

R. HAMZA

**Application et Étude de cas. Modélisation et
optimisation d'un procédé industriel pour la
clarification des jus de pommes**

*Revue française d'automatique, d'informatique et de recherche
opérationnelle. Recherche opérationnelle*, tome 24, n° 2 (1990),
p. 201-213.

http://www.numdam.org/item?id=RO_1990__24_2_201_0

© AFCET, 1990, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

MODÉLISATION ET OPTIMISATION D'UN PROCÉDÉ INDUSTRIEL POUR LA CLARIFICATION DES JUS DE POMMES (*)

par R. HAMZA (1)

Résumé. — *Ce travail présente l'élaboration d'un modèle mathématique permettant de piloter une nouvelle chaîne de clarification de jus de pommes et d'assurer une conduite optimale de la chaîne. Dans une première partie, nous déterminons les paramètres fondamentaux pour la modélisation à l'aide de techniques d'analyse des données et nous décrivons le modèle récurrent adapté à notre procédé. Nous abordons ensuite, dans une deuxième partie, l'optimisation de ce modèle qui conduit à la détermination des paramètres de commande optimaux.*

Mots clés : Clarification des jus de pommes; modèle mathématique; analyse en composantes principales; équations récurrentes; optimisation; moyennes mobiles.

Abstract. — *We propose a mathematical model allowing to pilot a new clarification chain of apple juices and ensuring the optimal working of the process. In a first part we determine the fundamental model parameters by using the data analysis techniques, and we describe the recurrent model adapted to our process. Then we study the optimal control of this model which allows to determine the optimal command parameters.*

Keywords : Clarification of apple juices; mathematical model; recurrent equations; optimization techniques.

NOTATION

t	= Temps (le pas de temps est de 20 secondes).
T_0	= Température initiale du liquide (°C).
T_1	= Température à la sortie du premier échangeur (°C).
T_2	= Température à la sortie du deuxième échangeur (°C).
V_0	= Viscosité initiale du liquide (constante).
V_1	= Viscosité à la sortie du premier échangeur (millivolt).
V_2	= Viscosité à la sortie du deuxième échangeur (millivolt).
CJ	= Concentration enzymatique dans le jus.
DL	= Débit du liquide (litres/heures).

(*) Reçu novembre 1989.

(1) Université Paris-VI, MEDIMAT, 15, rue de l'École de Médecine, 75270 Paris Cedex 06.

1. INTRODUCTION

Dans le domaine industriel, les problèmes de produits finis (amélioration de la qualité, constance de la qualité, ...) impliquent de pouvoir trouver une relation qui met en évidence le phénomène de production.

Néanmoins la complexité de cette relation peut être telle qu'on peut préférer à une description de détail, une interprétation statistique globale. Mais une telle recherche devra faire intervenir les paramètres nécessaires et suffisants pour cette interprétation.

Dans ce travail, nous montrerons comment, moyennant des montages expérimentaux un peu particuliers, il est possible d'obtenir une commande optimale (continue et contrôlable) pour éliminer les macromolécules de pectine qui empêchent toute filtration du jus brut.

2. IDÉES DE BASE ET HYPOTHÈSES DE DÉPART DES INDUSTRIELS

Comme le processus est basé sur une réaction enzymatique on peut lui associer l'équation de Michaelis-Menten [5, 9].



E = Enzyme libre; $[ES]$ = Complexe enzyme substrat; R = Rapide; S = Substrat; P = Produit de la réaction; L = Lente.

Dans cette équation et d'après des résultats biochimiques cités en [8] la première partie de la réaction est plus rapide que la deuxième, d'où l'idée des industriels de n'utiliser que la première partie de la réaction. Dans notre problème l'objectif est d'éliminer les molécules de pectine les plus longues.

Il faut d'abord enlever ces molécules en agissant sur la température. Or ces molécules sont de formes variables et il arrive que le comportement des deux molécules (celle de l'enzyme et de pectine) soit différent.

Il y a donc nécessité de piloter la réaction. Le pilotage peut être effectué en deux étapes :

(1) Amener la pectine dans son état de réceptivité maximale à l'égard des enzymes.

(2) Ajouter alors les enzymes et continuer à chauffer en surveillant la filtration.

