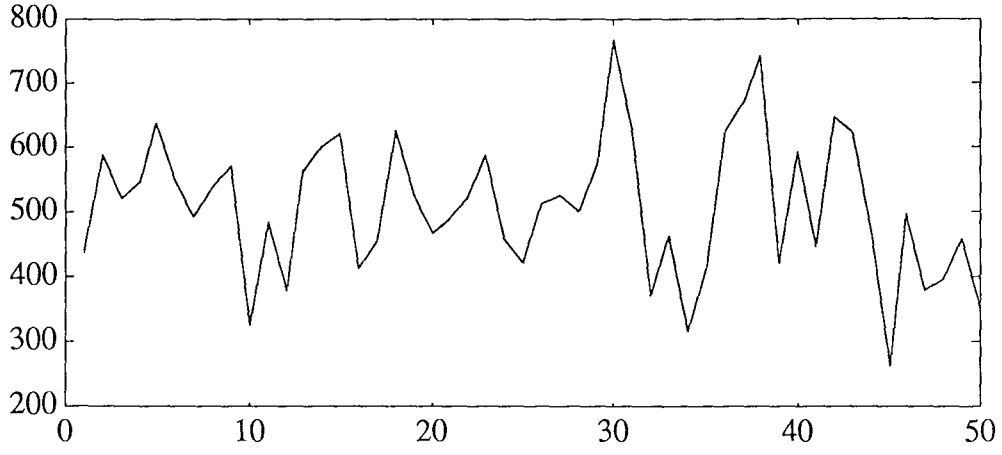



Figure IV-21 : évolution chronologique du débit de la voie 6

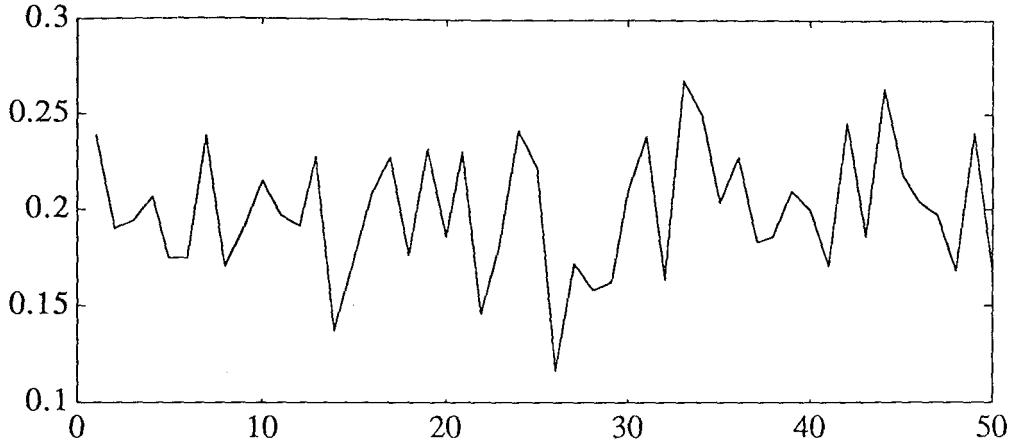



Figure IV-22 : évolution chronologique de la teneur 1 de la voie 2

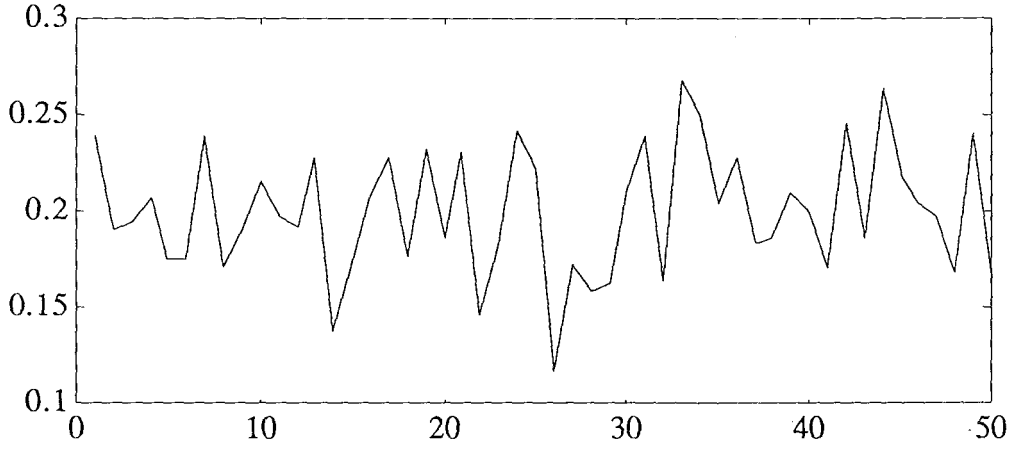



Figure IV-23 : évolution chronologique de la teneur 2 de la voie 4

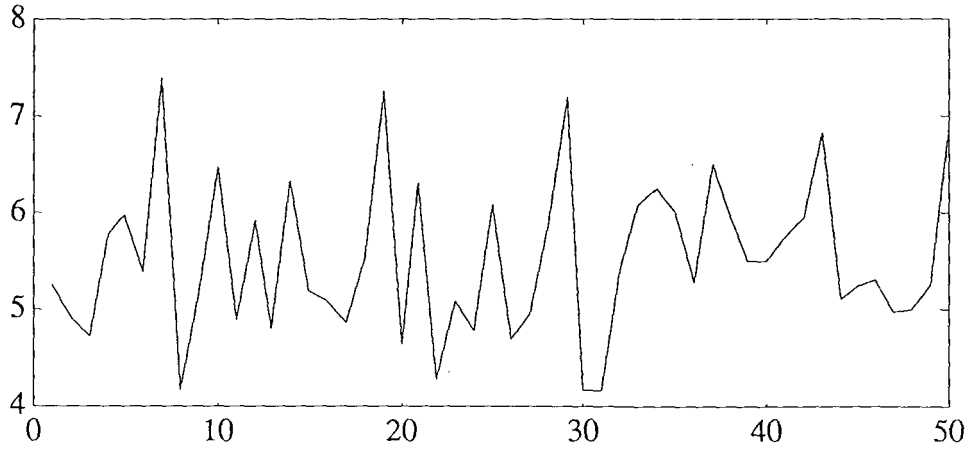



Figure IV-24 : évolution chronologique de la teneur 3 de la voie 6

| | débit massique | | teneur 1 | | teneur 2 | | teneur 3 | |
|---|---|---|---|---|---|---|---|---|
| voie | estimé | écart type | estimée | écart type | estimée | écart type | estimée | écart type |
| 1 | 2106 | 399.1 | 0.65 | 0.10 | 2.13 | 0.32 | 29.74 | 4.38 |
| 2 | 221 | 41.8 | 0.54 | 0.09 | 2.16 | 0.36 | 23.27 | 3.93 |
| 3 | 1886 | 434.2 | 0.66 | 0.09 | 2.14 | 0.29 | 30.51 | 4.18 |
| 4 | 600 | 112.6 | 1.66 | 0.24 | 4.56 | 0.66 | 37.44 | 5.38 |
| 5 | 1286 | 253.9 | 0.19 | 0.04 | 1.01 | 0.20 | 27.27 | 5.30 |
| 6 | 171 | 38.5 | 0.61 | 0.08 | 3.92 | 0.56 | 34.97 | 4.95 |
| 7 | 1115 | 206.1 | 0.13 | 0.02 | 0.57 | 0.09 | 26.09 | 4.22 |
| 8 | 249 | 49.2 | 0.20 | 0.03 | 0.72 | 0.12 | 41.42 | 6.76 |
| 9 | 1364 | 270.3 | 0.14 | 0.02 | 0.60 | 0.10 | 28.89 | 4.77 |
| 10 | 759 | 139.4 | 0.19 | 0.03 | 0.86 | 0.12 | 47.51 | 6.81 |
| 11 | 605 | 118.5 | 0.08 | 0.01 | 0.28 | 0.04 | 5.53 | 0.79 |
| 12 | 510 | 106.8 | 0.18 | 0.03 | 0.92 | 0.16 | 50.48 | 8.54 |

Tableau IV-16 : estimés des débits et des teneurs

Le tableau IV-16 montre que les résultats obtenus concernant les estimées des grandeurs réelles et des écarts-type sont voisins des grandeurs réelles. Ceci prouve l'adaptation de l'algorithme aux systèmes bilinéaires partiellement observés.

## IV-3 RÉCONCILIATION DE DONNÉES DES SYSTEMES MULTI-LINÉAIRES

Les systèmes n-linéaires sont utilisés pour décrire les modèles de connaissance de processus physiques fonctionnant en régime statique dans lesquels il est nécessaire de prendre en compte les débits massiques partiels et totaux. Ces modèles sont supposés exacts, c'est une condition nécessaire pour pouvoir estimer les grandeurs réelles ainsi que leurs variances.

### IV-3-1 FORMULATION DU PROBLEME

La forme générale d'un modèle multi-linéaire est définie par l'expression suivante :

$$R_j = M \prod_{i=1}^{j} X_j \qquad \text{pour } j = 1, ..., n \qquad \text{(IV-83)}$$

où n est l'ordre maximal du modèle,

$X_j$ le vecteur de mesures de dimension v.

Comme dans le paragraphe précédent, nous supposons que les erreurs de mesure sont distribuées selon une loi normale de moyenne nulle et de variance constante. Cette hypothèse nous conduit à minimiser le critère $\Phi$ par rapport à $\hat{X}_j$ et $V_j$.

$$\Phi = \frac{m}{2} \mathrm{Log}(\prod_{j=1}^{n} |V_j|) + \frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{m} (\hat{X}_j - X_{ij})^T V_j^{-1} (\hat{X}_j - X_{ij})$$

sous la contrainte $M\prod_{i=1}^{j}\hat{X}_j = 0$ pour j = 1, ..., n (IV-84)

Le lagrangien associé à ce problème :

$$L = \frac{m}{2}\mathrm{Log}(\prod_{j=1}^{n} |V_j|) + \frac{1}{2}\sum_{j=1}^{n}\sum_{i=1}^{m} (\hat{X}_j - X_{ij})^T V_j^{-1} (\hat{X}_j - X_{ij}) + \sum_{j=1}^{n}\lambda_j^T M\prod_{i=1}^{j}\hat{X}_j \quad \text{(IV-85)}$$

est stationnaire pour j= 1,..., n quand :

$$\frac{\partial L}{\partial \hat{X}_j} = V_j^{-1}\sum_{i=1}^{m}(\hat{X}_j - X_{ij}) + \sum_{j=1}^{n}(M\prod_{k=1,i\neq j}^{j}\hat{X}_k)^T\lambda_k = 0 \quad \text{(IV-86-1)}$$

$$\frac{\partial L}{\partial V_j} = [m\, I_v - \mathrm{diag}(\sum_{i=1}^{m}(\hat{X}_j - X_{ij})(\hat{X}_j - X_{ij})^T)]\, V_j^{-1} = 0 \quad \text{(IV-86-2)}$$

$$\frac{\partial L}{\partial \lambda_j} = M\prod_{i=1}^{j}\hat{X}_j = 0 \quad \text{(IV-86-3)}$$

$\lambda_j$ est le vecteur des multiplicateurs de Lagrange et $I_v$ la matrice identité.

En général, ce problème n'admet pas de solution analytique. Nous proposons le calcul hiérarchisé basé sur la méthode sur la méthode de relaxation pour résoudre le système non linéaire précédent.

### IV-3-2 RÉSOLUTION HIÉRARCHISÉE

La méthode de résolution que nous proposons pour résoudre le système (IV-86) est similaire à celle utilisée dans le cas des systèmes bilinéaires. La démarche de calcul a été présentée par J. Ragot **[RAGO 89]** pour l'équilibrage de bilan des systèmes n-linéaires.

Pour simplifier la représentation des matrices contenant les estimées $\hat{X}$, nous posons :

$$M_{1.i} = M\,\Lambda_1\,\Lambda_2 \ldots \Lambda_i \quad \text{(IV-87)}$$

où $\Lambda_i$ est la matrice diagonale formée avec les estimées $\hat{X}_i$.

Les estimées peuvent alors s'écrire :

$$\hat{X}_{i+1} = [\, I - V_{i+1}M_{1.i}^T(M_{1.i}V_{i+1}M_{1.i}^T)^{-1}M_{1.i}][\bar{X}_{i+1} - V_{i+1}\Lambda_{i+2}\Lambda_{i+1}^{-1}V_{i+2}^{-1}(\hat{X}_{i+2} - \bar{X}_{i+2})] \quad \text{(IV-88)}$$

L'expression des multiplicateurs de Lagrange est donnée par :

$$\lambda_{i+1} = (M_{1.i}V_{i+1}M_{1.i}^T)^{-1}M_{1.i}[m\bar{X}_{i+1} - V_{i+1}\Lambda_{i+2}\prod_{k=1}^{i}\Lambda_k(M^T\lambda_{i+2} + \sum_{j=i+3}^{n}\prod_{k=i+3}^{j}\Lambda_k M^T\lambda_j)] \qquad \text{(IV-89)}$$

et les matrices de variance s'écrivent :

$$V_i = \frac{1}{m}\mathrm{diag}[\sum_{j=1}^{m}(\hat{X}_i - X_{ij})(\hat{X}_i - X_{ij})^T] \qquad \text{pour } i = 1, ..., n \qquad \text{(IV-90)}$$

### IV-3-2-1 Description de l'algorithme de résolution d'un système trilinéaire

La résolution du système non linéaire s'effectue alors selon une structure hiérarchisée par la méthode de relaxation à trois niveaux :

- Le premier niveau calcule $\hat{X}_1$ et $\lambda_2$ à partir de $V_1$ initialisée à la matrice identité et les variables $\hat{X}_2$ et $\hat{X}_3$ initialisées aux valeurs moyennes des mesures, puis calcule la matrice $V_1$.

- Le deuxième niveau calcule $\hat{X}_2$ et $\lambda_2$ à partir de $V_2$ et $\hat{X}_1$ fournis par le niveau 1, puis calcule la matrice $V_2$.

- Le troisième niveau calcule $\hat{X}_3$ et $\lambda_3$ à partir de $V_3$ et les vecteurs $\hat{X}_1$ et $\hat{X}_2$ fournis par le niveau 1, puis calcule la matrice $V_3$.

Le schéma de l'algorithme utilisé est présenté à la figure IV-25.

