UNIVERSITÉ DE MONTRÉAL

OUTILS TECHNOLOGIQUES EN GESTION DE LA QUALITÉ : TROIS EXEMPLES D'APPLICATION DES STATISTIQUES INDUSTRIELLES

HUGUES SIMARD

DÉPARTEMENT DE MATHÉMATIQUES ET DE GÉNIE INDUSTRIEL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE EN INGÉNIERIE

(GÉNIE INDUSTRIEL)

DÉCEMBRE 1996

© Hugues Simard, 1996.

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-26513-7

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

OUTILS TECHNOLOGIQUES EN GESTION DE LA QUALITÉ:
TROIS EXEMPLES D'APPLICATION DES STATISTIQUES INDUSTRIELLES

présenté par: SIMARD Hugues
en vue de l'obtention du diplôme de: Maîtrise en ingénierie
a été dûment accepté par le jury d'examen constitué de:

Mme LAPIERRE Sophie, Ph.D., présidente
M. CLOUTIER Guy, Doctorat, membre et directeur de recherche
M. GOURDEAU Richard, Ph.D., membre

À mes parents, pour leur support constant durant les moments difficiles et leurs encouragements sans lesquels je n'aurais pu mener ce projet à bien.

À Marie-Claude, pour avoir fait preuve de patience et pour m'avoir ramené sur terre à maintes reprises tout au long de ma maîtrise.

REMERCIEMENTS

J'aimerais remercier mon directeur de thèse, Guy Cloutier, dont les sages conseils m'ont guidé à travers les méandres bien souvent fictifs de mon existence et de mes projets. Je tiens de plus à souligner le support qu'il m'a accordé durant ces deux dernières années.

Je voudrais aussi exprimer ma gratitude à Gordon Broderick, du Centre de Technologies Noranda, pour m'avoir accordé la possibilité de participer à ses projets de recherches. Par la même occasion, je tiens a souligner l'apport constant d'idées novatrices et la collaboration au démarrage de projets de Bernard Clément. Grâce à ses connaissances et son réseau de relations, j'ai été immergé dans le monde industriel des statistiques.

J'aimerais souligner l'apport particulier de Sophie Lapierre et de Richard Gourdeau, siégeant sur le jury, pour leurs judicieux conseils et commentaires ayant permis d'épurer cet humble document.

Finalement, je profite de cette occasion pour souligner la contribution de Jean-Daniel Cusin, président de Cybernostic Inc., à la complétion des travaux. Son support a grandement facilité les dernières phases de correction et de remise du mémoire.

RÉSUMÉ

Si les outils organisationnels de la gestion de la qualité se sont taillé une place importante dans les pratiques de gestion des opérations, les statistiques industrielles, quant à elles, s'imposent plus difficilement au Canada comme outil d'intervention dans l'analyse et l'optimisation de procédés. Dans ce contexte, il devient important de familiariser les gestionnaires de procédés à un mode d'utilisation accessible, et tout de même efficace, des statistiques. Ce texte présente donc trois utilisations distinctes des statistiques comme outil d'analyse et de conception de produits ou de procédés.

Le premier cas présente l'analyse d'un procédé métallurgique: la réduction électrique d'ilménite. L'étude sert à déterminer l'ensemble de variables à inclure dans les campagnes d'analyse de la stabilité du processus. L'étude conduit à une distinction importante entre deux types de facteurs: les paramètres d'opération, sur lesquels les opérateurs interviennent directement, et les variables de contrôle, permettant le suivi du procédé et influencées par les premières. Le banc d'essais devrait donc n'inclure que les paramètres d'opération comme variables explicatives des fluctuations de la variabilité du processus.

La seconde étude de cas présente l'application des statistiques industrielles dans une usine de pâte et papier. L'expérience vise à mesurer l'effet de la variation des proportions d'un mélange de matières fibreuses et de leur capacité de rétention d'eau sur les qualités optiques et mécaniques du produit fini. Les modèles de régression construits ont donné des résultats satisfaisants dans l'étude de l'opacité et de la résistance en cisaillement (R^2 de 0,83 et 0,80, respectivement). Il n'a pas été possible d'évaluer un modèle fiable de l'énergie absorbée en tension.(R^2 de 0,54). Il semble nécessaire de revoir les plages de variation de ses facteurs explicatifs ou de considérer l'effet de nouveaux facteurs.

La troisième étude de cas présente la méthodologie suivie lors de la conception sur simulateur d'un outil servant à caractériser l'erreur interne d'un robot. D'après les résultats obtenus, la posture, exprimée comme sa valeur de conditionnement de la matrice de propagation, a un effet déterminant sur la justesse de l'identification de la dispersion angulaire. De plus, le montage conçu localement s'est révélé plus flexible quant au choix de postures minimisant l'erreur d'identification.

La réalisation des trois études illustre l'effet important des conditions environnantes sur la qualité de l'information obtenue: il s'est avéré plus facile d'obtenir un modèle fiable dans la dernière étude de cas, réalisée sur un simulateur. L'information contenue dans les essais sur les mélanges de pâtes a quant à elle souligné l'importance que l'on doit accorder à une planification de la prise de mesures basée sur l'objectif d'analyse. Sinon, ainsi qu'observé dans la première étude de cas, l'information qu'on peut en extraire est incomplète et imprécise.

ABSTRACT

Quality management principles have widely been recognized as powerful means to help industries reach their competitive goals. Nonetheless, the use of industrial statistics, which should be an integrated part of total quality management, has not been deployed as much as the other techniques of quality improvement. Since better knowledge on the process usually allows operating cost reductions, it is important to make managers statistically aware: that is to provide them with any easy but efficient way to use statistics. The following are three case studies showing the use of industrial statistics in design and process improvement projects.

The first case study presents the analysis of an ore reduction process (ilmenite). This study is the starting point of a series of on-line experiments and serves to identify the set of parameters to include in the experiments. Results show the importance of differentiating two kinds of variables: operating parameters, which are used to modify operating conditions, and control parameters, which allow controlling the operating conditions. Once it is done, it is essential to avoid the inclusion of control variables into the set of independent variables during the experiment.

The second case study shows an application of industrial statistics in the paper/pulp industry. We used an experimental design to evaluate the effects of variation in the pulp mixture on its quality characteristics (optical and mechanical). The method has proven to provide poor results for the tensile energy model ($R^2 = 0,54$). Although all factors should be significant, the tested values did not allow strong conclusions. Maybe other variables, like coarseness and average length of fiber should be considered in further analysis. Nonetheless, the estimated models for opacity and tearness showed good results (R^2 of 0,83 and 0,80, respectively).

The third case study presents the design of a tool that will measure the internal error of a robot manipulator. The experiment is designed on a simulator. We have observed the effects of the encoders' precision, the measuring instrument's precision, the configuration of the measuring instruments and the configuration of the robot on the identification error of the angular variability. According to our results, the robot configuration has a dominant effect on the identification error. Furthermore, the proposed measuring setup allows for more flexibility, or a wider range of robot configurations that minimize the error of identification.

The results of the case studies pointed out the important contribution of the environmental conditions to the analysis: the study performed on a simulator has given much more accurate results than the on-line analysis. We also realized the importance of planning the data gathering. By not doing so, analysis results, like the ones from the first case study, are partial and rarely accurate.

TABLE DES MATIÈRES

DÉDICACE iv
REMERCIEMENTS v
RÉSUMÉ vi
ABSTRACT viii
TABLE DES MATIÈRES x
LISTE DES TABLEAUX xii
LISTE DES FIGURES xiii
LISTE DES ANNEXES xiv
INTRODUCTION 1
CHAPITRE I: MÉTHODOLOGIE SUIVIE LORS DES ANALYSES DE CAS 6
1.1. Les outils de statistiques industrielles 6
1.2. Terminologie utilisée 6
1.3. La cueillette de données 7
1.4. L'analyse des données 11
CHAPITRE II: ÉTUDE DU LIEN OBSERVÉ ENTRE LES CONDITIONS D'OPÉRATION D'UN FOUR DE FONDERIE ET LE NIVEAU D'OUVERTURE DE SES CONDUITS 18
2.1. Mise en situation 18
2.2. Description du procédé de réduction 19
2.3. Choix des variables à analyser 20
2.4. Plan de travail 23
2.5. Estimation et diagnostic 24
2.6. Conclusions et projections 25
CHAPITRE III: MODÉLISATION DU COMPORTEMENT D'UN MÉLANGE DE MATIÈRES FIBREUSES 29
3.1. Mise en situation 29

3.2. Choix des variables de réponse à analyser 30
3.3. Choix des paramètres à inclure dans le modèle 30
3.4. Modalités choisies et modèles considérés 32
3.5. Estimation et diagnostic 34
3.6. Conclusion 39
CHAPITRE IV: CALIBRATION D'UN INSTRUMENT DE MESURES SERVANT À CARACTÉRISER L'ERREUR DE RÉPÉTABILITÉ D'UN ROBOT 43
4.1. Mise en situation 43
4.2. Variable de réponse analysée 44
4.3. Choix des paramètres du modèle 48
4.4. Modalités choisies et modèles considérés 51
4.5. Estimation et diagnostic 54
4.6. Conclusion 56
CONCLUSION 61
RÉFÉRENCES BIBLIOGRAPHIQUES 66

LISTE DES TABLEAUX

Tableau 1.1: Terminologie en vigueur 6
Tableau 1.2: Matrice d'analyse de la variance 17
Tableau 2.1: Paramètres étudiés 22
Tableau 3.1: Description des variables de réponse étudiées 31
Tableau 3.2: Modalités choisies, fractions de pâtes 33
Tableau 3.3: Comparaison des trois méthodes d'estimation 35
Tableau 3.4: Estimation des modèles des caractéristiques modélisées 37
Tableau 3.5: Niveaux de modalités pour atteindre les valeurs maximales 41
Tableau 4.1: Modalité standardisées des paramètres du modèle 53
Tableau 4.2: Modèles de l'erreur d'identification estimés 56

LISTE DES FIGURES

Figure 1.1: Schéma d'une expérimentation 8
Figure 1.2: Schéma factoriel complet à trois modalités 10
Figure 1.3: Schéma fractionnaire selon Box-Benhken 11
Figure 1.4: Schéma factoriel complet à deux modalités 12
Figure 1.5: Distribution des résidus sur une échelle gaussienne 15
Figure 1.6: Résidus en fonction des valeurs estimées 16
Figure 2.1: Four à réduction 20
Figure 2.2: Niveaux d'ouverture des conduits d'évacuation 21
Figure 2.3: Résidus selon l'ouverture estimée 26
Figure 2.4: Paramètre d'opération et variable de contrôle 27
Figure 2.5: Facteurs de la deuxième phase 28
Figure 3.1: Diagramme cause à effet des variables considérées lors de l'analyse 31
Figure 3.2: Matrice Box-Benhken 33
Figure 3.3: Structure du plan d'expériences 34
Figure 3.4: Déviation des indices d'égouttage 36
Figure 3.5: Distribution des résidus selon TEA estimé 39
Figure 3.6: Distribution des résidus du modèle de TEA 40
Figure 4.1: Montage de comparateurs proposé par Langmoen 45
Figure 4.2: Montage de comparateurs proposé par Mooring 46
Figure 4.3: Montage de comparateurs proposé par Riley 47
Figure 4.4: Montage de comparateurs proposé par Acquaviva 48
Figure 4.5: Montage de comparateurs conçu à l'interne 49
Figure 4.6: Valeur de conditionnement selon la séquence d'essais 51
Figure 4.7: Structure du plan hiérarchisé 54
Figure 4.8: Distribution des résidus sur une échelle gaussienne 57
Figure 4.9: Progression de l'erreur selon la valeur estimée 58
Figure 4.10: Séquence des erreurs d'identification calculées 59
Figure 4.11: Erreur d'identification selon la valeur de conditionnements 60

LISTE DES ANNEXES

Annexe A: Liste des fonctions de statistiques programmées sous MatLab 68

Annexe B: Analyse de corrélation, étude de cas 1 70

Annexe C: Plan d'essais de l'étude du mélange de matières fibreuses 72

Annexe D: Analyse de régression, étude de cas 2 75

INTRODUCTION

«À partir de 1997, les entreprises désirant soumettre leur candidature pour remplir des mandats du gouvernement québécois devront être enregistrées à un des programmes ISO 9000.» Cette décision du gouvernement provincial, ainsi que les nombreux programmes de subvention mis sur pieds, illustrent l'importance qu'accordent les dirigeants politiques aux programmes d'amélioration continue et à la gestion de la qualité au Québec. La "vague qualité" balaye d'ailleurs tous les pays industrialisés. Devant les bonds fulgurants de productivité réalisés par les entreprises nipponnes grâce à ces outils, les entreprises occidentales ont dû emboîter le pas pour faire face à la très forte compétition.

La nouvelle approche face à la qualité met désormais les clients en relation avec le produit et le processus de production. L'avis du client prend une importance marquée: ses exigences servent maintenant de barème pour définir les caractéristiques de conception des produits. De surcroît, les programmes d'assurance de la qualité, comme ISO 9000, permettent d'assainir la gestion des entreprises. Bien qu'elle semble à prime abord une charge administrative excessive, l'implantation d'ISO 9000 permet généralement d'identifier les points critiques de l'activité de production (par conséquent, elle oriente les actions de réduction des coûts).

La pratique actuelle, notamment au Québec, est dominée par l'usage de stratégies réactives orientées vers la détection et la correction des problèmes (ou défauts) existants[1]. Les stratégies pro-actives de la qualité sont quant à elles orientées vers la prévention des problèmes en vue de parer les coûts de la non-qualité. Cette stratégie requiert non

---

[1] À ce jour, environ 5 à 10 % des entreprises localisées au Québec utilisent des stratégies d'intervention en qualité. Cette donnée fut présentée par la firme de consultants en gestion RDS lors d'un congrès organisé par l'Ordre des Ingénieurs du Québec, en mars 1994.

seulement la connaissance physique des liens de cause à effet entre les divers paramètres du processus et les caractéristiques décrivant la qualité, mais aussi une analyse de risques, la réalisation d'expériences sur le procédé de production ainsi que du jugement pour justifier les actions entreprises (Kolarik, 1995)

Dans ses écritures comptables, un administrateur présente aisément les économies réalisées suite à la correction des défauts. Les stratégies réactives se justifient facilement à l'aide de tels montants, précis et indiscutables, justifiés a posteriori. Au contraire, la justification d'une stratégie pro-active, dont les montants restent à dépenser et dont les bénéfices demeurent incertains, s'avère plus délicate. Cependant, l'usage des méthodes pro-actives peut mener à des cycles de développement rapides des produits ou procédés. Ces méthodes cherchent à éviter les coûts de la non-qualité alors que les méthodes réactives tendent à les réduire.

Les perspectives d'amélioration de la qualité sont proportionnelles à la pertinence de nos connaissances des éléments analysés, et à l'énergie et l'enthousiasme dont nous faisons preuve dans nos interventions (Kolarik, 1995). Ainsi, la réussite de l'implantation d'un programme de gestion de la qualité ne dépend pas seulement des outils, si puissants soient-ils, mais aussi de la volonté et de la foi des gestionnaires et employés.

D'autre part, le succès d'une démarche d'amélioration de la qualité repose sur l'acquisition et l'application judicieuse (économiquement) de nos connaissances sur les produits et procédés. Les statistiques industrielles, dont les plans d'expériences, servent à ces fins. Elles forment une gamme d'outils puissants, tant pour la conception de nouveaux produits (prototypes) que pour l'analyse de la production (modélisation, contrôle et optimisation de procédés et de produits).

L'expérimentation sert à identifier et à quantifier les relations de cause à effet pour fournir des informations fiables aux décideurs qui dirigent les activités de production. Un plan d'expériences est une étude systématique basée sur des principes statistiques. La solidité de la démarche permet d'appuyer les conclusions obtenues et de justifier les interventions sur le processus.

L'objectif de ce mémoire est de présenter l'application de certains outils statistiques fréquemment utilisés dans des programmes d'amélioration continue. La démarche suivie comprend la résolution de trois études de cas où les statistiques industrielles servent d'alternative aux méthodes classiques de modélisation physico-chimiques.

Définir un projet touchant à toutes les facettes d'une analyse statistique pose un problème important: les interventions en entreprise n'ont jamais été suffisamment prolongées pour couvrir un si large problème. Par ailleurs, l'immensité de la tâche que représente un projet touchant à la fois à la conception d'un produit, à la stabilisation d'un procédé, et à la définition des relations de cause à effet entre les facteurs d'influence et les caractéristiques du produit nous a contraint à ne toucher qu'à certains aspects dans chaque étude de cas. Ainsi, le premier cas présente la phase de préparation, les analyses préliminaires et les recommandations sur la planification des essais en usine visant à stabiliser un procédé. La seconde étude de cas présente l'analyse d'une série d'essais réalisés dans le cadre d'une étude de l'influence des recettes et des facteurs d'influence sur la qualité d'un produit. Finalement, la troisième étude de cas illustre l'utilisation de plans d'expériences dans la conception d'un prototype.

La première étude de cas présente l'analyse exploratoire d'un procédé métallurgique. Le projet vise à restreindre le nombre de caractéristiques du procédé à inclure dans une série d'essais réalisée ultérieurement sur un four à réduction. À

l'origine, 300 variables décrivent l'état de fonctionnement de la machine. Cette étude sert de préparation à une campagne d'essais sur le procédé visant l'amélioration des conditions et du niveau de production. L'étude préliminaire comprend l'échantillonnage et l'analyse de données déjà recueillies sur bases de données. Nous avons appliqué principalement deux outils: l'analyse de corrélation et l'étude de modèle de régression.

La seconde étude de cas présente la démarche suivie dans l'analyse de caractéristiques d'un mélange de pâtes de papier, et de l'effet de variations dans le mélange. Le double objectif consiste à fournir à l'entreprise un modèle de prévision des caractéristiques de qualité, et de l'utiliser éventuellement comme outil dans la formation d'employés. Les modèles de régression avec inclusion de composantes non-linéaires et les méthodes avancées de sélection de variables ont servi à la synthèse d'une expérience réalisée en usine, selon une matrice de 60 essais. Les mêmes outils ont permis de réaliser la troisième étude de cas, la caractérisation de l'erreur machine d'un manipulateur robotisé.

La dernière étude de cas illustre l'application de l'analyse statistique dans la conception d'un instrument de mesure servant à caractériser l'erreur interne d'un robot. L'analyse, réalisée sur un simulateur, mènera à la conception et à la fabrication de l'outil de mesures performant le mieux selon les critères établis. Ce cas présente l'application des plans hiérarchiques (hierarchical designs) ou plans factoriels mixtes. Ces travaux sont réalisés dans le cadre de travaux de recherches au département de mathématiques et de génie industriel.

Les chapitres 2, 3 et 4 présentent la démarche suivie lors de la résolution de ces trois études de cas. Le chapitre 1 expose les principes régissant la prise de données et l'estimation des modèles des chapitres suivants ainsi que la terminologie employée tout au long de l'étude. Les conclusions étant particulières à chacune des analyses, nous

présentons à la fin du texte un retour sur les résultats et sur les perspectives à envisager pour d'autres applications des statistiques industrielles.

Ce chapitre de conclusion soulèvera les principaux points de comparaison des études de cas. Cette comparaison permettra d'établir une série de mises en garde et une procédure générale d'analyse.

CHAPITRE I
MÉTHODOLOGIE SUIVIE LORS DES ANALYSES DE CAS

1.1 Les outils de statistiques industrielles

La reconstruction industrielle du Japon a permis aux méthodes statistiques développées entre le 18e et le 20e siècle de jouer un rôle essentiel dans l'étude et l'optimisation de produits ou de procédés. Les brillants succès des Japonais ont incité les entreprises européennes et américaines à considérer l'application des statistiques dans leurs activités de production. Ces outils polyvalents permettent des interventions rapides et structurées sur les processus. On peut désormais améliorer produits et procédés sans connaître toutes les lois physiques et chimiques du phénomène étudié. Ce chapitre présente les outils statistiques et les bases de la méthodologie qui ont permis la réalisation des trois études de cas présentées dans les chapitres qui suivent.

1.2 Terminologie utilisée

La revue bibliographique nous a permis de constater le manque d'uniformité dans la définition des termes couramment utilisés en statistiques industrielles. Le tableau 1.1 présente les mots et définitions que nous avons considérés les plus justes.

Tableau 1.1: Terminologie en vigueur

Termes	Définitions
bruit, erreur	partie non expliquée du modèle
bruit blanc	variation purement aléatoire
bruit coloré	partie non expliquée du modèle comportant une composante déterministe

Tableau 1.1: Terminologie en vigueur (suite)

caractéristique, réponse	variable dépendante ou objectif
environnement	ensemble des variables, contrôlées ou non, exclues du modèle estimé
matrice d'essais	combinaison des modalités prévues pour l'ensemble du plan d'expérience
modalité	niveau, valeur imposée à un paramètre du plan d'expérience
paramètre	variable explicative dans un modèle de régression

1.3 La cueillette de données

Les observations analysées à l'aide des outils statistiques sont idéalement recueillies dans un environnement contrôlé. Dans le cas contraire, la contribution du bruit à la variation (perturbations aléatoires, figure 1.1) est artificiellement amplifiée.

Par conséquent, l'analyse de données historiques pose un problème important. En général, les mesures des paramètres n'ont pas été effectuées en fonction de l'étude et ont fluctué suivant les variations de l'environnement. Dans ce cas, on peut convenir de conditions environnementales qui définissent les plages de variation des paramètres à l'intérieur desquelles on considère un fonctionnement normal. Ces plages servent de filtres lors de la cueillette d'information sur la banque de données.

En revanche, lors de prise d'une données par expérimentation, nous maintenons constantes les conditions environnementales et imposons des variations contrôlées aux sujets de l'expérience. L'objectif de la planification d'expériences est d'évaluer l'effet sur une caractéristique attribuée à des perturbations imposées par des variations contrôlées de paramètres (figure 1.1). L'analyse vise à quantifier la contribution des paramètres à la

variation de la caractéristique mesurée. Les plans d'expériences peuvent servir à l'étude et l'amélioration de procédés, de produits existants ou encore à la conception de produits (développement de prototypes).

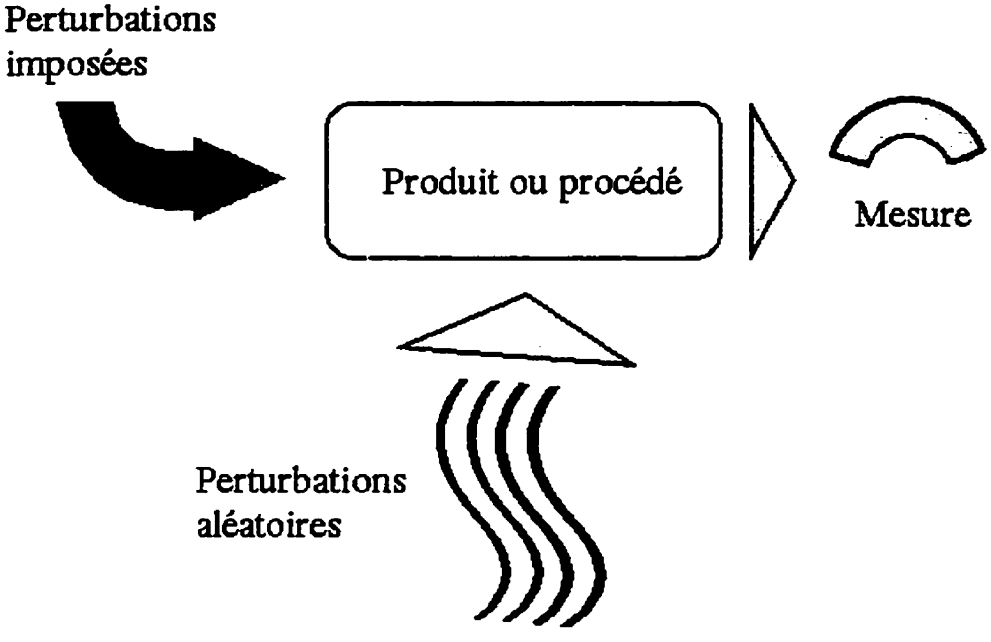


Figure 1.1: Schéma d'une expérimentation

En outre, la qualité de l'information recueillie dépend du plan choisi (Barreau, 1995). Le choix d'un plan d'expériences implique l'étude des contraintes physiques, technologiques et économiques. Il faut prendre en compte le choix de la caractéristique à optimiser et le nombre de paramètres susceptibles de l'influencer, et déterminer le nombre de modalités à exploiter. Le choix du plan dépend aussi du niveau de connaissance que nous avons du produit ou du procédé à analyser. La combinaison d'essais changera selon la réalisation d'une analyse exploratoire ou d'une optimisation. Une bonne connaissance du procédé permet aussi d'anticiper les effets croisés importants (interactions entre les facteurs) et de les inclure dans le modèle de régression. Finalement,

le coût de l'expérimentation et la capacité des intervenants à maintenir la qualité des produits limitent l'importance des bancs d'essais.

La documentation à ce sujet regorge de matrices d'expériences déjà préparées. Box, Hunter & Hunter (1978) présente les schémas factoriels complets et fractionnaires ainsi que l'analyse de surface de réponse. Les auteurs proposent de plus l'utilisation de l'«opération évolutive», une application particulière des plans factoriels. John (1971) explore la problématique particulière de l'analyse de mélanges[2] et expose la méthodologie à suivre lors de ce type d'expériences.

Dans les étude de cas des chapitres 2 à 4, on utilise deux types de matrices d'expériences: les plans factoriels et les plans de Box-Behnken (Draper, 1981). Dans un plan factoriel, un expérimentateur choisit un nombre fixe de niveaux ou de valeurs cibles pour différents paramètres. Ensuite, il réalise des mesures sur la variable de réponse en incluant les combinaisons des niveaux fixés. L'appellation «plans factoriels» sert souvent à décrire les plans à deux niveaux pour toutes les variables (2^k, pour k variables). L'utilisation de ce type de plans d'essais et de ses variantes est répandue car elle permet de recueillir une information précise en peu d'essais. Particulièrement, les plans factoriels fractionnaires se révèlent efficaces lors d'analyses exploratoires. Ils permettent d'inclure un grand nombre de paramètres en négligeant les effets de corrélation potentiels. Ce type de plans permet d'opter pour une stratégie d'analyse superficielle mais globale, plutôt que précise, sur un nombre restreint de paramètres qui peut ne pas inclure ceux qui ont une influence sur la mesure. Le nombre d'essais requis croît très rapidement lorsque nous considérons plus de deux niveaux par paramètre. Par exemple, un plan construit sur seulement quatre variables où les nombres de modalités respectives sont trois, trois, cinq et sept exigera 315 essais.

---

[2]: L'intervalle de variation des facteurs est limité: les fractions massiques ou volumiques des éléments du mélange à varier sont situées entre 0 et 100 % et la somme des fractions doit être égale à 100 %.

Les plans fractionnaires, comme les matrices de Box-Behnken ou de Box-Wilson (Draper, 1981), deviennent dès lors très intéressants.