Les deux hypothèses sur lesquelles se basent les industriels sont les suivantes :

(a) La vitesse de formation du complexe $[ES]$ suit la loi d'Arrhénius [12].

$V = C \cdot E \exp(-A/RT)$; A , R et C = constantes; E = concentration en enzymes; T = température absolue (°K),
d'où des vitesses pouvant devenir très importantes quand la température augmente.

(b) L'activité enzymatique décroît en fonction du temps pour une température de chauffage donnée suivant la loi exponentielle.

$E = E_0 \exp(-T/T_0)$; T = température de chauffage; T_0 = constante de temps; E_0 = concentration initiale en enzymes.

3. EXPÉRIMENTATION

Le pilote expérimental a été conçu selon le schéma décrit dans la figure 1. Le chauffage et le préchauffage sont réalisés grâce à deux échangeurs à plaques. L'injection d'enzyme en continu est obtenue par une pompe doseuse située en amont d'un petit mélangeur statique.

Les mesures sont effectuées en continu à partir des capteurs de température, de débit, de viscosité reliés à un micro-ordinateur (Apple IIe).

Les mesures sont stockées toutes les 20 secondes. Les vannes de vapeur et la pompe doseuse sont pilotées directement par un micro-ordinateur.

4. ANALYSES STATISTIQUES

4.1. Analyse en composantes principales

N'ayant ni modèle préalable ni idée sur les conditions expérimentales optimales nous avons effectué sur l'ensemble des données acquises une série d'analyses en composantes principales (A.C.P.) [3].

On a d'abord procédé sur chaque essai à une A.C.P. sur l'ensemble des observations et des variables soit T_0 , T_1 , T_2 , V_1 , V_2 , DL et CJ . L'examen des résultats nous a permis de faire plusieurs remarques :

Le cercle de corrélation (fig. 2) montre que DL et CJ sont parfaitement anticorrélés (ceci traduit le fait que les industriels font en sorte que le produit $DL \times CJ$ soit constant) et que V_1 s'oppose à T_1 et T_2 sur l'axe 2.