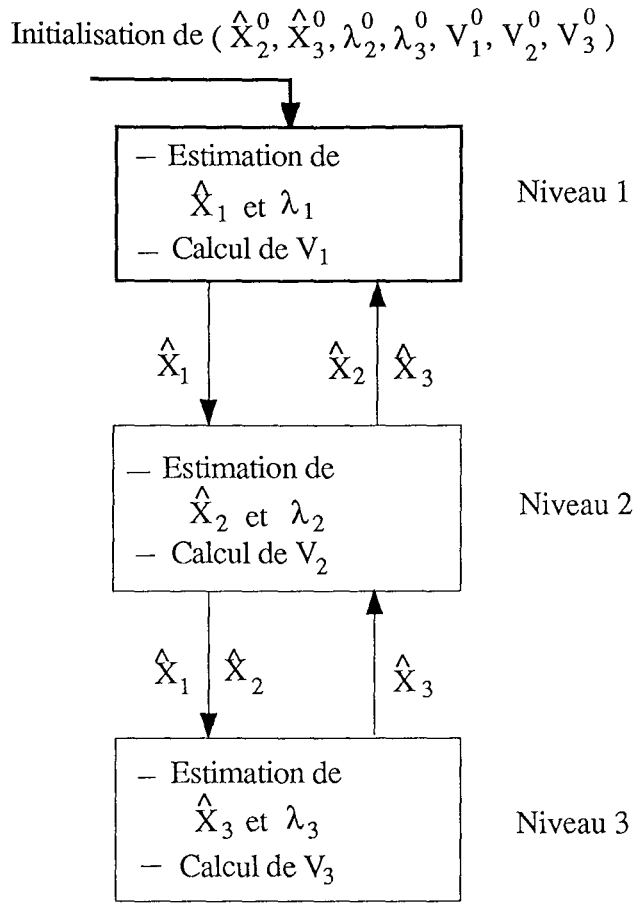



Figure IV-25 : algorithme de relaxation

Etape d'initialisation :

$$V_1^0 = V_2^0 = V_3^0 = I$$

$$\hat{X}_2^0 = \frac{1}{m} \sum_{j=1}^{m} X_{2j}$$

$$\hat{X}_3^0 = \frac{1}{m} \sum_{j=1}^{m} X_{3j}$$

et $\lambda_2^0 = \lambda_3^0 = 0$

Le test d'arrêt est effectué sur la norme carrée des dérivées premières du lagrangien autour de $10^{-5}$.

### IV-3-2-2 Application de la méthode à un système trilinéaire

Pour une raison de simplification, nous avons appliqué l'algorithme à un système trilinéaire dont les grandeurs caractéristiques sont : débit, teneur et densité.

Deux exemples de petite et moyenne dimension ont été choisis.

### Application à un réseau de petite dimension

Considérons le réseau de transport de matière représenté par la figure IV-26 comportant quatre nœuds et huit voies :

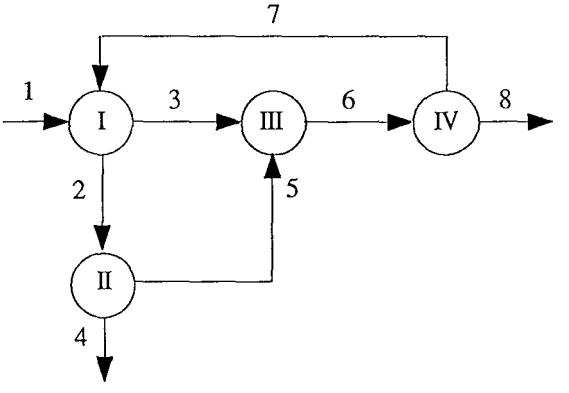



Figure IV-26 : réseau de transport de matière

Dans une première application, nous simulerons des mesures évoluant autour d'un point de fonctionnement. Les évolutions chronologiques des débits, des teneurs et des densités sont représentées par les figures IV-27, IV-28 et IV-29.

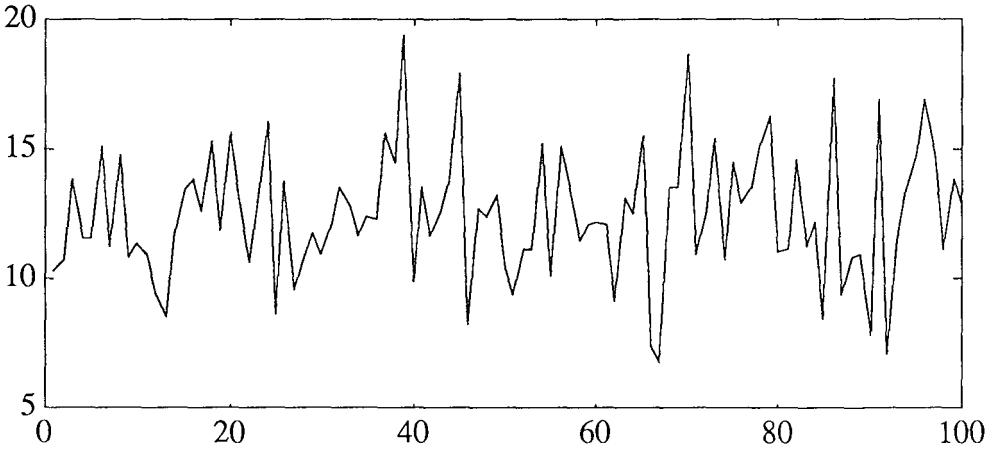



Figure IV-27 : évolution chronologique du débit de la voie 3

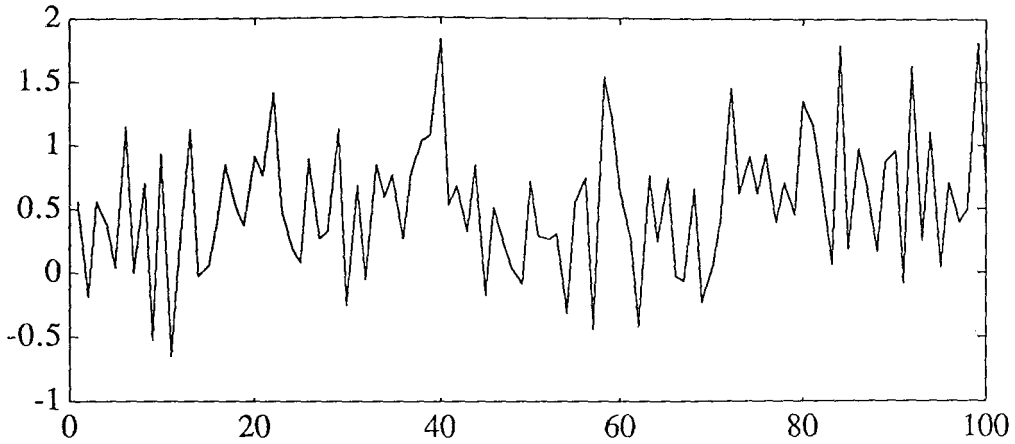



Figure IV-28 : évolution chronologique de la teneur de la voie 2

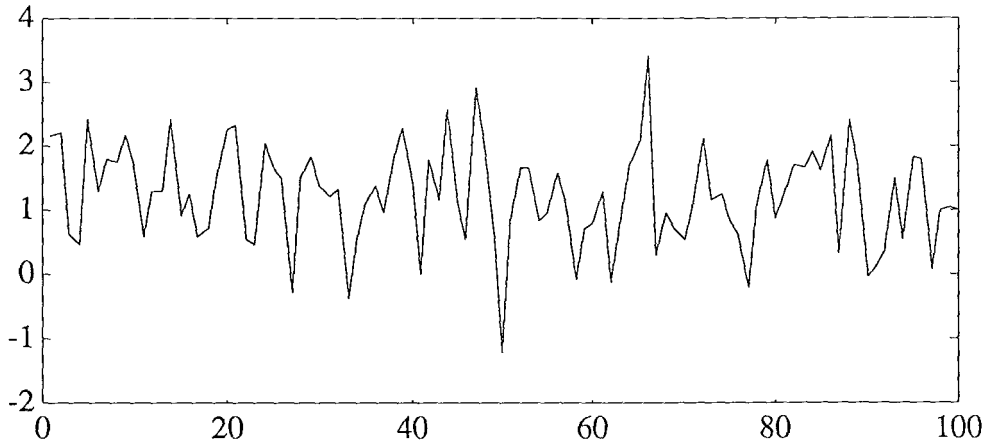



Figure IV-29 : évolution chronologique de la densité de la voie 5

Les résultats de l'estimation des grandeurs réelles et de leurs écarts-type sont présentés dans le tableau IV-17.

| voie | débits | teneurs | densités | écarts-type débits | écarts-type teneurs | écarts-type densités |
|---|---|---|---|---|---|---|
| 1 | 15.00 | 0.75 | 1.50 | 7.500 | 0.375 | 0.750 |
| 2 | 7.50 | 0.54 | 1.20 | 3.750 | 0.270 | 0.600 |
| 3 | 12.50 | 0.63 | 1.67 | 6.250 | 0.315 | 0.835 |
| 4 | 3.50 | 0.25 | 1.10 | 1.750 | 0.125 | 0.550 |
| 5 | 4.00 | 0.79 | 1.23 | 2.000 | 0.397 | 0.615 |
| 6 | 16.50 | 0.67 | 1.55 | 8.250 | 0.335 | 0.775 |
| 7 | 5.00 | 0.14 | 1.70 | 2.500 | 0.068 | 0.850 |
| 8 | 11.50 | 0.90 | 1.54 | 5.750 | 0.451 | 0.770 |

Tableau IV-17 : grandeurs réelles et leurs écarts-type

L'application de l'algorithme nous permet d'estimer les grandeurs réelles et leurs écarts-type présentés dans le tableau IV-18.

| voie | débits estimés | teneurs estimées | densités estimées | écarts-type débits | écarts-type teneurs | écarts-type densités |
|---|---|---|---|---|---|---|
| 1 | 15.01 | 0.75 | 1.50 | 7.425 | 0.371 | 0.743 |
| 2 | 7.49 | 0.54 | 1.20 | 3.712 | 0.267 | 0.594 |
| 3 | 12.50 | 0.63 | 1.67 | 6.187 | 0.312 | 0.826 |
| 4 | 3.50 | 0.25 | 1.09 | 1.732 | 0.124 | 0.544 |
| 5 | 3.99 | 0.79 | 1.22 | 1.980 | 0.393 | 0.608 |
| 6 | 16.50 | 0.66 | 1.54 | 8.167 | 0.332 | 0.767 |
| 7 | 4.99 | 0.13 | 1.69 | 2.475 | 0.066 | 0.841 |
| 8 | 11.50 | 0.90 | 1.53 | 5.692 | 0.446 | 0.762 |

Tableau IV-18 : grandeurs estimées et leurs écarts-type

Pour la deuxième application relative au même réseau, nous avons simulé trois zones de fonctionnement statique. Sur les figures IV-30, IV-31 et IV-32, la première zone est comprise entre l'observation 1 et 39, la deuxième entre 41 et 70 et la troisième entre 71 et 100.

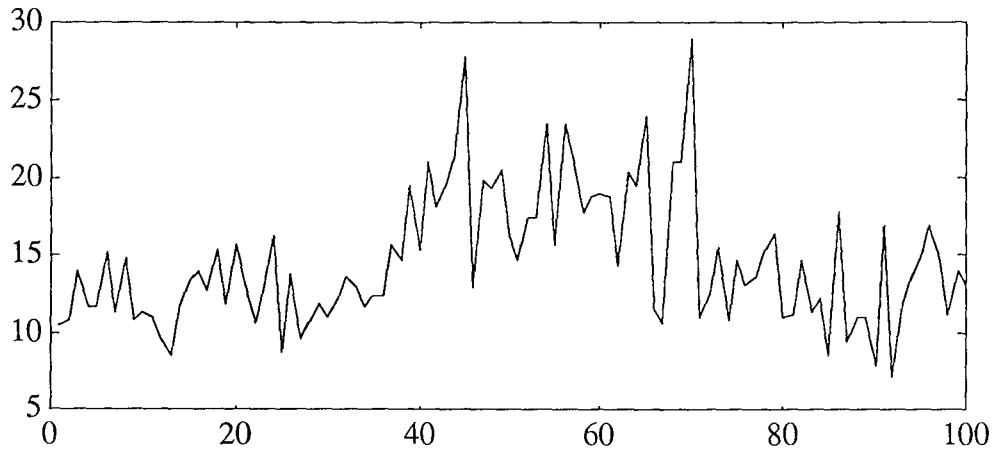



Figure IV-30 : évolution chronologique du débit de la voie 3

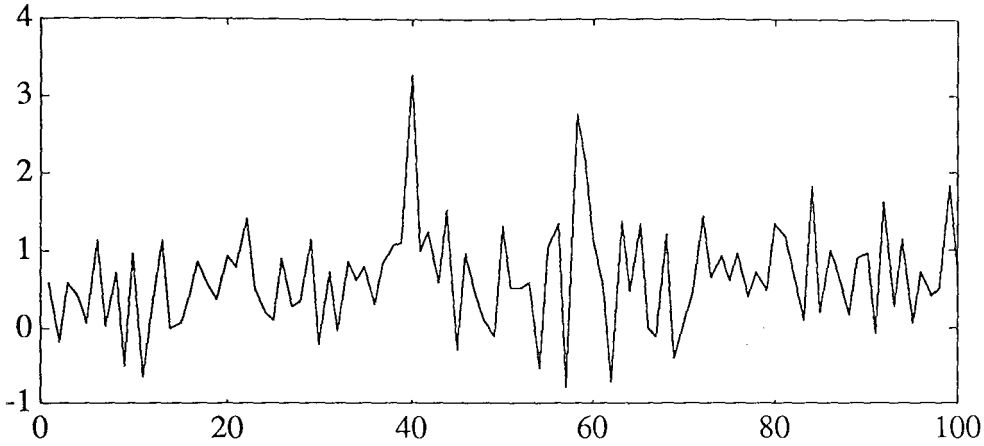



Figure IV-31 : évolution chronologique de la teneur de la voie 2

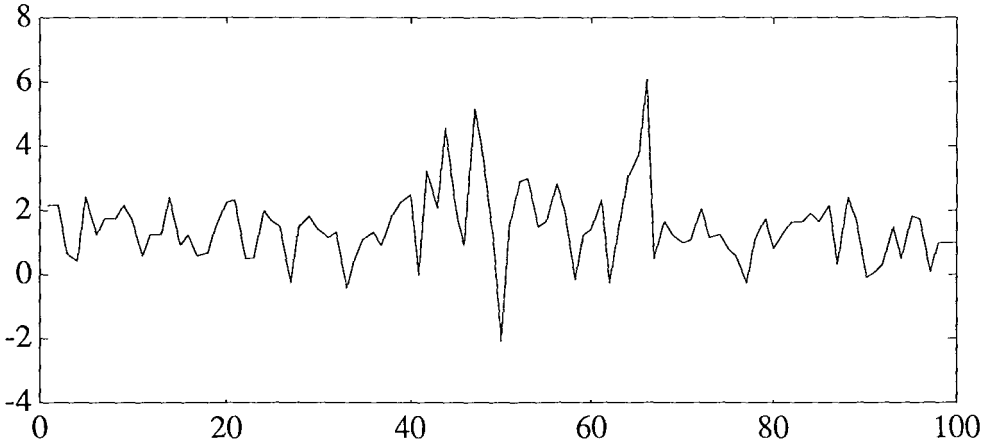



Figure IV-32 : évolution chronologique de la densité de la voie 5

Les résultats de la simulation des grandeurs réelles sont représentés dans le tableau IV-19.