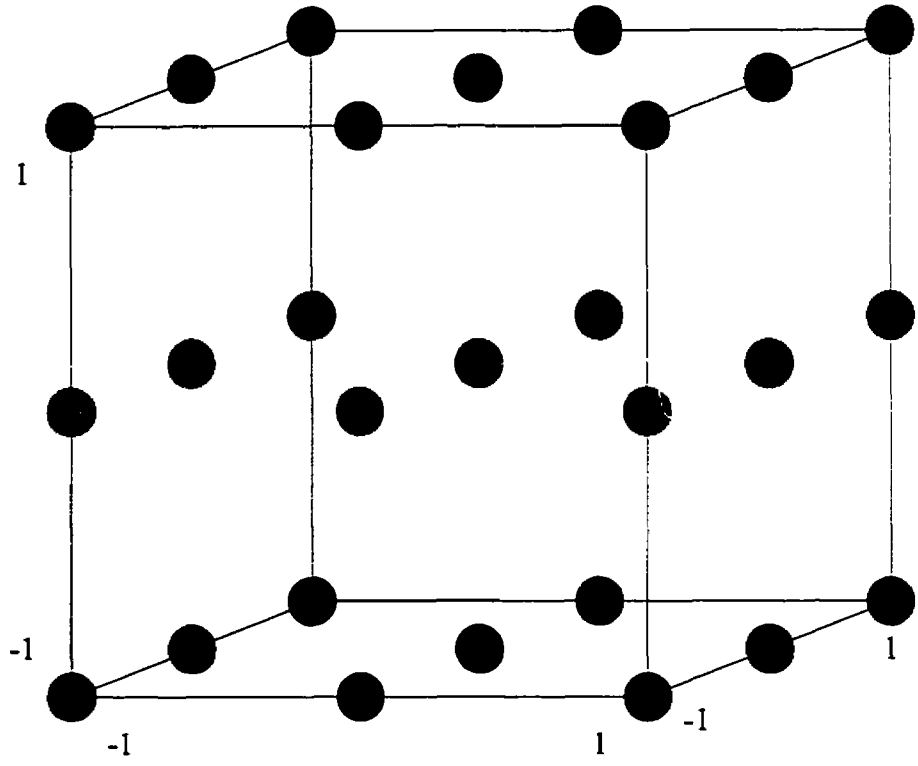


Figure 1.2: Schéma factoriel complet à trois modalités

Par exemple, un plan complet à trois paramètres et trois niveaux pour chacun des facteurs représente 27 essais. La matrice Box-Behnken à trois paramètres ne comprend que 13 essais. Si on se représente dans l'espace les structures des essais, le plan complet est celui où tous les coins et arêtes du cube ainsi que les centres de surface (figure 1.2) sont explorés. Le plan Box-Behnken ne comprend que les centres des arêtes du cube et son centre (figure 1.3). Par ailleurs, un plan factoriel 2^k n'exigerait que la réalisation de 8 essais (figure 1.4).

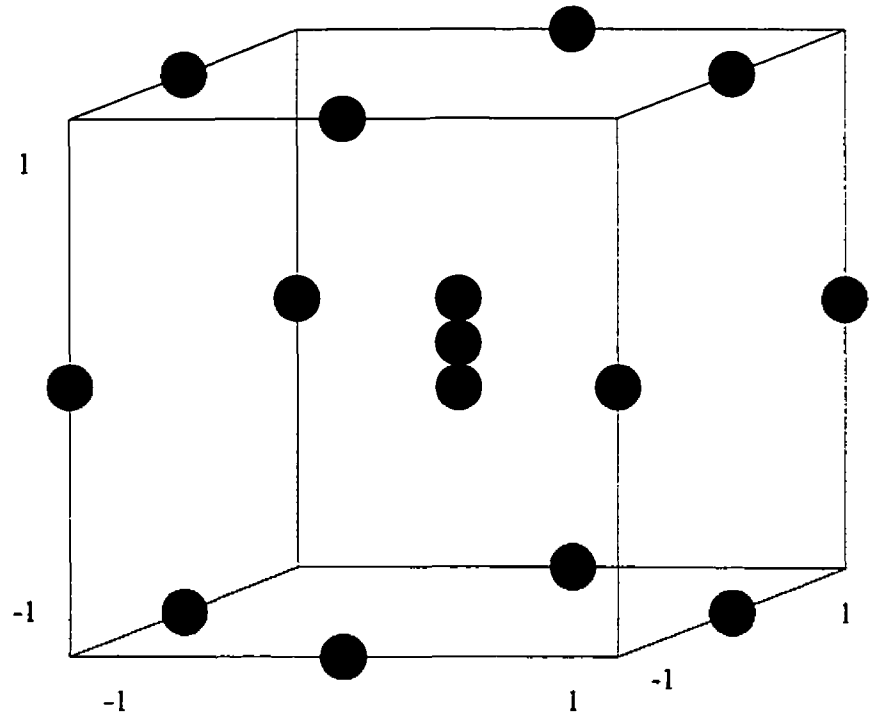


Figure 1.3: Schéma fractionnaire selon Box-Benhken

D'autres matrices ont été proposées pour permettre l'emploi de cinq modalités (Box-Wilson), de structure mixte et d'analyse de la variabilité (Taguchi), ainsi que les modèles à deux modalités non-orthogonaux (Plackett-Burman). L'application des plans de Taguchi (Lorenzen, 1993) est largement répandue en Europe et aux États-Unis grâce à sa simplicité.

1.4 L'analyse des données

Face à un problème inconnu, il est préférable de réaliser d'abord une analyse exploratoire, afin de se familiariser avec le problème et d'identifier les éléments pertinents. L'analyse de corrélation permet de discriminer les facteurs corrélés à la caractéristique étudiée de ceux ayant une incidence marginale (Draper, 1981).

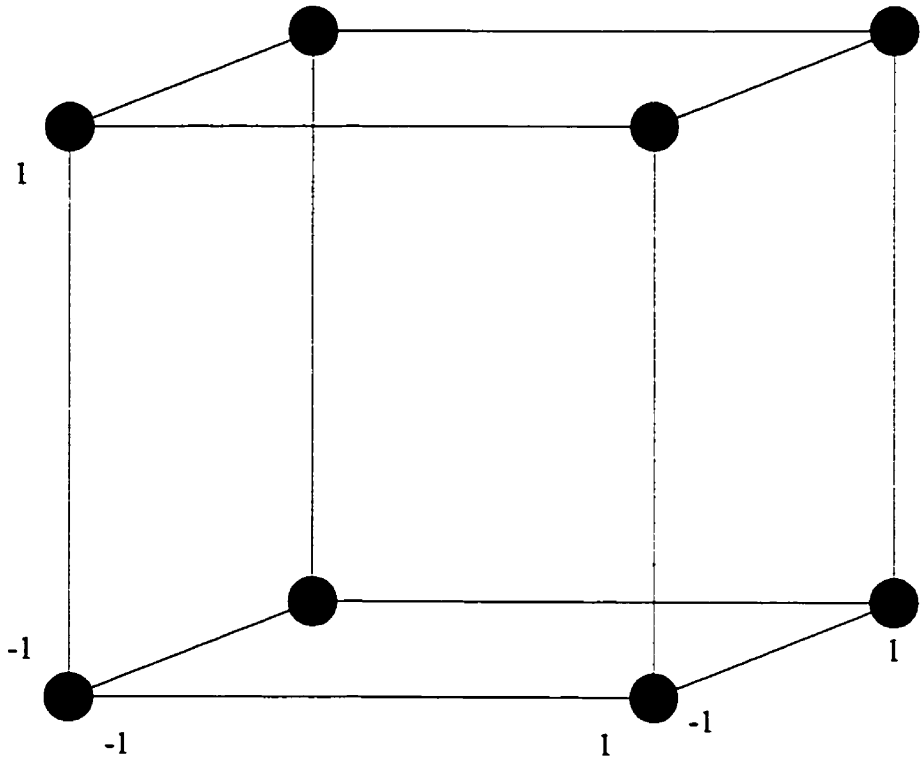


Figure 1.4: Schéma factoriel complet à deux modalités

Cette information peut servir à un premier élagage dans les facteurs explicatifs. La corrélation de deux variables, X et Y, se calcule selon l'équation 1.1, où μ_X, μ_Y représentent les moyennes des observations.

$$cor(X,Y) = \frac{\Sigma(X_i - \mu_X)(Y_i - \mu_Y)}{\sqrt{\Sigma(X_i - \mu_X)^2 \ \Sigma(Y_j - \mu_Y)^2}} \tag{1.1}$$

Cette analyse permet d'identifier les variables très corrélées pour les retirer du modèle. Cette quasi-colinéarité rend la matrice de covariance des observations singulière ou quasi-singulière. Dans ce cas, l'estimation par moindres carrés est sujette à des erreurs numériques ou tout simplement impossible.

La technique d'estimation de modèle utilisée est la régression linéaire multi-variée (Multivariate Least Square Modelling). Les modèles prennent la forme présentée à l'équation 1.2 (***X*** contient les observations des paramètres explicatifs, ***Y*** est le vecteur des observations de la variable de réponse et ***B*** le vecteur des coefficients du modèle).

$$Y = XB \tag{1.2}$$

On estime les coefficients ***B*** selon l'équation 2.3 où σ^2_B correspond à la variance (erreur) de l'estimation des coefficients estimés et *MSe* est la variance résiduelle (Mean Square error).

$$B = (X^T X)^{-1} X^T Y, \ \sigma^2_B = diag((X^T X)^{-1}) x MSe \tag{1.3}$$

Ce type de modèle peut: soit présenter un ensemble de paramètres ***X*** indépendants, soit inclure des combinaisons non-linéaires de ces derniers. Par exemple, $X = [X_1 \ X_2 \ X_1^2 \ I]$ (I est une matrice-colonne identité) serait la composition d'une matrice de paramètres servant à estimer un modèle quadratique ($y = B_1x_1 + B_2x_2 + B_{12}x_1^2 + cte$). Lorsque nous estimons les modèles associés aux plans d'expériences, on retrouve souvent des compositions de deux variables modélisant l'effet croisé des paramètres. Par exemple, $X = [X_1 \ X_2 \ X_1X_2 \ I]$ servirait à estimer un modèle de premier ordre avec l'interaction de deuxième niveau X_1X_2. Cette transformation des composants non linéaires permet de les analyser comme un système linéaire.

L'estimation des modèles par des méthodes itératives permet d'obtenir des modèles simples, modélisant au mieux la variable de réponse, en évaluant a priori la signification des éléments considérés (Draper, 1981). L'algorithme de régression pas à pas et la fonction de sélection vers l'avant intègrent un à un les facteurs selon leurs niveaux de signification estimés. L'algorithme sélectionne la variable la plus significative à partir d'un test de Fisher sur les contributions marginales, calculées à l'équation 1.4.

$$F = \frac{m(Y^T(I - X_g)x(X^T X - X^T X_g X)^{-1} X^T (I - X_g)Y(m-r-1)}{Y'(I - [X_{\text{mod}}\ X]([X_{mod}; X^T][X_{\text{mod}}\ X])^{-1}[X_{mod}; X^T]} \tag{1.4}$$

Dans cette formule, X_{mod} représente l'ensemble des observations des variables ajoutées au modèle, X les observations des paramètres pas encore inclus dans le modèle alors que X_g est un résultat intermédiaire calculé à partir des variables incluses dans le modèle (équation 1.5).

$$X_g = X_{\text{mod}}(X_{\text{mod}} X_{\text{mod}}^T)^{-1} X_{\text{mod}}^T, \quad X_g = 0 \; si \; X_{\text{mod}} = \varnothing \tag{1.5}$$

La variable Y exprime les observations de la variable dépendante. Si la contribution marginale maximale est supérieure au niveau minimal fixé, elle est ajoutée dans l'ensemble X_{mod}. La fonction de sélection vers l'avant construit d'abord l'ensemble des variables significatives puis estime un modèle sur l'ensemble résultant.

L'algorithme de régression pas à pas estime quant à lui un modèle sur l'ensemble des variables du modèle à chaque itération. Lorsqu'un des coefficients est non significatif, l'algorithme rejette le paramètre associé de l'ensemble X_{mod}. Un paramètre inclus à la première itération peut, selon cet algorithme, être exclu à la deuxième passe ou plus tard. Si par exemple, au terme de trois itérations, le modèle correspond à celui décrit à l'équation 1.6 (partie gauche), l'ajout de la variable X_4 à la quatrième étape peut rendre la variable X_1 non significative (partie droite). L'algorithme rejette la variable X_1 et ne la considère plus lors des prochaines étapes de modélisation.

$$y = B_1 X_1 + B_2 X_2 \rightarrow y = B_2 X_2 + B_4 X_4 \tag{1.6}$$

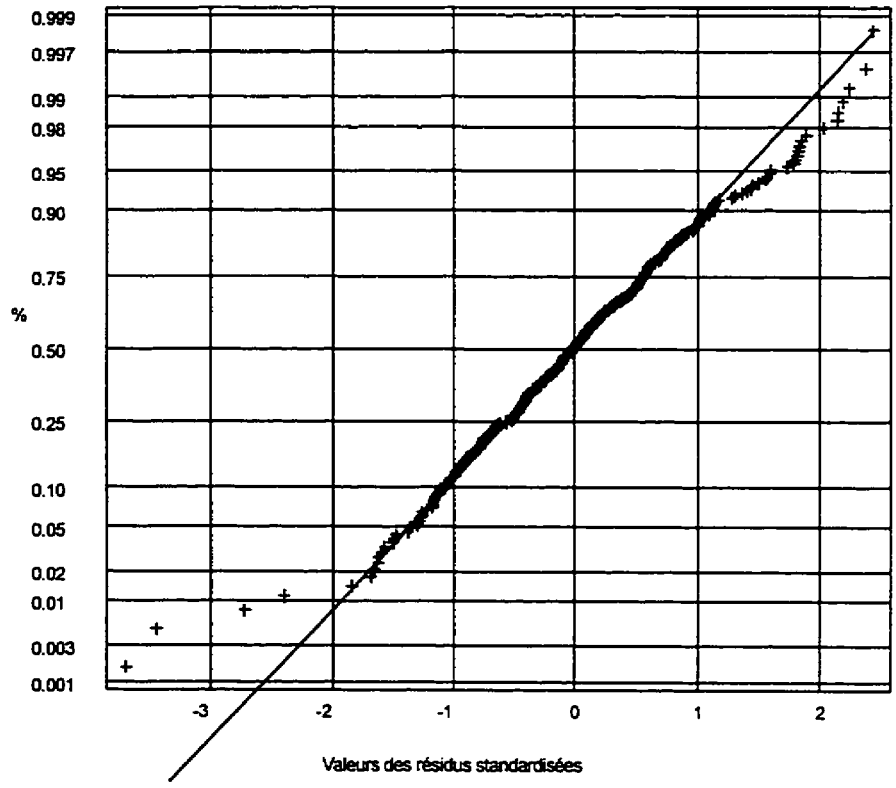


Figure 1.5: Distribution des résidus sur une échelle gaussienne

Suivent alors différents tests visant à valider les estimations obtenues. L'analyse des résidus (comportement non expliqué et distribution), l'analyse de la variance du modèle (niveau d'explication, manque d'ajustement, etc.) et le coefficient d'ajustement du modèle R^2 servent d'indices de la qualité du modèle évalué (Box, 1978).

Par ailleurs, on rejette un modèle si les résidus (la partie non expliquée par le modèle) ne correspondent pas à un bruit blanc. De plus, les modèles de régression sont estimés en supposant une distribution des résidus suivant une loi normale de moyenne 0. Ce comportement s'illustre facilement sur des graphiques à coordonnées gausso-arithmétiques (Normal Probability Plots). Dans ce système d'axe, une distribution

normale est représentée par une ligne droite (figure 1.5). Il existe par ailleurs des indices numériques (l'aplatissement et la symétrie) qui permettent de caractériser la ressemblance entre la distribution des résidus et celle d'une loi normale (John, 1971). Cependant, nous jugeons satisfaisante pour les fins de l'étude l'appréciation visuelle. D'autre part, l'observation d'un comportement non-aléatoire dans un graphique présentant les résidus comme une fonction des valeurs estimées (figure 1.6) indiquerait qu'un effort de modélisation supplémentaire est nécessaire.

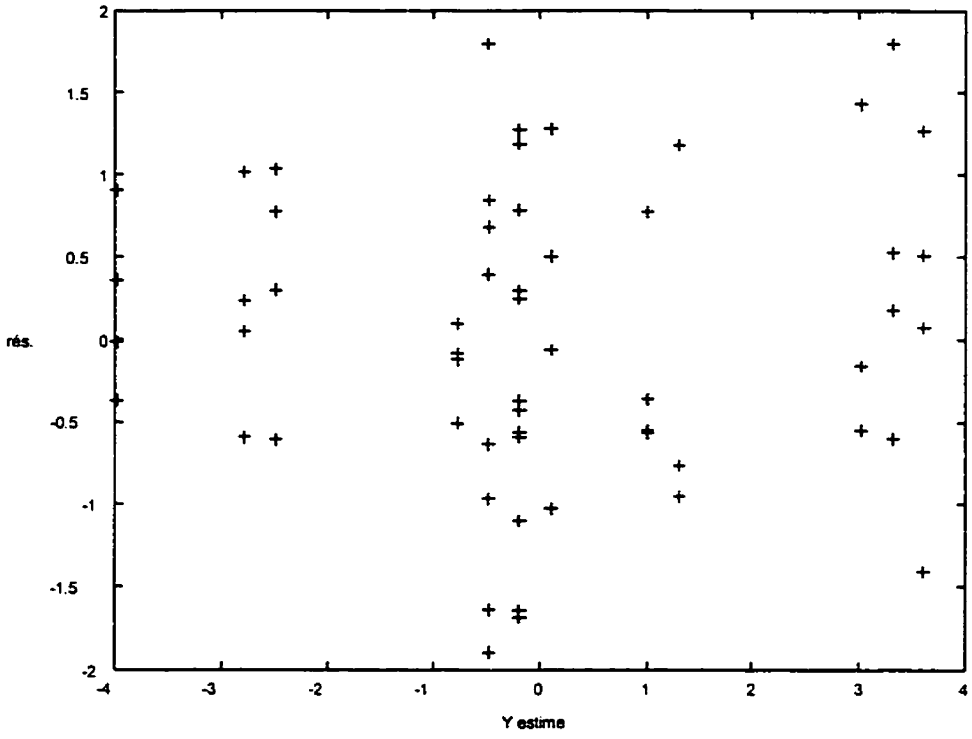


Figure 1.6: Résidus en fonction des valeurs estimées

Finalement, l'analyse de la variance permet de comparer la variance résiduelle à la variation captée par le modèle. Une table d'analyse de variance présente habituellement le

nombre de degrés de liberté associé à chaque élément et le carré moyen des erreurs (variance).

On compare les variances à l'aide du test de Fisher. Lorsque plusieurs mesures de la variable dépendante pour une même combinaison des paramètres indépendants sont disponibles, nous pouvons évaluer l'erreur d'ajustement du modèle. L'erreur de mesure s'évalue à partir de ces répétitions selon l'équation 1.7, où d_j est le nombre de répétition pour le j-ième ensemble de répétition et p est le nombre d'ensembles.

$$MS_e = \sum_{j=1}^{p} \sum_{i=1}^{d_j} \frac{(x_{ij} - \bar{x}_j)^2}{d_j - 1} \tag{1.7}$$

L'erreur due au manque d'ajustement correspond à la différence entre l'erreur résiduelle du modèle et l'erreur de mesure. Le tableau 1.2 présente un exemple de table d'analyse de la variance. On évalue le niveau de signification des constantes de Fisher (F) calculées à l'aide de la fonction cumulative de la distribution de Fisher. Les analyses des chapitres suivants ne présentent en général que la dernière colonne (les valeurs de F calculées, partie ombrée du tableau).

Tableau 1.2: Matrice d'analyse de la variance

	SCe	*MSe*	deg. liberté	Test *F*
Modèle	80 000	10 000	8	20
Erreur d'ajustement	32 000	941	34	1.88
Erreur de mesure	3 000	500	6	–
Total	115 000	–	50	–

CHAPITRE II
ÉTUDE DU LIEN OBSERVÉ ENTRE LES PARAMÈTRES D'OPÉRATION D'UN FOUR DE FONDERIE ET LE NIVEAU D'OUVERTURE DE SES CONDUITS D'ÉVACUATIONS

2.1 Mise en situation

Les industries métallurgiques et minières représentent une part importante de l'activité manufacturière québécoise. Le volume de production de ces entreprises et leur position sur le marché international exigent un suivi et une amélioration continuelle du procédé et de la qualité des produits offerts. L'entreprise où se déroule l'étude comporte à ces fins un département de recherche et développement. Ce département utilise des analyses statistiques pour le suivi du procédé, afin d'en prévoir les irrégularités, et pour les analyses chimiques.

L'entreprise exploite des fours de réduction qui extraient de la fonte et de l'oxyde de titane d'un minerai provenant du Québec et de l'Inde. Le minerai est d'abord nettoyé et enrichi, puis envoyé aux fours de réduction. Les produits extraits par coulée sont finalement concassés - dans le cas du titane- ou transformés et moulés - en ce qui concerne la fonte. Les interventions sur le procédé demeurent très rares et généralement limitées aux opérations précédant ou suivant les opérations de réduction. Les seules analyses statistiques réalisées à l'usine de réduction ont été jusqu'à maintenant limitées à l'étude de données historiques.

Ce projet sert de point de départ à une large étude menée sur les fours de réduction. L'ensemble des travaux comprend une série d'essais visant à identifier les mécanismes permettant de contrôler les conditions de fonctionnement. Ce chapitre présente la

méthodologie et les principaux résultats de l'étape préliminaire. Cette analyse exploratoire est réalisée sur un ensemble de données recueillies sur trois ans. Elle vise à limiter les étapes subséquentes aux éléments ayant statistiquement le plus d'influence sur le fonctionnement des fours.

Aux sections 2.2 et 2.3, nous décrivons le procédé et les variables sur lesquelles nous nous penchons lors de l'étude. Les sections 2.4 et 2.5 présentent les outils statistiques utilisés et les résultats de la modélisation réalisée. La section 2.6 revient sur les principaux points et expose les lignes directrices de la phase d'analyse suivante.

2.2 Description du procédé de réduction

La réduction électrique permet de séparer le minerai en deux produits, la fonte (C1 sur la figure 2.1) et l'oxyde de titane (C2). Le minerai enrichi est mélangé à du charbon et envoyé dans le plasma créé entre les trois paires d'électrodes. Le carbone (charbon) capte l'oxygène de l'oxyde de fer, à potentiel de réduction supérieur à celui du titane, puis sort par les conduits d'évacuation sous forme de monoxyde de carbone (A sur la figure 2.1). Le composé C1 devient alors plus dense que C2 et forme une couche sous l'oxyde de titane. On pratique des trous de coulée dans la paroi du four pour extraire les produits. L'opérateur peut modifier le chargement et le niveau de fonctionnement du four (voltage, puissance, bilan d'énergie et de masse).

En fonctionnement normal, il se forme sur les parois une couche plus ou moins épaisse d'oxyde de titane solide (partie hachurée). L'épaisseur de cette couche dépend des conditions d'opération. Une couche légère isole et protège le réfractaire du mur. Trop épaisse, elle obstrue les conduits d'évacuation et ralentit la transformation du minerai. Le monoxyde de carbone émis est récupéré pour servir de carburant, nécessaire au traitement (enrichissement) du minerai. L'obstruction freine la cadence, favorise les débordements et

parfois même l'explosion du four. Ces anomalies occasionnent des pertes de plusieurs millions de dollars par année. L'objectif consiste à identifier, dans un premier temps, les paramètres des fours qui ont un lien significatif avec l'obstruction des conduits d'évacuation et, dans un second temps, de proposer une stratégie d'essais sur les fours.

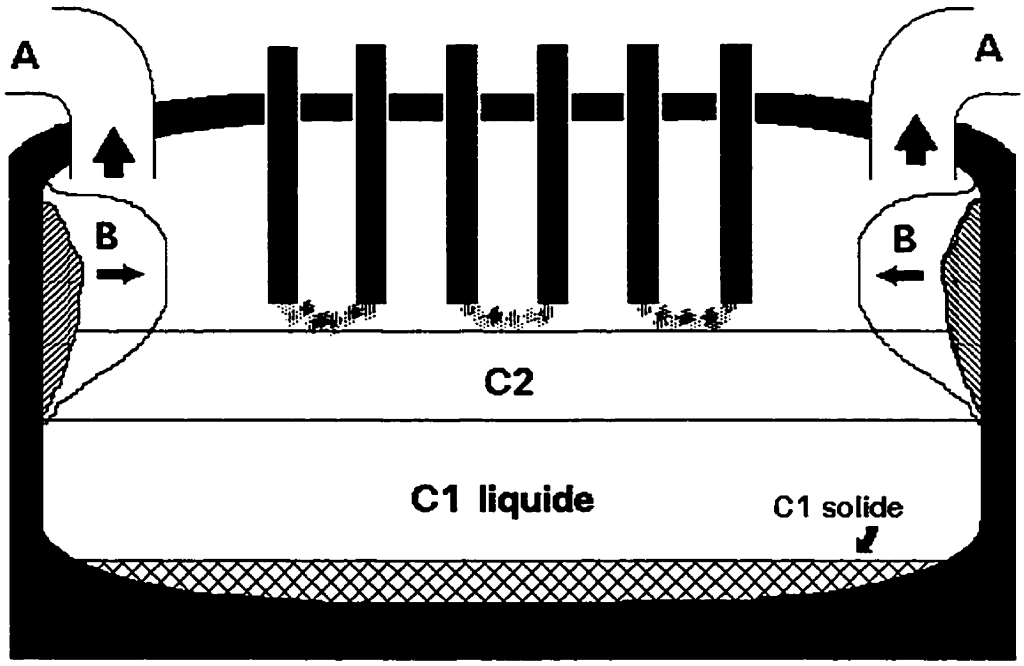


Figure 2.1: Four à réduction

2.3 Choix des variables à analyser

Nous analysons l'ouverture des conduits d'évacuation, caractéristique exprimée en pourcentage de l'ouverture maximale. Cependant, la mesure de l'ouverture reste approximative: l'opérateur effectue une évaluation visuelle deux fois par jour (voir figure 2.2; a) représente un blocage à 50 %, b à 70 %, c à 100 % et d à 0 %). Les valeurs conservées sont le résultat de la moyenne arithmétique des deux évaluations. Au moment de l'étude, aucune mesure plus fiable n'existait. Le type de mesure de la variable de réponse

fixe ainsi la période de mesure: les données sont des moyennes quotidiennes des caractéristiques et paramètres considérés.