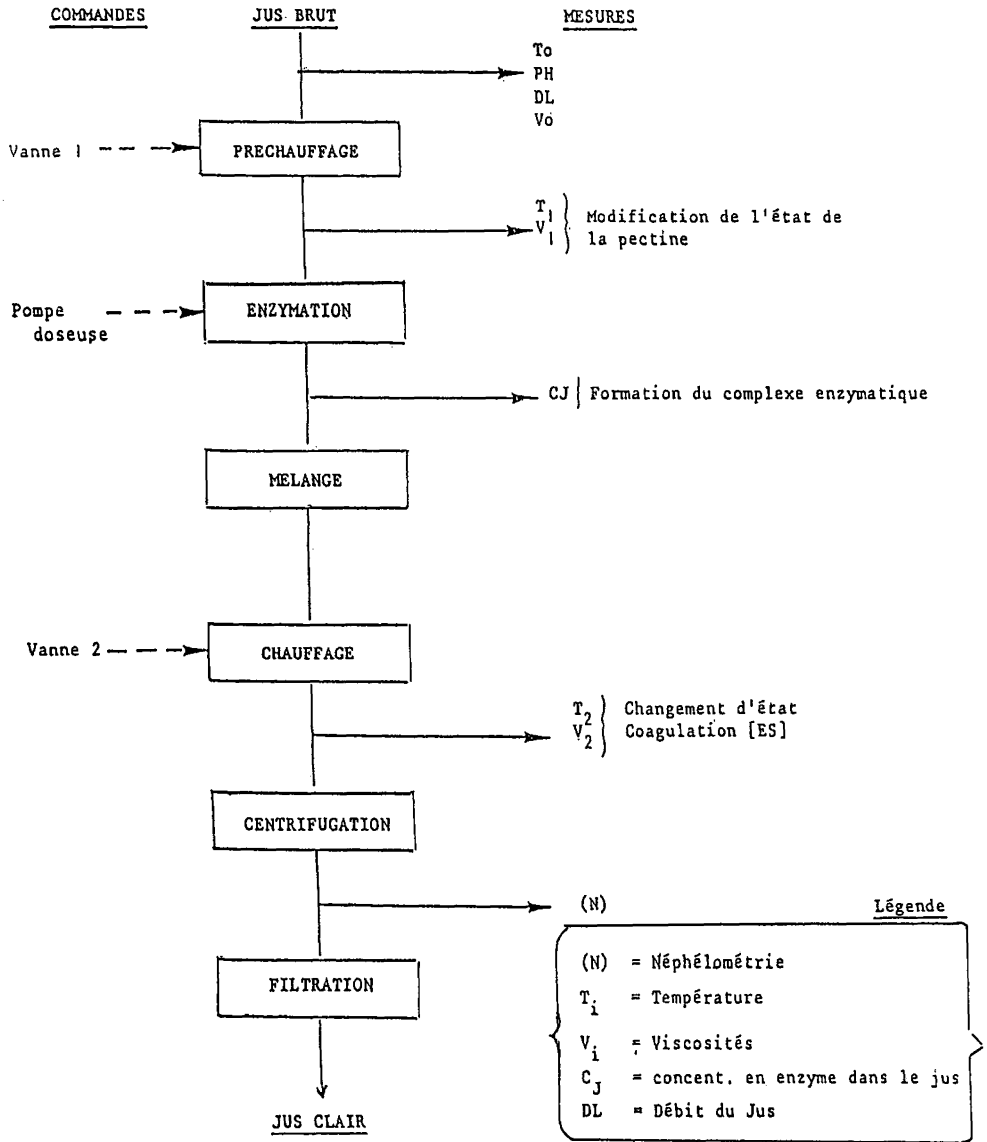


Figure 1.

CERCLE DES CORRELATIONS

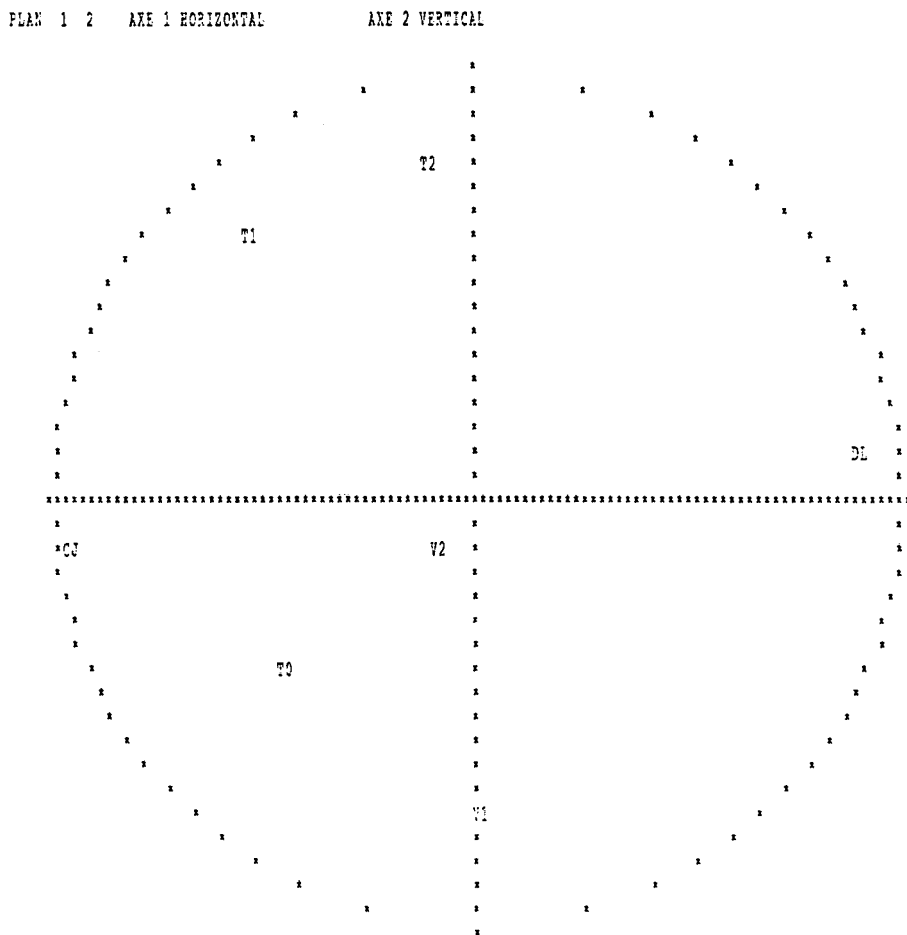


Figure 2.

