

C. BOULIER

G. COCHART

**Les problèmes de prévisions pour des
produits liés aux loisirs**

*Revue française d'automatique, d'informatique et de recherche
opérationnelle. Recherche opérationnelle*, tome 4, n° V3 (1970),
p. 15-29.

http://www.numdam.org/item?id=RO_1970__4_3_15_0

© AFCET, 1970, 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/>

LES PROBLEMES DE PREVISIONS POUR DES PRODUITS LIES AUX LOISIRS

par C. BOULIER et G. COCHART (1)

Résumé. — *Les modèles de prévisions, développés et expérimentés ci-après touchent une gamme de produits directement liés aux loisirs.*

Une des originalités d'un des modèles repose dans le fait que nous avons essayé d'assimiler « des pointes mobiles » correspondant aux fêtes (Pâques, Pentecôte, etc...) qui sont elles-mêmes mobiles, à des impulsions qui s'amortissent dans le temps.

L'estimation des différents paramètres a été faite par la technique statistique proprement dite.

INTRODUCTION

Les séries chronologiques sont plus ou moins faciles à prévoir. Dans ce cadre, nous allons présenter une méthode de prévisions relatives à des réceptions hebdomadaires de produits, en l'occurrence des bobines, connaissant la prévision annuelle de ces dernières.

Les prévisions de traitement se font indépendamment des ventes, même avec un certain décalage, on ne peut déduire l'un de l'autre car :

a) Il y a absence de prévisions hebdomadaires au niveau des ventes de ces produits.

b) Et même si de telles prévisions existaient, des études statistiques ont montré qu'elles seraient inutilisables. En effet, il se produit des stocks chez les différents agents intermédiaires (grossistes, revendeurs, clients, amateurs, ...).

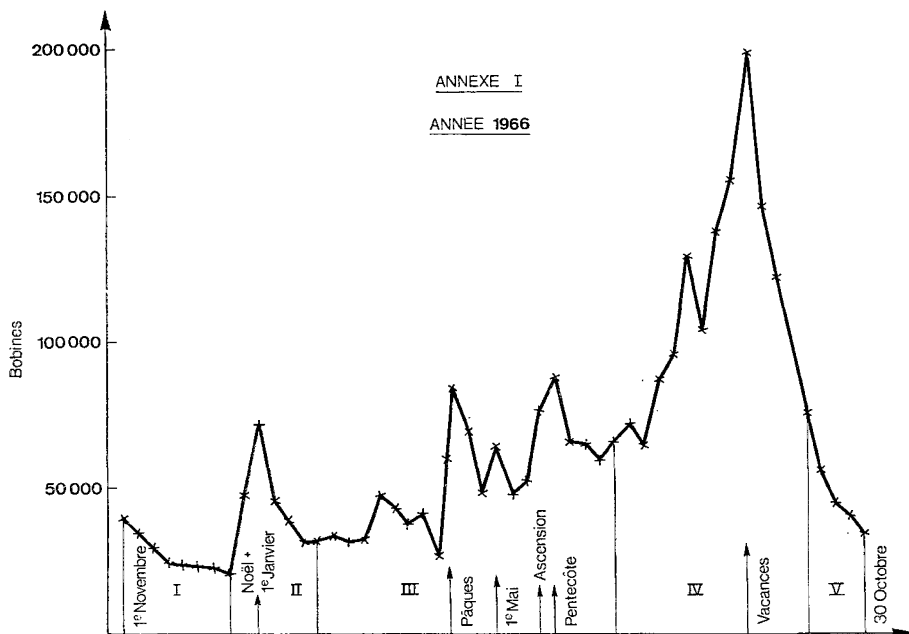
Ceci a été appliqué au sein d'une grande unité de production pour laquelle la connaissance de la prévision était déterminante pour les besoins en personnels et matériels.

(1) Groupe de Recherche Opérationnelle Kodak-Pathé.

1. Caractéristiques des réceptions

Nous avons à faire ici à une gamme de produits dont l'utilisation est intimement liée à la politique des loisirs ; plus particulièrement des vacances, et par voie de conséquence, directement liée à la clientèle.

Dès lors, examinons une courbe annuelle représentant en réel les réceptions des bobines au cours de l'année 1966 (annexe I). Nous avons une courbe saisonnière très particulière.