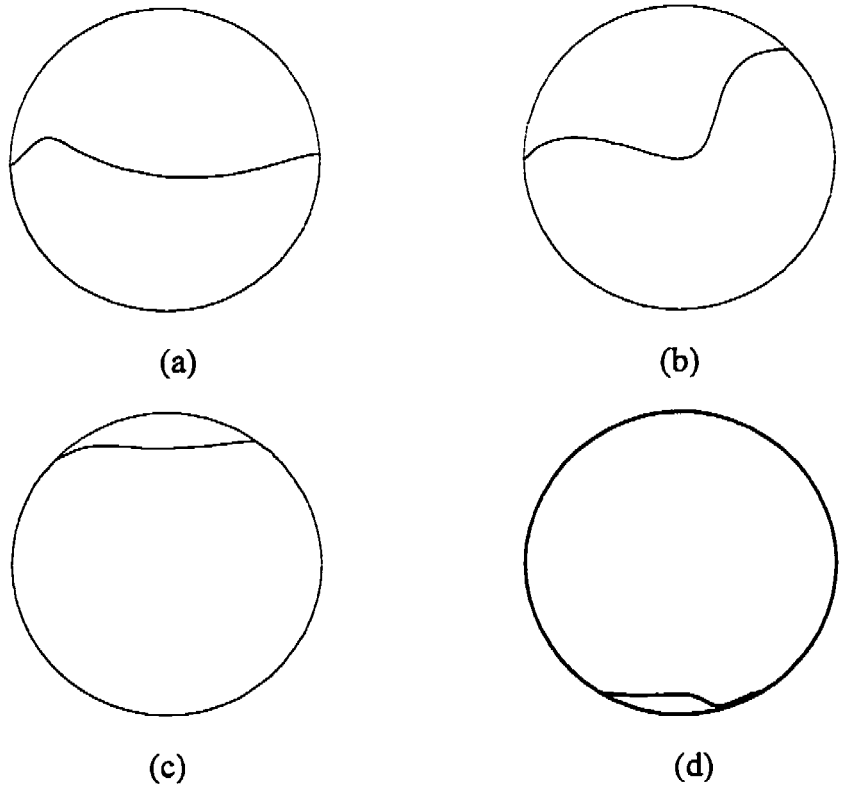


Figure 2.2: Niveaux d'ouverture des conduits d'évacuation

Puisque le nombre de paramètres apparaissant dans la base de données est très élevé (plus de 300), l'analyse statistique a été précédée d'une séance de concertation générale (brainstorming) à laquelle des opérateurs, techniciens, ingénieurs et chercheurs de l'entreprise ont participé. Le résultat de cette séance sert de point de départ à l'analyse. Après consultation, nous avons limité l'étude aux paramètres inclus dans le tableau 2.1. Toutes les variables sont centrées et réduites selon leurs moyennes et écart types respectifs pour éviter le biais sur l'évaluation de l'influence des variables et pour respecter la confidentialité des informations. Nous présentons tout de même la nature des variables au tableau, afin de faciliter la compréhension. L'analyse ne couvre que l'effet des paramètres d'opération des fours. Elle exclue l'information concernant le traitement du minerai.

Tableau 2.1: Paramètres étudiés

Variables	Unités
Puissance d'opération	MW
C1 non réduit restant dans C2	%
Pureté de C2	%
Utilisation du four (100 % = 24 h)	%
Réajustement du mélange (minerai ajouté)	%
Rapport charbon/minerai	%
Excès d'énergie	KWh/lb
Épaisseur de C2 solide sur les parois	pouces
Hauteur totale de la phase liquide	pouces
Température de C1 à la sortie du four	F
Température de C2 à la sortie du four	F
Rendement massique	%

On calcule l'excès d'énergie à partir de l'énergie théorique requise pour réduire la quantité de minerai introduite dans le four. Le réajustement du mélange permet de contrebalancer les déséquilibres thermiques ou chimiques dans le four. Le rendement massique résulte du rapport entre les matières brutes introduites dans le four (charbon et minerai) et les produits extraits du four.

2.4 Plan de travail

L'analyse est réalisée sur les mesures compilées par le service informatique. La compilation des données par l'entreprise inclut toutes les mesures réalisées sur le procédé, peu importe le niveau de production. Pour compenser l'absence de contrôle sur les paramètres, nous ne considérons dans les calculs que les données correspondant à des lectures réalisées en fonctionnement normal, soit lorsque :

- le four est utilisé à au moins 80 % de sa disponibilité maximale (24 heures);
- l'épaisseur de la couche solide sur l'oxyde de titane est de moins de 15";
- le niveau d'énergie fourni correspond au niveau nécessaire, considérant la quantité de minerai chargée dans le four (l'excès d'énergie, est inférieur, en valeur absolue, à 0,02 KWh/lb de minerai);
- le four est arrêté pour moussage moins de 4 % de la journée. Le moussage est l'épaississement de la scorie survenant lorsque le monoxyde de carbone ne réussit pas à s'en échapper (le titane est trop visqueux);
- la température de C1 à la sortie est située entre 2 400° F et 2 900° F.

Certaines des variables servant à décrire le fonctionnement stable figurent aussi dans la liste des paramètres à analyser. Les plages permises pour ces paramètres définissent le domaine de validité des conditions expérimentales

Nous réaliserons tout d'abord une analyse de la corrélation entre les paramètres d'entrée. Les résultats de cette analyse nous permettrons ensuite de choisir adéquatement le modèle de régression à tester. Il sera finalement validé par une analyse des résidus et par l'examen de sa justification physico-chimique.

2.5 Estimation et diagnostic

La matrice des corrélations estimées figure à l'annexe B. En observant les résultats, on constate que les températures de C1 et C2 sont très corrélées avec la puissance d'opération et le pourcentage d'utilisation. Ces corrélations élevées (entre 0,80 et 0,92) indiquent qu'une baisse de la température des mélanges se synchronise à une baisse de la puissance ou du taux d'utilisation. Ce résultat statistique s'accorde à la compréhension intuitive du système: à basse puissance ou arrêté, la température interne du four baisse. La baisse des températures à la sortie du four (celles mesurées) coïncide, par conservation d'énergie, à celle de la température interne.

Finalement, nous observons un lien très fort entre la puissance d'opération et le taux d'utilisation. En interrogeant le personnel du département d'informatique, nous avons appris que la valeur de la puissance d'opération figurant dans la base de données résulte du calcul d'une moyenne quotidienne. Ainsi, la moyenne obtenue s'abaissera durant une période d'arrêt prolongé (la puissance est nulle pendant les périodes d'arrêt).

Par conséquent, ces résultats nous suggèrent la suppression des températures comme variables explicatives. L'effet sur l'ouverture des conduits qu'auraient ces facteurs s'explique entièrement par celui de la puissance et de l'utilisation. À prime abord, nous aurions tendance à négliger aussi le taux d'utilisation. Cependant, on considère un four qui fonctionne à puissance réduite comme étant utilisé à 100 %. Nous tiendrons donc compte du haut niveau de corrélation entre ces variables lors de l'estimation du modèle, en incluant un paramètre d'interaction. Bien que certains paramètres ne semblent pas avoir de lien significatif avec l'ouverture des conduits d'évacuation, nous conservons tout de même dans l'ensemble de facteurs les éléments restants.

Ces variables servent à l'élaboration d'un modèle qui sera estimé selon la méthode présentée au chapitre 1. Nous estimons un modèle de premier ordre avec interaction (sans composante quadratique). Nous considérons les variables du tableau 2.1 en excluant les températures des liquides C1 et C2. Le modèle résultant est présenté à l'équation 2.1 (l'astérisque en exposant rappelle que les variables sont normalisées).

$$tour = 68,18 - 2,69(Mpur^{*}) + 14,44(Puis^{*}) \qquad (2.1)$$

Le coefficient de ce modèle est très faible (0,28). Il est impossible d'évaluer un coefficient de manque d'ajustement parce que nous ne possédons aucune série d'essais répétés. L'analyse graphique des résidus montre cependant une tendance bien identifiée qui est non expliquée par le modèle. Le graphique présenté à la figure 2.3 illustre une tendance linéaire entre les résidus et les valeurs estimées.

2.6. Conclusions et projections

L'analyse des données conservées dans la base de l'entreprise par un modèle de régression n'a pas conduit à l'élaboration d'un modèle de prédiction fiable. Le coefficient de régression du modèle présentant le niveau d'ouverture des tours en fonction du réajustement du mélange et de la puissance d'opération est faible (0,28). De plus, l'analyse des résidus expose une tendance marquée que le modèle n'arrive pas à expliquer.

Tel que prévu, l'absence de contrôle sur l'environnement a amplifié la perturbation aléatoire. L'imprécision des données passées pose aussi un problème important: les entrées «oubliées» par le personnel sont estimées par le service informatique. De plus, une grande partie de l'information recueillie est diluée dans le calcul des moyennes quotidiennes (or, toutes les données étudiées sont des moyennes journalières). En conclusion, il faudra lors de la prochaine étape utiliser des mesures et une méthode d'échantillonnage permettant d'éviter

le traitement réalisé par le système informatique. La variable de réponse devrait être une mesure que l'on peut prendre dans les minutes suivant les modifications imposées au four.

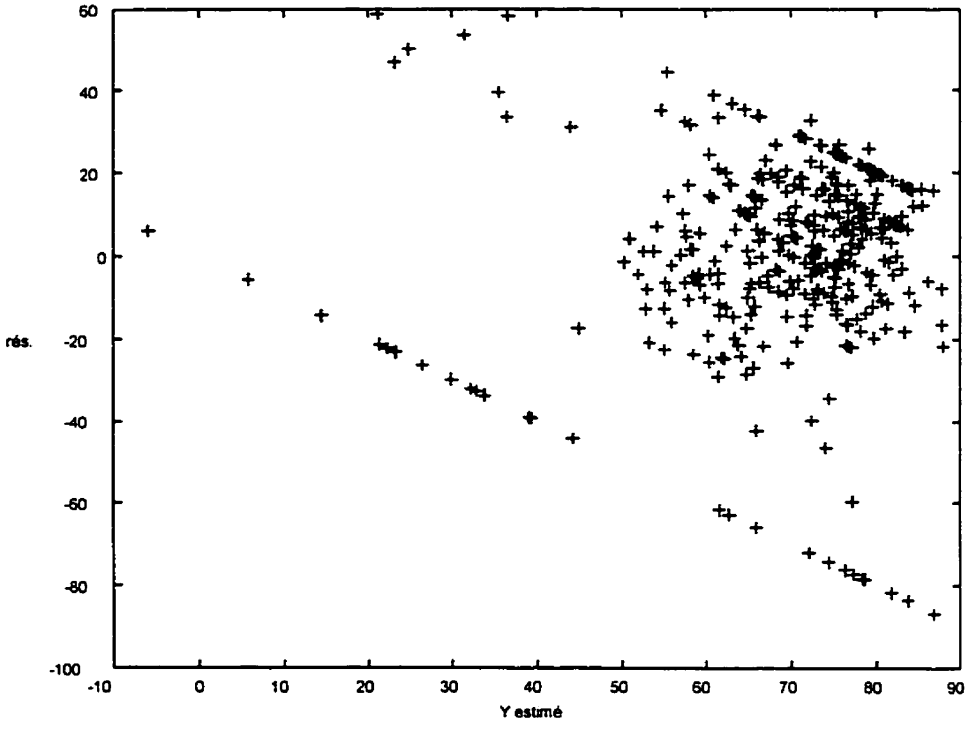


Figure 2.3: Résidus selon l'ouverture estimée

Nous remarquons de plus que les analystes de l'entreprise ne distinguent pas les éléments servant au contrôle du procédé de ceux par lesquels ils interviennent. Ils incluent dans les facteurs causaux des variables qui, pour les ingénieurs de production, servent de variable de réponse. L'étude aurait pu par exemple inclure les températures des phases liquides. Cependant, l'opérateur du four n'a pas de contrôle sur cette donnée; c'est en fait par des modifications du niveau de puissance, du bilan de masse et d'énergie qu'il contrôle la température interne.

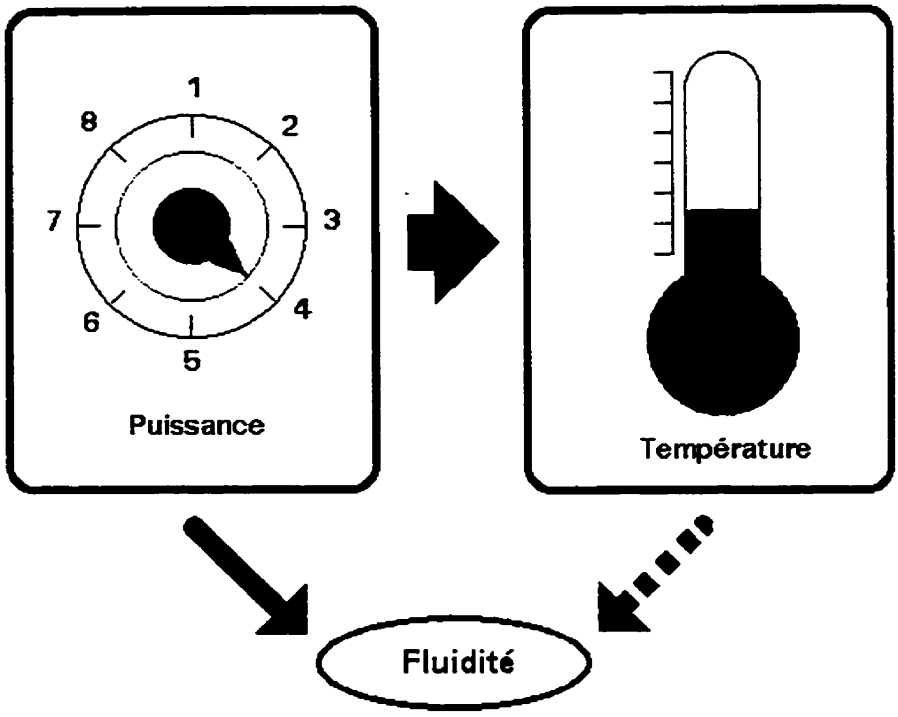


Figure 2.4: Paramètre d'opération et variable de contrôle

Il convient donc, lors de la prochaine étape, de distinguer deux classes de variables: les paramètres d'opération et les variables de contrôle. La figure 2.4 présente la distinction entre ces deux types d'information. Cet exemple illustre que la viscosité d'un liquide est fonction de sa température. Cependant, un opérateur ne modifie pas la température; il fixe le niveau de puissance en fonction des températures observée et souhaitée.

Les éléments sur lesquels l'opérateur intervient directement pour modifier les conditions de fonctionnement sont appelés les paramètres d'opération. Les variables de contrôle servent quant à elles de sondes transmettant des informations sur les conditions de fonctionnement. Nous agissons sur le procédé par l'entremise de variables comme la quantité de minerai, de charbon, le niveau de puissance et le voltage alors que nous observons le résultat de nos actions sur la température, l'épaisseur des accumulations solides sur les parois, la pureté de C2, etc.

Le plan d'essais réalisé sur le four (prochaine étape du projet) devrait donc inclure tous les paramètres d'opération pour déterminer ceux qui servent à contrôler le niveau d'ouverture des conduits d'évacuation. La figure 2.5 présente une ébauche des paramètres à considérer. Étant donné le nombre imposant de variables à traiter, il serait préférable de limiter l'analyse à un plan factoriel fractionnaire à deux niveaux (Box, 1978).

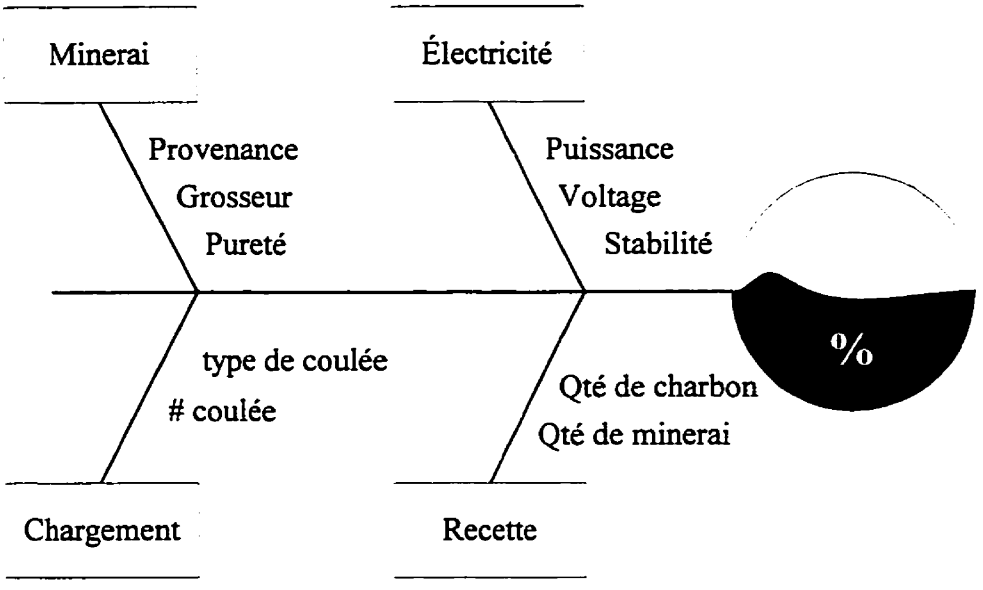


Figure 2.5: Facteurs de la deuxième phase

CHAPITRE III
MODÉLISATION DU COMPORTEMENT D'UN MÉLANGE DE MATIÈRES FIBREUSES

3.1 Mise en situation

Dans l'industrie papetière, les outils classiques d'analyse statistique dans le domaine de la qualité se butent à la quantité et la variété de l'information. On décrit la qualité d'une pâte par une quinzaine de caractéristiques généralement corrélées entre elles. Cependant, ces outils ont été développés pour l'étude de processus uni-variés ou d'ensembles de variables indépendantes. Par contre, les méthodes multi-variées comme la planification et l'analyse d'expériences s'adaptent bien à l'analyse de processus complexes. Ce chapitre présente une analyse réalisée sur un mélange de pâtes produit par une usine située en Abitibi. L'étude vise à quantifier la relation entre les caractéristiques mesurées sur le papier produit et la variation de la proportion, et de la qualité des composants du mélange.

Les modèles obtenus serviront à l'entreprise comme outils de planification des mélanges à partir des exigences des clients. Les modèles développés pourront également servir à la formation des opérateurs de l'usine et à leur montrer l'effet d'une variation des paramètres d'opération sur le papier.

Dans un premier temps, nous définissons les variables de réponses étudiées ainsi que les paramètres que nous inclurons dans le modèle. Par la suite, nous exposerons le type de plan d'expérience choisi et les résultats de la modélisation. Le choix du plan et les essais ayant déjà été réalisés lors du début du projet, nous supposerons que les éléments n'ayant pas été mesurés étaient contrôlés lors des essais.

3.2 Choix des variables réponses à analyser

L'expérience permet de mesurer seize caractéristiques du mélange de pâtes produit. Ces mesures caractérisent l'aspect de la feuille produite ainsi que ses propriétés mécaniques. Nous limitons l'analyse à trois caractéristiques: l'énergie absorbée en tension, la déchirure et l'opacité du papier. La définition de ces variables est présentée au tableau 3.1.

3.3 Choix des paramètres à inclure dans le modèle

Le mélange se compose de quatre types de pâte: mécanique de meule, de fibres recyclées, à très haut rendement et kraft blanchie. À cause de contraintes expérimentales (exposées à la section 3.4), l'analyse est limitée aux effets que pourraient avoir les trois dernières. Nous étudions aussi l'effet que pourrait avoir l'indice d'égouttage (ou de la capacité de rétention d'eau) des pâtes de fibres recyclée et de celles à très haut rendement. Nous mesurons de plus la variation dans les longueurs de fibres et autres caractéristiques des pâtes. Le graphique de la figure 3.1 illustre les liens de cause à effet que nous considérons lors de la planification et de la réalisation des essais. Dans les catégories décrivant les caractéristiques, les deux premières lettres indiquent le nom de la catégorie (AV pour globale, HR pour pâte à très haut rendement et FR pour la fibre recyclée). Les lettres suivantes indiquent le nom de la caractéristique mesurée; LM? pour les longueurs moyennes de fibres (? remplace le caractère indiquant le type de pondération utilisé lors du calcul), COR décrit la grosseur de la fibre, FLX, la flexibilité de la fibre et SPS, sa surface spécifique. Nous retrouvons l'indice d'égouttage dans les catégories «caract. HR» et «caract. FR» (HRCSF, FRCSF).

Tableau 3.1: Description des variables de réponse étudiées

Variables*	Définition
Énergie absorbée en tension (TEA)	Aire sous la courbe de la fonction Tension/allongement, exprimée en gramme par centimètres (g/cm). Calculée à partir de l'allongement à la rupture et de la résistance en tension.
Déchirure (DE)	Résistance en cisaillement de la feuille ($mN \times m^2/g$).
Opacité (OP)	Portion de lumière bloquée par la feuille (%). Calculée à partir des valeurs mesurées de diffusion et d'absorption.

*: Par besoin de protection des renseignements confidentiels, les valeurs numériques réelles seront remplacées par les valeurs standardisées.

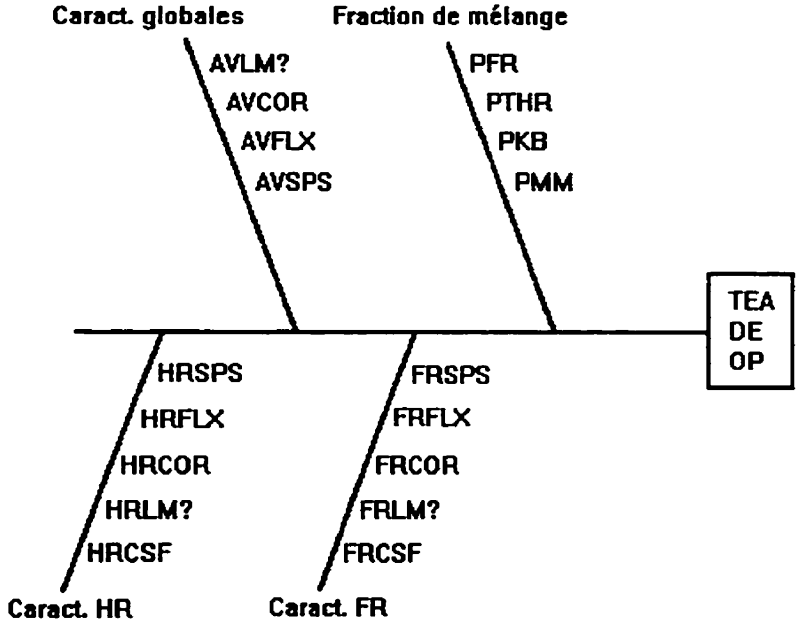


Figure 3.1: Diagramme cause à effet des variables considérées lors de l'analyse

3.4 Modalités choisies et modèles considérés

Nous désirons observer l'effet de la variation des proportions de pâtes (fractions de mélange). Selon les commentaires émis sur les effets de ces variables, il est légitime de les croire non-linéaires. Il faut au moins trois niveaux différents dans les variations imposées aux proportions de pâtes pour le vérifier. De plus, nous cherchons à savoir si une variation du degré d'égouttage aurait aussi un effet significatif sur les caractéristiques étudiées. Nous considérons deux niveaux pour le vérifier. Ces essais étant coûteux, nous utilisons une matrice d'expérience qui minimise le nombre d'essais tout en permettant d'étudier les composantes non-linéaires des fractions de mélange.

Ce type d'analyse est habituellement classé dans la catégorie de plans d'expériences pour mélanges. Ces plans sont construits en respectant la contrainte sur le domaine de la fonction. On ne peut avoir plus de 100 % d'un des éléments du mélange et le total des proportions doit être égal à 100 %. Nous contournons ce problème en n'incluant dans le modèle que les proportions de pâtes de fibres recyclées, kraft blanchie et à très haut rendement. La proportion de pâte mécanique de meule a la fonction d'excipient et sert à compléter le mélange (amener le total des proportions à 100 %). Les résultats de la modélisation ne demeureront valides que pour un sous-espace contraint, mais ce détour nous offre plus de flexibilité dans le choix de plans.

Les travaux antérieurs ont permis aux chercheurs d'identifier des relations non-linéaires entre les fractions de mélange et les caractéristiques de la feuille. Cette conclusion pousse à considérer plus de deux modalités pour ces paramètres. Nous avons donc opté pour une structure d'essai de type Box-Behnken qui comprend trois modalités par variable. Les modalités choisies sont présentées au tableau 3.2.

Tableau 3.2: Modalités choisies, fractions de pâtes

Variables	Modalité 1 (poids: -1)	Modalité 2 (poids: 0)	Modalité 3 (poids: 1)
PFR	0 %	10 %	20 %
PTHR	30 %	40 %	50 %
PKB	0 %	5 %	10 %

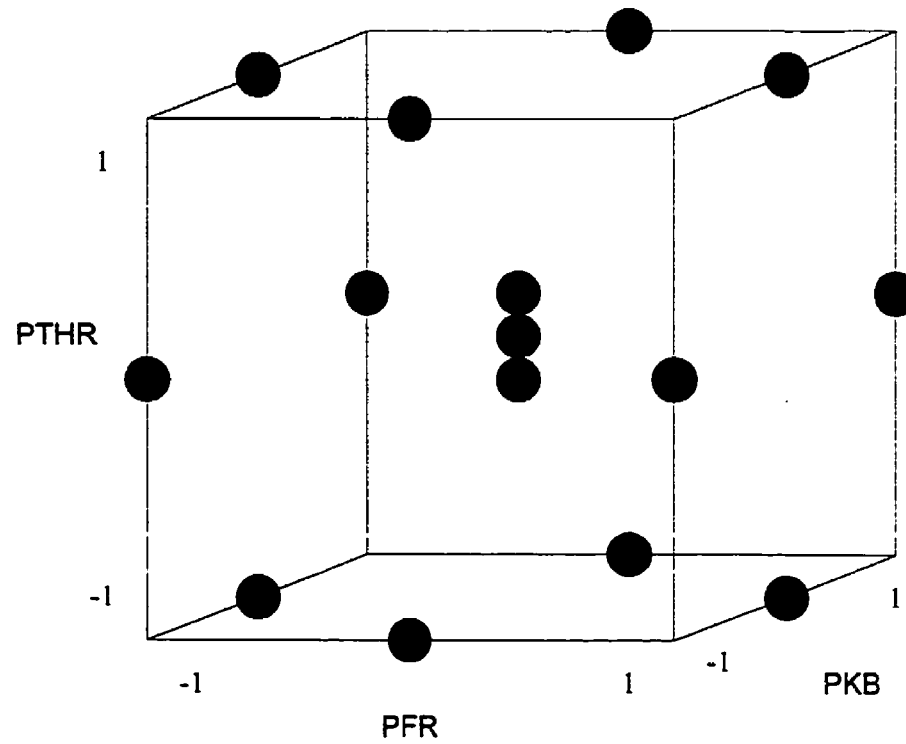


Figure 3.2: Matrice Box-Behnken

La figure 3.2 représente graphiquement la structure d'essais (la matrice résultante est présentée à l'annexe C). Nous remarquons qu'au moins une des variables se situe à la modalité médiane lors d'un essai. Les trois points superposés indiquent à quelles modalités nous réalisons les trois réplications servant au calcul de l'erreur d'ajustement.

Dans le plan proposé, nous combinons cette structure est associée à une matrice de deux variables (factoriel complet, quatre essais) pour inclure les indices d'égouttage. La figure 3.3. illustre la structure globale du plan. La matrice complète comprend 60 essais dont quatre séries de trois réplications. La prise en considération de toutes les combinaisons possibles nécessite 108 essais, sans compter les réplications.