Le plan principal (*fig. 3*) met en évidence deux nuages de points très compacts qui correspondent respectivement aux points où le jus est clair et aux points où le jus est trouble.

CJ et *DL* étant parfaitement anticorrélés il suffit de garder l'une de ces deux variables, de plus T_0 et *PH*, étant pratiquement constants, ont été supprimés de l'analyse. Une A.C.P. réalisée avec l'ensemble des variables

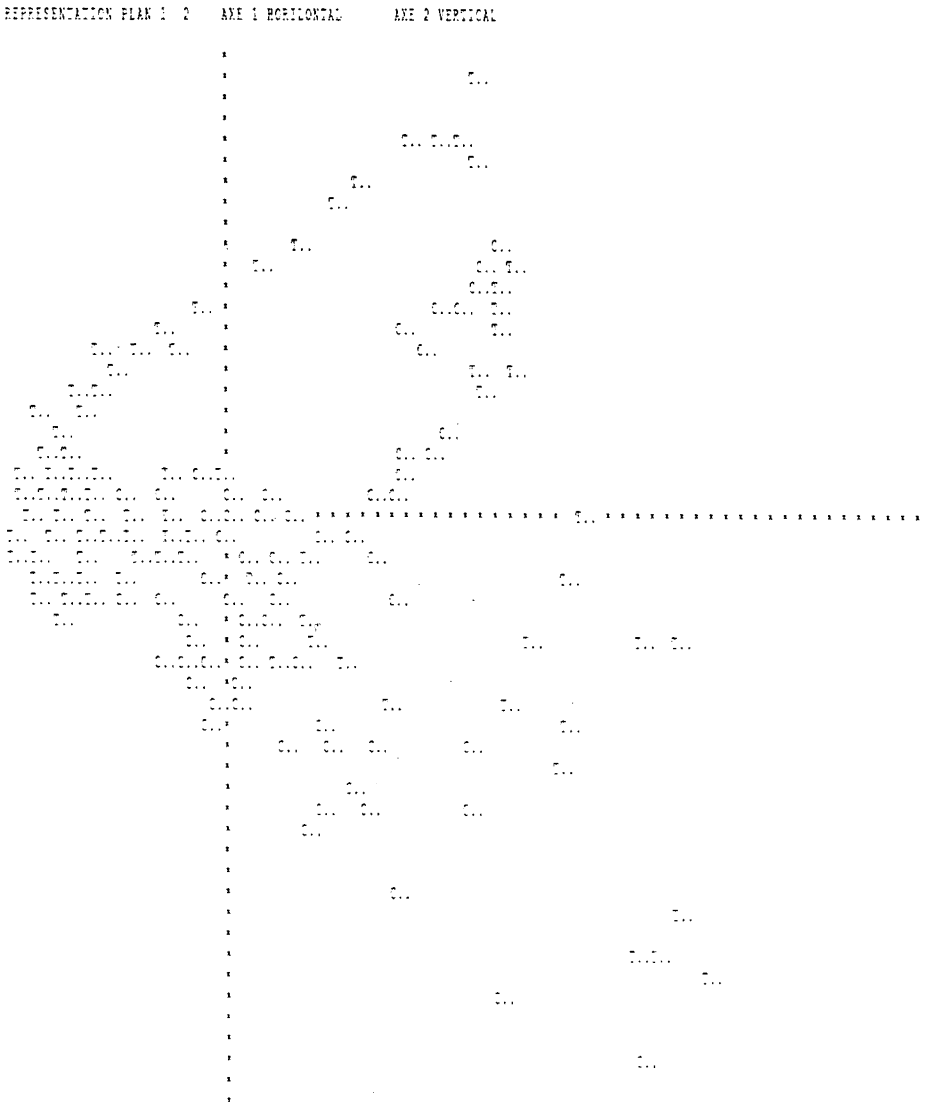
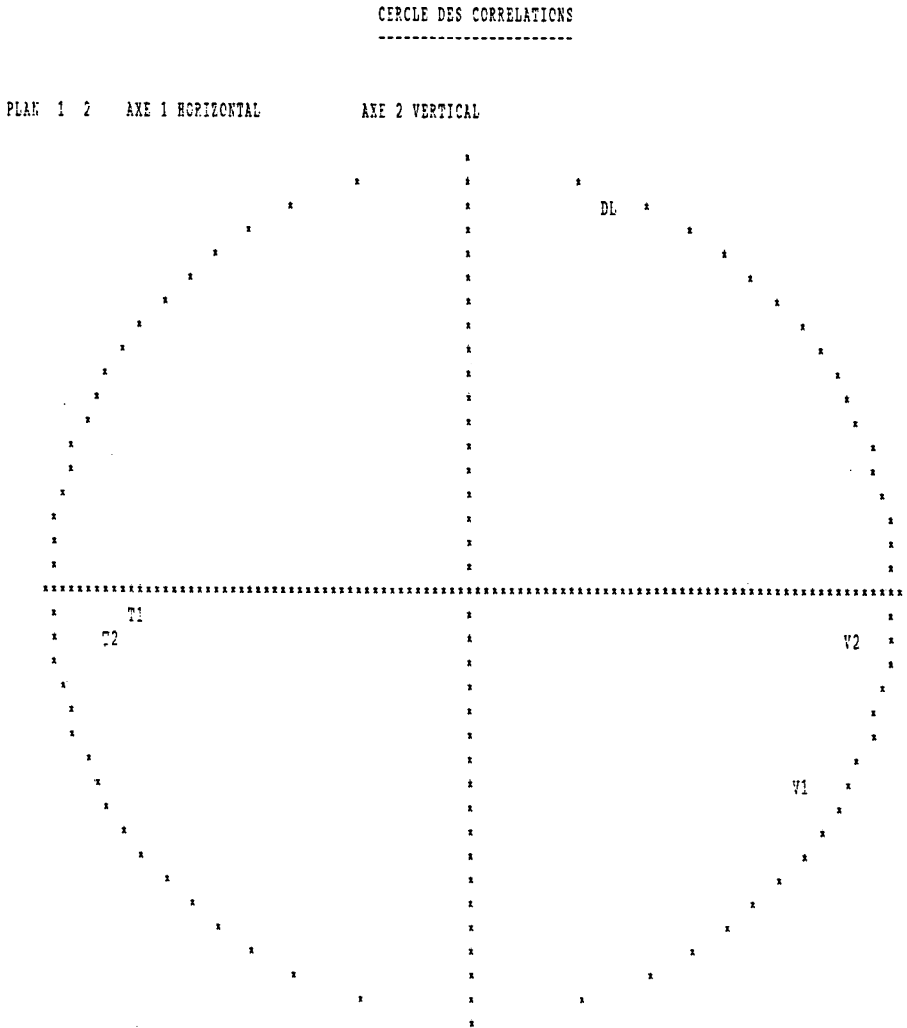
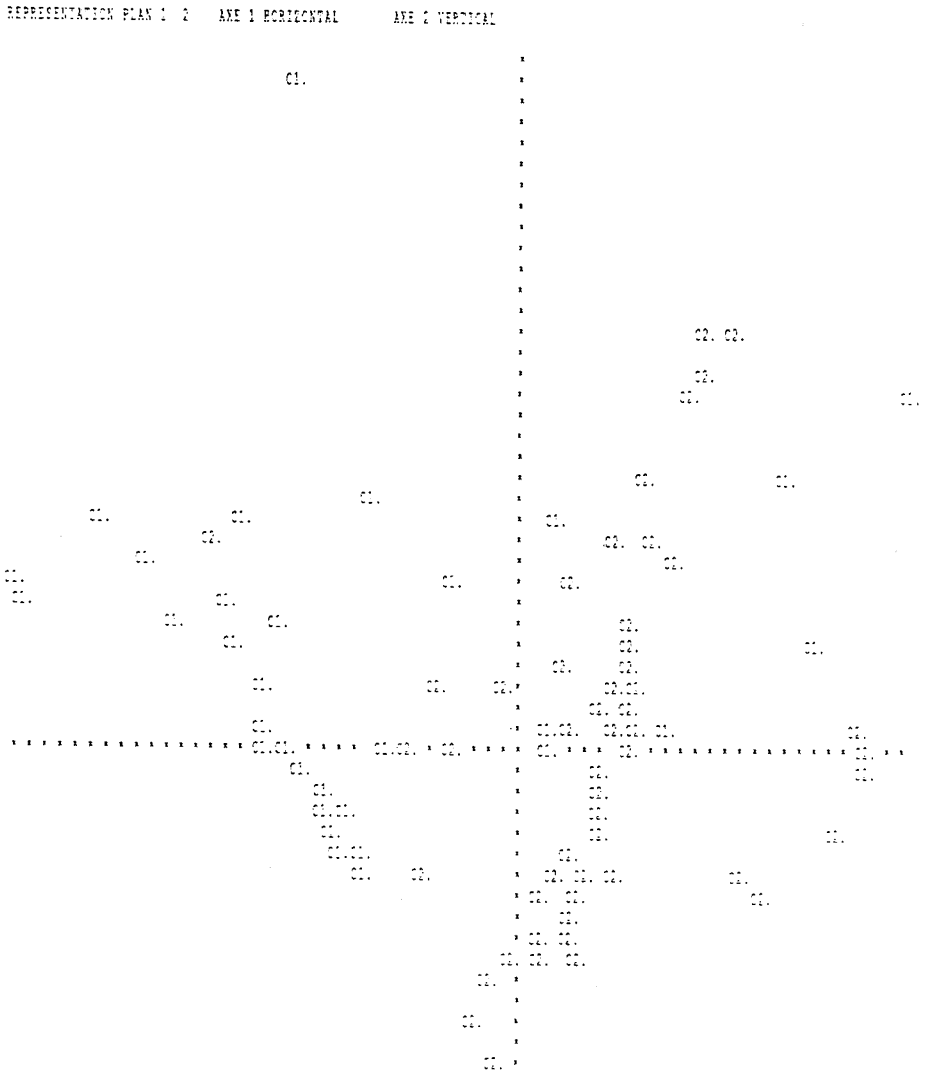