| | débits massiques | | | teneurs | | | densités | | |
|---|---|---|---|---|---|---|---|---|---|
| voie | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 |
| 1 | 15.00 | 17.85 | 15.00 | 0.75 | 0.89 | 0.75 | 1.50 | 1.78 | 1.50 |
| 2 | 7.50 | 8.92 | 7.50 | 0.54 | 0.64 | 0.54 | 1.20 | 1.43 | 1.20 |
| 3 | 12.50 | 14.87 | 12.50 | 0.63 | 0.75 | 0.63 | 1.67 | 1.98 | 1.67 |
| 4 | 3.50 | 4.16 | 3.50 | 0.25 | 0.29 | 0.25 | 1.10 | 1.31 | 1.10 |
| 5 | 4.00 | 4.76 | 4.00 | 0.79 | 0.94 | 0.79 | 1.23 | 1.46 | 1.23 |
| 6 | 16.50 | 19.63 | 16.50 | 0.67 | 0.79 | 0.67 | 1.55 | 1.84 | 1.55 |
| 7 | 5.00 | 5.95 | 5.00 | 0.14 | 0.16 | 0.14 | 1.70 | 2.02 | 1.70 |
| 8 | 11.50 | 13.68 | 11.50 | 0.90 | 1.07 | 0.90 | 1.54 | 1.83 | 1.54 |

Tableau IV-19 : grandeurs réelles

Les écarts-type réels des débits, des teneurs et des densités sont :

$$\sigma_1^* = [2.956 \quad 2.075 \quad 2.676 \quad 1.419 \quad 1.531 \quad 3.029 \quad 1.686 \quad 2.505]$$

$$\sigma_2^* = [0.642 \quad 0.555 \quad 0.585 \quad 0.381 \quad 0.660 \quad 0.613 \quad 0.273 \quad 0.716]$$

$$\sigma_3^* = [0.916 \quad 0.819 \quad 0.967 \quad 0.791 \quad 0.851 \quad 0.932 \quad 0.972 \quad 0.921]$$

L'application de l'algorithme nous permet d'estimer les grandeurs réelles. Les résultats sont présentés dans le tableau IV-20.

| | débits massiques | | | teneurs | | | densités | | |
|---|---|---|---|---|---|---|---|---|---|
| voie | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 | zone 1 | zone 2 | zone 3 |
| 1 | 14.69 | 18.17 | 15.26 | 0.81 | 0.79 | 0.74 | 1.51 | 1.77 | 1.52 |
| 2 | 7.50 | 9.10 | 7.40 | 0.50 | 0.62 | 0.52 | 1.20 | 1.43 | 1.21 |
| 3 | 12.38 | 14.97 | 12.64 | 0.60 | 0.71 | 0.61 | 1.67 | 1.98 | 1.67 |
| 4 | 3.45 | 4.18 | 3.58 | 0.24 | 0.27 | 0.28 | 1.10 | 1.34 | 1.15 |
| 5 | 4.05 | 4.92 | 3.81 | 0.78 | 0.91 | 0.74 | 1.22 | 1.45 | 1.23 |
| 6 | 16.43 | 19.89 | 16.45 | 0.64 | 0.79 | 0.66 | 1.54 | 1.80 | 1.55 |
| 7 | 5.20 | 5.90 | 4.78 | 0.14 | 0.15 | 0.14 | 1.73 | 1.98 | 1.72 |
| 8 | 11.23 | 13.98 | 11.67 | 0.91 | 1.027 | 0.92 | 1.51 | 1.83 | 1.50 |

Tableau IV-20 : grandeurs estimées

Les écarts-type estimés sont :

$$\sigma_1 = [2.973 \quad 2.073 \quad 2.647 \quad 1.428 \quad 1.533 \quad 3.038 \quad 1.672 \quad 2.477]$$

$$\sigma_2 = [0.619 \quad 0.536 \quad 0.586 \quad 0.376 \quad 0.658 \quad 0.607 \quad 0.272 \quad 0.707]$$

$$\sigma_3 = [0.876 \quad 0.835 \quad 0.904 \quad 0.815 \quad 0.824 \quad 0.945 \quad 0.971 \quad 0.953]$$

## Application à un réseau de dimension moyenne

Nous considérons le réseau de transport de matière (figure IV-16) comportant sept nœuds et treize voies sur lesquelles nous avons généré des mesures de débits, de teneurs et de densités. Les évolutions chronologiques de certaines mesures sont représentées dans les figures IV-33, IV-34 et IV-35.

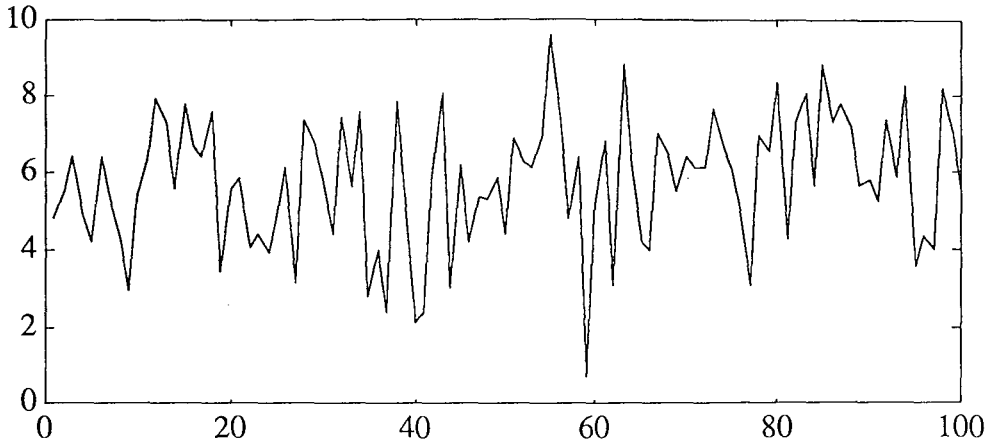



Figure IV-33 : évolution chronologique du débit de la voie 4

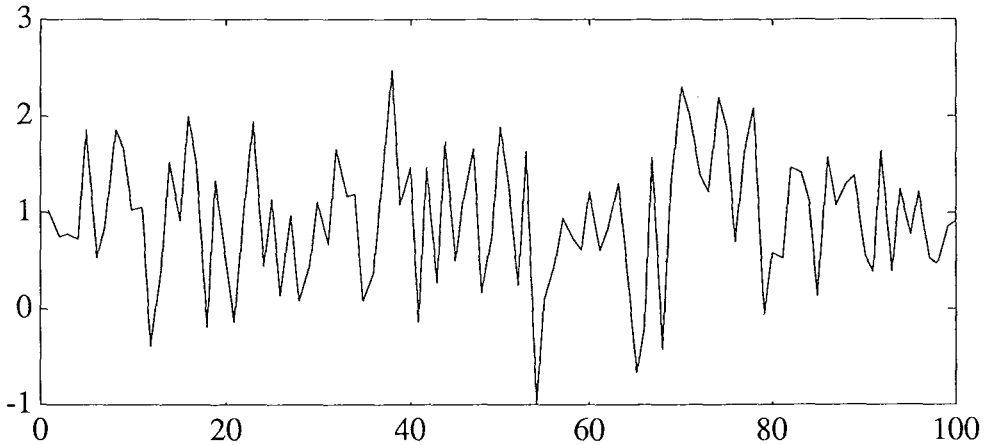



Figure IV-34 : évolution chronologique de la teneur de la voie 6

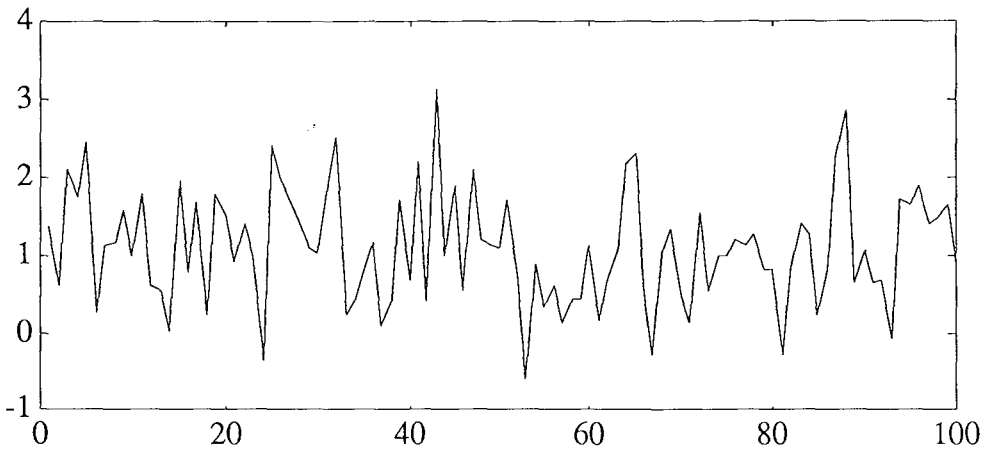



Figure IV-35 : évolution chronologique de la densité de la voie 8

Les résultats de la simulation des grandeurs réelles et de leurs écarts-type ainsi que leurs estimées sont donnés respectivement par les tableaux IV-21 et IV-22.

| voie | débits | teneurs | densités | écarts-type débits | écarts-type teneurs | écarts-type densités |
|---|---|---|---|---|---|---|
| 1 | 25.50 | 0.87 | 1.65 | 3.571 | 0.659 | 0.908 |
| 2 | 15.75 | 0.84 | 1.59 | 2.806 | 0.648 | 0.891 |
| 3 | 9.75 | 0.92 | 1.74 | 2.208 | 0.678 | 0.932 |
| 4 | 5.75 | 0.99 | 1.85 | 1.696 | 0.706 | 0.962 |
| 5 | 10.00 | 0.75 | 1.39 | 2.236 | 0.612 | 0.833 |
| 6 | 15.50 | 0.94 | 1.78 | 2.784 | 0.688 | 0.943 |
| 7 | 3.25 | 0.52 | 0.60 | 1.275 | 0.509 | 0.547 |
| 8 | 8.75 | 0.74 | 1.09 | 2.091 | 0.607 | 0.738 |
| 9 | 4.50 | 0.61 | 1.60 | 1.500 | 0.552 | 0.894 |
| 10 | 4.25 | 0.55 | 1.80 | 1.458 | 0.524 | 0.945 |
| 11 | 4.50 | 0.91 | 0.68 | 1.500 | 0.676 | 0.583 |
| 12 | 9.00 | 0.76 | 1.05 | 2.121 | 0.617 | 0.724 |
| 13 | 24.50 | 0.88 | 1.55 | 3.500 | 0.663 | 0.880 |

Tableau IV-21 : grandeurs réelles

| voie | débits | teneurs | densités | écarts-type débits | écarts-type teneurs | écarts-type densités |
|---|---|---|---|---|---|---|
| 1 | 25.50 | 0.88 | 1.64 | 3.553 | 0.656 | 0.904 |
| 2 | 15.74 | 0.83 | 1.58 | 2.792 | 0.645 | 0.887 |
| 3 | 9.73 | 0.90 | 1.73 | 2.197 | 0.674 | 0.928 |
| 4 | 5.71 | 0.97 | 1.88 | 1.687 | 0.702 | 0.957 |
| 5 | 9.99 | 0.74 | 1.41 | 2.225 | 0.609 | 0.829 |
| 6 | 15.53 | 0.94 | 1.75 | 2.769 | 0.685 | 0.939 |
| 7 | 3.21 | 0.51 | 0.61 | 1.268 | 0.507 | 0.545 |
| 8 | 8.64 | 0.73 | 1.08 | 2.081 | 0.604 | 0.734 |
| 9 | 4.49 | 0.60 | 1.58 | 1.492 | 0.549 | 0.890 |
| 10 | 4.24 | 0.56 | 1.81 | 1.450 | 0.522 | 0.944 |
| 11 | 4.51 | 0.92 | 0.67 | 1.493 | 0.672 | 0.580 |
| 12 | 9.03 | 0.75 | 1.04 | 2.111 | 0.614 | 0.703 |
| 13 | 24.57 | 0.89 | 1.53 | 3.483 | 0.659 | 0.876 |

Tableau IV-22 : grandeurs estimées

Pour les exemples que nous avons présentés, les résultats obtenus par cette méthode montrent que les grandeurs estimées sont très proches des grandeurs réelles.

## CONCLUSION

Nous avons présenté une méthode d'estimation de variance des erreurs de mesure et des grandeurs réelles basée sur la minimisation d'un critère issu de la fonction de vraisemblance sous contraintes multi-linéaires.

Un algorithme de relaxation a été mis en œuvre pour résoudre le système non linéaire issu de l'application de la technique des multiplicateurs de Lagrange. L'implantation de cet algorithme sur un calculateur est simple à réaliser.

Dans un premier temps, nous avons développé une méthode d'estimation de variance des systèmes linéaires qui se montre plus avantageuse que celles développées par Almasy [**ALMA 84**] et Darouach [**DARO 89**]. La convergence du calcul est plus rapide et la méthode demande moins de campagnes de mesure (moins d'observations): dans la plupart des exemples traités, une centaine d'observations est largement suffisante pour estimer la variance des erreurs de mesure qui est souvent plus proche de la variance réelle. Nous avons ensuite étendu cette méthode aux systèmes linéaires partiellement observés.