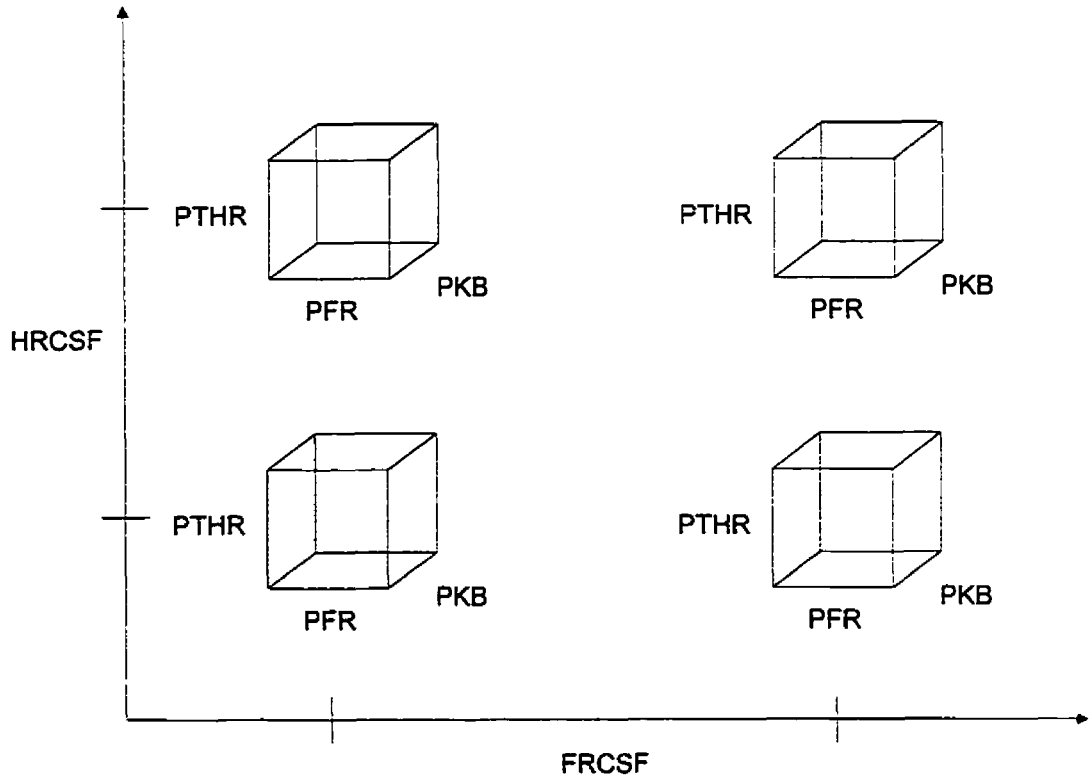


Figure 3.3: Structure du plan d'expériences

3.5 Estimation et diagnostic

On utilise trois méthodes d'estimation de modèles. Tout d'abord, nous évaluons un modèle selon la méthode de régression linéaire et calculons le niveau de signification des coefficients obtenus. On rejette les paramètres du modèle lorsque leur coefficient est jugé non-significatif. Par la suite, nous appliquons les algorithmes de sélection vers l'avant et de régression pas à pas.

À la lumière des premières estimations, nous constatons que les résultats des trois méthodes d'estimation s'équivalent. Les différences observées au tableau 3.3 s'expliquent par le fait que la méthode pas à pas construit des modèles en général plus simples (moins de coefficients).

Tableau 3.3: Comparaison des trois méthodes d'estimation*

Méthode	R^2	probabilité de manque d'ajustement
Régression linéaire	0.843	0.956
Sélection vers l'avant	0.843	0.956
Régression pas à pas	0.827	0.958

*: Modèle quadratique avec interactions de premier niveau, mesure de l'opacité

La suite de l'analyse ne présente que les résultats obtenus par régression pas à pas. À niveau d'explication équivalent, les modèles comportant moins de coefficients sont plus intéressants. La simplicité des modèles facilite leur intégration comme outil de planification; Les utilisateurs rejettent généralement les solutions trop complexes.

L'observation des mesures révèle que la durée d'entreposage a eu un effet important sur l'indice d'égouttage des pâtes de fibres recyclées et à très haut rendement. Les valeurs mesurées diffèrent de façon importante des valeurs visées (437, 575 ml pour la variable HRCSF et 95, 115 ml pour FRCSF, voir figure 3.4). La lecture des autre mesures réalisées sur ces deux pâtes (la grosseur des fibres, la flexibilité et la longueur moyenne des fibres) nous permettent de supposer que seule la mesure de l'indice d'égouttage a été perturbée. C'est sous cette hypothèse que nous continuons l'étude. Les prochaines campagnes d'essais devront cependant être organisées de façon à éviter de tels décalages.

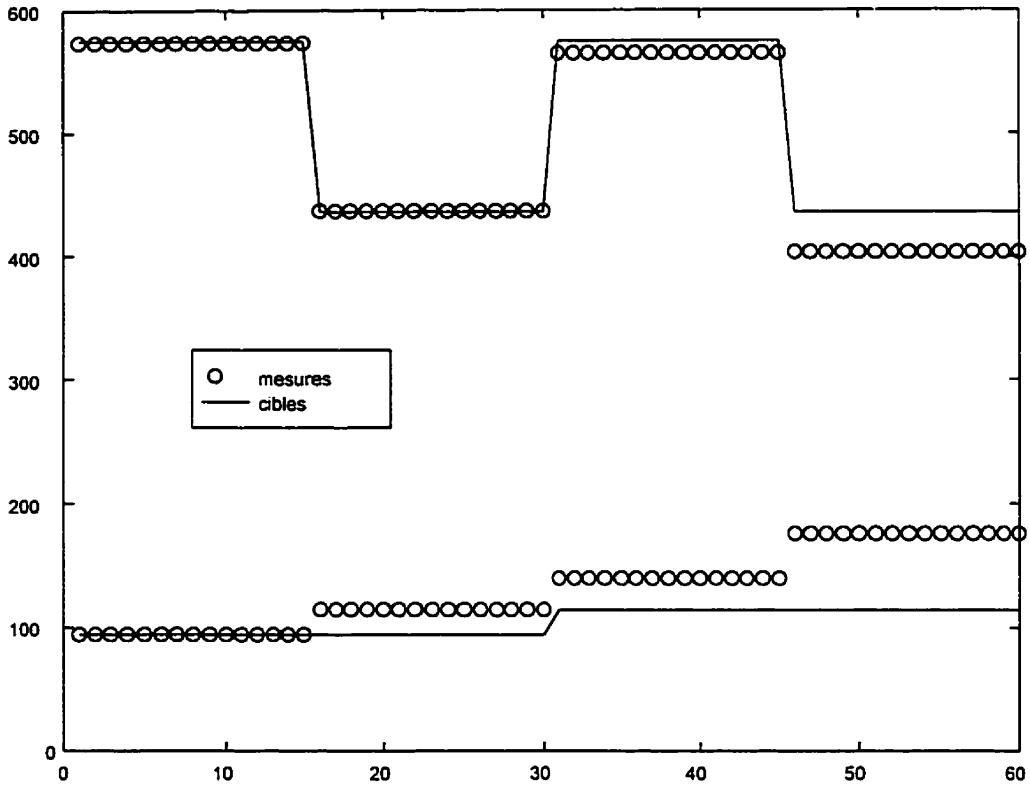


Figure 3.4: Déviation des indices d'égouttage

Nous avons estimé les paramètres d'un modèle dont la forme générale figure à l'équation 3.1. Les produits croisés (X_iX_j) servent à modéliser les interactions de premier niveau entre les facteurs étudiés. Les calculs sont réalisés à l'aide d'outils statistiques programmés sous MatLab (Voir description des outils, Annexe A).

$$Y = \sum_{i=1}^{5} a_i X_i + \sum_{i=1}^{5} a_j X_j^2 + \sum_{i=1}^{4} \sum_{j=i+1}^{5} a_{ij} X_i X_j \tag{3.1}$$

L'estimation de modèles de l'énergie absorbée en tension (TEA) a donné de piètres résultats. Cependant, cette variable a fait l'objet dans le passé de différentes

analyses qui ont mené à l'élaboration de modèles fiables. Compte tenu du faible coefficient de manque d'ajustement (environ 27 %), nous considérons donc un manque d'adéquation dans les conditions environnementales ou dans le choix des modalités. Bien que les résultats sur la signification des paramètres identifiés soient probants, il serait préférable de reprendre les essais pour cette mesure. Durant cette campagne, il faudrait accorder une attention particulière à l'environnement et reconsidérer la nature et la plage de variation des facteurs explicatifs.

L'observation des valeurs de l'indice de déchirure (DE) révèle que la mesure prise au 5ième essai du quatrième bloc est aberrante. L'ordre de grandeur de cette valeur ne respecte pas celui de l'ensemble des observations (au moins le triple de la moyenne des 59 autres observations). Cette valeur a été interpolée à l'aide d'un modèle linéaire construit avec les valeurs correctes avant de procéder à l'estimation du modèle final. Cette interpolation n'a toutefois pas altéré la qualité du modèle obtenu.

Tableau 3.4: Estimation des modèles des caractéristiques modélisées

Caractéristiques	R^2	C.M.A.	Modèles
TEA	0.540	27 %	34.69 + 1.48(PFR) + 1.77(PTHR) +... 2.31(PKB) - 1.26(HRCSF) + ... 1.68(PFR)^2 - 1.21(PTHR*FRCSF)
DE	0.796	74 %	8.86 + 0.72(PTHR) + 0.81(PKB) +... 0.13(FRCSF) + 0.20(HRCSF)
OP	0.826	95 %	94.67 -0.85(PTHR) -0.67(PKB) -... 0.36(HRCSF) +0.20(FRCSF)

R^2: Coefficient de régression;

C.M.A.: probabilité de manque d'ajustement.

Nous acceptons le modèle de l'opacité, bien que le coefficient de manque d'ajustement atteigne le seuil de signification (95 %). Cette décision est appuyée par l'analyse des résidus; leur distribution et leur comportement face à l'augmentation des valeurs estimées nous indiquent que la caractéristique est bien modélisée. Le manque d'ajustement peut être amplifié par une ou plusieurs valeurs anormales dans la série d'essais. Le tableau 3.4 expose le sommaire des résultats de l'estimation finale des trois modèles associés aux caractéristiques étudiées. Les résultats exacts et les graphiques d'analyse des résidus des variables DE et OP figurent à l'annexe D.

Le graphique des résidus en fonction des valeurs estimées de la figure 3.5 ne présente pas de tendance réfutant la justesse du modèle. L'étendue des variations ne progresse pas selon la valeur de TEA estimée et les résidus sont distribués autour de zéro, sans tendance marquée. Étant donné le comportement aléatoire des résidus, nous n'avons aucune indication de la nécessité d'enrichir le modèle. D'après les données étudiées, il est superflu d'effectuer une transformation (ln, y^2, etc.) de la variable de réponse. Le modèle est jugé acceptable au point de vue du comportement des résidus.

L'évaluation visuelle de la distribution des résidus n'indique pas d'écart important par rapport à un comportement gaussien. Le graphique de la figure 3.6 nous permet d'admettre que les résidus de la variable TEA fluctuent selon une loi de distribution normale (la ligne droite sur le graphique) dans 80 % des cas. Les observations sont similaires dans les graphiques présentés en annexe. De plus, aucune observation inhabituelle n'est remarquée sur les vecteurs de résidus: en fait, les résidus sur cette échelle suivent à peu près un droite, ce qui correspond au comportement d'une distribution selon une loi normale (figure 3.6).

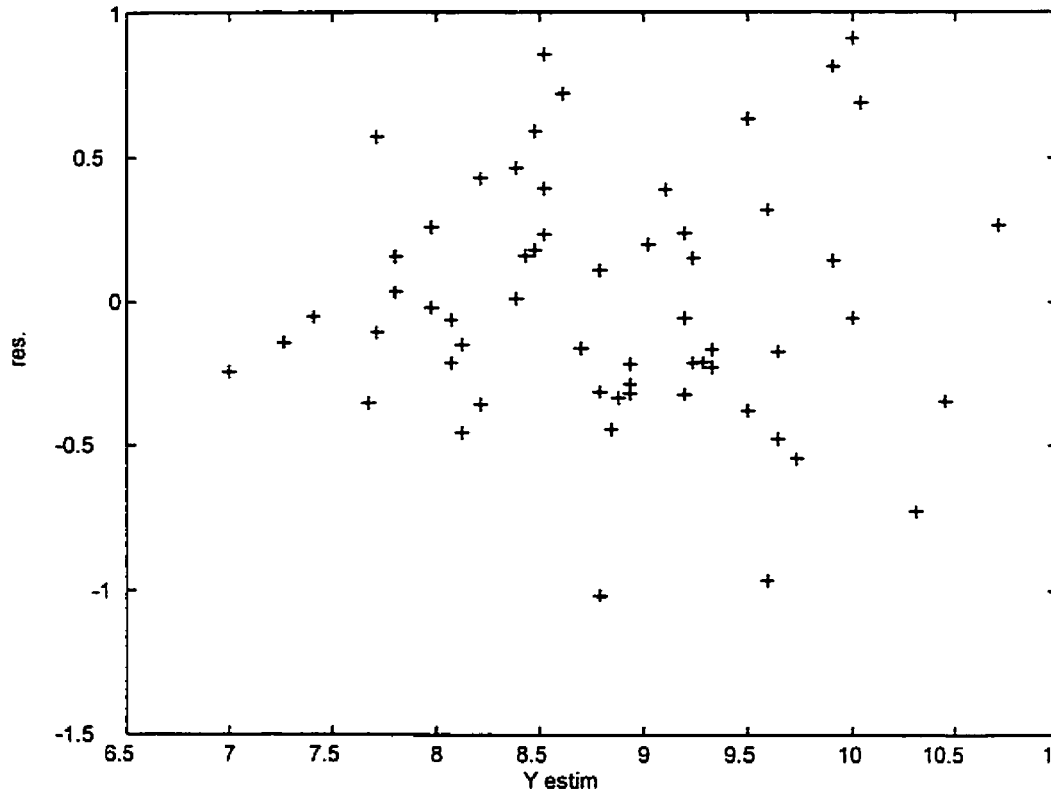


Figure 3.5: Distribution des résidus selon TEA estimé

3.6. Conclusion

Les résultats obtenus lors de la modélisation de l'indice de déchirement ne nous permettent pas d'établir un modèle fiable expliquant cette variable. Le coefficient de régression faible (R^2=0,54) indique une grande variabilité inexpliquée non reliée au manque d'ajustement (C.M.A.=27 %). La difficulté rencontrée lors de l'estimation peut être due à des perturbations importantes de l'environnement. Cependant, les analyses des autres caractéristiques ne semblent pas ou peu affectées par l'environnement. Il faudrait, lors de prochaines études, reconsidérer les modalités des facteurs. Il est possible que la plage de variation ne soit pas assez importante pour permettre de distinguer les fluctuations aléatoires de celles commandées par les facteurs. Nous pourrions également

envisager, lors de campagnes d'essais ultérieures, d'autres facteurs comme par exemple la grosseur et la longueur des fibres.

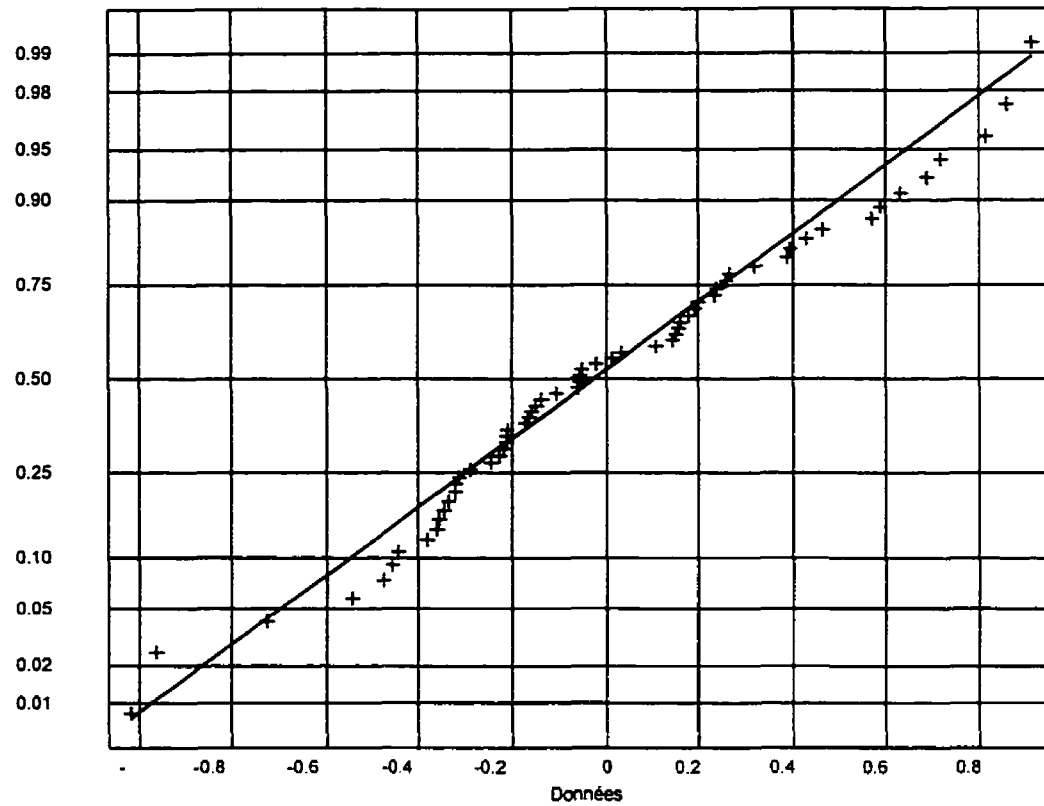


Figure 3.6: Distribution des résidus du modèle de TEA

L'estimation du modèle exprimant le niveau de déchirure (DE) comme une fonction des proportions du mélange a donné des résultats fiables (R^2=0,80, C.M.A.=74 %). Ce modèle linéaire présente la déchirure (résistance en cisaillement) comme une fonction de la proportion de pâte à très haut rendement, de pâte Kraft blanchie et des indices d'égouttage étudiés (HRCSF, FRCSF). Cependant, la valeur observée à l'essai 50 (no. 5, bloc 4) a été remplacée par une interpolation, étant donné

l'amplitude de l'écart avec les autres valeurs. Cette erreur serait imputable à un manque d'attention du technicien lors de la transcription des mesures. Selon ce modèle, nous devons utiliser les valeurs du tableau 3.5 pour atteindre la déchirure maximale.

La méthode suivie a permis d'estimer un modèle explicatif adéquat pour l'opacité (R^2=0,83, C.M.A.=95 %). Cette caractéristique apparaît affectée par les mêmes facteurs que la résistance en cisaillement (tous les facteurs initiaux, sauf la proportion de fibres recyclées). Nous observons toutefois un indice de manque d'ajustement élevé : 95 %, le niveau limite de signification acceptable. Devant un cas limite, il faut se pencher sur le comportement des résidus avant de décider de la validité du modèle. Selon les graphiques présentés à l'annexe D, aucune tendance marquée n'a été négligée par le modèle. Une série supplémentaire d'essais pourrait préciser ces résultats en étudiant l'effet des variables qui ont été maintenues sous contrôle durant cette analyse. Les modalités sur lesquelles fixer les paramètres sont présentées au tableau 3.5 (de même que les niveaux nécessaires pour maximiser la résistance en tension, bien que ce modèle ne soit pas satisfaisant).

Tableau 3.5: Niveaux de modalités pour atteindre les valeurs maximales

Facteurs	TEA *	DE	OP
PFR	20 %	**	**
PTHR	50 %	50 %	30 %
PKB	10 %	10 %	10 %
FRCSF	**	115 ml	95 ml
HRCSF	437 ml	575 ml	575 ml

* : Bien que le modèle soit insatisfaisant, la signification des paramètres nous permet quand même de conclure sur le niveau à atteindre pour maximiser l'énergie absorbée en tension.

**: Non significatifs.

Les niveaux auxquels fixer les facteurs sont différents pour la résistance en cisaillement et pour l'opacité. Pour maximiser DE, il faut fixer les facteurs à leurs modalités maximales alors que l'opacité exige que les proportions de fibres et l'indice d'égouttage de la fibre à haut rendement soient à leur niveau minimal. Devant des objectifs de production conflictuels (maximiser l'opacité et la résistance en cisaillement), les responsables devront choisir le meilleur compromis.

Le grand nombre de caractéristiques servant à la description de la qualité du papier demeure un problème important. Si l'évaluation des relations cause à effet des facteurs et des caractéristiques physiques permet la préparation de stratégies d'intervention, elle laisse le problème du suivi du procédé sans réponse. L'humain normal ne peut généralement pas suivre plus de quatre ou cinq sources d'information visuelle. Dans le cas de procédés complexes, comme celui faisant l'objet de cette étude, il faut présenter une information synthétisée.

Dans cette optique, nous prévoyons étudier l'application de l'analyse par composantes principales. Cette méthode est utilisée pour modéliser la variation d'un grand nombre de caractéristiques à partir d'un nombre limité de variables. Il serait intéressant d'évaluer le même type de modèles que ceux présentés dans cette étude, mais en considérant comme variable réponse le résultat de la modélisation par composante principale.

CHAPITRE IV
CONCEPTION D'UN INSTRUMENT DE MESURES SERVANT À CARACTÉRISER L'ERREUR DE RÉPÉTABILITÉ D'UN ROBOT

4.1 Mise en situation

La caractérisation de la précision d'une machine est une opération essentielle réalisée entre le moment de l'acquisition et de la mise en fonction dans l'usine. Les méthodes systématiques d'évaluation des machines par le fournisseur lui permettent de différencier les produits qu'il offre. Cette évaluation de la performance de la machine ne correspond toutefois pas à celle réalisée dans le contexte des activités de l'acheteur. Les mesures du fabricant ne sont généralement pas établies sous les mêmes conditions que celles imposées par l'entreprise cliente, ce qui oblige les acheteurs à ausculter leur nouvelle acquisition avant de la mettre en fonction. La connaissance des capacités réelles de la machine permet aussi de planifier son utilisation à moyen terme. La connaissance réelle de la capacité permet donc de déterminer si les équipements achetés répondent aux besoins de précision actuels et d'envisager d'autres applications dans l'avenir.

Cette étape prend de l'importance lors de l'achat d'un manipulateur robotisé. L'erreur machine cumulée n'est pas linéaire et change selon la posture dans laquelle le robot se trouve (Ramsli, 1991). Il faut donc établir un protocole qui permette de quantifier l'erreur du robot dans le sous-espace dans lequel il réalisera sa tâche. Nous choisissons donc d'évaluer la dispersion angulaire des articulations plutôt que l'erreur en position et en orientation de la tâche. Cette mesure ne varie pas selon les conditions d'utilisation et permet d'évaluer la répétabilité sur la tâche. Cette étude vise donc à déterminer les conditions sous lesquelles l'estimation de la répétabilité des articulations

est la plus facilement réalisée. L'étude mènera à la fabrication de l'outil répondant le mieux à cet objectif.

L'étude est réalisée en deux phases. Tout d'abord, nous réalisons l'analyse statistique sur un simulateur programmé sous MatLab. Ensuite, nous concevons, à la lumière des résultats obtenus, l'outil de mesure qui présente les meilleures performances sur le simulateur et l'utilisons pour estimer en laboratoire la contribution de chaque articulation à l'erreur de répétabilité du robot.

4.2 Variable de réponse analysée

La répétabilité est définie selon la norme ISO 9283 comme étant la dispersion en position (dans l'espace cartésien) et en orientation (rotations autour des axes X, Y et Z). Cependant, on montre, tant théoriquement qu'expérimentalement, que la répétabilité dépend en grande partie de la posture du robot (Ramsli, 1991). Il est donc préférable de se définir une mesure caractérisant la répétabilité du robot qui ne soit pas affectée par la posture de celui-ci: les erreurs angulaires des articulations. L'erreur angulaire suit une distribution uniforme caractérisée par la précision des encodeurs. Étant donné que cette valeur est constante, c'est sur son identification que les efforts d'analyse seront concentrés.

Le torseur des petits déplacements permet d'exprimer la répétabilité selon la norme ISO 9283. Le torseur est lié à l'erreur angulaire δq par la matrice jacobienne (équation 4.1). En considérant l'influence d'autres facteurs internes, la réciproque de l'équation 4.1 permet l'extraction et l'estimation de δq à partir de l'erreur de répétabilité mesurée et exprimée sous la forme du torseur. Il est important de se rappeler que nous ne mesurons pas directement le torseur de petits déplacements. On le calcule à partir des

données mesurées $d(q)$, reliées au torseur par la matrice $T(f)$, qui explique l'organisation spatiale des capteurs regroupés en instrument de mesures.

$$\delta\chi = J(q)\delta q \tag{4.1}$$

En incluant la matrice de configuration $T(f)$ et la matrice de l'effet des réducteurs $R(q)$ ainsi qu'en tenant compte de la précision des instruments de mesures (erreur d'arrondi), les mesures $d(q,f)$ s'expriment comme une fonction de δq telle qu'indiquée à l'équation 2.2.

$$d(q, f) = T(f)\, J(q)\, \delta q \tag{4.2}$$

où

$$\delta q = R(q)\, \delta e \tag{4.2-1}$$

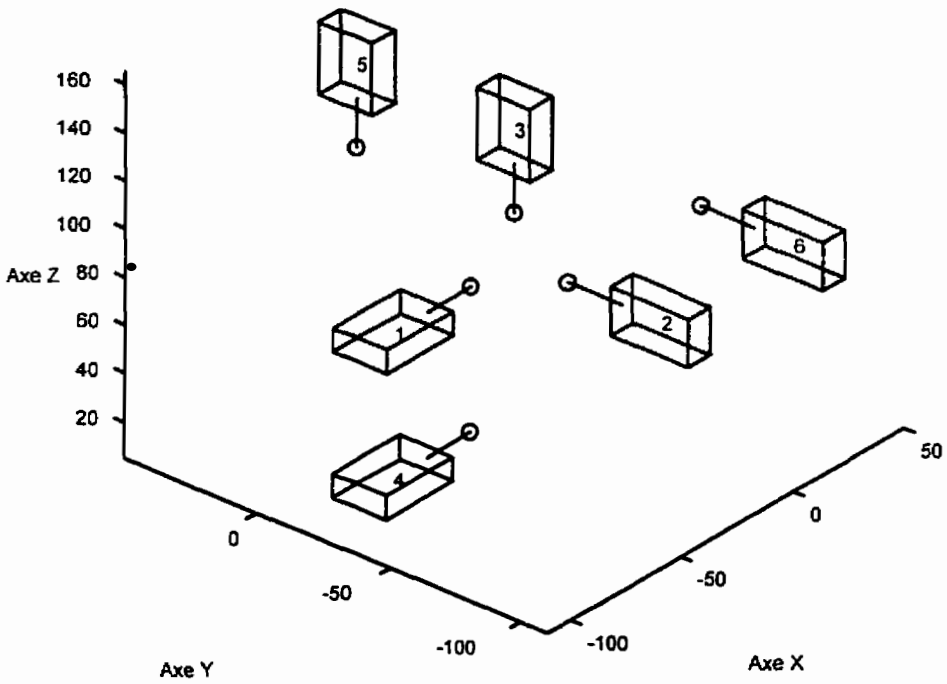


Figure 4.1: Montage de comparateurs proposé par Langmoen

Nous appelons le produit $T(f)J(q)R(q)$ la matrice de propagation des variations $\rho(q,f)$. En multipliant les deux côtés de l'équation par leurs transposées respectives, nous obtenons l'expression de la matrice de covariance des mesures telle que:

$$\rho(q,f) = T(f)\, J(q)\, R(q)\, v\, R(q)^T\, J(q)^T\, T(f)^T \tag{4.3}$$

où v est la matrice de covariance des encodeurs.