Figure 3. — C, points où le jus est clair;
T, points où le jus est trouble.

restantes (T_1 , T_2 , V_1 , V_2 , DL) montre d'ailleurs que l'on retrouve les mêmes résultats que précédemment. Il a semblé intéressant également de faire une A.C.P. sur l'ensemble des points où le jus est clair uniquement. Le cercle de



corrélations (fig. 4) permet de constater que T_1 et T_2 sont corrélées ainsi que V_1 et V_2 et que ces deux groupes s'opposent sur l'axe 1. Le plan principal



**Figure 5. — C_1 , points où le jus est clair avec $PH=3,2$;
 C_2 , points où le jus est clair avec $PH=3,65$.**

(fig. 5) est très intéressant car on y remarque deux groupes distincts de points C_1 et C_2 correspondant chacun à une des deux valeurs prises par le PH .

L'A.C.P. a été réalisée à l'aide du logiciel STAT.ITCF.

4.2. Programme STAT.TV

Le programme STAT.TV, que nous avons mis au point, prépare l'obtention des développements récurrents en étendant la méthode des moyennes mobiles [4], [13]. STAT.TV considère les séries chronologiques comme des processus stochastiques de loi inconnue. Or, d'après un théorème de Paul-Levy [11], toute variable aléatoire peut être définie par ses moments à condition qu'elle possède une fonction caractéristique. D'autre part, il existe des estimateurs sans biais pour l'espérance mathématique, la variance et le coefficient de corrélation linéaire, même quand la distribution ou la densité de probabilité est inconnue.

Ces estimateurs sont d'usage courant ce sont :

- pour l'espérance mathématique la moyenne E des valeurs au besoin pondérées;
- pour la variance

$$V = \sum_{i=1}^n (y_i - E y)^2 / (n - 1);$$

- pour la corrélation linéaire

$$R = \sum_{i=1}^n [(y_i - E y)(x_i - E_x)] / S(y) S(x) / n;$$

avec

y_i = les valeurs connues de la variable aléatoire Y ;

x_i = les valeurs connues de la variable aléatoire X ;

$S(X)$ = l'écart-type de la variable X ;

$S(Y)$ = l'écart-type de la variable Y .

Cette expression s'étendant aux coefficients d'autocorrélation, d'après Wannacott et Wannacott [13], qui assimilent les relations de récurrence linéaire à des régressions. Ces expressions peuvent être pondérées. Il suffit alors de les appliquer aux groupements élémentaires de taille n utilisés pour les moyennes mobiles. On obtiendra des suites de valeurs fonction du temps. Il y a pourtant une précaution à prendre, elle est indiquée par Alex Rosengard [11], le test de Bartlett, qui sert à comparer plusieurs populations par leurs variances, est valide seulement dès que chacun des effectifs est supérieur ou égal à 6.