Dans un deuxième temps, nous avons estimé la variance des erreurs de mesure des systèmes bilinéaires complètement et partiellement observés. Le même algorithme de relaxation utilisant le principe de l'itération directe a été appliqué pour estimer cette variance que dans le cas des systèmes linéaires.

Nous avons ensuite généralisé cette technique au cas de systèmes décrits par des équations n-linéaires (fréquemment rencontrés dans les domaines chimiques et minéralogiques).

# CHAPITRE V

## DETECTION DE CAPTEURS DEFAILLANTS PAR LA METHODE DU RAPPORT DE VRAISEMBLANCE GENERALISE

## INTRODUCTION

Les techniques présentées dans les deux chapitres précédents permettent d'obtenir une estimation des grandeurs vraies qui vérifient des équations de contrainte de type linéaire ou multi-linéaire. La validité de la méthode d'estimation repose sur des hypothèses affectant la distribution des erreurs de mesures et la structure du modèle. Si ces hypothèses ne sont pas vérifiées, les grandeurs validées perdent leur signification statistique.

La nécessité d'améliorer la cohérence des mesures collectées sur un processus demeure un facteur important pour sa commande et son suivi. Cette cohérence est liée à la validité des mesures qui sont souvent entachées d'erreurs de toute nature (cf. chapitre I).

Pour détecter des défauts tels que les biais de faibles amplitudes dûs à des instruments de mesure ou des fuites, Narasimhan [**NARA 87**] utilise la méthode du rapport de vraisemblance généralisé GLR introduit par Willsky [**WILL 74**].

Nous avons constaté que ce test de rapport de probabilité d'hypothèses appliqué directement à la détection de défauts ne donne pas de résultats satisfaisants. Pour améliorer les performances de ce test, nous avons appliqué d'abord le test des résidus de bilans normalisés [**MAH 82**] pour détecter les nœuds en défaut, seules les voies appartenant à ces nœuds sont prises en compte par le test GLR.

Pour montrer l'efficacité de cette méthode, plusieurs exemples ont été simulés avec la présence d'un (ou plusieurs) défaut(s) sur une seule campagne de mesure. Ensuite, nous l'avons étendue à des exemples comportant plusieurs observations.

La méthode présentée fait appel à un seuil de détection. Ce seuil peut être défini à partir de considérations statistiques sur le niveau de signification des tests proposés.

## V-1 DÉTECTION DE BIAIS DES SYSTEMES COMPLETEMENT MESURÉS

### V-1-1 PRÉSENTATION DU PROBLEME

Notre étude sera restreinte au cas des systèmes linéaires en régime statique. De plus, nous formulons les hypothèses préliminaires suivantes :

- toutes les grandeurs sont mesurées,
- le modèle utilisé est exact,
- les erreurs de mesures sont non correlées et distribuées selon une loi normale centrée,
- la variance de ces erreurs est connue.

En l'absence de défaut l'équation de mesure s'écrit :

$$X = X^* + \varepsilon \qquad \text{(V-1)}$$

Le modèle du procédé est pris sous la forme :

$$MX^* = 0 \qquad \text{(V-2)}$$

avec

$X$ : vecteur des mesures de dimension v,
$X^*$ : vecteur des grandeurs réelles de dimension v,
$M$ : matrice d'incidence de dimension (n x v),
$\varepsilon$ : vecteur des erreurs aléatoires de mesures de dimension v.

Dans le cas où certaines mesures sont biaisées, les erreurs de mesure sont distribuées selon une loi normale de moyenne b et de matrice de variance V.

Dans une première phase nous considérons le cas où une seule mesure est en défaut et nous l'étendons ensuite aux cas de plusieurs défauts simultanés. Dans ce cas, l'expression (V-1) devient :

$$X = X^* + \varepsilon + be_i \qquad \text{(V-3)}$$

où $e_i$ est un vecteur de dimension v contenant 1 en position i et zéro ailleurs et b est l'amplitude du biais inconnu.

Le problème de détection, localisation et correction de ce biais consiste à déterminer la mesure i en défaut et à estimer l'amplitude du biais.

### V-1-2 PROPRIÉTÉS STATISTIQUES DES RÉSIDUS DES ÉQUATIONS DE BILAN

Soit un réseau redondant formé de v voies et n nœuds et décrit par sa matrice d'incidence M. Le vecteur des mesures X, de dimension v, ne vérifie pas en général les équations de bilan :

$$R = MX \qquad \text{(V-4)}$$

R est un vecteur de dimension n représentant les écarts de fermeture (ou résidus) des équations de bilan.

En l'absence de biais, les erreurs de mesure sont distribuées selon une loi normale centrée et de variance V.

A partir des équations (V-2) et (V-4), nous exprimons R en fonction de $\varepsilon$ :

$$R = M\varepsilon \qquad \text{(V-5)}$$

L'espérance mathématique des erreurs de mesures étant supposée nulle, nous en déduisons :

$$Esp(R) = MEsp(\varepsilon) = 0 \qquad \text{(V-6)}$$

La matrice de variance-covariance H, de dimension n x n, des résidus R s'écrit :

$$H = MEsp(\varepsilon\varepsilon^T)M$$

donc

$$H = MVM^T \qquad \text{(V-7)}$$

En présence de biais entachant la mesure i, le vecteur R s'exprime en fonction de $\varepsilon$ et de l'amplitude du biais :

$$R = M\varepsilon + Mbe_i \qquad \text{(V-8)}$$

De la même manière nous calculons l'espérance mathématique de R :

$$Esp(R) = bf_i \qquad \text{(V-9)}$$

avec

$$f_i = Me_i$$

La matrice de variance-covariance H des résidus R s'écrit :

$$H = Esp[(R - Esp(R))(R - Esp(R))^T]$$

donc

$$H = MVM^T \qquad \text{(V-10)}$$

Dans les deux cas, la matrice de variance-covariance des résidus est la même.

### V-1-3 INTRODUCTION DES HYPOTHESES DE DÉCISION

Pour parvenir à prendre une décision sur les caractéristiques statistiques d'une population, nous sommes souvent amenés à utiliser l'information que donne un échantillon de cette population.

Pour cela il est commode de poser deux hypothèses exhaustives par rapport auxquelles on teste l'échantillon :

- soit une hypothèse $H_0$ selon laquelle la population présente les caractéristiques statistiques A, confirmant l'absence de biais,

- et une hypothèse $H_1$ telle que la population présente les caractéristiques statistiques B, confirmant la présence de biais.

On teste ensuite les caractéristiques statistiques d'un échantillon en fonction de A et B afin de déterminer laquelle des hypothèses $H_0$ ou $H_1$ est vraie.

Pour la détection de défaillance sur un système, nous pouvons poser une hypothèse $H_0$ telle que le système ait un fonctionnement normal avec les caractéristiques statistiques A et une hypothèse $H_1$ telle que le système ait un fonctionnement anormal avec les caractéristiques statistiques B.

Dans notre cas, il s'agit de tester si les mesures confirment ou contredisent une opinion *a priori* concernant la valeur moyenne de la distribution des erreurs de mesure R.

Les hypothèses citées précédemment peuvent être explicitées sous la forme suivante :

$H_0$ : Esp(R) = 0 est l'hypothèse où aucun défaut n'est présent,

$H_1$ : Esp(R) = b $f_i$ est l'hypothèse où un défaut est présent.

Le rapport de vraisemblance de $H_1$ par rapport à $H_0$ pour R est défini par :

$$\eta = \frac{\text{Proba}(R/H_1)}{\text{Proba}(R/H_0)} \qquad \text{(V-11)}$$

En l'absence de biais :

$$\text{Proba}(R/H_0) = \frac{|H|^{-1/2}}{\sqrt{2\pi}} \exp(-\frac{1}{2} R^T H^{-1} R) \qquad \text{(V-12)}$$

En présence de biais :

$$\text{Proba}(R/H_1) = \frac{|H|^{-1/2}}{\sqrt{2\pi}} \exp(-\frac{1}{2} (R - b\, f_i)^T H^{-1} (R - b\, f_i)) \qquad \text{(V-13)}$$

A l'aide des équations (V-11), (V-12) et (V-13) nous exprimons $\eta$ en fonction de R et b :

$$\eta = \frac{\exp(-\frac{1}{2} (R - b\, f_i)^T H^{-1} (R - b\, f_i))}{\exp(-\frac{1}{2} R^T H^{-1} R)} \qquad \text{(V-14)}$$

Nous cherchons à maximiser $\eta$ par rapport à b, ceci revient à maximiser son logarithme népérien comme suit :

$$T_i = 2\,\text{Log}(\eta)$$

$$T_i = R^T H^{-1} R - (R - b\, f_i)^T H^{-1} (R - b\, f_i) \qquad \text{(V-15)}$$

En développant l'expression (V-15), nous trouvons :

$$T_i = 2b f_i^T H^{-1} R - b^2 f_i^T H^{-1} f_i \qquad \text{(V-16)}$$

### V-1-4 ESTIMATION DE L'AMPLITUDE DU BIAIS

Pour estimer l'amplitude du biais, nous calculerons le maximum de $T_i$ par rapport à $\hat{b}$, où $\hat{b}$ représente l'estimée de b :

$$\frac{\partial T_i}{\partial \hat{b}} = 2 f_i^T H^{-1} R - 2\hat{b}\, f_i^T H^{-1} f_i = 0$$

d'où

$$\hat{b} = (f_i^T H^{-1} f_i)^{-1} (f_i^T H^{-1} R) \qquad \text{(V-17)}$$

En remplaçant $\hat{b}$ dans l'équation (V-16) et puisque l'expression $(f_i^T H^{-1} f_i)$ est scalaire, nous obtenons :

$$T_i = \frac{d_i^2}{c_i} \qquad \text{(V-18)}$$

avec

$$d_i = f_i^T H^{-1} R \qquad \text{(V-19)}$$

$$c_i = f_i^T H^{-1} f_i \qquad \text{(V-20)}$$

Pour détecter et localiser la présence d'un biais nous utilisons le test du GLR défini par l'expression suivante :

$$T = \sup_i T_i \qquad \text{(V-21)}$$

Après avoir calculé l'ensemble des $T_i$, le maximum T est comparé à un seuil dont on fera l'étude par la suite. Si T est inférieur à cette valeur limite, les grandeurs mesurées sont acceptées et aucun capteur n'est reconnu défaillant. Par contre, s'il est supérieur au seuil, on suspecte individuellement toutes les mesures. A l'aide de la technique précédente, les grandeurs mesurées sont modifiées et un nouveau résidu est calculé.

## Détermination du seuil de détection

En l'absence de biais, $d_i$ suit une loi normale de moyenne nulle et de variance $c_i$ calculée à partir de l'expression (V-20).

$$d_i \sim N(0,c_i) \qquad \text{(V-22)}$$

On déduit donc que $T_i$ suit une loi du Khi 2 à v degrés de liberté.

Nous utilisons les mêmes arguments que Mah [**MAH 82b**] afin de détecter la présence d'un biais. Ainsi, nous rejetons l'hypothèse $H_0$ et nous concluons qu'il y a un biais dans la $i^{ème}$ voie si :

$$\sup_i T_i > \chi^2_{1,1-\beta} \qquad \text{(V-23)}$$

où

$$\beta = 1 - (1-\alpha)^{1/v} \qquad \text{(V-24)}$$

$\alpha$ est le seuil de confiance que l'on impose.

Ainsi, pour $\alpha = 0.2$ et $v = 6$ on aura :

$\beta = 0.0365$

A partir de la valeur de ß nous cherchons la valeur $\chi^2_{1,1-\beta}$ à l'aide de la table du Khi2, on obtient :

$$\chi^2_{1,1-\beta} = 0.238$$

## V-1-5 DÉTECTION DE PLUSIEURS DÉFAUTS

Le traitement précédent permet la détection et la localisation d'un seul défaut. Dans le cas où le processus présente plusieurs biais, nous proposons une technique basée sur l'élimination en série des biais détectés.

Nous introduisons le vecteur **Ns** contenant les nœuds suspects déterminés par la méthode des résidus normalisés. Celle-ci est basée sur l'analyse des résidus R calculés à l'aide de l'expression (V-4) ; on obtient les résidus normalisés par l'expression suivante :

$$R_{n_j} = (H_{jj})^{-1/2} R_j \qquad \text{pour } j = 1, ..., n \qquad \text{(V-25)}$$

où $H_{jj}$ sont les termes diagonaux de la matrice de covariance H.

Dans cette méthode nous identifions un biais en appliquant le test GLR donné par les expressions (V-16) et (V-21). Si un biais est identifié, nous estimons son amplitude à l'aide de l'expression (V-17), puis nous le supprimons de la mesure biaisée. Nous procédons de la même façon pour toutes les voies jusqu'à l'élimination de tous les défauts. Si le réseau est formé de **v** voies, le calcul est donc effectué **v** fois.

### Description de la procédure de détection

La procédure de détection et de localisation de défauts peut être décomposée selon les étapes suivantes :

#### étape 1- Recherche de nœuds suspects et de voies suspectes

On calcule le vecteur des résidus normalisés à l'aide de l'équation (V-25) ; les composantes de ce vecteur sont comparées à un seuil fixé par l'utilisateur. Si aucune composante de ce vecteur ne dépasse ce seuil, on passe à l'étape 4. Sinon, on localise les nœuds suspects (vecteur Ns). Toutes les voies appartenant à ces nœuds sont considérées suspectes (Vs est le vecteur des voies suspectes).

#### étape 2- Application du test GLR

Pour toutes les voie suspectées, on calcule le rapport de vraisemblance donné par les équations (V-18), (V-19) et (V-20). La plus grande valeur du vecteur d'éléments $T_i$ ( pour i = 1, ..., nombre de voies suspectes ) est comparée à un seuil donné par la relation (V-23). Si cette valeur est supérieure à ce seuil, alors la voie associée est déclarée défaillante.

**étape 3- Estimation de l'amplitude du biais**

On calcule l'amplitude du biais de la voie dont la mesure est biaisée par l'expression (V-17). On compense la mesure par élimination de ce biais, puis on recommence la procédure à partir de l'étape 1 pour détecter d'éventuels nouveaux défauts.

## V-1-6 RÉCONCILIATION DE DONNÉES

Après avoir localisé les biais et estimé leurs amplitudes, nous sommes amenés à estimer les grandeurs réelles qui vérifient les équations de bilan (V-2).