En admettant l'indépendance physique des encodeurs, v devient une matrice diagonale. En trouvant l'expression numérique de v, nous pouvons, à partir de l'équation 4.3, estimer l'erreur de répétabilité sur la tâche, quelque soit la posture du robot. v se calcule à partir de l'équation 4.4.

$$v = R(q)^{-1}\, J(q)^{-1}\, T(f)^{-1}\, cov\;\; T(f)^{T^{-1}}\, J(q)^{T^{-1}}\, R(q)^{T^{-1}} \tag{4.4}$$

où *cov* est la covariance des données mesurées.

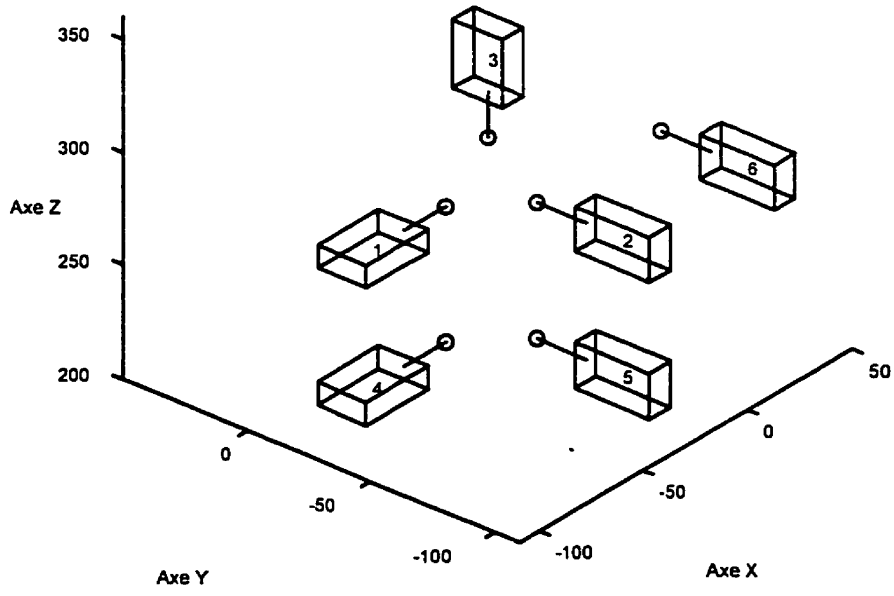


Figure 4.2: Montage de comparateurs proposé par Mooring

Nous considérons, lors de la simulation, une structure robotisée dont nous connaissons déjà la précision des encodeurs. L'objectif de la première étape est de déterminer les meilleures conditions d'estimation: celles minimisant l'écart entre la valeur connue et l'estimation de notre montage. La variable de réponse du modèle est donc l'erreur d'estimation de cette dispersion, exprimée en pourcentage. Comme cette réponse s'exprime sous forme d'un vecteur-colonne (six erreurs pour six articulations), nous limiterons l'analyse à l'erreur d'identification maximale.

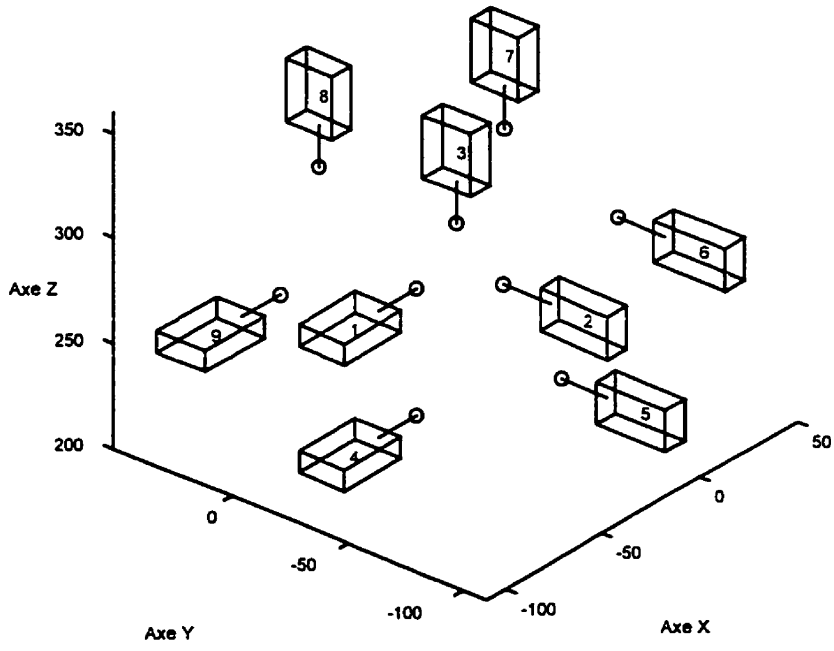


Figure 4.3: Montage de comparateurs proposé par Riley

4.3 Choix des paramètres du modèle

Le simulateur nous fournit un environnement expérimental exempt de perturbations extérieures à la résolution machine près, en double précision. Ainsi, seules les variations des paramètres produiront une variation de la réponse. Nous avons limité l'étude à l'influence de quatre facteurs:

- la précision des comparateurs (instruments de mesure);
- la configuration des comparateurs (disposition dans l'espace);
- la résolution des encodeurs des articulations (l'erreur interne du robot);
- la posture du robot.

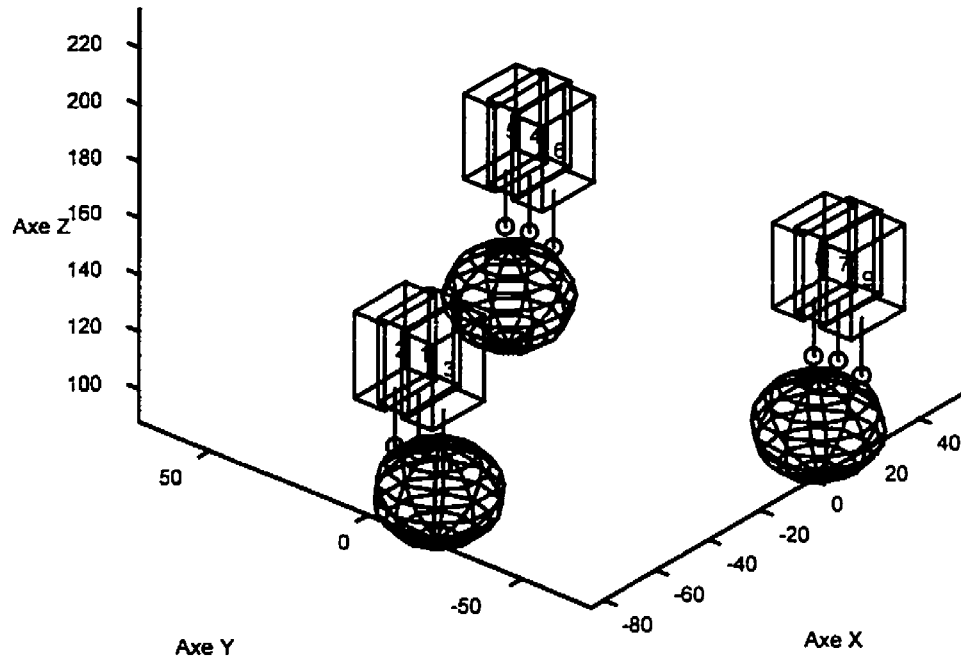


Figure 4.4: Montage de comparateurs proposé par Acquaviva

Les instruments de mesure modélisés sont des comparateurs à tige rentrante (Mitsutoyo M275436). Les comparateurs du laboratoire sont précis au centième de millimètre près. A priori, un instrument de mesure plus précis devrait nous permettre de meilleures estimations de la position du poignet du robot. Nous avons donc considéré trois modalités de précision: au dixième (1/10), au centième (1/100) et au millième (1/1 000) de millimètre.

Plusieurs auteurs ont utilisé ce type de comparateurs dans l'analyse de répétabilité de manipulateurs robotisés. Nous avons identifié quatre montages:

- la structure de Langmoen (1984), à six comparateurs (figure 4.1);
- la structure de Mooring (1986), à six comparateurs (figure 4.2);
- la structure de Riley (1987), à neuf comparateurs (figure 4.3);
- la structure d'Acquaviva (1985), à neuf comparateurs (figure 4.4).

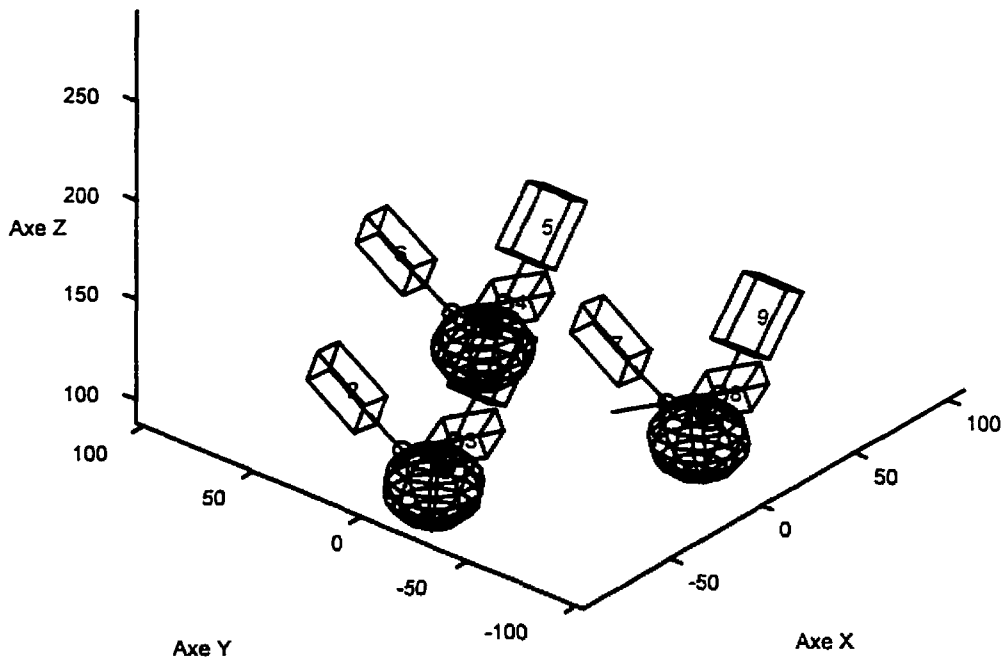


Figure 4.5: Montage de comparateurs conçu à l'interne

La configuration proposée par Mooring est une application du repère isostatique de Kelvin (un appui plan, un appui droite et un appui point). Riley y ajoute trois comparateurs de façon à avoir un montage hyperstatique, ce qui permet d'évaluer la position et l'orientation d'un objet par une régression (méthode des moindres carrés). Finalement, Acquaviva a suggéré une configuration de neuf comparateurs agencés sur un seul plan. Le système est composé de trois groupes de trois comparateurs organisés selon un triangle équilatéral. Les comparateurs sont stimulés par trois sphères fixées au poignet du robot. Inspiré de cette dernière suggestion, nous proposons un système de 9 comparateurs organisés selon des ensembles de tétraèdres réguliers. Les points de contact de trois comparateurs sur une sphère ont une relation tétraédrique avec le centre de la sphère et ces centres respectent le même type de relation avec le poignet du robot. (figure 4.5).

Nous cherchons en sommes à identifier la dispersion des articulations ou leurs erreurs angulaires. Nous voulons donc étudier la sensibilité de notre modèle à diverses précisions d'encodeurs (ordre de grandeur des erreurs angulaires). Le robot du laboratoire a en réalité des encodeurs de 16 bits aux trois premières articulations et 14 aux trois dernières (on simule le comportement du robot en utilisant les dimensions et niveaux de précisions inscrits dans le manuel d'utilisation du robot afin de nous assurer des résultats s'approchant de la réalité). Nous considérons trois niveaux: un robot à six encodeurs de 14 bits, le robot réel et un robot à six encodeurs de 16 bits.

Compte tenu qu'il existe un nombre infini de positions possibles, on décrira la posture du robot par le conditionnement de la matrice de propagation des variations. Le conditionnement se définit comme le rapport entre la plus grande et la plus petite valeur singulière d'une matrice (max(valeurs singulières)/min(valeurs singulières)). Cette valeur est un indice du «niveau de singularité» de la matrice. Un conditionnement très élevé nous indique que la matrice est quasi-singulière, ce qui cause des erreurs dans l'évaluation

numérique de son inverse (le plus petit conditionnement possible - 1 - correspond à celui d'une matrice orthogonale).

4.4 Modalités choisies et modèles considérés

On estime le modèle à partir des valeurs standardisées des paramètres du modèle. Par exemple, la précision des comparateurs varie du centième au dix millième de

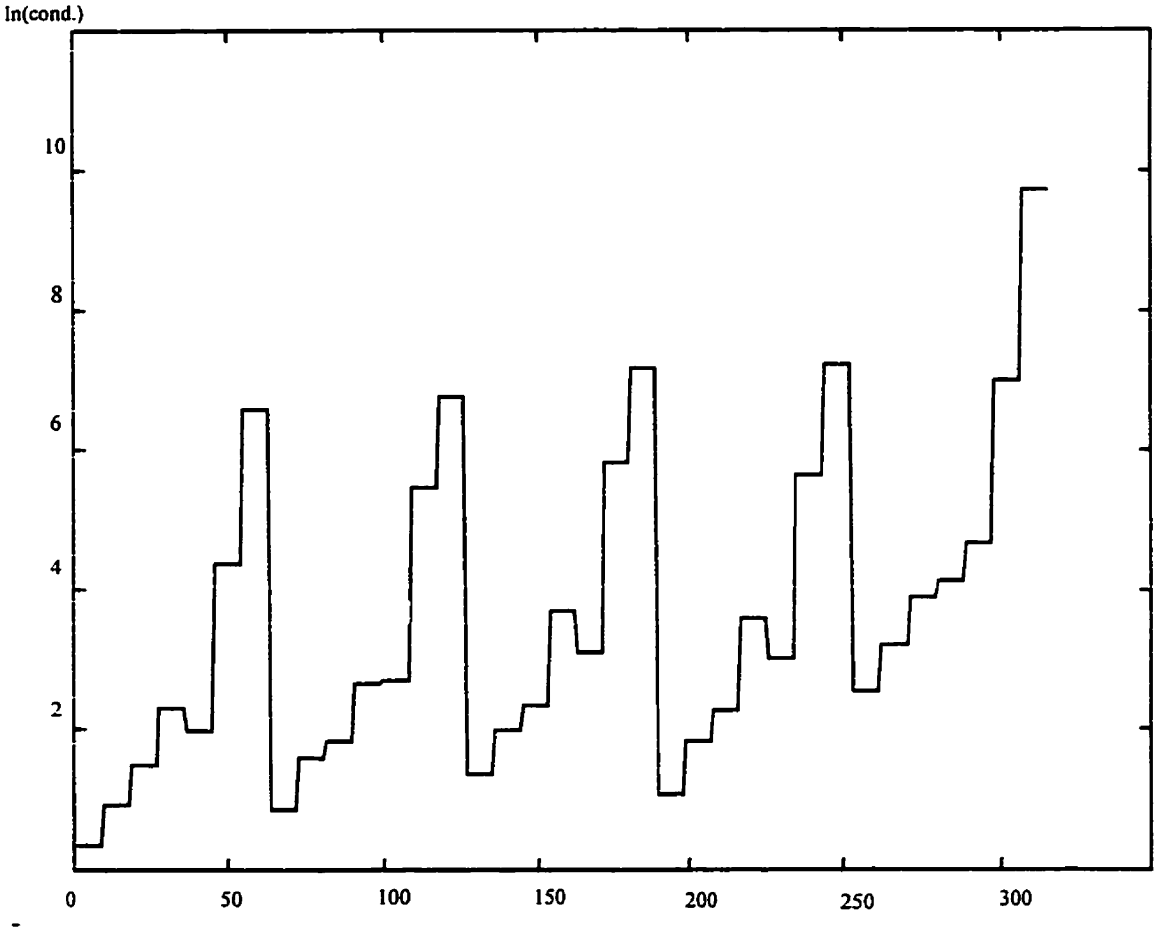


Figure 4.6: Valeur de conditionnement selon la séquence d'essais

millimètre mais nous utilisons uniquement les valeurs -1, 0 et 1 pour la modélisation (ceci afin d'éviter de masquer la contribution d'une variable importante par un facteur fluctuant entre des valeurs très élevées).

Le tableau 4.1 présente les paramètres et leurs modalités associées. Nous considérons sept postures différentes, auxquelles sont associées des niveaux de conditionnements croissants. Cependant, le conditionnement est recalculé pour chaque outil et chaque niveau de précision des instruments de mesure et du robot. La figure 4.6 présente le graphique des valeurs de conditionnement utilisées lors des essais sur simulateur.

L'organisation des paramètres et le fait que l'analyse soit réalisée sur un support informatique nous permet d'inclure toutes les combinaisons possibles des différents niveaux. Nous réalisons donc un plan hiérarchisé (hierachical design) comprenant 315 combinaisons de modalités (7 x 3 x 3 x 5). La figure 4.7 illustre la structure du plan.

Étant donné que les modalités de chaque paramètre demeurent les mêmes pour toutes les combinaisons, cette structure de plans nous permet de considérer les effets croisés des paramètres d'entrée sur l'erreur d'identification. On réalise les essais sous MatLab et l'erreur d'identification de chaque combinaison est estimée par la moyenne de 1 000 répétitions.

Tableau 4.1: Modalité standardisées des paramètres du modèle

Variables	Valeurs considérées	Valeurs standardisées
Montage	Langmoen Mooring Riley Acquaviva Interne	-2 -1 0 1 2
Précisions du robot	6 joints 14 bits 3 joints 14 bits, 3 joints 16 bits 6 joints 16 bits	-1 0 1
Précisions des comparateurs	1/100 1/1 000 1/10 000	-1 0 1
Valeurs de conditionnement (posture du robot)	associés à sept configurations, recalculé pour les différents montages et précisions.	-3 -2 -1 0 1 2 3

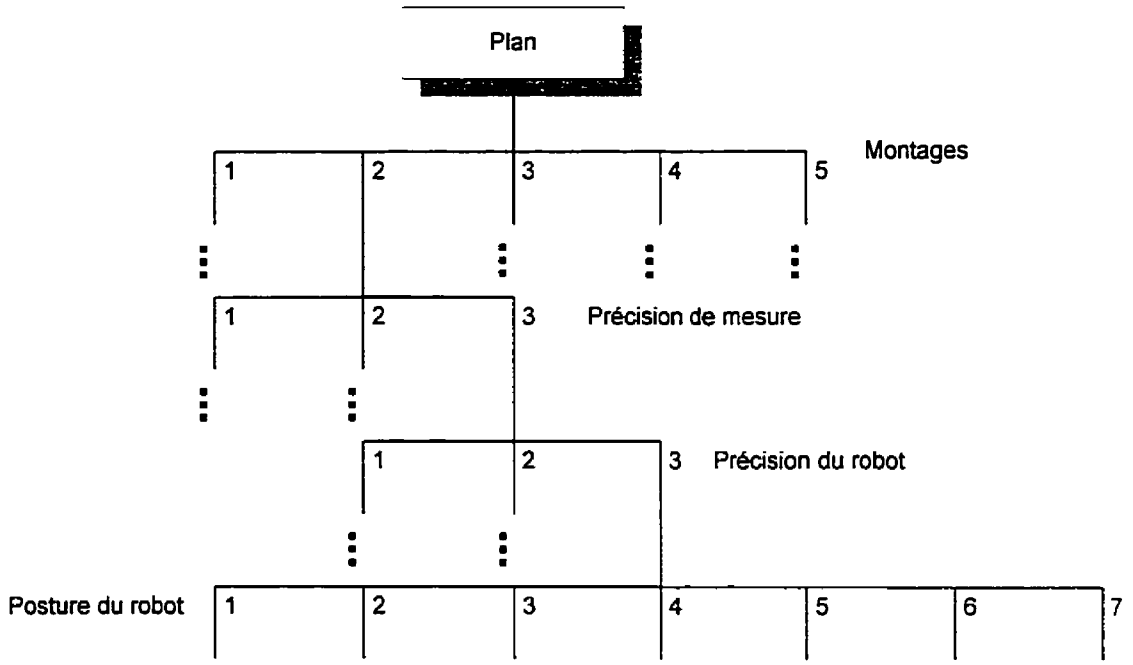


Figure 4.7: Structure du plan hiérarchisé

4.5 Estimation et diagnostic

La figure 4.8 présente le signal émis par la variable de réponse, selon les essais réalisés. On estime les modèles de régression à l'aide d'un Toolbox d'analyse multivariée développé sous MatLab. Nous cherchons à évaluer l'effet linéaire et quadratique des paramètres testés. Nous observerons de plus les effets croisés de premier niveau (interaction 2 à 2) des variables principales. L'équation générale du modèle est la suivante (K représente l'ordonnée à l'origine):

$$Y = \sum_{i=1}^{4}(a_i X_i + a_{2i} X_i^2) + \sum_{i=1}^{3}\sum_{j=i+1}^{4} a_{ij} X_i X_j + K \qquad (4.5)$$

Nous réalisons deux estimations: une première évaluation générale, suivie d'une élimination des paramètres non significatifs et une réévaluation finale du modèle. Les résultats de la modélisation figurent au tableau 4.2.

La progression des valeurs de conditionnement s'approche plus d'une fonction exponentielle que d'une fonction quadratique. Nous évaluons un second modèle dans lequel nous remanions les niveaux des modalités afin d'utiliser la transformation logarithmique du conditionnement (voir le résultat, tableau 4.2).

Les deux modèles estimés présentent un excellent niveau d'explication de la variable de réponse (coefficients de régression de 0,92 et 0,95). Cependant, Le modèle 2 est plus parcimonieux que le modèle 1 et obtient un meilleur coefficient de régression.

Deux variables apparaissent dominantes à la lumière des résultats. Les deux modèles accordent de l'importance aux valeurs de conditionnement associées aux postures du robot (x2). L'erreur d'identification est aussi liée à la précision des instruments de mesure (x4). Le modèle 1 se distingue du modèle 2 par la présence dans la liste des facteurs à effets principaux du type de montage (x1). Le modèle 2 ne présente l'effet du type de montage qu'en interaction avec la valeur de conditionnement.

La distribution statistique des résidus du modèle 2 s'apparente (visuellement) bien à une distribution normale. L'hypothèse implicite de multi-normalité du modèle est respectée.

Le graphique de la figure 4.9 (résidus en fonction de la valeur estimée) présente un comportement qui apparaît non aléatoire. Cependant, aucune transformation ou ajout au modèle n'a permis d'atténuer cette tendance. Compte tenu de la valeur élevée du coefficient de régression, de la distribution des résidus et considérant que les décisions basées sur le modèle n'impliqueront que des données situées à l'intérieur du domaine de la fonction, nous préférons la structure du modèle 2 pour expliquer la variation de l'erreur d'identification.

Tableau 4.2: Modèles de l'erreur d'identification estimés

	Modèle 1	Modèle 2
Équations	2.5491 +0.6761 x1 +1.3681 x2 +0.8789 x3 -3.1015 x4 +0.1858 x1^2 +0.4400 x2^2 +0.3412 x3^2 +0.9766 x4^2 +0.1353 x1x2 +0.3271 x1x4 +0.1243 x2x3 -0.4424 x2x4 -0.4717 x3x4	3.9883 +5.7667 x2 +0.8789 x3 -3.1015 x4 +3.4867 x2^2 +0.3412 x3^2 +0.9766 x4^2 +0.4224 x1x2 +0.6293 x2x3 -2.1704 x2x4 -0.4717 x3x4
R^2	0.9214	0.9596
Test de Fisher sur le modèle	271.5958	722.4197

4.6 Conclusion

En observant les valeurs d'erreur d'identification obtenues lors des essais, nous remarquons une progression similaire à celle des valeurs de conditionnement (figure 4.10). Les modèles estimés ont confirmé ce phénomène. L'usage de facteurs standardisés (ramenés à une même échelle) nous permet d'utiliser la valeur numérique des coefficients

pour comparer les facteurs. Or, les valeurs de conditionnement dominent en importance face aux autres données. Les décisions concernant la conception de l'outil de mesure (ici, on parlera plutôt de conditions d'utilisation) seront donc grandement influencée par les valeurs de conditionnement.

L'analyse des modèles de régression vise à identifier la structure et les conditions idéales à l'identification de l'erreur interne du robot. L'expérience a permis de vérifier l'effet de la posture du robot, de la précision de ses encodeurs, du type de configuration

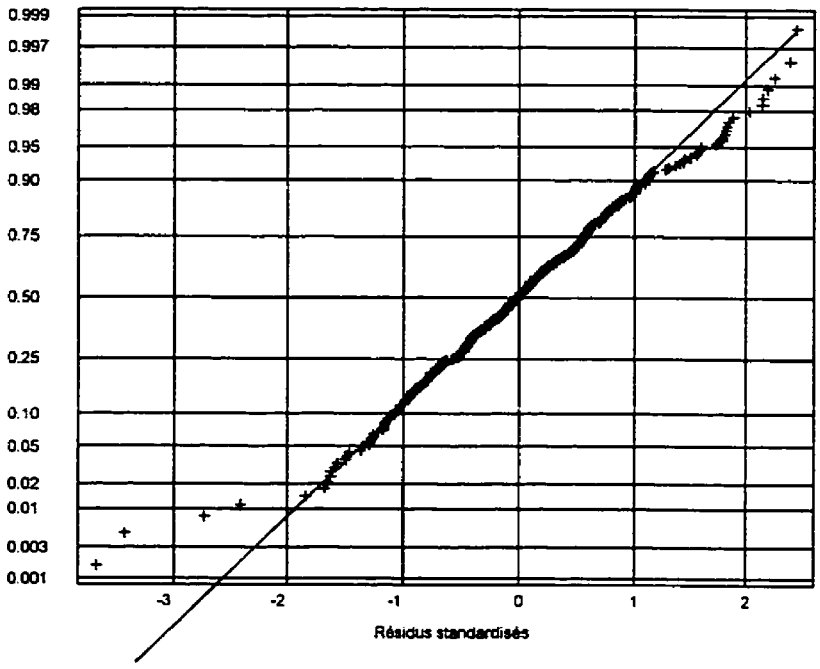


Figure 4.8: Distribution des résidus sur une échelle gaussienne

et de la précision des instruments de mesures (comparateurs). Cependant, deux de ces variables, soient les précisions du robot et des instruments de mesure, ne peuvent être physiquement changées. Le laboratoire est actuellement équipé de comparateurs précis au

centième de millimètre et d'un robot CRS-PLUS, ayant des encodeurs de 16 bits aux trois premières articulations et de 14 bits aux trois dernières.