C'est pourquoi STAT.TV travaille avec des groupements élémentaires de taille 6.

En possession des suites des valeurs dans le temps :

- fonctions d'espérance mathématique qui décrivent la tendance du processus;
- fonctions de variance et d'écart-type qui nous donnent la suite des valeurs de fluctuation;
- fonctions d'autocorrélation qui permettent de voir dans quelle mesure l'état actuel du processus dépend de son passé immédiat;
- fonctions de corrélation qui permettent de savoir si les processus X et Y sont associés, dans quels intervalles ils le sont et dans quelle mesure.

Ainsi en examinant ces fonctions on aura un outil probabiliste qui permettra d'écrire les meilleures relations de récurrence possibles, simples et multiples.

5. INTERPRÉTATION DES RÉSULTATS

Les résultats nous ont permis de dégager la période de réaction, elle dure 320 secondes. Cette réaction s'est produite dans la partie convexe de $T_2(t)$ et la partie concave de $V_2(t)$.

Les valeurs d'autocorrélations mobiles de T_1 et T_2 sont toutes positives et considérables mais oscillantes ce qui prouve l'existence d'une relation récurrente dont l'ordre est au moins égal à deux.

Les valeurs des corrélations mobiles montrent que V_1 et T_1 sont associées ainsi que V_2 et T_2 , les tableaux d'autocorrélations dont les valeurs sont positives, importantes et oscillantes montrent qu'on peut trouver une relation de récurrence reliant V_1 et T_1 ainsi que V_2 à T_2 et V_1 .

Finalement cette analyse a suggéré les relations récurrentes entre les diverses grandeurs, récurrentes autonomes pour $T_1(t)$ et $T_2(t)$, récurrente $V_1(T_1, V_1)$ et enfin récurrente $V_2(T_2, V_1, V_2)$. Autrement dit l'analyse a permis la détermination de la structure du modèle.

6. MODÉLISATION

6.1. Un premier modèle

L'utilisation du matériel informatique pour réaliser l'automatisation du procédé impose de travailler avec des valeurs discrètes des différents signaux. Pour cela on utilise des modèles récurrents [7]. Ce choix a été confirmé par l'analyse STAT.TV.

On s'intéresse surtout aux données pour lesquelles la réaction s'est produite. Les équations du modèle ont les formes suivantes :

$$\begin{aligned}
 T_1(t+2) &= a_0 + a_1 T_1(t+1) + a_2 T_1(t) \\
 V_1(t+1) &= b_0 V_1(t) + b_1 T_1(t+1) + b_2 T_1(t) \\
 T_2(t+2) &= c_0 + c_1 T_2(t+1) + c_2 T_2(t) \\
 V_2(t+1) &= d_0 + d_1 V_2(t) + d_2 T_2(t+1) \\
 &\quad + d_3 T_2(t) + d_4 V_1(t+1) + d_5 V_1(t).
 \end{aligned} \tag{1}$$

Les coefficients ont été identifiés par la méthode linéaire de moindres carrés avec des erreurs relatives quadratiques moyennes qui ne dépassent pas 6 %.

6.2. Un second modèle

Ce modèle montre l'effet de la température sur la viscosité dans les deux échangeurs pour un jus clair. En effet c'est le modèle de Flucher [10] qui a la forme suivante :

$$\begin{aligned}
 \text{Ln}(V_1) &= A + \frac{B}{T_1 - T_0} \\
 \text{Ln}(V_2) &= C + \frac{D}{T_2 - T_1}.
 \end{aligned}$$

Les coefficients A , B , C et D ont été identifiés par la méthode de moindres carrés linéaires avec des erreurs relatives quadratiques moyennes qui ne dépassent pas 4 %.

7. OPTIMISATION

On remarque dans le premier modèle que :

T_1 est définie par $T_1(0)$, $T_1(1)$;

V_1 est définie par $T_1(0)$, $T_1(1)$ donc $V_1(1)$ est alors connu;

T_2 est définie par $T_2(0)$, $T_2(1)$;

V_2 est définie par $V_2(0)$, $T_2(0)$, $T_2(1)$, $V_1(0)$, $V_1(1)$.