La relation (V-1) peut être écrite sous la forme suivante :

$$X = \hat{X} + \varepsilon + B \qquad \text{(V-26)}$$

avec

$$B = b\, e_i \qquad \text{(V-27)}$$

B est le vecteur des amplitudes des biais estimées de dimension v et i le numéro de la voie dont la mesure est biaisée.

d'où

$$X - B = \hat{X} + \varepsilon \qquad \text{(V-28)}$$

La technique d'équilibrage de bilan matière permet d'obtenir l'estimée $\hat{X}$ en fonction de la mesure biaisée.

$$\hat{X} = P(X - B) \qquad \text{(V-29)}$$

avec

$$P = I - VM^T(MVM^T)^{-1}M \qquad \text{(V-30)}$$

P est la matrice de projection de dimension (v x v).

## V-1-7 ESSAIS EN SIMULATION

### V-1-7-1 Application de la détection en mono-observation

Pour illustrer l'application de la méthode précédente, considérons le réseau de la figure V-1 formé de quatre nœuds et huit voies dont les mesures et les écarts-type des erreurs de mesure sont donnés dans le tableau V-1.

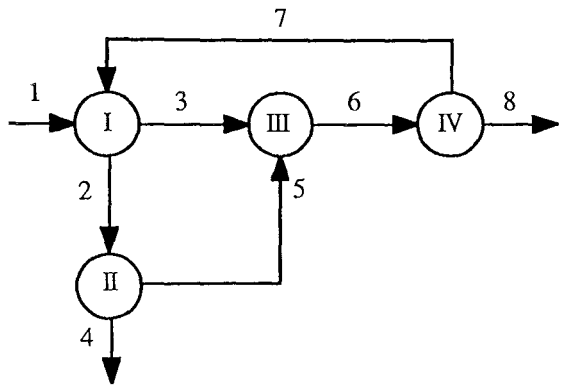



Figure V-1 : circuit élémentaire de transport de matière

Ainsi la matrice d'incidence associée à ce réseau est :

$$M = \begin{bmatrix} 1 & -1 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 \end{bmatrix}$$

| voie | mesure | écart-type |
|---|---|---|
| 1 | 15.20 | 2.32 |
| 2 | 8.31 | 1.12 |
| 3 | 13.42 | 1.87 |
| 4 | 3.25 | 0.52 |
| 5 | 5.70 | 0.60 |
| 6 | 19.75 | 2.47 |
| 7 | 5.91 | 0.75 |
| 8 | 12.90 | 1.72 |

Tableau V-1 : valeurs des mesures et leurs écarts-type

Pour cet exemple, nous allons considérer deux cas, dans le premier, nous simulerons un biais sur la voie 1, puis nous décrirons toutes les étapes de calcul qui permettent de localiser et de détecter le capteur défaillant ainsi que l'amplitude du biais qui entache la mesure délivrée par ce capteur. Dans le second cas, nous simulerons plusieurs défauts d'amplitudes variables sur toutes les voies.

Dans un premier temps, nous simulons un biais de l'ordre de quatre fois l'écart-type de la mesure de la voie 1 qui devient :

$X_1 = 24.50$

Les différentes étapes de localisation et de détection du biais avec estimation de son amplitude sont les suivantes :

**Etape 1** : calcul des résidus et résidus normalisés.

Le vecteur $R_n$ des résidus normalisés calculé à partir de l'équation (V-25) est :

$$R = \begin{bmatrix} 8.68 \\ -0.64 \\ -0.63 \\ 0.93 \end{bmatrix} \quad \text{et} \quad R_n = \begin{bmatrix} 2.69 \\ -0.47 \\ -0.20 \\ 0.30 \end{bmatrix}$$

La matrice de variance-covariance des résidus et son inverse sont :

$$H = \begin{bmatrix} 10.41 & -1.26 & -3.51 & -0.56 \\ -1.26 & 1.90 & -0.36 & 0 \\ -3.51 & -0.36 & 10.00 & -6.12 \\ -0.56 & 0 & -6.12 & 9.66 \end{bmatrix}$$

$$H^{-1} = \begin{bmatrix} 0.15 & 0.12 & 0.10 & 0.07 \\ 0.12 & 0.632 & 0.11 & 0.08 \\ 0.10 & 0.11 & 0.23 & 0.15 \\ 0.07 & 0.08 & 0.15 & 0.20 \end{bmatrix}$$

Les vecteurs Ns des nœuds suspects et Vs des voies suspectes sont :

$Ns = [\ 1\ ]$ et $Vs = [\ 1\ \ 2\ \ 3\ \ 7\ ]$

**Etape 2** : calcul du test G.L.R.

Le vecteur des $T_i$ du test GLR donnés par l'expression (V-18) est :

$T_i = [\ 0.583\ \ 0.189\ \ 0.0313\ \ 0.061\ ]$

**Etape 3** : localisation du défaut et estimation de l'amplitude du biais.

Pour un seuil de confiance $\alpha = 0.05$ et donc $\beta = 0.017$, la valeur du seuil, donnée par la table du Khi2, est égale à 0.31.

La voie 1 est en défaut puisque $\sup(T_i)$ dépasse la valeur du seuil calculée précédemment.

L'amplitude estimée du biais $\hat{b}$ donnée par la relation (V-17) est :

$\hat{b} = 8.204$

## Suppression du biais des mesures

Le nouveau vecteur des résidus, après suppression du biais, s'écrit :

$$R = \begin{bmatrix} 0.48 \\ -0.64 \\ -0.63 \\ 0.93 \end{bmatrix} \quad \text{et} \quad R_n = \begin{bmatrix} 0.15 \\ -0.47 \\ -0.20 \\ 0.30 \end{bmatrix}$$

Une fois que le biais de la voie 1 a été supprimé, nous pouvons déduire, à partir des relations (V-29) et (V-30), les estimées $\hat{X}$ données par le tableau V-2.

| voie | mesure | estimée | terme correctif | terme correctif normalisé |
|---|---|---|---|---|
| 1 | 24.50 | 16.30 | -8.20 | -3.65 |
| 2 | 8.31 | 8.74 | 0.43 | 0.38 |
| 3 | 13.42 | 13.51 | 0.09 | 0.05 |
| 4 | 3.25 | 3.16 | -0.09 | -0.18 |
| 5 | 5.70 | 5.58 | -0.11 | -0.19 |
| 6 | 19.75 | 19.10 | -0.65 | -0.26 |
| 7 | 5.91 | 5.96 | 0.05 | 0.06 |
| 8 | 12.90 | 13.14 | 0.24 | 0.14 |

Tableau V-2 : valeurs des estimées après élimination du biais

## Etude de la robustesse de la méthode

A partir d'un vecteur de grandeurs vraies et d'un vecteur d'erreurs aléatoires distribuées normalement, reliés par l'équation (V-1), nous avons simulé le vecteur de mesures des voies de la figure V-2. Nous avons ensuite biaisé chaque voie d'une amplitude variant successivement de -50%, -40%, -30%, -20%, 20%, 30%, 40% et 50% de la mesure initiale.

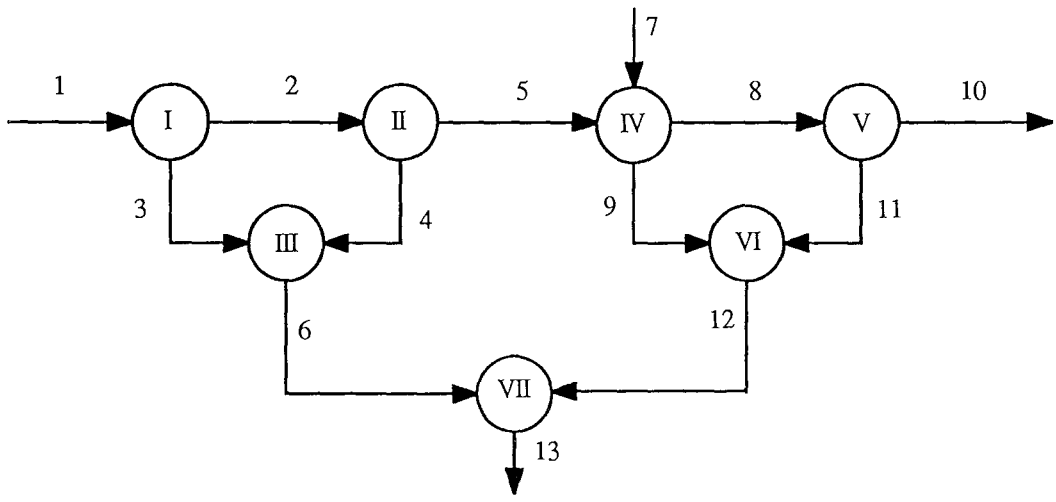



Figure V-2 : réseau de transport de matière

Le tableau V-3 présente les mesures des voies et leurs écarts-type avant de biaiser la mesure de chaque voie avec des amplitudes différentes :

| voie | mesure | écart-type |
|---|---|---|
| 1 | 25.05 | 2.525 |
| 2 | 15.22 | 1.984 |
| 3 | 9.48 | 1.561 |
| 4 | 6.55 | 1.199 |
| 5 | 7.69 | 1.581 |
| 6 | 14.79 | 1.968 |
| 7 | 3.85 | 0.901 |
| 8 | 8.14 | 1.479 |
| 9 | 4.21 | 1.061 |
| 10 | 4.75 | 1.031 |
| 11 | 4.30 | 1.061 |
| 12 | 8.85 | 1.500 |
| 13 | 25.37 | 2.475 |

Tableau V-3 : valeurs des mesures et leurs écarts-type

Les résultats des tests sont rassemblés dans le tableau V-4. A travers ces résultats, nous constatons que, sur le réseau de la figure V-2, la détection se fait à plus de 83.6%.

| voie | bonne détection % | fausse détection % | aucune détection % |
|---|---|---|---|
| 1 | 87.5 | 0 | 12.5 |
| 2 | 100 | 0 | 0 |
| 3 | 75 | 25 | 0 |
| 4 | 50 | 0 | 50 |
| 5 | 100 | 0 | 0 |
| 6 | 87.5 | 12.5 | 0 |
| 7 | 75 | 25 | 0 |
| 8 | 100 | 0 | 0 |
| 9 | 75 | 12.5 | 12.5 |
| 10 | 75 | 25 | 0 |
| 11 | 87.5 | 0 | 12.5 |
| 12 | 100 | 0 | 0 |
| 13 | 75 | 0 | 25 |

Tableau V-4 : résultats numériques de détection

Comme il était prévisible, les performances du test dépendent de la taille du réseau et du rapport entre le biais et l'amplitude de la mesure. Toutefois ces performances sont nettement supérieures à celles d'autres méthodes de détection telles que la méthode des résidus de bilan normalisés et des termes correctifs normalisés **[FAYO 87]** et celle de l'espace de parité **[AITO 89]**.

### V-1-7-2 Application de la détection en multi-observations

Nous avons appliqué la procédure du GLR au réseau de la figure V-2. Les exemples ont été simulés sur 100 observations et nous avons juxtaposé trois plages de régimes statiques.

#### 1$^{er}$ cas : simulation d'un seul défaut

Dans un premier temps, un défaut en échelon a été introduit entre les observations 30 et 40 sur la voie 2 avec une amplitude de 40% de la mesure. L'évolution du débit en fonction du nombre d'observation est donnée à la figure V-3.

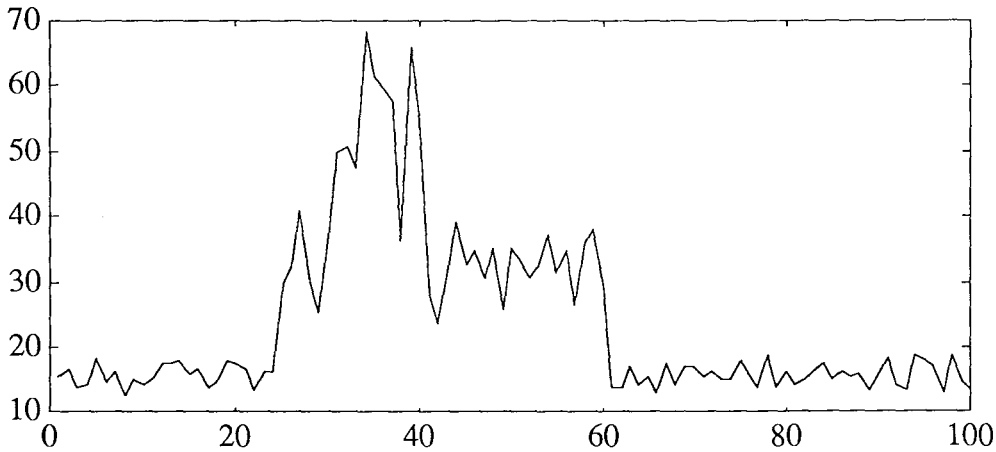



Figure V-3 : évolution chronologique du débit massique de la voie 2

Après l'application de la procédure de détection, nous avons localisé le moment d'apparition du biais sur le débit de la voie 2 sans aucun retard, à partir de la 30$^{ème}$ observation.

#### 2$^{ème}$ cas : simulation de deux défauts

Dans un deuxième temps, nous avons simulé deux défauts en échelon. Nous avons introduit un défaut sur la voie 2 avec une amplitude de 40% de la mesure entre les observations 30 et 40 (figure V-3), sur la voie 6 un défaut avec une amplitude de 30 % de la mesure entre les instants 30 et 35. L'évolution chronologique du débit de la voie 6 est présentée à la figure V-4.

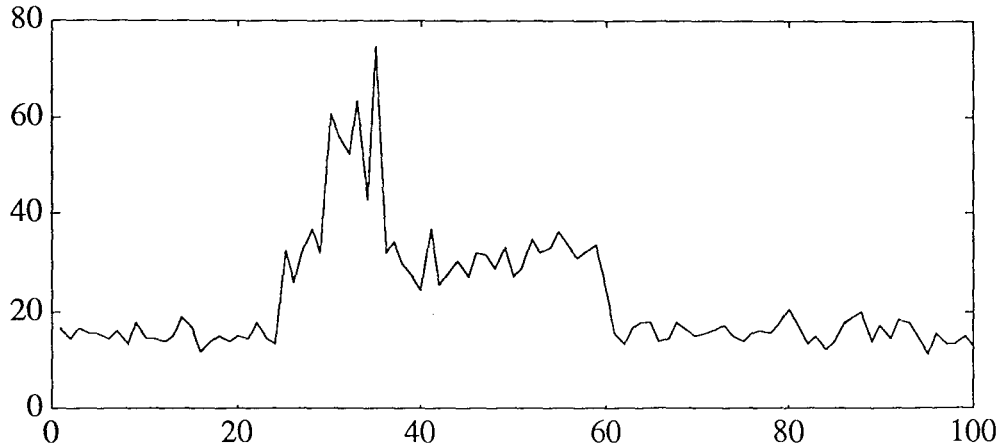



Figure V-4 : évolutions chronologiques du débit de la voie 6

Nous avons constaté, sur cet essai, que la présence de défaut sur deux voies simultanément et en présence de plusieurs zones de fonctionnement statique ne détériore pas les performances de l'algorithme de détection basé sur le test GLR.