Le modèle final devient donc (Er est l'erreur d'identification, X_1 et X_2 le type de montage et la valeur de conditionnement):

$$Er = 8.0664 + 7.9371 X_2 + 3.4867 X_2^2 + 0.4224 X_1 X_2 \tag{4.6}$$

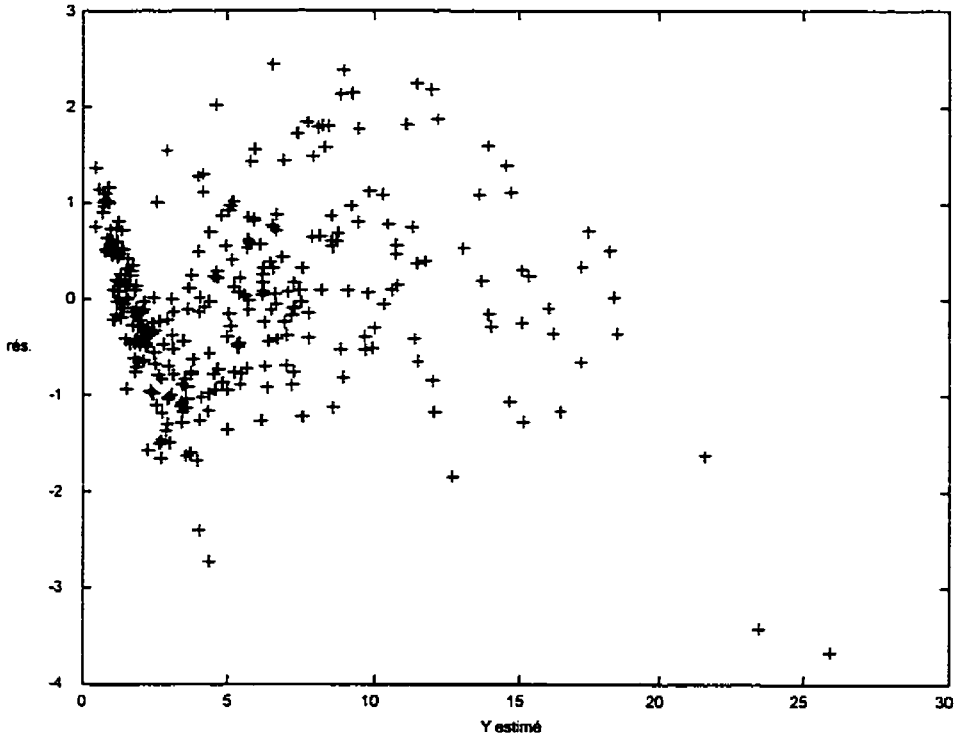


Figure 4.9: Progression de l'erreur selon la valeur estimée

Ce modèle exprime surtout l'erreur d'identification en fonction de la valeur de conditionnement. Cependant, il semble que les montages aient un rôle déterminant dans l'identification de la répétabilité angulaire. L'erreur d'identification est invariante pour

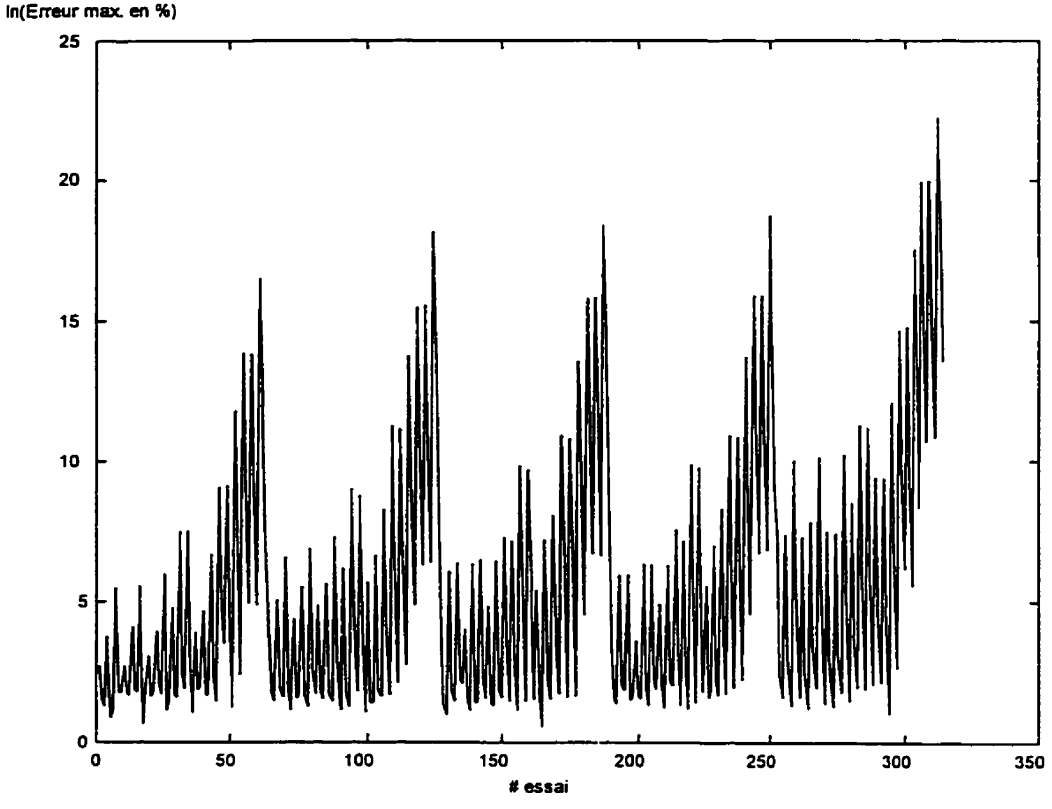


Figure 4.10: Séquence des erreurs d'identification calculées

une plage donnée de valeurs de conditionnement (de la valeur minimale 1 à un certain pallier, voir figure 4.11). C'est l'étendue de cette plage qui distingue les montages. Selon ce résultat, la configuration que nous proposons est plus performante que les autres étudiées. Nous suggérons par conséquent de fabriquer ce montage (modalité 2 dans l'expérience).

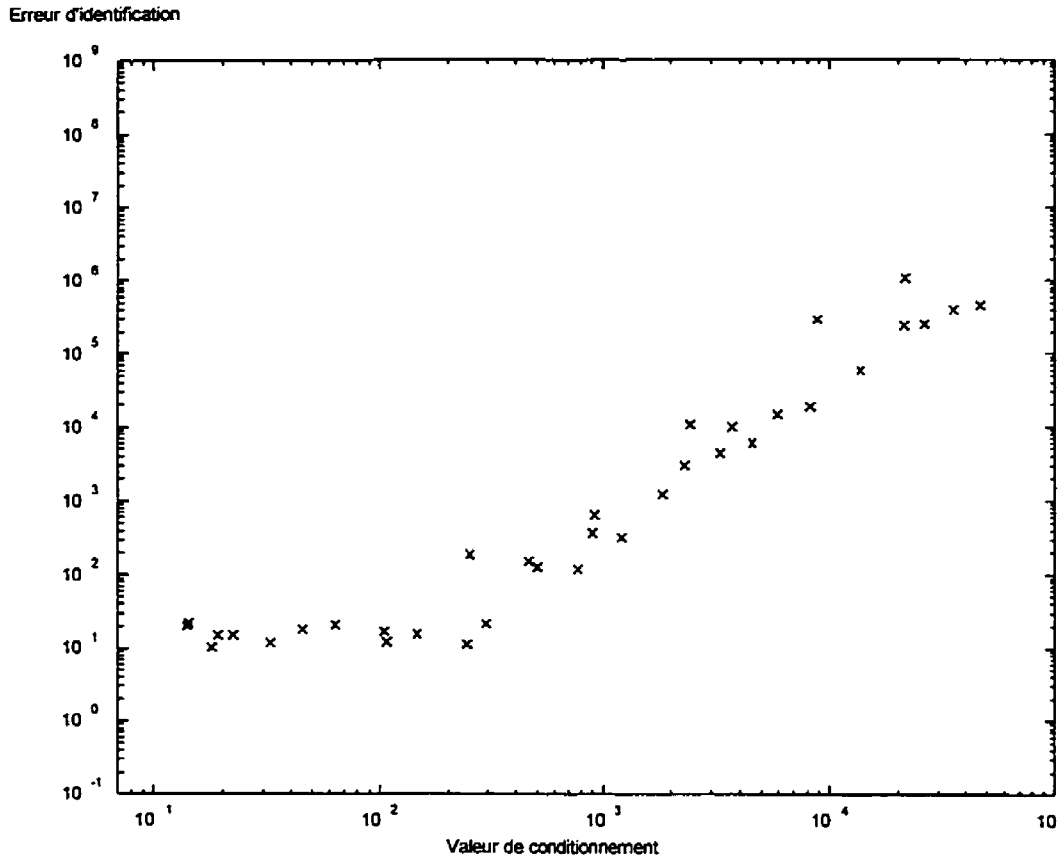


Figure 4.11: Erreur d'identification selon la valeur de conditionnements

CONCLUSION

Les statistiques utilisées à l'analyse et l'optimisation de processus donnent en général de bons résultats. L'analyse statistique doit cependant être appuyée sur des bases solides. On doit commencer par se documenter sur les activités de façon à cibler les problèmes prioritaires. L'analyse statistique restera inefficace, voire inutile, tant et aussi longtemps que subsisteront des problèmes de circulation d'information ou de retard dans les de prises de décisions. C'est à ce niveau que l'implantation d'un programme d'assurance de la qualité (ISO 9000) est la plus utile.

Une fois les problèmes techniques prioritaire ciblés, l'information recueillie guide le choix des outils mathématiques à utiliser. L'étape suivant la documentation des activités en est une de préparation des essais (donc définition de l'environnement expérimental, choix des facteurs à étudier et de la méthode à suivre). Il est essentiel de considérer les effets à court terme de l'analyse sur la production autant que les bénéfices que peuvent apporter les essais. Pour assurer l'efficacité de l'étude et la justesse des résultats, l'analyste devra avoir le maximum de contrôle sur les conditions d'opération. Ce n'est qu'une fois ces phases préparatoires complétées que les outils statistiques peuvent servir efficacement.

Le chapitre III présentait une application des statistiques non pas comme outil d'optimisation mais bien comme méthodologie d'analyse exploratoire. La faiblesse des résultats obtenus s'explique principalement par le peu de contrôle imposé au procédé durant la prise de mesures. Les données ont été mesurées sans but précis. Bien que l'on considère le niveau de fonctionnement normal des appareils, rien n'a été fait en réalité pour assurer une constance dans les conditions d'opération d'une journée à l'autre.

Nous avons, dans ce cas, relevé deux sources d'erreurs importantes: bon nombre des données entrées dans le système informatique sont conservées sous une forme synthétisée (moyenne, maximum, etc.), ce qui en rend l'analyse limitée, et plusieurs informations fournies par les opérateurs sont fragmentaires ou erronées.

Cette analyse a tout de même eu deux conséquences: elle a illustré à la direction de la société les limites de l'analyse sur des données passées, alors que cette forme d'étude était la seule permise, et elle a permis de définir les classes de variables à inclure dans la suite de l'étude.

Il y avait, sur ce dernier point, une importante confusion chez les utilisateurs en ce qui concerne les paramètres ayant un effet sur la caractéristique mesurée. On ne distinguait pas, dans l'entreprise, les paramètres d'opération (éléments servant à modifier ou maintenir les conditions d'opération) des variables de contrôle (relevés permettant d'effectuer le suivi du procédé). Ce malentendu a compliqué les analyses de procédé parce qu'on y incluait des facteurs et des variables affectées par ces derniers. En fait, sur les 300 variables de la base de donnée, moins d'une vingtaine décrivent les paramètres d'opération. Les campagnes d'essais sur le processus gagneront en efficacité si les analystes se concentrent sur ces facteurs indépendants.

La seconde étude présente une application des plans d'expériences visant à modéliser et éventuellement optimiser un procédé. Contrairement aux dirigeants de l'entreprise métallurgique, ceux de l'usine de papier sont plus familiers avec les statistiques industrielles. L'analyse des procédés chimiques par des méthodes statistiques est d'ailleurs à un niveau de maturation plus avancé que dans l'industrie lourde.

Cette étude a permis d'établir des modèles fiables pour deux des variables (opacité, résistance en cisaillement). Les modèles obtenus des autres variables (R^2 de 0,80

et 0,83 pour DE et OP, respectivement) sont des modèles linéaires, sans interaction. Les caractéristiques modélisées le sont à partir de la même série de facteurs (les proportion de pâte à haut rendement, Kraft blanchie et les indices d'égouttage de la fibre recyclée et à haut rendement). Seule la modélisation de l'énergie absorbée en tension n'a pas donné de bons résultats. Il est possible que la plage de variation des facteurs n'ait pas été assez importante pour affecter cette mesure ou que des erreurs de mesures ait eu plus d'influence sur cette caractéristique.

Cette analyse en usine avait fait l'objet d'une rigoureuse préparation et la base de données analysée a été constituée en imposant des variations contrôlées aux paramètres. Ce contexte se prête donc mieux à des analyses poussées que celui de l'étude de cas précédente. Néanmoins, l'intervention humaine qu'est la prise de mesure ne fut pas infaillible. L'entreposage prolongé des échantillons a perturbé les valeurs d'égouttage, ce qui laisse présager que d'autres caractéristiques auraient pu être affectées. La manipulation en usine devra, dans ce cas faire l'objet de contrôles plus rigoureux.

Dans la dernière étude de cas, les plans d'expériences ont servi d'outil de conception d'un instrument de mesures. Le développement du prototype sur a été réalisé sur simulateur, ce qui a permis une étude plus large dans un environnement expérimental "parfait". En effet, les modèles étudiés peuvent être beaucoup plus flexibles que ne le seraient les constructions physiques. Par ailleurs, les conditions environnementales sont exemptes de variation non expliquées. La simulation nous garantie que les seules variations non aléatoires sont celles que nous imposons. Dans un tel environnement, il devient aisé d'isoler les perturbations aléatoires et d'identifier un modèle fiable.

La performance de modélisation du chapitre V l'illustre bien. Les deux modèles testés explique une très grande partie de la variation de l'erreur d'identification (R^2 0,92 et 0,95). Cependant, il aurait été intéressant de réaliser une série d'essais suivant les

même modalités afin de quantifier l'erreur résiduelle et ainsi mesurer le manque d'ajustement tel que calculé au chapitre IV. Les prochaines modélisations sur simulateur devraient inclure cette procédure, d'autant plus que le coût des essais ne se chiffre qu'en temps d'utilisation de l'outil informatique. Étant donnés les résultats atteints à la suite de la transformation logarithmique des valeurs de conditionnement, l'information sur le manque d'ajustement nous aurait peut-être montré l'inadéquation du modèle initial et nous aurait alors forcé à envisager l'utilisation du logarithme des valeurs de conditionnement.

La capacité d'obtenir des modèles représentatifs dans les études des chapitres IV et V illustre bien l'influence de l'environnement sur les résultats. Au chapitre V, seules les variations imposées à chaque essai avaient une influence sur la variable de réponse, alors qu'à la précédente étude de cas les effets conjugués du temps, de l'environnement et des paramètres négligés ont affecté les résultats. S'il fut possible dans les deux cas d'établir certaines conclusions, il faut tout de même éviter de commettre l'impair de généraliser les résultats sans vérification. En sommes, la prudence est de mise quant à la précision des informations fournies par les modèles statistiques et aux conclusions qui en découlent. Cependant, ce type d'outil permet, et permit dans les deux cas, d'extraire puis d'utiliser une information difficilement accessible autrement.

La conception de systèmes "juste à temps", l'utilisation de logiciels de gestion de fabrication et, plus récemment, l'entretien productif (total productive maintenance) ont connu une popularité presque instantanée dans les milieux industriels. Ces méthodes sont attrayantes parce qu'elles permettent d'améliorer la productivité des activités de production. Malheureusement, la popularité de ces méthodes pourtant intéressantes s'est estompée rapidement. La plupart des utilisateurs qui avaient perçu cette série d'outils comme la panacée réglant tous leurs problèmes ont été déçus des résultats.

L'application des statistiques industrielles s'ajoute à cette liste de méthodes d'amélioration de la productivité. Ces outils sont polyvalents et, avec l'avancement des connaissances et le développement d'applications informatiques, ont un avenir prometteur. Cependant, l'analyse statistique seule, réalisée dans un environnement instable est inutile. Les utilisateurs devront faire preuve de prudence dans l'application et prendre bien soins de préparer le sujet avant l'analyse. Comme mentionné au début de ce chapitre, l'analyse doit être précédée d'une phase préparatoire, afin d'identifier les éléments à inclure dans l'étude et la description de l'environnement expérimental.

Par ailleurs, il est important de rappeler que la non qualité dans une entreprise est trop souvent la conséquence de problèmes organisationnels. Un manque de communication entre l'administration et les opérateurs, une circulation inefficace de l'information, des structures hiérarchiques lourdes associées à des décisions centralisées, entre autres, causent des problèmes plus importants que la mise au point des machines et l'optimisation des produits. Réaliser une analyse dans un contexte d'organisation inefficace permet rarement des bénéfices significatifs.

Dans l'implantation d'un programme de gestion de la qualité, l'utilisation des outils techniques, incluant les statistiques industrielles, ne vient qu'après l'assainissement des systèmes d'information (décisions) et des politiques d'entreprises (rôle et participation de chacun). C'est cet équilibre entre les gens et les équipements qui permet une réelle augmentation de la productivité dans un contexte de gestion de la qualité.

RÉFÉRENCES BIBLIOGRAPHIQUES

ACQUAVIVA, S., BELFORTE, G., IPPOLITO, R., ROMITI, A. (1985). A new measuring system for Robot Precision. Proceedings of the Annual British Robot Association Conference, Birmingham, 13-22.

BARREAU, A., et al. (1995). Sélection d'un plan d'expériences: Étude de l'adéquation des contraintes. Acte de conférence du Congrès International de Génie Industriel, Montréal, 695-703.

BOX, G. E. P., HUNTER, W. G., HUNTER, J. S. (1978) Statistics for Experimenters. John Wiley & Sons. 655 p.

BRODERICK, G., PARIS, J., VALADE, J. L. (1995). A composite representation of pulp quality. Chemometrics and intelligent laboratory systems, 29, 19-28.

DRAPER, N.R. AND SMITH, H. (1981). Applied Regression Analysis, 2nd ed. John Wiley & Sons. 710 p.

JOHN, PETER W. M. (1971). Statistical Design and Analysis of Experiments. Macmillan, New York. 356 p.

KOLARIK, WILLIAM J. (1995). Creating Quality: Concepts, Systems, Strategies, and tools. Mc Graw Hill, New York, 925 p.

KRISHNAIAH, P.R., KANAL, L.N. (1982) handbook of statistics 2. North-Holland, Amsterdam, 903 p.

LANGMOHEN, R., LIEN, T. K., RAMSLI, E. (1984). Testing of industrial robots, Proceedings of the 14th International Symposium on Industrial Robots, 201-207.

LORENZEN, THOMAS J., ANDERSON, VIRGIL L. (1993). Design of Experiments: A No-Name Approach, Marcel Dekker inc., New York, 414 p.

MOORING, B. W., PACK, T. J. (1986). Determination and specification of robot repeatability. Proceedings 1986 International conference on robotics and automation, CA, USA. 1017-1023.

RAMSLI, E. (1991). Probability Distribution of Rpeatability of industrial Robots. The International Journal of Robotic Research, 10, 3, 276-283.

RILEY, D. D. Robot Calibration and Performance Specification Determination. Robot 11 - Proceedings of the 17th international symposium on industrial robots. IL, USA. 10-1 - 10-15.

ANNEXE A: LISTE DES FONCTIONS DE STATISTIQUES PROGRAMMÉES SOUS MatLab

Multireg: Fonction générale d'estimation et d'analyse;

Stepwise: Fonction d'estimation de modèles selon la méthode pas à pas;

For_sec: Fonction de sélection de paramètres selon la procédure de sélection vers l'avant;

Ridgereg: Fonction d'estimation de modèles selon la méthode « Ridge »;

Pca_mod: Fonction de modélisation des composantes principales;

Residual: Fonction d'analyse des résidus - distribution, détection de tendances inexpliquées;

Anovamod: Fonction d'analyse de variance, incluant le test de manque d'ajustement;

Tboxcox: Fonction d'estimation du paramètre de transformation Box-Cox;

Cr_m_std: Fonction de standardisation des données, selon moyenne et écart-type échantillonnaux (ou imposés);

Cr_mean: Fonction de centrage des données selon leur moyenne

Hyp_test: Fonction de test d'hypothèse sur les paramètres du modèle (surtout appelée par la fonction Stepwise);

Cette série de fonctions utilise les fonctions suivantes, faisant partie du Statistical Toolbox de MatLab:

Fcdf: Estimation de la valeur p pour des valeurs de F et des degrés de liberté connus;

Finv: Estimation de la valeur F pour des valeurs de p et des degrés de liberté connus;

Tinv: Estimation de la valeur t pour des valeurs de p et de degré de liberté connus;

Norminv: Estimation de la valeur z prise sur une distribution normale: N(0,1) pour une valeur de p donnée;

Normplot: Présentation du graphique gausso-arithmétique (Normal probability plot);

Regress: Estimation d'un modèle selon la méthode des moindres carrés;

ANNEXE B: ANALYSE DE CORRÉLATION, ÉTUDE DE CAS 1

Annexe B: Analyse de corrélation, étude de cas 1

	tours	l	k	j	i	h	g	f	e	d	c	b
a) Rapport C/M	-0.30	-0.40	-0.24	-0.55	-0.47	-0.52	-0.34	-0.05	-0.56	-0.52	-0.13	0.04
b) Excès d'énergie	-0.13	-0.15	-0.17	-0.31	-0.28	-0.26	-0.01	0.16	-0.21	-0.21	-0.08	
c) Réajustement (mélange)	-0.08	0.00	0.00	0.12	0.14	-0.16	-0.02	-0.07	0.10	0.03		
d) Puissance d'opération	0.52	0.92	0.89	0.64	0.56	0.60	0.75	0.39	0.88			
e) Utilisation	0.41	0.74	0.87	0.65	0.59	0.63	0.79	0.42				
f) Hauteur, phase liq	0.08	0.41	0.57	-0.11	-0.18	0.00	0.77					
g) Épaisseur, C2 solide	0.30	0.67	0.79	0.39	0.38	0.41						
h) C1 non réduit dans C2	0.34	0.39	0.39	0.81	0.67							
i) Rendement massique	0.29	0.38	0.36	0.83								
j) Pureté de C2	0.37	0.47	0.41									
k) température de C1	0.41	0.88										
l) température de C2	0.47											

ANNEXE C: PLAN D'ESSAIS DE L'ÉTUDE DU MÉLANGE DE MATIÈRES FIBREUSES

PFR	PTHR	PKB	HRCSF	FRCSF	GRAM	DENS	ALL	TEA	ECL	DE	BLN	OP	L	A	B
20	50	5	573	95	60.01978	.379872	1.909	37.75	2.612472	10.13287	58.59	92.04	87.88	.18	12.53
20	30	5	573	95	61.51214	.3973652	1.893	37.3	2.334499	7.856566	58.5	95.07	87.62	.3	12.15
00	50	5	573	95	59.65738	.3563762	1.77	33.29	2.316562	9.121322	57.81	93.09	87.9	.52	13.3
00	30	5	573	95	59.81149	.3664919	1.778	32.07	2.098259	8.006502	57.53	94.99	87.45	.78	12.77
20	40	10	573	95	61.49243	.3896858	1.867	36.98	2.610077	9.910612	59.13	92.98	87.96	.32	12.14
20	40	0	573	95	60.89155	.3839316	1.794	33.67	2.315592	7.957253	58.17	94.47	87.55	.37	12.37
00	40	10	573	95	60.59857	.3792151	1.932	36.77	2.351541	8.62751	57.21	94.1	87.42	1.14	13.05
00	40	0	573	95	61.17492	.3694138	1.684	31.19	2.255826	8.23844	55.69	95.31	86.71	1.04	13.31
10	50	10	573	95	60.32715	.3872089	1.987	41.12	2.633972	9.581858	57.57	92.22	87.46	.89	12.77
10	50	0	573	95	60.90191	.3664375	1.864	37.47	2.527014	8.533011	56.59	93.98	86.97	1.01	12.87
10	30	10	573	95	60.62715	.3911429	1.998	40.78	2.457645	8.540625	57.33	94.44	87.1	.93	12.37
10	30	0	573	95	60.98003	.378288	1.761	32.55	2.192521	7.122317	57.13	95.64	87.08	.73	12.54
10	40	5	573	95	60.23096	.3816918	1.832	34.76	2.395778	8.471759	58.57	93.92	87.98	.35	12.71
10	40	5	573	95	58.65995	.3745846	1.773	31.35	2.417322	7.767796	58.23	93.4	87.84	.36	12.83
10	40	5	573	95	59.0002	.3664609	1.705	28.95	2.257619	8.89315	58.92	93.39	88.12	.5	12.63
20	50	5	437	115	60.55959	.4080835	2.078	46.76	3.043283	9.389623	54.49	94.05	85.69	1.23	12.67
20	30	5	437	115	62.50632	.3996568	1.909	39.04	2.56774	7.962549	55.93	96.06	86.21	.36	12.15
00	50	5	437	115	61.4052	.3896269	2.04	43.12	2.672412	9.025233	54.68	94.91	86.19	1.19	13.35
00	30	5	437	115	63.25705	.383376	1.776	34.37	2.303301	7.838285	56.27	96.4	86.89	.91	12.98
20	40	10	437	115	61.318	.4211401	2.115	46.38	2.901269	9.099482	56.04	93.95	86.4	.89	12.36
20	40	0	437	115	60.84074	.3971328	1.838	36.84	2.716929	8.283691	55.6	95.28	86.27	.91	12.56
00	40	10	437	115	60.28929	.4024652	1.947	38.74	2.721876	9.161052	55.54	94.49	86.51	1.23	13.06
00	40	0	437	115	60.81315	.3903283	1.85	35.5	2.45835	7.606289	55.11	95.81	86.41	1.15	13.31
10	50	10	437	115	60.59921	.3899563	2.02	39.94	2.780564	10.7299	55.87	93.15	86.55	1.18	12.78
10	50	0	437	115	59.53773	.3811634	1.942	37.07	2.556362	8.591484	55.06	94.75	86.24	1.14	13.07
10	30	10	437	115	59.87203	.3803814	1.894	33.57	2.420162	9.329726	56.82	95.16	86.84	.93	12.36
10	30	0	437	115	61.0399	.3781902	1.775	30.95	2.214945	6.755452	56.48	96.27	86.79	.73	12.63
10	40	5	437	115	59.83359	.3860232	1.944	35.75	2.520324	9.377678	55.39	94.69	86.32	1.14	2.87
10	40	5	437	115	59.7916	.370456	1.795	29.92	2.450177	8.754448	55.86	94.95	86.57	1.08	12.84
10	40	5	437	115	60.20082	.3849158	1.715	30.18	2.480032	8.913875	55.26	95.3	86.3	1.13	12.96
20	50	5	565	140	61.51226	.3576294	1.932	37.34	2.578348	10.72368	55.96	93.56	86.56	.86	12.75
20	30	5	565	140	61.19273	.3686309	1.886	34.57	2.256804	9.066826	56.64	95.39	86.69	.74	12.28
00	50	5	565	140	59.67481	.3629855	1.832	32.57	2.374536	10.05471	56.1	93.97	86.94	1.02	13.22
00	30	5	565	140	60.31144	.3607143	1.771	30.48	2.090814	8.658178	56.62	95.64	87.08	.84	12.99
20	40	10	565	140	60.56839	.3661934	1.964	38.81	2.544231	10.91152	56.68	93.6	86.78	.69	12.41
20	40	0	565	140	61.42065	.2184234	1.807	32.68	2.341232	8.849248	56.48	95.14	86.79	.58	12.63
00	40	10	565	140	60.40635	.3678828	1.908	35.78	2.413653	9.943337	56.86	94.28	87.26	.81	13.03
00	40	0	565	140	61.7192	.360931	1.775	31.4	2.171124	8.399678	56.47	95.85	87.12	.81	13.21
10	50	10	565	140	59.66238	.3660269	1.811	32.72	2.584543	10.98254	56.5	92.89	86.93	.87	12.84
10	50	0	565	140	60.51291	.35554	1.698	29.23	2.462285	9.490212	56.02	94.47	86.81	.78	13.11
10	30	10	565	140	59.00048	.3720081	1.869	33.38	2.215236	9.07395	56.53	95.05	86.73	1.11	12.46
10	30	0	565	140	60.37981	.3750299	1.744	31.35	2.196098	7.317854	55.64	96.13	86.37	1.03	12.7
10	40	5	565	140	60.02853	.3629295	1.943	36.48	2.307236	9.138114	55.52	94.86	86.33	1.25	12.75
10	40	5	565	140	60.67783	.366412	1.85	33.82	2.244642	9.433388	56.16	94.74	86.65	1.07	12.68
10	40	5	565	140	59.77614	.3583702	1.821	33.05	2.320324	8.872208	55.53	94.75	86.29	1.23	12.68
20	50	5	404	176	61.34691	.383898	2.019	43.41	2.712443	9.473734	54.64	94.18	85.79	1.72	12.68
20	30	5	404	176	60.07203	.3754502	1.87	35.19	2.350512	8.640445	56.15	95.1	86.5	.92	12.43
00	50	5	404	176	62.23051	.3794543	1.854	36.91	2.585549	9.167764	55.24	95.28	86.38	1.53	13.08
00	30	5	404	176	60.08584	.373204	1.866	33.64	2.193528	7.855042	56.09	95.83	86.76	1.26	12.92