On considère la fonctionnelle $J = \sum_{t=1}^n V_2(t)$ qu'on minimisera par rapport à $V_2(0)$, $T_1(0)$, $T_1(1)$, $T_2(0)$, $T_2(1)$, $V_1(0)$, au temps de déclenchement de la réaction.

Pour cela on a utilisé la méthode déterministe d'optimisation globale Aliénor [1, 2].

Les résultats sont les suivants :

$$\begin{aligned} T_1(0) &= 33,523 & T_2(1) &= 65,639 \\ T_1(1) &= 33,958 & V_2(0) &= 387,766 \\ T_2(0) &= 65,217 & V_1(0) &= 346,598 & t_d &= 6. \end{aligned}$$

La réaction s'est donc déclenchée au bout de $6 \times 20 = 120$ secondes.

Remarque : On peut aussi utiliser la méthode du simplexe [6] puisqu'on part d'un système linéaire : On peut écrire.

$$J = \sum_{t=1}^n \{V_2(t) + O \cdot V_1(t) + O \cdot T_1(t) + O \cdot T_2(t)\}$$

avec pour contraintes les équations linéaires de (I) et les autres du type $X_{\min} < X < X_{\max}$.

Autrement dit le problème se formule sous forme d'un programme linéaire de $4 \times n$ variables (n = nombre de pas de temps).

Cette méthode a un inconvénient majeur, le programme obtenu va être très gros et ne se résoud pas sur n'importe quelle machine. Un microcalculateur sera insuffisant.

8. CONCLUSION

L'application du programme STAT.TV a permis de dégager une bonne solution du problème industriel posé. Elle a permis aussi de modéliser le problème en indiquant les équations récurrentes qu'il fallait chercher et leurs domaines d'application.

De plus on a pu calculer les températures T_1 et T_2 pour obtenir le minimum de viscosité V_2 et l'instant t_d correspondant. On a alors constaté que les fonctions obtenues étaient valides jusqu'à $2,4 t_d$: le modèle présente un coefficient de sécurité de 2,4.

Enfin, on peut noter que ce travail partant de l'analyse des données et aboutissant à l'optimisation du procédé devrait permettre des économies importantes aux industriels. Il démontre la possibilité et l'utilité d'une coopération université-industrie lorsque le souci des applications concrètes reste au premier plan.

BIBLIOGRAPHIE

1. Y. CHERRUAULT, *Mathematical Modelling in Biomedicine*, Optimal control of Bio-medical Systems, D. Reidel, Publishing Company, 1986.
2. Y. CHERRUAULT et A. GUILLEZ, *Algorithmes originaux pour la simulation numérique et l'optimisation*, Industries aliment. et agric. n° 10, octobre 1987.
3. J. M. BOUROCHE et G. SAPORTA, *L'analyse des données*, Collec. Que sais-je?, P.U.F., 1980.
4. B. GRAIS, *Méthodes Statistiques-2*, Dunod, Paris, 1974.
5. J. P. KERNEVEZ, *Enzyme Mathematics*, North-Holland, 1980.
6. M. MINOUX, *Programmation mathématique; théorie et algorithmes*, tome 1, Dunod, Paris, 1983.
7. K. NAJIM et G. MURATET, *Pratique de la régulation numérique des processus industriels*, Masson, Paris, 1983.
8. A. L. LEHNINGER, *Biochimie*, Seconde édition, Flammarion médecine-sciences, 1977, p. 185-195.
9. Lee A. SEGEL, *Mathematical Models in Molecular and Cellular Biology*, Cambridge University Press, 1980.
10. M. A. RAO, R. S. SHALLENBERGER et H. J. COOLEY, *Effect of Temperature on Viscosity of Fluids with high Sugar Content*, 1986.
11. A. ROSENGARD, *Probabilités et Statistiques dans la recherche scientifique*, Dunod et C.E.A., 1971.
12. G. D. SARVACOS, *Effect of Temperature on Viscosity of Fruit Juices and Purees*, Journal of food sciences, vol. 35, 1970, p. 122-125.
13. R. WANNACOTT et F. WANNACOTT, *Statistiques*, Economica, 1988.