### 3$^{ème}$ cas : simulation de trois défauts

Dans un troisième temps, nous avons simulé trois défauts sous forme de trois échelons d'amplitudes différentes. Nous avons introduit un défaut sur la voie 2 avec une amplitude de 40% de la mesure entre les observations 30 et 40 (figure V-3), sur la voie 6 un défaut avec une amplitude de 30 % de la mesure entre les observations 30 et 35 (figure V-4) et sur la voie 12 un défaut avec une amplitude de 40% de la mesure entre les observations 30 et 35. La figure V-5 représente l'évolution temporelle de la voie 12.

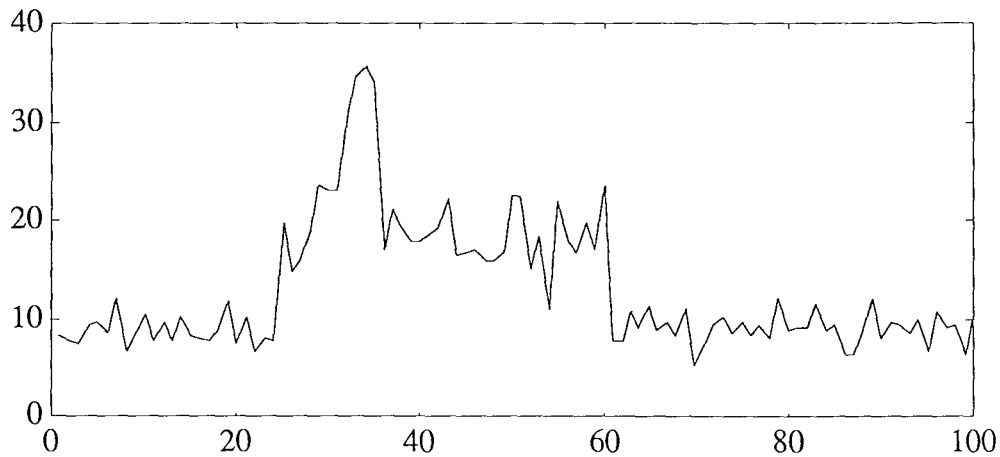



Figure V-5 : évolution temporelle du débit de la voie 12

Après avoir appliqué le test G.L.R, nous pouvons conclure que la détection des trois biais s'est effectuée à 100% à partir de la 33$^{ème}$ observation. Au travers les différentes simulations que nous avons effectuées, nous avons constaté que les performances de l'algorithme proposé sont assez peu sensibles au nombre de voies en défaut.

## V-2 DÉTECTION DE DÉFAUT DANS LE CAS DE DONNÉES MANQUANTES

### V-2-1 FORMULATION DU PROBLEME

Considérons la matrice d'incidence M représentative d'un réseau constitué de n nœuds et v voies et comportant m points de mesure ( $m \leq v$). Celle-ci peut alors être décomposée, après permutation éventuelle des colonnes, en deux sous-matrices $M_1$ et $M_2$ dont les colonnes respectives correspondent aux voies mesurées et aux voies non mesurées.

$$M = [\, M_1 \mid M_2 \,]$$

avec

$M_1$ sous-matrice de dimension n x m,
$M_2$ sous-matrice de dimension n x (v-m).

En tenant compte de cette notation, l'équation de bilan (V-2) s'écrit :

$$R = \left[\, M_1 \;\; M_2 \,\right] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

d'où :

$$R = M_1 X_1 + M_2 X_2 \qquad \text{(V-31)}$$

avec

$X_1$ vecteur des voies mesurées de dimension m,
$X_2$ vecteur des voies non mesurées de dimension v-m.

On peut partitionner $M_1$ et $M_2$ comme suit :

$$M = \begin{array}{c} \left[\begin{array}{c|c|c} M_{11} & M_{21} & M_{22} \\ \hline M_{12} & M_{23} & M_{24} \end{array}\right] \begin{array}{l} \updownarrow r \\ \updownarrow n\text{-}r \end{array} \\ \begin{array}{ccc} \longleftrightarrow m & \longleftrightarrow r & \longleftrightarrow v\text{-}m\text{-}r \end{array} \end{array} \qquad \text{(V-32)}$$

où $M_{21}$ est la sous-matrice régulière de rang maximal de la matrice $M_2$. Les équations de bilan peuvent alors s'écrire :

$$M_{21}^{-1} M_{11} \hat{X}_1 + \hat{X}_{21} + M_{21}^{-1} M_{22} \hat{X}_{22} = 0 \qquad \text{(V-33)}$$

$$M_{12} \hat{X}_1 + M_{12} \hat{X}_{21} + M_{24} \hat{X}_{22} = 0 \qquad \text{(V-34)}$$

Par combinaison linéaire des expressions (V-33) et (V-34) [**MAQU 87**], nous obtenons la partie redondante de la matrice d'incidence écrite sous la forme :

$$A\hat{X}_1 = 0 \qquad \text{(V-35)}$$

avec

$$A = M_{12} - M_{23} M_{21}^{-1} M_{11} \qquad \text{(V-36)}$$

### V-2-2 RÉCONCILIATION DE DONNÉES

On cherche à minimiser le critère quadratique $\Phi$ par rapport $\hat{X}_1$ :

$$\Phi = \| \hat{X}_1 - X_1 \|^2_{V_1^{-1}} \qquad \text{(V-37)}$$

sous la contrainte $A\hat{X}_1 = 0$

où $V_1$ est la matrice de variance de $X_1$ de dimension m x m.

En appliquant la méthode des multiplicateurs de Lagrange, nous obtenons les estimées $\hat{X}_1$ :

$$\hat{X}_1 = P_1 X_1 \qquad \text{(V-38)}$$

avec

$$P_1 = I - V_1 A^T (A V_1 A^T)^{-1} A \qquad \text{(V-39)}$$

Dans la méthode du G.L.R, la matrice A remplacera la matrice d'incidence M et $X_1$ le vecteur de mesure X.

### V-2-3 EXEMPLE DE SIMULATION NUMÉRIQUE

Nous considérons le réseau de la figure V-6, comportant huit nœuds et quinze voies pour lequel nous avons simulé les débits des voies (1, 2, 4, 6, 8, 10 et 14).

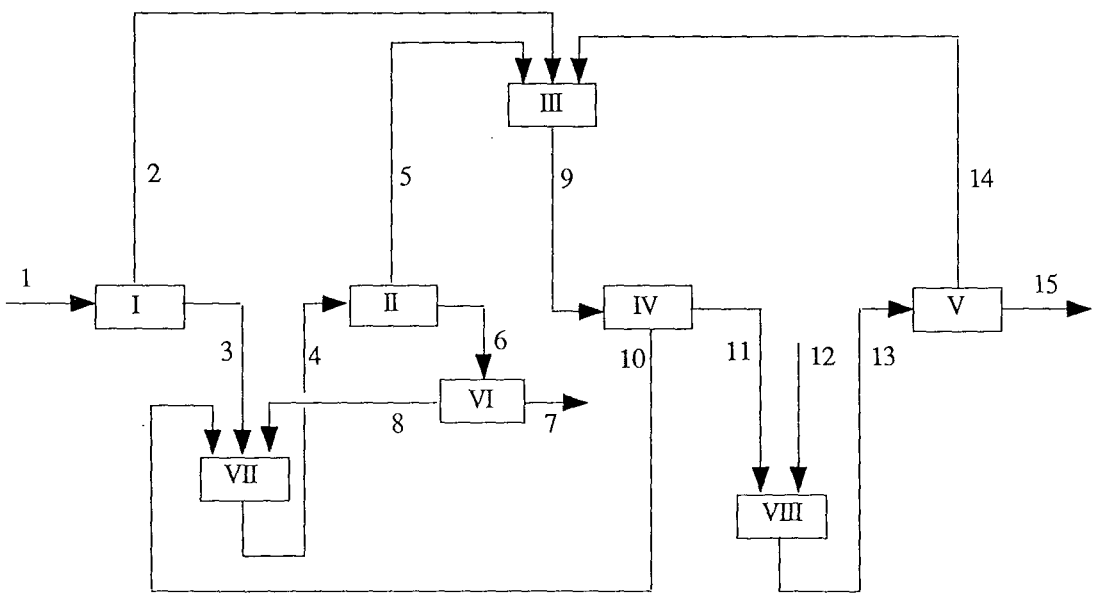



Figure V-6 : réseau de transport de matière

La matrice d'incidence M est :

| voie | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| II | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| III | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | 0 |
| IV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 |
| V | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 |
| VI | 0 | 0 | 0 | 0 | 0 | 1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| VII | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| VIII | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | -1 | 0 | 0 |

Par permutation de ses colonnes, on réordonne tout d'abord la matrice M de manière à isoler la partie mesurée de celle qui ne l'est pas.

| voie | 1 | 2 | 4 | 6 | 8 | 10 | 14 | 3 | 5 | 7 | 9 | 11 | 12 | 13 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | 1 | -1 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| II | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 |
| III | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | -1 | 0 | 0 | 0 | 0 |
| IV | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 |
| V | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 |
| VI | 0 | 0 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 |
| VII | 0 | 0 | -1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| VIII | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | -1 | 0 |

partie mesurée ($M_1$) partie non mesurée ($M_2$)

Après l'application de la procédure de recherche de la partie régulière $M_{21}$, nous réordonnons la matrice d'incidence pour obtenir une forme canonique équivalente avec la permutation des lignes de la matrice précédente, la nouvelle matrice est la suivante :

| voie | 1 | 2 | 4 | 6 | 8 | 10 | 14 | 3 | 5 | 7 | 9 | 11 | 12 | 13 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | -1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| II | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VI | 0 | 0 | 0 | -1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| III | 0 | -1 | -1 | 1 | 0 | 0 | -1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| IV | 0 | -1 | -1 | 1 | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| VIII | 0 | 1 | 0 | -1 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -1 |
| V | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | -1 |
| VII | 1 | -1 | -1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

A partir de cette matrice nous pouvons déduire la matrice A qui contient les équations de redondance :

$$A = [\ 1 \quad -1 \quad -1 \quad 0 \quad 1 \quad 1 \quad 0\ ]$$

En suivant la démarche utilisée dans le paragraphe précédent consistant à faire varier l'amplitude du biais successivement de -50%, -40%, -20%, 20%, 30%, 40% et 50% de la mesure initiale effectuée sur les voies mesurées, nous avons obtenu plus de 75% de détection.

## CONCLUSION

Une méthode systématique de détection et localisation de biais appliquée aux processus physiques linéaires en régime de fonctionnement statique a été mise en œuvre. En outre, nous avons considéré, dans un premier temps, que toutes les voies sont mesurées et que les erreurs de mesures sont centrées et de variances constantes. Dans un deuxième temps, nous avons réalisé la détection sur des systèmes linéaires partiellement observés.

Dans les deux cas, la procédure de détection se déroule avant l'étape de validation. Une fois les erreurs systématiques localisées, nous estimons les amplitudes de leurs biais puis on les supprime des mesures.

L'efficacité de la méthode proposée peut être comparée à d'autres méthodes de détection telles que celles citées en annexe.

# CONCLUSION

Dans cette étude, nous avons développé des méthodes qui permettent de réconcilier des données issues d'un procédé industriel en fonctionnement statique, décrit par des équations linéaires ou multi-linéaires structurellement exactes. Ces méthodes supposent que la matrice de variance-covariance des erreurs de mesure est inconnue.

Dans un premier temps, nous avons traité le problème d'estimation de variance connue partiellement. Cette connaissance provient du rapport de précision entre les différents capteurs. Les résultats d'estimations, obtenus par l'application de l'algorithme à itération directe, sont très proches des valeurs réelles.

Dans un deuxième temps, nous avons supposé que la matrice de variance est inconnue. L'algorithme à itération directe a été appliqué successivement aux systèmes linéaires, bilinéaires et multi-linéaires. Nous avons constaté que la convergence du calcul, sur de nombreux essais de simulation, demande moins d'observations que certaines méthodes que nous avons citées. En effet, dans la plupart de ces essais, quelques dizaines d'observations suffisent pour que les grandeurs estimées soient plus proches des grandeurs réelles. Du point de vue numérique, cette méthode converge rapidement.

Ces méthodes ont été appliquées sur des signaux simulés. Il nous semble nécessaire de valider les algorithmes que nous avons mis en œuvre sur des exemples industriels.

Dans le cas des systèmes linéaires un test de détection basé sur le rapport de vraisemblance de deux hypothèses a été utilisé pour la localisation des capteurs défaillants. Cette technique basée sur une compensation en série des mesures suspectes est plus avantageuse que les méthodes employant une élimination directe de ces mesures telle que la méthode des termes correctifs normalisés .

Notre étude s'est révélée riche d'enseignements et nous a permis de trouver des solutions au problème d'estimation de variance des erreurs de mesure basée sur l'hypothèse de la normalité des erreurs de mesure. Cependant, certains points particuliers méritent, selon nous, des développement théoriques :

- étendre l'algorithme de relaxation à des systèmes non-linéaires en régime de fonctionnement statique (par exemple en linéarisant les équations de contraintes),

- étendre les différents algorithmes aux systèmes complètement observés évoluant en régime dynamique,

- établir des relations récursives pour l'algorithme de relaxation qui seront bien adaptés au traitement en ligne,

- étudier l'influence de l'amplitude des bruits de mesure sur l'estimation de la variance.

Les points cités précédemment constituent à notre sens des axes de recherche qu'il conviendra d'explorer.

# ANNEXES

# - ANNEXE A -

## DÉFINITIONS ET PROPRIÉTÉS DE CERTAINS PRODUITS MATRICIELS

### A-1 PRODUIT DE HADAMARD

Le produit de Hadamard d'un vecteur X de composantes $x_i$ et d'un vecteur Y de composantes $y_i$ est le vecteur Z de composantes $z_i$ tel que :

$$z_i = (X * Y)_i = x_i y_i \qquad \text{(A-1)}$$

Par extension, si A et B sont deux matrices de dimensions respectives (p x q) et (q x r),

$a_{ij}$ le $(i,j)^{\text{ième}}$ élément de A,

$B_j^T$ la $j^{\text{ième}}$ ligne de la matrice B.