20	40	10	404	176	158.3859	1.013994	1.954	38.53	1.007034	3.689232	55.8	94.31	86.18	1.51	12.18
20	40	0	404	176	60.69195	.3841263	1.667	30.27	2.455021	7.97308	55.44	95.62	86.12	1.07	12.46
00	40	10	404	176	59.61105	.3826127	1.91	36.02	2.499537	9.191583	55.78	94.64	86.6	1.62	12.91
00	40	0	404	176	60.83267	.3792561	1.878	36.44	2.296463	7.665752	55.01	95.91	86.29	1.66	13.15
10	50	10	404	176	59.49503	.3878424	2.008	39.92	2.800234	10.1062	55.37	93.65	86.21	1.69	12.64
10	50	0	404	176	59.89092	.3800185	1.881	35.92	2.484517	8.394105	54.75	94.93	85.99	1.76	12.91
10	30	10	404	176	60.03862	.385357	1.959	37.85	2.371807	9.220209	56.46	95.17	86.7	1.17	12.44
10	30	0	404	176	60.16311	.3760194	1.761	30.75	2.117577	7.354645	55.97	96.01	86.61	1.06	12.8
10	40	5	404	176	60.36212	.3934949	1.98	39.56	2.546299	8.609318	55.08	94.95	86.13	1.5	12.81
10	40	5	404	176	60.13791	.3925451	1.758	32.51	2.534914	8.641416	55.47	94.92	86.32	1.39	12.76
10	40	5	404	176	60.73411	.3908244	1.964	40.47	2.616322	8.711597	55.46	95.38	86.23	1.36	12.61

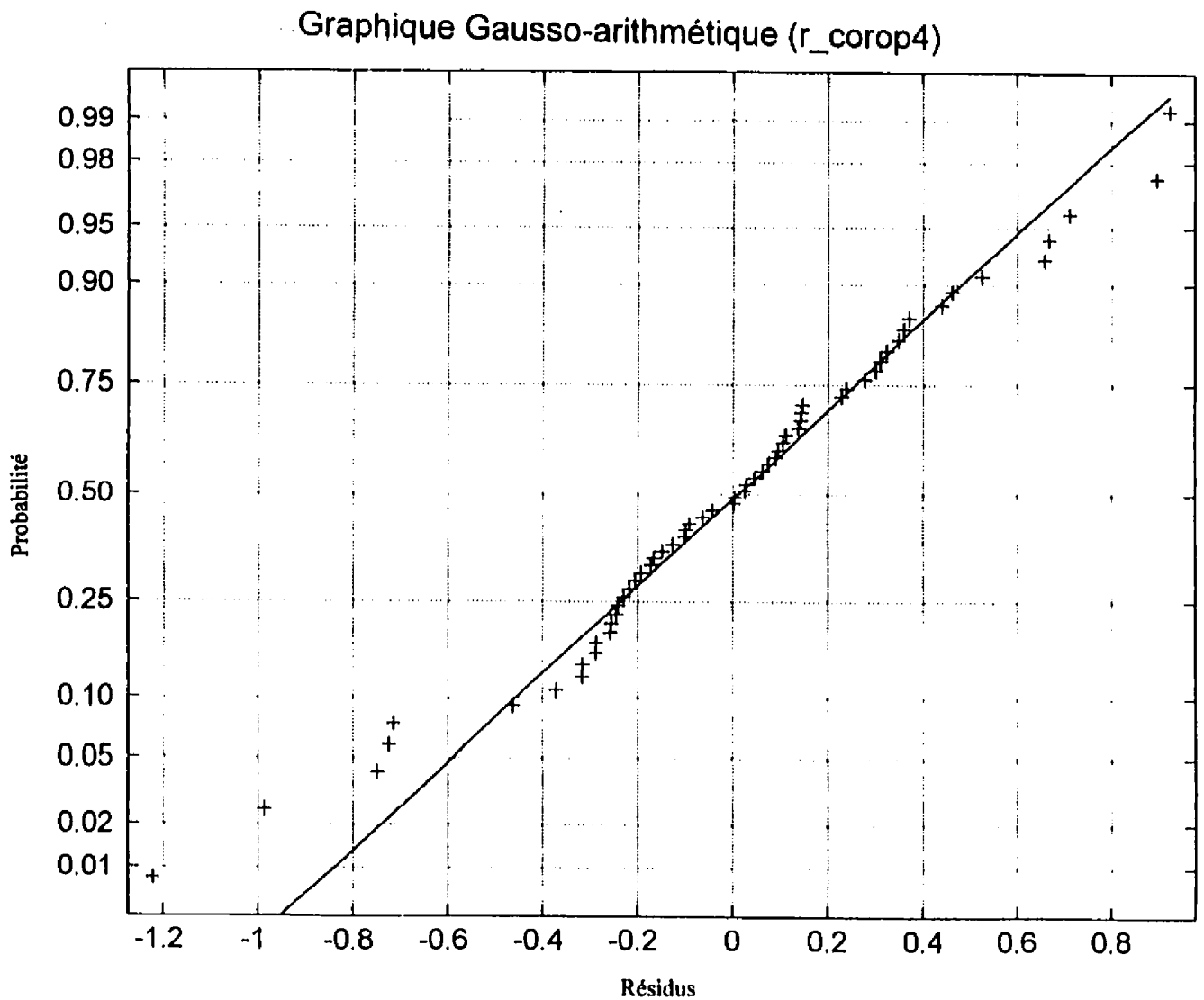
ANNEXE D: ANALYSE DE RÉGRESSION, ÉTUDE DE CAS 2

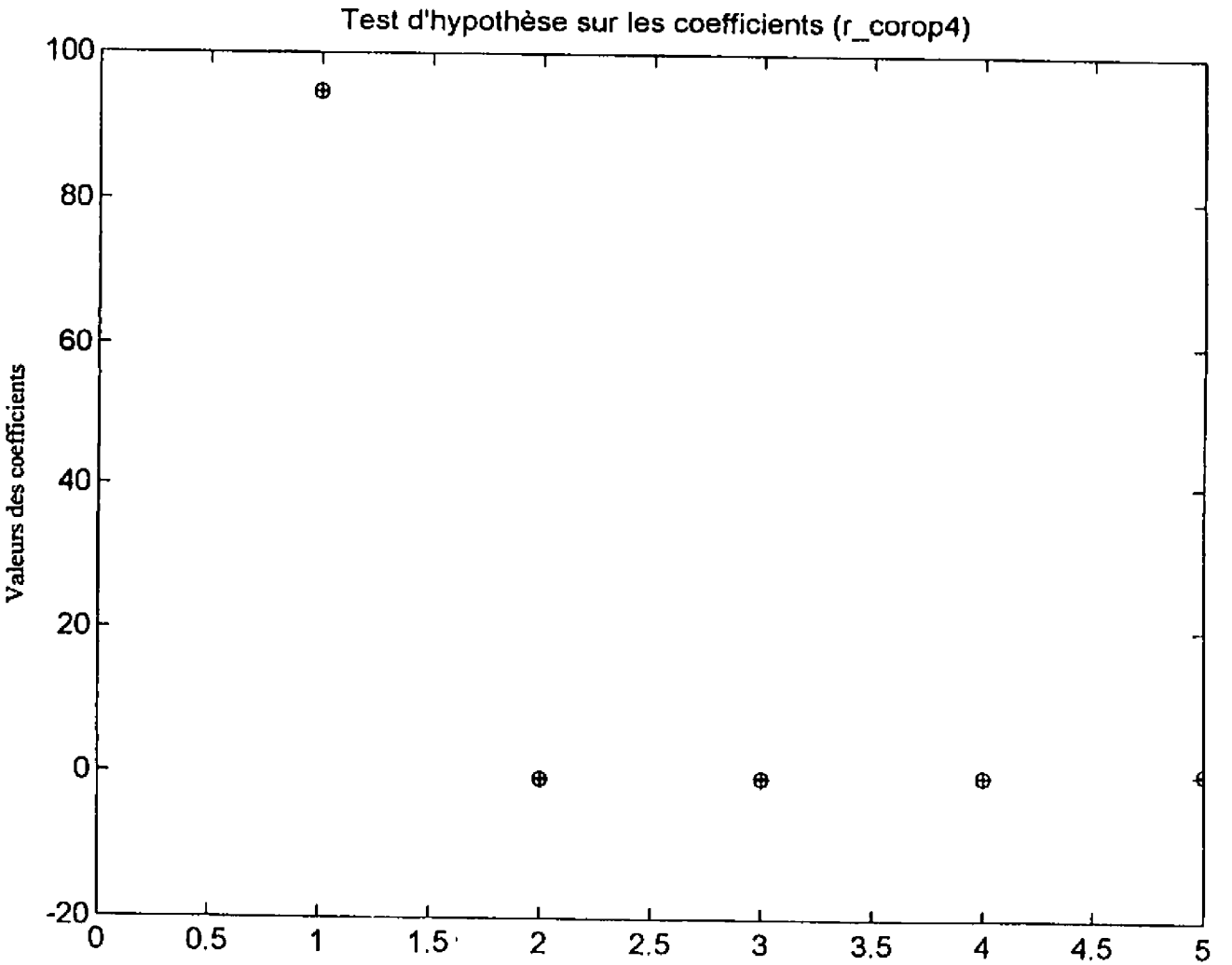
Variable : OP

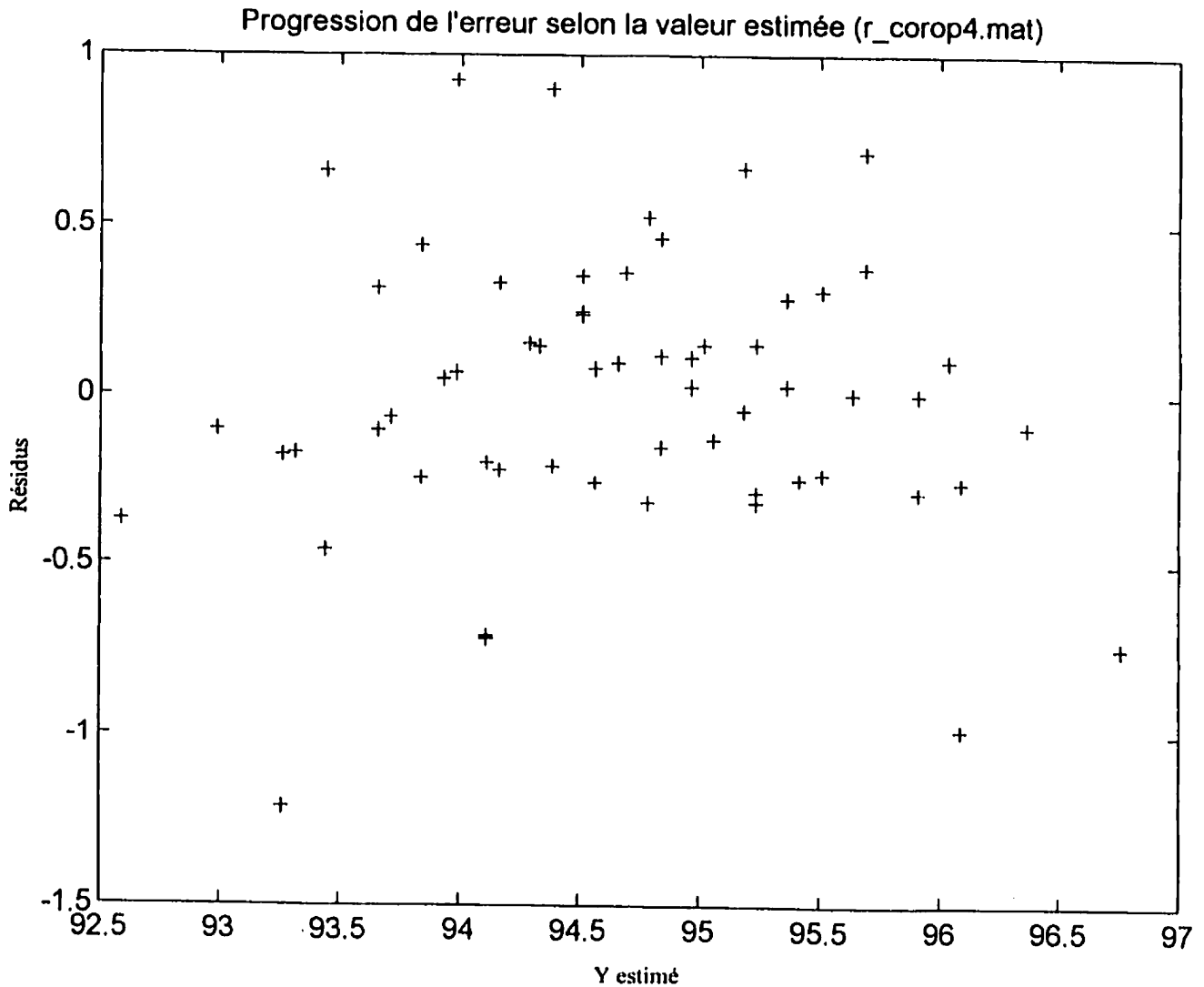
Modèle : Linéaire

Note : Modèle final

```
» load r_corop4
» B

B =

   94.6743
   -0.8509
   -0.6716
   -0.3623
    0.1990

» BINT

BINT =

   94.5637   94.7849
   -1.0024   -0.6995
   -0.8230   -0.5201
   -0.4729   -0.2517
    0.0884    0.3096

» N_mod

N_mod =

PTHR
PKB
HRCSP OR
FRCSP OR

» F_lof

F_lof =

    3.1683
    0.9562

» R_F

R_F =

    0.8264   65.4566         0
```


Graphique Gausso-arithmétique (r_corop4)
Probabilité
0.99
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
0.01
-1.2
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
Résidus


Test d'hypothèse sur les coefficients (r_corop4)
Valeurs des coefficients
100
80
60
40
20
0
-20
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

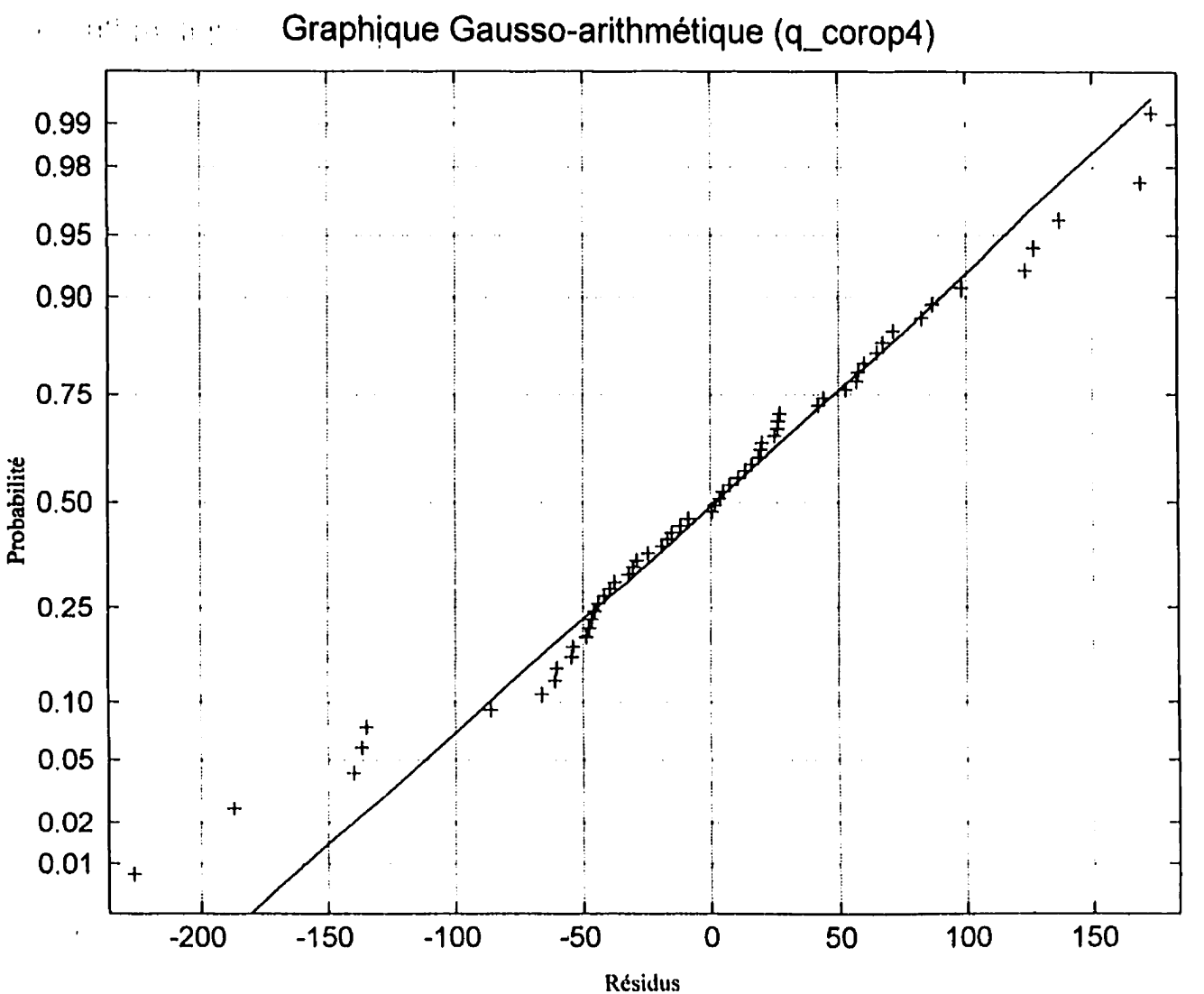

Progression de l'erreur selon la valeur estimée (r_corop4.mat)
Résidus
Y estimé
1
0.5
0
-0.5
-1
-1.5
92.5
93
93.5
94
94.5
95
95.5
96
96.5
97

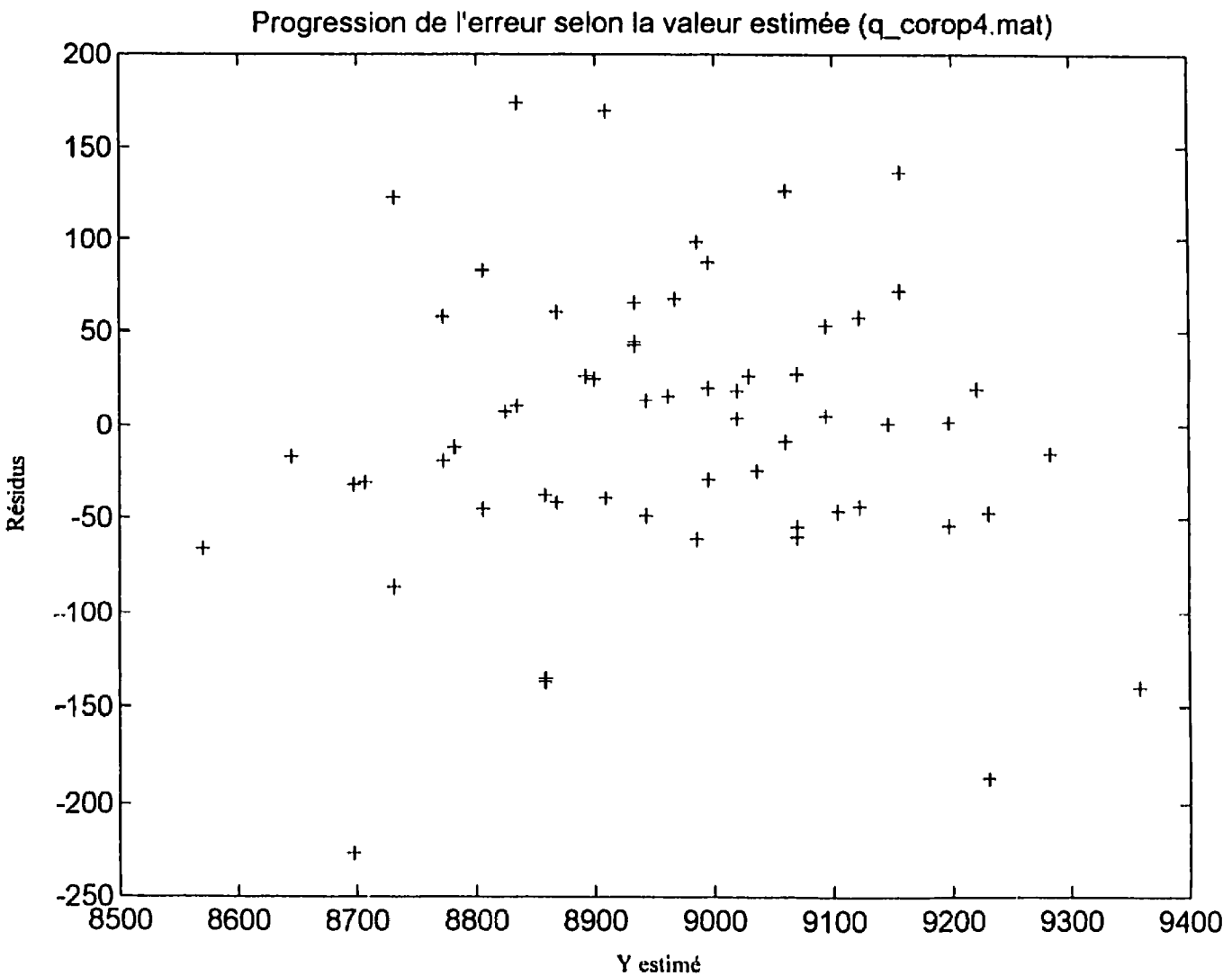
Variable : OP

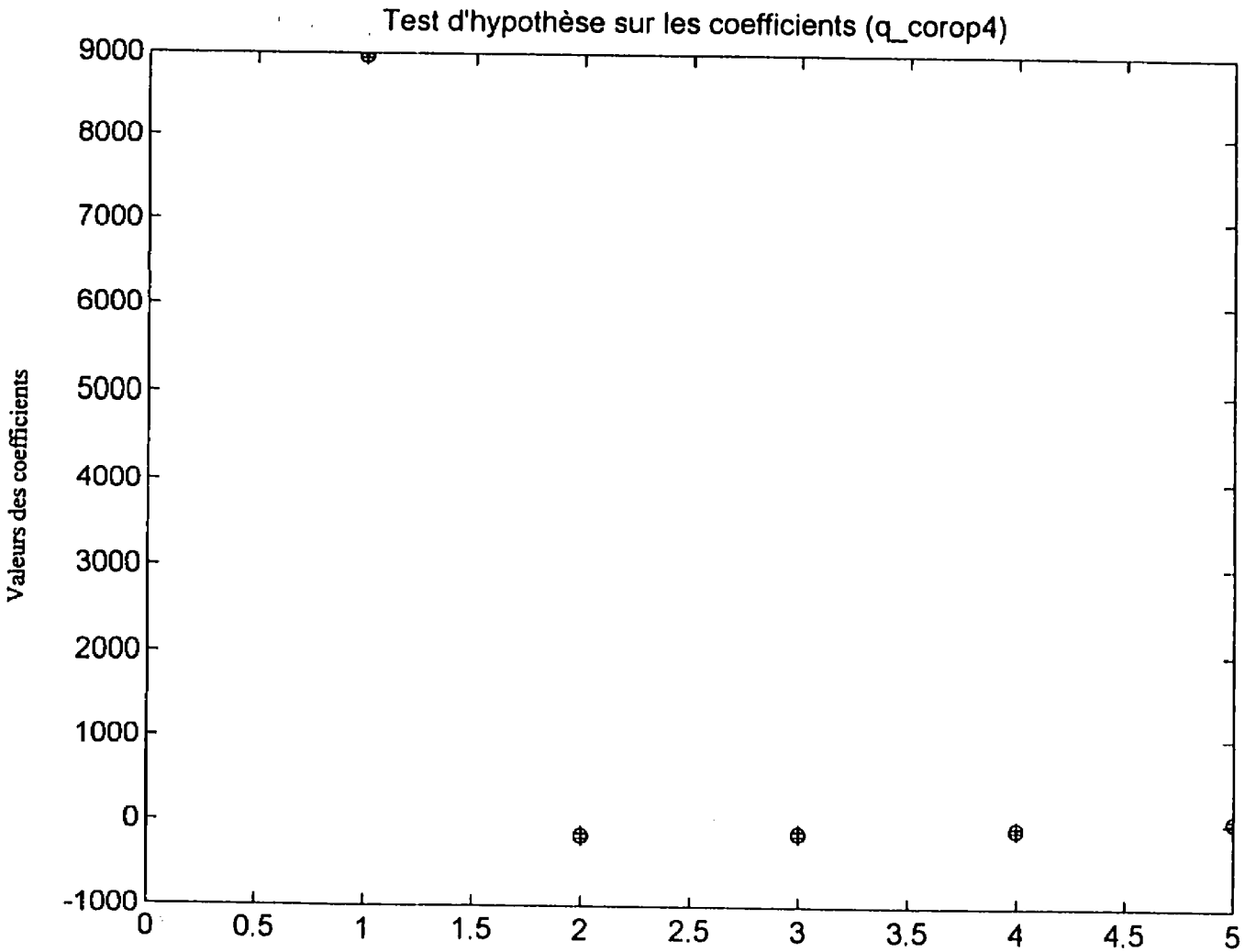
Modèle : Quadratique

Note :

```
» load q_corop4
» B

B =

  1.0e+003 *

    8.9642
   -0.1608
   -0.1270
   -0.0684
    0.0374

» BINT

BINT =

  1.0e+003 *

    8.9434    8.9850
   -0.1893   -0.1323
   -0.1555   -0.0985
   -0.0892   -0.0476
    0.0166    0.0582

» N_mod

N_mod =

PTHR
PKB
HRCSF
FRCSF

» F_lof

F_lof =

    3.1326
    0.9548

» R_F

R_F =

    0.8277   66.0394         0
```