La $(i,j)^{\text{ième}}$ partition du produit $A * B$ est définie par :

$$(A * B)_{ij} = a_{ij} B_j^T \qquad \text{(A-2)}$$

La matrice résultat, $A * B$ est de dimension p x (q x r)

### Propriétés du produit *

Si A est une matrice et X et Y sont deux vecteurs dont les dimensions sont cohérentes, on a :

$$A\,(X * Y) = A\,(Y * X) = (A * Y)\,X = (A * X)\,Y \qquad \text{(A-3)}$$

En effet :

$$(A\,(X * Y))_i = \sum_j a_{ij} X_j Y_j = \sum_j a_{ij} Y_j X_j = (A\,(Y * X))_i$$

$$= (\sum_j (a_{ij} X_j)) Y_j = ((A * X) Y)_i$$

$$= (\sum_j (a_{ij} Y_j)) X_j = ((A * Y) X)_i$$

## A-2 PRODUIT DE KRONECKER DÉCRIT PAR L'OPÉRATEUR vec(.)

Soit $A = \{a_{ij}\}$ une matrice de dimension n x m et B une matrice de dimension (s x t), alors le produit de Kronecker de A et B, noté $A \otimes B$, est défini par :

$$A \otimes B = \{a_{ij} \,.\, B\} \qquad \text{(A-4)}$$

avec

$a_{ij}$ . B le produit matriciel de B par le scalaire $a_{ij}$.

La $j^{ème}$ colonne de la matrice A est notée $a_j$ :

$$A = (a_1, a_2, ...., a_m) \qquad \text{(A-5)}$$

Le vecteur colonne (m . n) s'écrit :

$$\text{vec}(A) = \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_m \end{bmatrix}$$

vec(A) est obtenu en mettant les colonnes de la matrice A l'une au dessous de l'autre.

**Propriétés du produit ⊗**

$$\text{vec}(A.B.C) = (C^T \otimes A) \,.\, \text{vec}(B) \qquad \text{(A-7)}$$

par conséquent :

$$\begin{aligned} \text{vec}(A \,.\, B) &= (I \otimes A) \,.\, \text{vec}(B) \\ &= (B^T \otimes I) \,.\, \text{vec}(A) \\ &= (B^T \otimes A) \,.\, \text{vec}(I) \qquad \text{(A-8)} \end{aligned}$$

vec(.) est une transformation linéaire, avec des scalaires quelconque $c_1$ et $c_2$, et les matrices $A_1$ et $A_2$ de même dimension :

$$\text{vec}(c_1A_1 + c_2A_2) = c_1.\text{vec}(A_1) + c_2.\ \text{vec}(A_2) \qquad \text{(A-9)}$$

# - ANNEXE B -

## RAPPEL DES TESTS DE DETECTION UTILISES DANS L'EQUILIBRAGE DE BILAN MATIERE APPLICATION AUX DONNEES MANQUANTES

### B-1 DÉTECTION DE DÉFAUT PAR LA MÉTHODE DES RÉSIDUS DE BILAN NORMALISÉS

Dans le cas de données manquantes, nous utilisons le critère résiduel qui permet de porter un jugement global sur la qualité des grandeurs mesurées avant de les valider. Pour cela, nous posons :

$$R_1 = A\,X_1 \qquad \text{(B-1)}$$

avec

- $R_1$ : vecteur des résidus de bilan de dimension r,
- $A$ : matrice de dimension (r x m) contenant la partie redondante de la matrice d'incidence M,
- $X_1$ : vecteur des grandeurs mesurées de dimension m,
- $r$ : rang de la matrice A,
- $m$ : nombre de points de mesure.

En l'absence de biais, le vecteur $R_1$ suit une loi normale de moyenne nulle et de variance $H_1$ définie par :

$$H_1 = AV_1A^T \qquad \text{(B-2)}$$

où $V_1$ est la matrice de variance des erreurs de mesure associée à $X_1$ de dimension (m x m).

Le critère résiduel utilisé par Almasy **[ALM 75]** est le suivant :

$$\Phi_R = R_1^T\,H_1^{-1}\,R_1 \qquad \text{(B-3)}$$

Ce critère suit une loi du Khi2 dont le nombre de degrés de liberté est égal au rang de la matrice A. Pour localiser et détecter les biais, nous normalisons le vecteur des résidus $R_1$ comme suit :

$$R_{n1} = [\mathrm{diag}(H_1)]^{-1/2}R_1 \qquad \text{(B-4)}$$

où $\mathrm{diag}(H_1)$ est la matrice diagonale extraite de $H_1$ et $R_{n1}$ le vecteur des résidus normalisés.

Si une des composantes du vecteur $R_{n1}$ est supérieure à un seuil donné par la loi normale, alors la voie correspondant à cette composante est biaisée.

## B-2 DÉTECTION DE DÉFAUT PAR LA MÉTHODE DES RÉSIDUS DE MESURE NORMALISÉS

Cette approche nécessite la réconciliation de données qui est effectuée en optimisant un critère quadratique sous contraintes linéaires.

$$\min_{\hat{X}_1} \Phi = \frac{1}{2} \| \hat{X}_1 - X_1 \|^2_{V_1^{-1}} \qquad \text{(B-5)}$$

$$\text{sous contrainte } A \hat{X}_1 = 0$$

Ce critère est minimum pour :

$$\hat{X}_1 = (I - V_1 A^T (A V_1 A^T)^{-1} A) X_1 \qquad \text{(B-6)}$$

Les termes correctifs e sont définis comme suit :

$$e = \hat{X}_1 - X_1 \qquad \text{(B-7)}$$

Transformons ce vecteur de dimension m par :

$$d = V_1^{-1} e \qquad \text{(B-8)}$$

Le vecteur d est normalement distribué, de moyenne nulle et de variance W.

$$W = A^T (A V_1 A^T)^{-1} A$$

Pour pouvoir comparer les différentes composantes du vecteur d; nous définissons pour chaque composante la variable $z_i$ qui suit une loi normale centrée réduite.

$$z_i = (W_{ii})^{-1/2} d_i \qquad \text{pour } 1 \leq i \leq m \qquad \text{(B-9)}$$

Si $|z_i|$ est supérieur à un seuil donné par la loi normale alors la mesure de la voie i est biaisée.

## B-3 DÉTECTION DE DÉFAUT PAR UTILISATION DE LA MATRICE DE PROJECTION [CROW 83]

L'équation de contrainte ou de bilan d'un système linéaire est définie par :

$$M X^* = 0 \tag{B-10}$$

Après permutation des colonnes de la matrice d'incidence en deux sous-matrices $M_1$ et $M_2$ dont les colonnes respectives correspondent aux voies mesurées et aux voies non mesurées. L'équation (B-10) devient :

$$M_1 X_1^* + M_2 X_2^* = 0 \tag{B-11}$$

Nous prémultiplions l'équation (B-11) par une matrice de projection $P_1$ de dimension (n-r) x n , nous aurons :

$$P_1 M_1 X_1^* + P_1 M_2 X_2^* = 0 \tag{B-12}$$

telle que :

$$P_1 M_2 = 0 \tag{B-13}$$

L'équation (B-12) devient :

$$A X_1^* = 0 \tag{B-14}$$

avec

$$A = P_1 M_1 \tag{B-15}$$

Le problème de validation de données dans le cas de données manquantes s'écrit :

$$\min_{\hat{X}_1} \Phi = \frac{1}{2} \| \hat{X}_1 - X_1 \|^2_{V_1^{-1}}$$

$$\text{sous la contrainte } A \hat{X}_1 = 0 \tag{B-16}$$

Pour résoudre ce problème d'optimisation sous contraintes linéaires, nous proposons d'abord une procédure de calcul de la matrice de projection $P_1$. Cette procédure se décompose en trois étapes :

**étape 1** : Recherche de la partie régulière de la matrice $M_2$

Après permutations des colonnes de la matrice $M_2$, on trouve :

$$M_2 = [\, M_0 \quad 0 \,] \tag{B-17}$$

où $M_0$ est une matrice de dimension n x (v-m-r)

avec

$$\text{rang}(M_0) = \text{rang}(M_2) = r$$

**étape 2** : Décomposition de la matrice $M_0$

On permute les lignes de la matrice $M_0$ pour obtenir sa partie régulière telle que :

$$M_0 D = \begin{bmatrix} M_{01} \\ M_{02} \end{bmatrix} \qquad \text{(B-18)}$$

où

$M_{01}$ est une matrice régulière de dimension r x r,
$M_{02}$ une matrice de dimension (n-r) x r,
D une matrice de permutation de dimension n x n.

**étape 3** : Calcul de la matrice de projection

$$P_1 = [\, - M_{02}M_{01}^{-1} \quad I_r \,]\, D \qquad \text{(B-19)}$$

où $I_r$ est une matrice identité de dimension r x r.

## Exemple d'application

L'exemple de la figure B-1 qui représente le réseau utilisé par Rosenberg [**ROSE 87**] est constitué de dix nœuds et seize voies et un nœud d'environnement. Sur ce réseau les débits de dix voies ont été simulés.

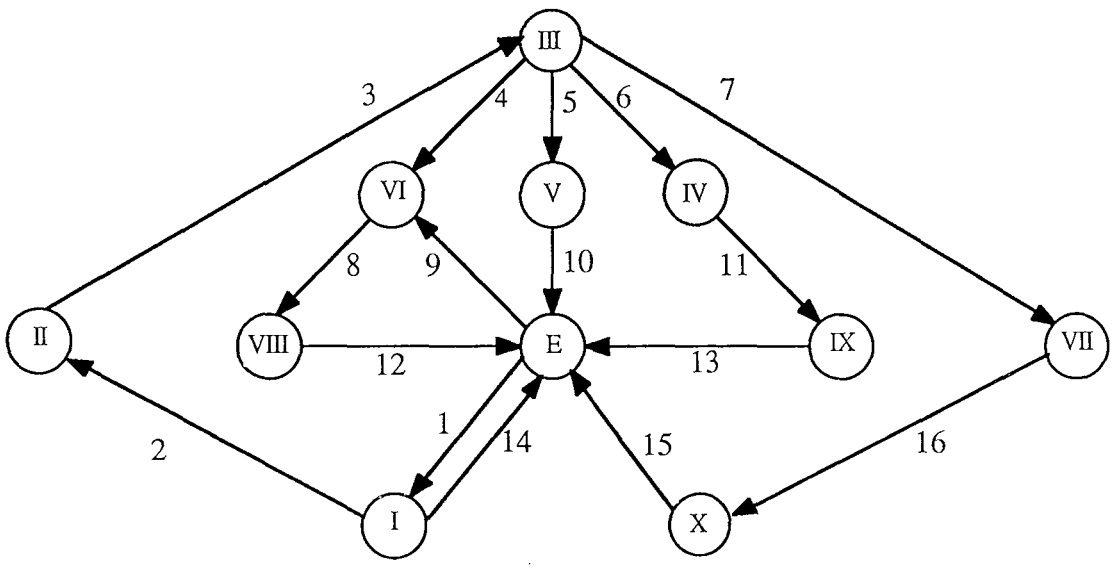



Figure B-1 : réseau de transport de matière

Les vecteurs des voies mesurées et non mesurées sont :

$$X_1 = [\ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10]^T$$

$X_2 = [11, 12, 13, 14, 15, 16]^T$

Ainsi sa matrice d'incidence est :

| voie | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 0 | 0 |
| II | 0 | 1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| III | 0 | 0 | 1 | -1 | -1 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| IV | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -1 | -1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| V | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 | 0 |
| VI | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 | 0 |
| VII | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 |
| VIII | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | -1 | 0 | 0 | 0 | 0 |
| IX | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | -1 | 0 | 0 | 0 |
| X | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | -1 | 1 |

partie mesurée ($M_1$) partie non mesurée ($M_2$)

**étape 1** : Recherche de la partie régulière de la matrice $M_2$

Après permutation des colonnes de la matrice $M_2$, nous obtenons la matrice $M_0$ de même rang que la matrice $M_2$, regroupant uniquement les colonnes 11, 12, 13, 14 et 16 :

$$M_0 = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

**étape 2** : Décomposition de la matrice $M_0$ en $M_{01}$ et $M_{02}$

$$M_{01} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

et

$$M_{02} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

La matrice de permutation D est égale à :

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

**étape 3** : Calcul de la matrice de projection

$$P_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Finalement, nous calculons à partir de l'équation (B-15) la matrice A :

$$A = \begin{bmatrix} 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Après avoir obtenu la matrice A, on peut appliquer les deux tests de détection utilisés précédemment.

# REFERENCES

[AITO 89] A. AITOUCHE, F. KRATZ, D. MAQUIN, J. RAGOT
"Détection et localisation d'erreurs de mesure par la technique de l'espace de parité", Conférences Capteurs 89, pp. 533-538, 1989.

[ALMA 75] G.A. ALMASY, T. SZTANO
"Checking and correction of measurements on the basis of linear system model", Problems of Control and Information Theory, vol. 4, n°1, pp. 57-69, 1975.

[ALMA 76] G.A. ALMASY, T. SZTANO
"Empirical models satisfying balance equations", Institute for Computing and Automation, Hungarian Academy of Sciences, 1976.

[ALMA 84] G.A. ALMASY, R.S. MAH
"Estimation of measurement error variances from process data", Industrial Engineering Chemistry Process Design Development, vol. 23, pp. 779-784, 1984.

[ASCH 82] G. ASCH
*Les capteurs en instrumentation industrielle*, Edition Dunod, 1982.

[BARD 74] Y. BARD
*Non linear parameter estimation*, Academic Press, 1974.

[BREP 84] J.C. BREPSON
"Validation de données et identification de systèmes dynamiques par équilibrage de bilan-matière", Thèse de 3ème cycle, Université de Nancy I, janvier 1984.

[CEA 78] COMMISSARIAT A L'ENERGIE ATOMIQUE
*Statistique appliquée à l'exploitation des mesures*, Edition Masson, 1978.

[CERR 80] M. CERR
*Instrumentation industrielle*, vol. 1, Techniques et Documentation, 1980.

[CHAF 87] H. CHAFOUK, J. RAGOT, J. FAYOLLE, D. MAQUIN
"Validation de données à variance inconnue", 16th IASTED international conference on identification, modeling and simulation, Paris, 1987.

[CIAR 82] P.G. CIARLET
*Introduction à l'analyse numérique matricielle et l'optimisation*, Edition Masson, 1982.

[COHE 82] I. COHEN, R. HANUS
"Influence de la fiabilité des capteurs sur la précision des bilans-matière, ou de l'influence des bilans-matière sur la précision et la fiabilité des capteurs", Congrès SIM, Besançon, 1982.

[CROW 83] C.M. CROWE, Y.A. GARCIA CAMPOS, A. HRYMAK
"Reconciliation of process flow rates by matrix projection, part 1 : the linear case", AIChE Journal, vol. 29, n° 6, pp. 881-888, 1983.

[CROW 88] C.M. CROWE
"Recursive identification of gross error in linear data reconciliation", AIChE Journal, vol. 34, n° 4, pp. 541-550, 1988.

[CROW 89] C.M. CROWE
"Observability and redundancy of process data for steady-state reconciliation", Chemical Engineering Science, vol. 44, n° 12, pp. 2909-2917, 1989.

**[DARO 83]** M. DAROUACH
"Aide au diagnostic de l'état des systèmes. Validation de mesures par équilibrage de bilan-matière", Thèse de Docteur-Ingénieur, Université de Nancy I, mars 1983.

**[DARO 86a]** M. DAROUACH, J. RAGOT, D. MAQUIN, J. FAYOLLE
"Calcul hiérarchisé et décomposition des systèmes linéaires et bilinéaires appliqué à la validation de données", APII, vol. 20, pp. 405-432, 1986.

**[DARO 86b]** M. DAROUACH, J. RAGOT, J. FAYOLLE, D. MAQUIN
"Validation de mesures par équilibrage hiérarchisé de bilans. Concepts et méthodologie", International Journal of Mineral Processing, vol. 17, pp. 273-285, 1986.

**[DARO 86c]** M. DAROUACH
"Observabilité et validation des données des systèmes de grande dimension, application à l'équilibrage des bilans de mesures", Thèse d'état, Université de Nancy I, juin 1986.

**[DARO 88]** M. DAROUACH, J. RAGOT, J. FAYOLLE, D. MAQUIN
"Data validation in large scale steady-state linear systems", 12th world congress on scientific computation, IMACS' 88, Paris, 18-22 juillet 1988.

**[DARO 89]** M. DAROUACH, J. RAGOT, M. ZASADZINSKI, G. KRZAKALA
"Maximum likehood estimator of measurement error variances in data reconciliation", Congrès IFAC AIPAC' 89, Nancy, 3-5 juillet 1989.

**[FAYO 87a]** J. FAYOLLE, D. MAQUIN, M. DAROUACH, J. RAGOT
"Analyse interactive de réseaux de transport de matière", Revue de l'Industrie Minérale, Les Techniques, pp. 274-280, 1987.

**[FAYO 87b]** J. FAYOLLE
"Validation de données et diagnostic de procédés industriels", Thèse de l'Université de Nancy I, octobre 1987.

**[GASP 77]** P. GASPART
"Amélioration de la précision des mesures industrielles redondantes", Congrès IASTED MECO'77, Zurich, 1977.

**[GERT 81]** J. GERTLER
"Estimating the noise covariance matrix for input-output filters", IFAC 8th world Congress, Kyoto 1981.

**[GOMO 89]** W. GOMOLKA
"Strategies of process data reconciliation in complex system described by sets of bilinear balance equations", Congrès IFAC AIPAC' 89, Nancy, 3-5 juillet 1989.

**[HAMA 86]** M. HAMAD
"Validation de mesures et détection de capteurs défaillants dans un système de contrôle commande", Thèse de l'Université des Sciences et Techniques de Lille, juin 1986.

**[HERT 82]** J.E. HERTEL, R.N. CLARK
"Instrument failure detection in partially observable systems", IEEE Transactions on Aerospace and Electric Systems, vol. 18, n° 3, 1982.

[HODO 79] D. HODOUIN, M. EVERELL
"A hierarchical procedure for adjustement and material balancing on mineral processes data", Research report 79-14, groupe de recherche en automatisation appliquée à l'industrie minérale, Université de Laval (Québec), août 1979.

[HODO 86] D. HODOUIN, C. BAZIN, F. FLAMENT
"Propagation of measurement errors in mass balance calculation of mineral processing data", Proceedings of APCOM 1986, Pennsylvania State University, april 1986.

[HODO 88] D. HODOUIN, M. GARON, M. REMILLARD, M. THERIEN
"Assessment of precious metals distribution in lac Mattagami flotation plant by computer mass balance calculation", Cim Bulletin, vol. 81, n° 919, pp. 62-69, 1988.

[HODO 89] D. HODOUIN, F. FLAMENT
"New developments in material balance calculations for mineral processing industry", Society of Mining Engineers annual meeting, pp. 89-147, Las Vegas, 1989.

[HOLL 89] W. HOLLY, R. COOK, C.M. CROWE
"Reconciliation of mass flow rate measurements in a chemical extraction plant", The Canadian Journal of Chemical Engineering, vol. 67, pp. 595-601, 1989.

[IORD 84] C. IORDACHE, R.S.H. MAH, A.C. TAMHANE
"Detection of gross errors in process data using measurement test", AIChE spring meeting, paper 43c, Anaheim, 1984.

[IORD 85] C. IORDACHE, R.S.H. MAH, A.C. TAMHANE
"Performance studies of the measurement test for detection of gross errors in process data", AIChE Journal, vol. 31, n° 7, pp. 1187-1200, 1985.

[JAEC 76] J.L. JAECH
"Large sample tests for Grubbs's estimators of instrument precision with more than two instruments", Technometrics, vol. 18, n° 2, pp. 127-133, 1976.

[JAEC 79] J.L. JAECH
"Extension of Grubbs's method when relative biases are not constant", Nuclear Material Management Journal, vol. 1, pp. 76-79, 1979.

[KRZA 80] G. KRZAKALA, J. RAGOT, C. HUMBERT
"Application du calcul hiérarchisé au traitement d'essais multiples d'un processus industriel à modèles linéaires connus", RAIRO Automatique, vol. 14, n° 4, pp. 317-341, 1980.

[KRZA 81] G. KRZAKALA
"Contribution à l'extension du calcul hiérarchisé. Application à l'équilibrage des bilans de mesure et à la commande des systèmes cascades", Thèse de Doctorat d'état, Université de Nancy I, 1981.

[KRZA 82] G. KRZAKALA, P. JACQUEMIN, J. RAGOT
"Méthode de décentralisation hiérarchisée pour l'équilibrage des bilans de mesure", Congrès IASTED MECO' 82, Tunis, 1-3 septembre 1982.

[LAGU 82] D. LAGUITTON, D. HODOUIN, C. BAZIN, F. FLAMENT, J. WILSON
"Utilisation par les exploitants canadiens de programmes de bilan-matière", Revue de l'Industrie Minérale, les Techniques, pp. 581, 1982.

**[LAGU 89]** D. LAGUITTON, J. LEUNG
"Insert the diskette and type MATBAL : 15 years of progress in material balance computation", Congrès IFAC AIPAC'89, vol. 2, pp. 124-128, Nancy, 3-5 juillet 1989.

**[LEBA 82]** L. LEBART, A. MORINEAU, J.P. FENELON
*Traitement des données statistiques. Méthodes et programmes*, Edition Dunod, 1982.

**[LEGR 80]** J. LEGRAS
*Algorithmes et programmes d'optimisation non linéaire avec contraintes*, Edition Masson, Paris, 1980.

**[MAH 76]** R.S.H. MAH, G.M. STANLEY, D.M DOWNING
"Reconciliation and rectification of process flow and inventory data", Industrial Engineering Chemistry Process Design Development, vol. 15, n° 1, pp. 175-183, 1976.

**[MAH 82a]** R.S.H. MAH
"Design and analysis of process performance monitoring systems", Chemical Process Control II, D.E. Seborg, editor Engineering Foundation, New York , 1982.

**[MAH 82b]** R.S.H. MAH, A.C. TAMHANE
"Detection of gross errors in process data ", AIChE Journal, vol. 28, n° 5, pp. 828-830, 1982.

**[MAQU 86]** D. MAQUIN, J. RAGOT, J. FAYOLLE, M. DAROUACH
"Observabilité des systèmes linéaires à mesures manquantes", Revue de l'Industrie Minérale, Les Techniques, pp. 53-56, 1986.

**[MAQU 87]** D. MAQUIN
"Observabilité, diagnostic et validation de données des procédés industriels", Thèse de l'Université de Nancy I, novembre 1987.

**[MAQU 88]** D. MAQUIN, J. RAGOT, M. DAROUACH, J. FAYOLLE
"Validation des mesures par équilibrage hiérarchisé de bilans-matière", International Journal of Mineral Processing, vol. 23, pp. 241-252, 1988.

**[MAQU 89]** D. MAQUIN, M. DAROUACH, J. RAGOT
"Observabilité et validation de mesures des systèmes bilinéaires", Congrès IFAC AIPAC'89, Nancy, 3-5 juillet 1989.

**[NARA 84]** S. NARASIMHAN
"Detection of steady-states and its application to process data reconciliation", Thesis, Northwestern University, Evanston, 1984.

**[NARA 86]** S. NARASIMHAN, R.S.H. MAH, A.C. TAMHANE, J.W. WOODWARD, J.C. HALE
"A composite statistical test for detecting changes of steady-states", AIChE Journal, vol. 32, n° 9, pp. 1409-1418, 1987.

**[NARA 87]** S. NARASIMHAN, R.S.H. MAH
"Generalized likelihood ratio method for gross error identification", AIChE Journal, vol. 33, n° 9, pp. 1514-1521 , 1987.

**[NEUI 83]** M. NEUILLY
Méthodes d'étalonnage par additions successives, Rapport CEA - R5220, 1983.

[PUN 72] L. PUN
*Introduction à la pratique de l'optimisation*, Edition Dunod, Paris 1972.

[RAGO 79] J. RAGOT, M. AUBRUN
"Equilibrage de bilans-matière dans l'industrie minérale", Revue de l'Industrie Minérale, pp. 93-98, 1979.

[RAGO 83] J. RAGOT, J.C. BREPSON, D. SAUTER
"Méthodologie d'étude des bilans matière", Revue de l'Industrie Minérale, les Techniques, pp. 572-580, décembre 1983.

[RAGO 84] J. RAGOT, M. DAROUACH, D. MAQUIN
"Mass-balance equilibration", Outline of contribution to the NATO Advanced Study Institue, Ankara, Turkey, août 1984.

[RAGO 89] J. RAGOT, D. MAQUIN, M. DAROUACH
"Data validation for very large scale systems in mineral industry", 6th IFAC symposium AMMMP, Buenos Aires , 4-8 septembre 1989.

[RAGO 90a] J. RAGOT, D. MAQUIN, H. CHAFOUK, A. AITOUCHE
"Data reconciliation with unknown variance", 7th IMEKO international symposium on technical diagnostics, Helsinki, 17-19 september 1990.

[RAGO 90b] J. RAGOT, M. DAROUACH, D. MAQUIN, G. BLOCH
*Validation de données et diagnostic*, Traité des Nouvelles Technologies, série Diagnostic et Maintenance, Edition Hermès, 1990.

[ROMA 81] J.A. ROMAGNOLI, G. STEPHANOPOULOS
"Rectification of process measurement data in the presence of gross errors", Chemical Engineering Science, vol. 36, n° 11, pp. 1849-1863, 1981.

[ROSE 87] J. ROSENBERG, R.S.H. MAH, C. IORDACHE
"Evaluation of schemes for detecting and identifying gross errors in process data", Industrial Engineering Chemistry Research, vol. 26, n° 3, pp. 555-564, 1987.

[SERT 86] R.W. SERTH, W. A. HEENAN
"Gross error detection and data reconciliation in steam-metering systems", AIChE Journal, vol. 32, n° 5, pp. 733-742, 1986.

[STAN 77] G.M. STANLEY, R.S.H. MAH
"Estimation of flows and temperatures in process networks", AIChE Journal, vol. 23, n° 5, pp. 642-650, 1977.

[STAN 82] G.M. STANLEY
"On line data reconciliation for process control", AIChE Annual meeting, Los Angeles, 1982.

[TAMH 82] A.C. TAMHANE
"A note of use of residuals for detecting an outlier in linear regression", Biometrika, vol. 69, n° 2, pp. 488-489, 1982.

[TAMH 85] A.C. TAMHANE, R.S.H. MAH
"Data reconciliation and gross error detection in chemical process networks", Technometrics, vol. 27, n° 4, pp. 409-422, 1985.

[WILL 74] A.S. WILLSKY, H.L. JONES
"A generalized likelihood ratio approach to state estimation in linear systems subject to abrupt changes", IEEE Conference Decision and Control, pp. 846, 1974.

[YAMA 88] K. YAMAMURA, M. NAKAJIMA, H. MATSUYAMA
"Detection of gross errors in process data using mass and energy balances", International Chemical Engineering, vol. 28, n° 1, pp. 91-98, 1988.

NOM DE L'ETUDIANT : CHAFOUK HOUCINE

NATURE DE LA THESE : DOCTORAT DE L'UNIVERSITE DE NANCY I

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le 22 NOV. 1990 / 2396

LE PRESIDENT DE L'UNIVERSITE DE NANCY I

UNIVERSITE NANCY I
LE PRESIDENT

M. BOULANGE

## Mots-clés

Validation de données, bilan matière, estimation de variance, biais de mesure, erreurs de mesure, observabilité.

## Résumé

L'objectif fondamental de cette thèse est de développer des algorithmes pour l'estimation des variances des erreurs de mesure. Les techniques proposées utilisent l'hypothèse de normalité des erreurs de mesure ce qui permet d'expliciter très simplement leur fonction de vraisemblance.

Les travaux antérieurs dans ce domaine sont peu nombreux ce qui justifie l'intérêt de ce travail. Au niveau pratique, l'approche proposée est également importante car dans la plupart des procédés industriels la méconnaissance de la variance des mesures est un fait très courant.

Dans ces différentes approches, nous avons particulièrement développé l'estimation de variances et de grandeurs réelles pour les systèmes linéaires et bilinéaires mesurés et partiellement mesurés et nous avons généralisé le calcul de ces estimations aux systèmes multi-linéaires globalement redondants. De nombreux exemples en simulation, ont permis de valider les concepts théoriques.