Graphique Gausso-arithmétique (q_corop4)
Probabilité
0.99
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
0.01
-200
-150
-100
-50
0
50
100
150
Résidus


Progression de l'erreur selon la valeur estimée (q_corop4.mat)
200
150
100
50
0
-50
-100
-150
-200
-250
Résidus
8500
8600
8700
8800
8900
9000
9100
9200
9300
9400
Y estimé

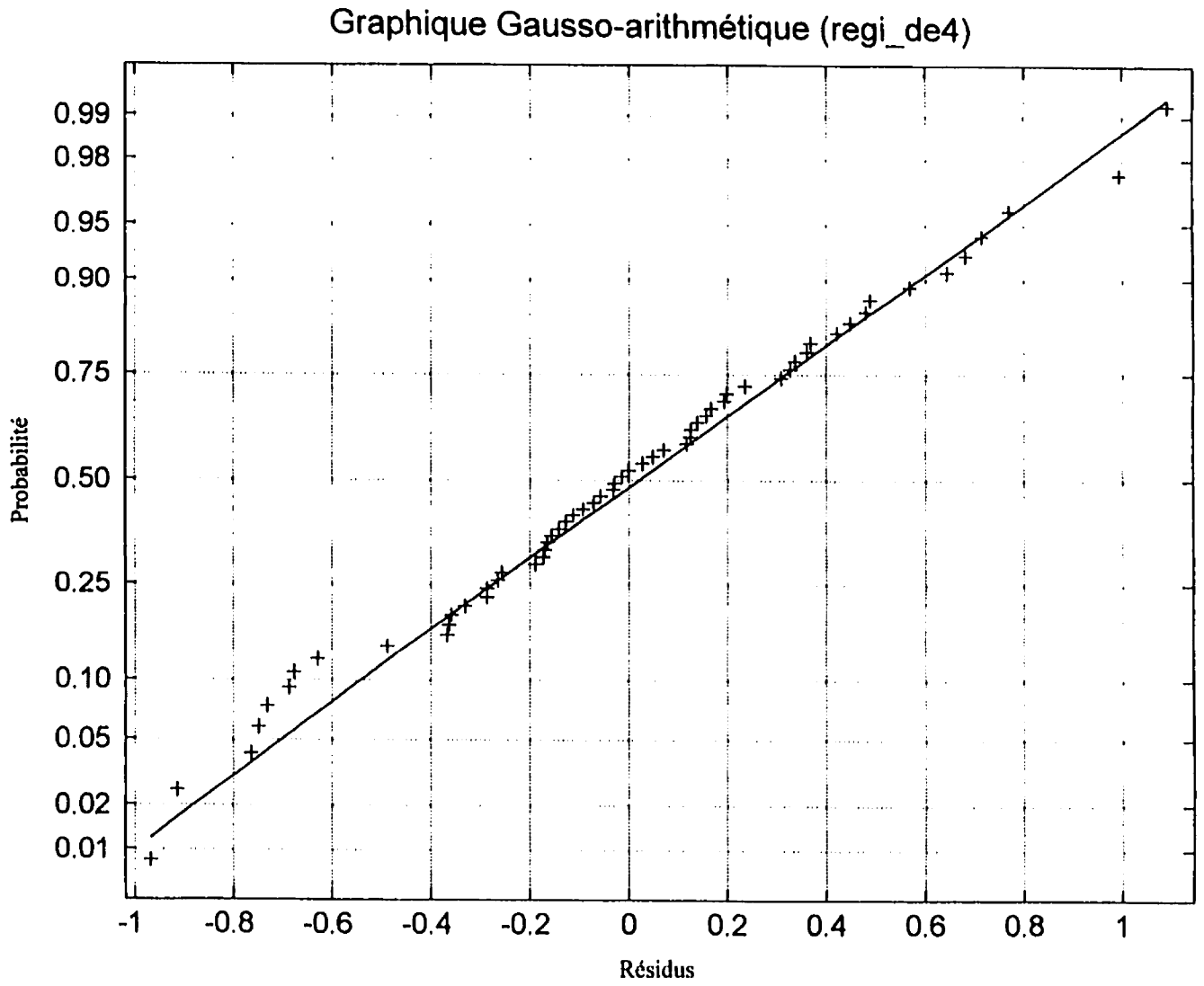

Test d'hypothèse sur les coefficients (q_corop4)
Valeurs des coefficients
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
-1000
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

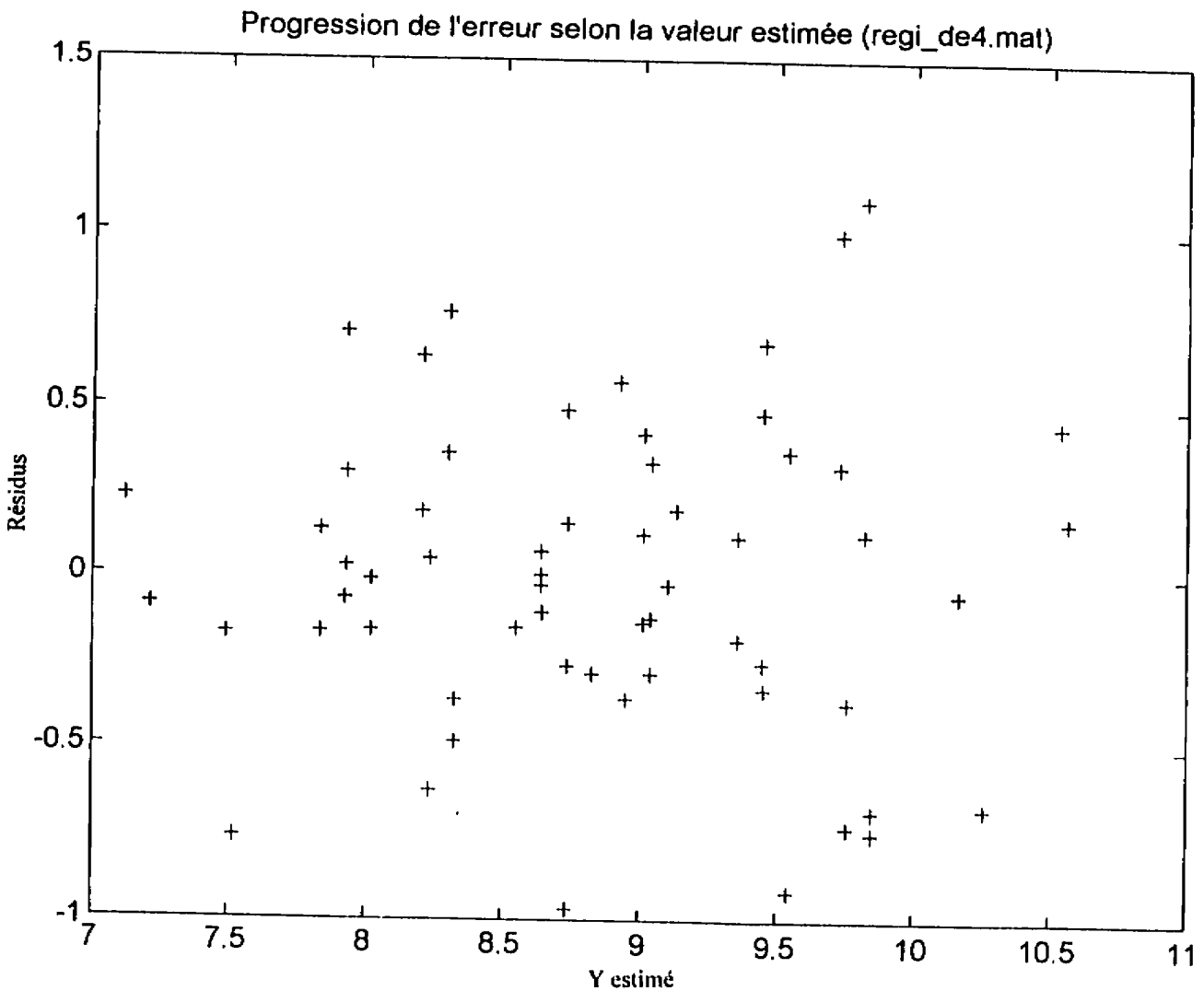
Variable : DE

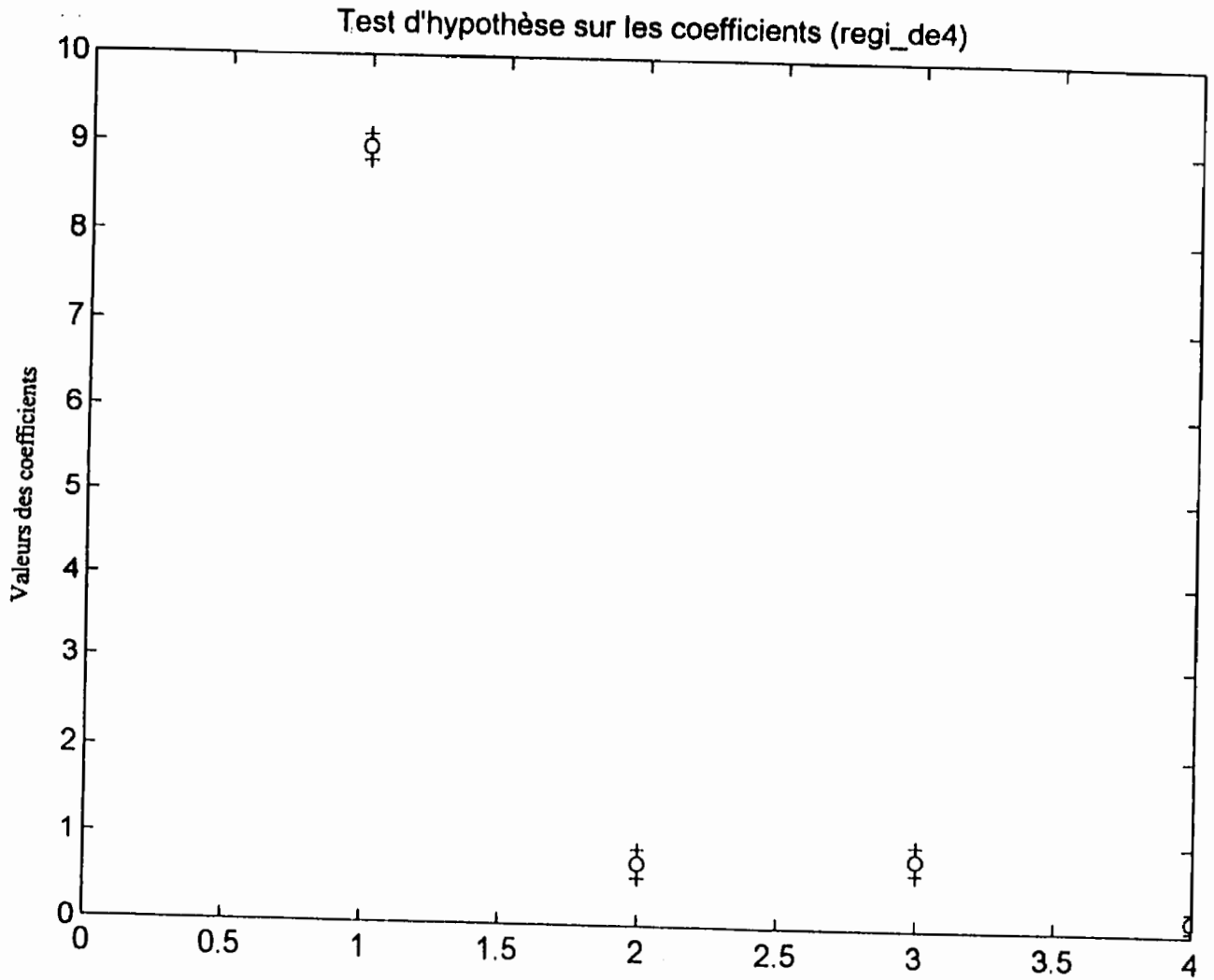
Modèle : Linéaire

Note : Modèle final

```
» load regi_de4
» B

B =

    8.9642
    0.7156
    0.8064
    0.1576

» BINT

BINT =

    8.8222    9.1062
    0.5503    0.8809
    0.6411    0.9717
    0.0475    0.2678

» N_mod

N_mod =

PTHR
PKB
x4x5

» F_lof

F_lof =

    1.8286
    0.8152

» R_F

R_F =

    0.7616   59.6398         0
```


Graphique Gausso-arithmétique (regi_de4)
Probabilité
0.99
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
0.01
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
Résidus


Progression de l'erreur selon la valeur estimée (regi_de4.mat)
Résidus
Y estimé
1.5
1
0.5
0
-0.5
-1
7
7.5
8
8.5
9
9.5
10
10.5
11

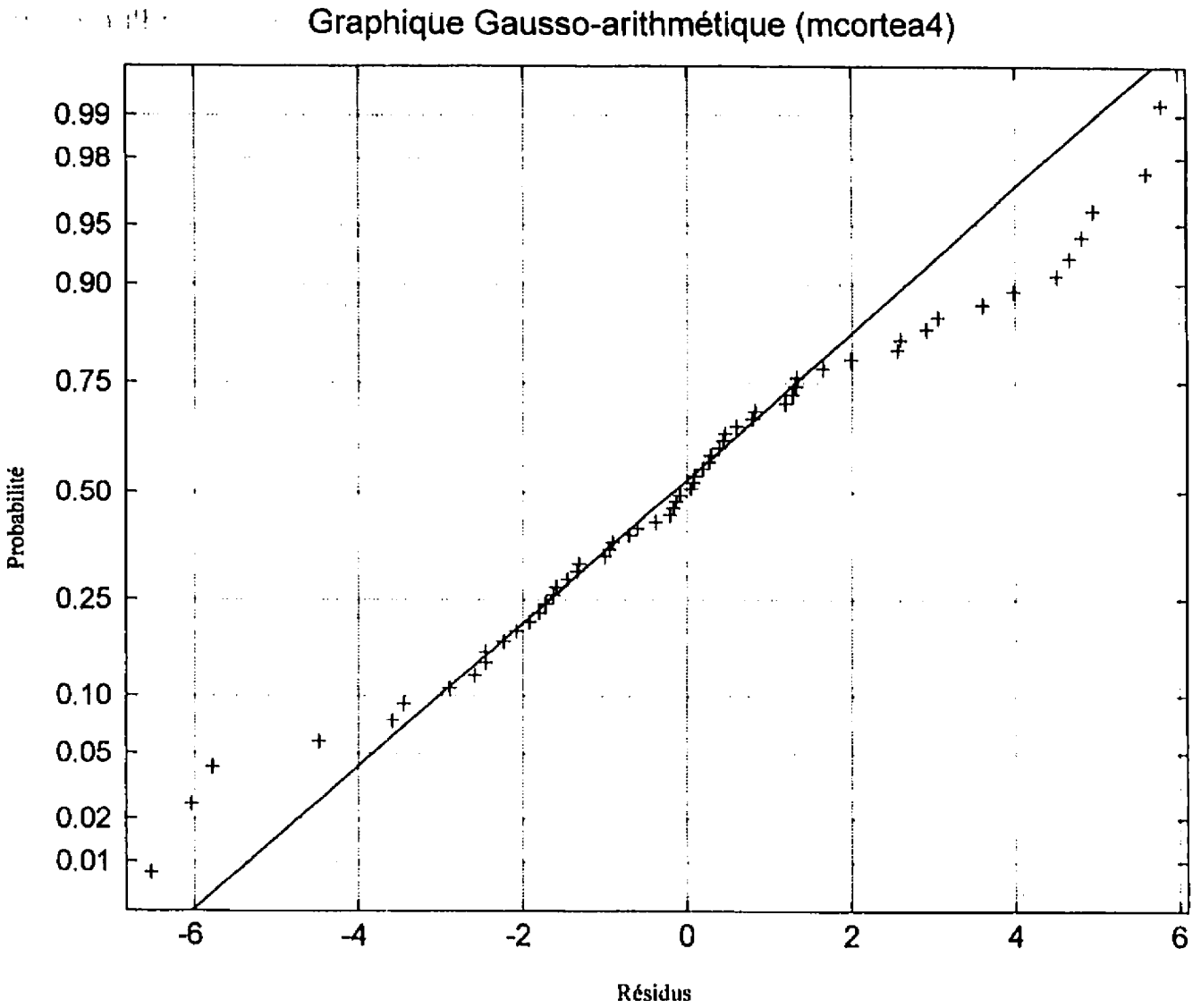

Test d'hypothèse sur les coefficients (regi_de4)
Valeurs des coefficients
10
9
8
7
6
5
4
3
2
1
0
0
0.5
1
1.5
2
2.5
3
3.5
4

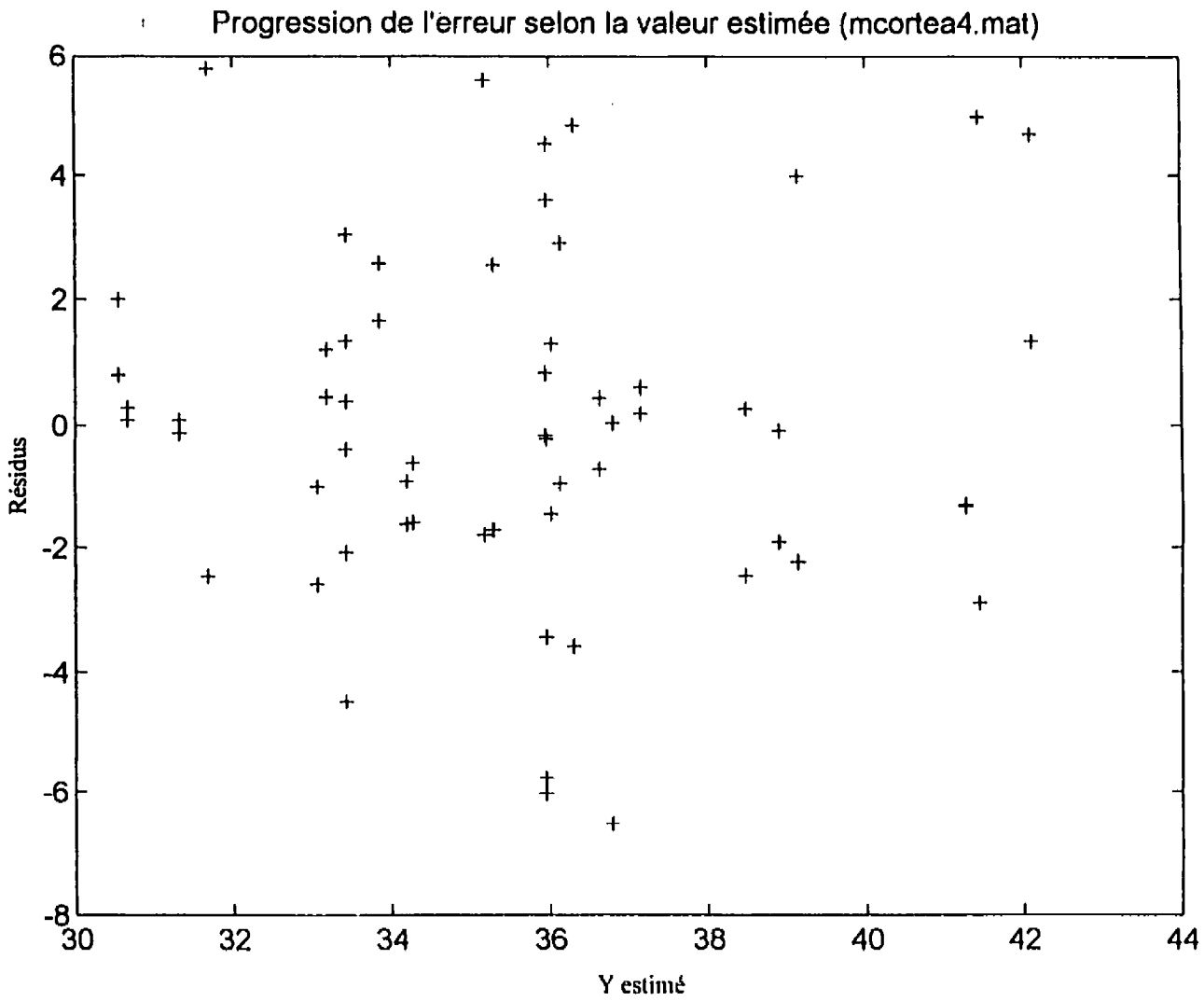
Variable : TEA

Modèle : Linéaire

Note : Modèle final

```
» load mcortea4
» B

B =

   34.6918
    1.4759
    1.7719
    2.3128
   -1.2643
    1.6773
   -1.2088

» BINT

BINT =

   33.5892   35.7944
    0.4445    2.5073
    0.7405    2.8033
    1.2814    3.3442
   -2.0176   -0.5111
    0.1675    3.1871
   -2.2401   -0.1774

» N_mod

N_mod =

PFR
PTHR
PKB
HRCOR
x1^2
x2x4

» F_lof

F_lof =

    0.7807
    0.2774

» R_F

R_F =

    0.5398   10.3611    0.0000
```


Graphique Gausso-arithmétique (mcortea4)
Probabilité
0.99
0.98
0.95
0.90
0.75
0.50
0.25
0.10
0.05
0.02
0.01
-6
-4
-2
0
2
4
6
Résidus


Progression de l'erreur selon la valeur estimée (mcortea4.mat)
Résidus
Y estimé
6
4
2
0
-2
-4
-6
-8
30
32
34
36
38
40
42
44

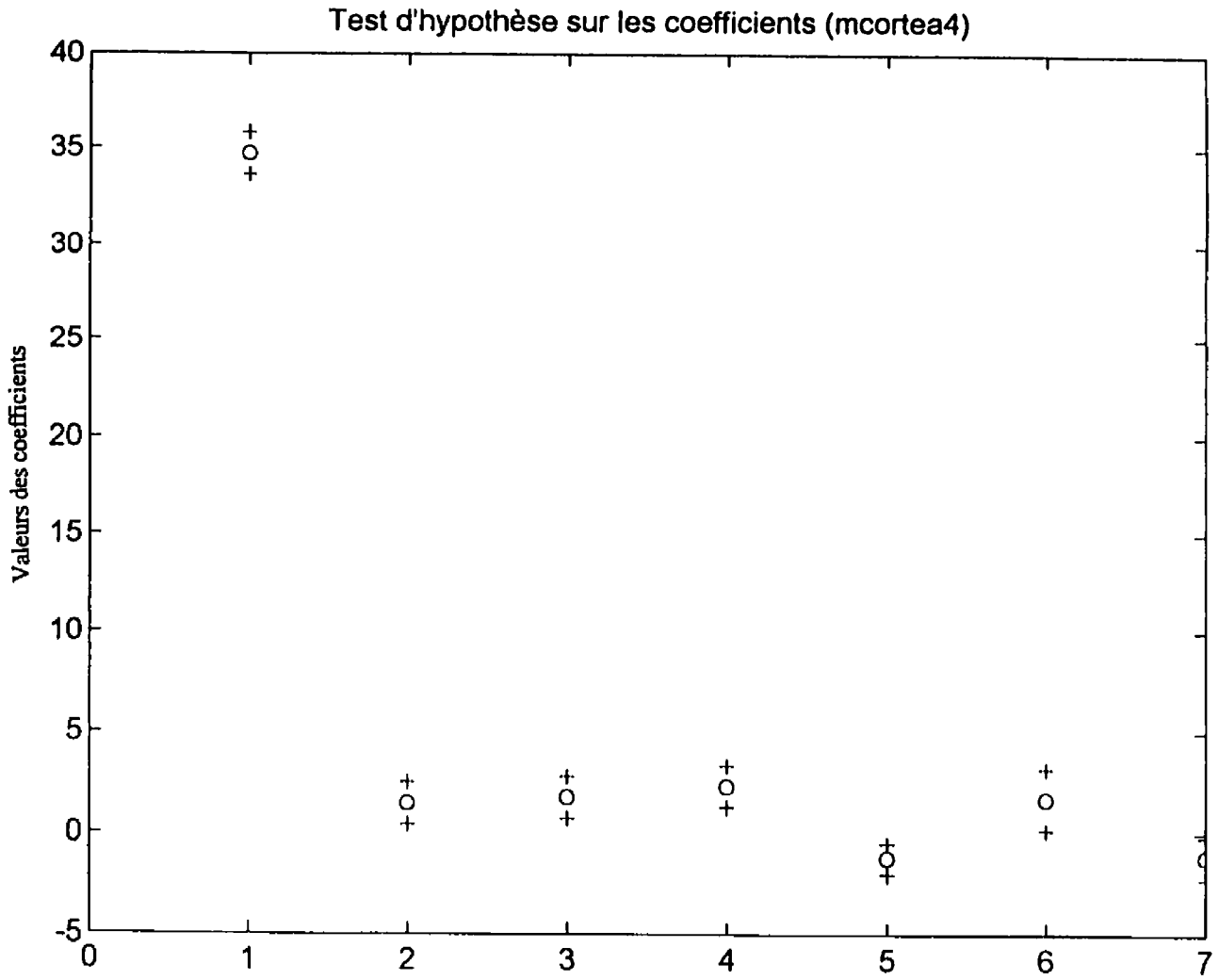
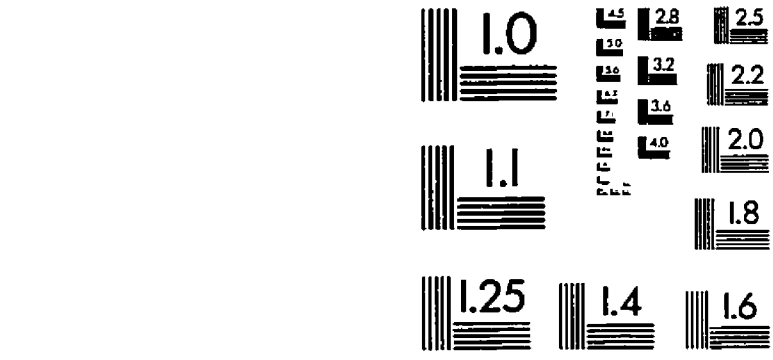
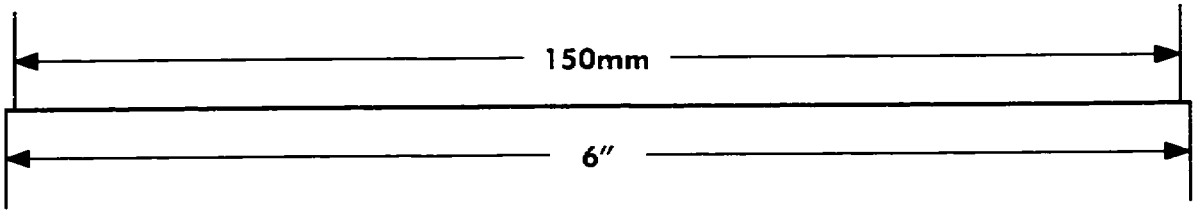

Test d'hypothèse sur les coefficients (mcortea4)
Valeurs des coefficients
40
35
30
25
20
15
10
5
0
-5
0
1
2
3
4
5
6
7

IMAGE EVALUATION TEST TARGET (QA-3)





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

© 1993, Applied Image, Inc., All Rights Reserved