Nous remarquons :

- une période relativement calme et basse en amplitude en novembre, décembre, février, mars ;
- une période très haute avec une pointe très accentuée au début de septembre ;
- des pointes aux périodes de fêtes comme Noël, Pâques, Pentecôte, etc...

Par suite, ces points soulignent deux choses :

- Il existe une différence énorme d'amplitude entre la semaine la plus haute qui se situe à la première semaine de septembre et la semaine la plus basse en décembre : le rapport est de 1 à 10.
- Nous avons des pointes accentuées aux fêtes mobiles.

En conséquence, seules des prévisions, au niveau hebdomadaire, auront un sens pratique.

Les prévisions mensuelles se déduiront immédiatement des précédentes.

2. Méthode utilisée

Pour réaliser au mieux cet objectif, il a fallu embrasser un historique assez riche en informations et utiliser la statistique.

a) Détermination des périodes

Nous avons dans un premier temps, découpé l'année en périodes homogènes qui caractérisaient la courbe des réceptions. Plus précisément, nous avons découpé l'année en cinq périodes de 8, 6, 21, 13 et 4 semaines, dont la production, en pourcentage par rapport au total, reste très stable, quelles que soient les années.

b) Détermination de coefficients hebdomadaires

Dans un deuxième temps, nous nous sommes attachés à déterminer des coefficients hebdomadaires propres à chaque période.

Toutefois, une période est de beaucoup plus longue que toutes les autres : c'est la troisième. Elle comprend 21 semaines ; elle a été découpée de la première semaine de février à la dernière semaine de juin, afin qu'elle englobât entièrement le phénomène dit des fêtes mobiles comme : les vacances scolaires de printemps, Pâques, 1^{er} Mai, Ascension et Pentecôte.

Nous avons assujetti les fêtes au moment où elles se produisaient, à des impulsions qui s'amortissent rapidement dans le temps. Elles sont superposées à un mouvement de fond, l'ensemble ayant un trend exponentiel.

Pour les autres périodes, le système est moins original ; nous avons utilisé le lissage dit I.N.S.E.E.

CHAPITRE I

ETABLISSEMENT DES PERIODES

1. Caractéristiques

Quand nous observons différentes courbes annuelles de l'historique en niveau hebdomadaire pour un produit donné, nous remarquons :

- un profil annuel assez semblable d'une année sur l'autre ;
- une pointe, la plus importante, se trouve toujours à la même époque de l'année ;
- des pointes assez importantes au moment des fêtes caractéristiques : Noël, Jour de l'An, Pâques, 1^{er} Mai, Ascension, Pentecôte ;
- des semaines très mouvementées dues à des causes tout à fait diverses.

Il était donc intéressant de pouvoir découper l'année en périodes dont les principales caractéristiques étaient d'avoir le même pourcentage de production indépendamment des années et couvrant le même phénomène. Dans le découpage annuel, ces périodes grouperont des nombres de semaines constantes.

2. Origine des temps

L'exercice de la Société s'étend du 1^{er} novembre au 30 octobre de l'année suivante.

Afin d'englober parfaitement le phénomène « fêtes de fin d'année », ce dernier est intégré tout entier dans la 2^e période. Par suite, la 1^{re} période s'étendra du début de l'exercice jusqu'à la semaine précédant Noël.

Le découpage s'est effectué comme ceci :

- La 1^{re} période précédant Noël : 8 semaines : c'est une période calme,
- La 2^e période de fin d'année : 6 semaines : elle intègre les fêtes,
- La 3^e période : 21 semaines : phénomène fêtes mobiles,
- La 4^e période : 13 semaines : phénomène grandes vacances,
- La 5^e période : 4 semaines : période de décroissance.

Remarque au sujet de la 3^e période : Pourquoi 21 semaines et pourquoi fêtes mobiles ?

Nous avons voulu englober systématiquement et intégralement toutes fêtes caractéristiques de cette époque.

D'autre part, les fêtes sont :

— soit mobiles par leur date,

— soit mobiles pour leur jour à l'intérieur de la semaine.

Ainsi, le 1^{er} Mai peut être la cause d'un long week-end lorsqu'il tombe un lundi ou un vendredi, et dès lors avoir des répercussions sur les réceptions de bobines.

3. Détermination des coefficients

Reprenant l'historique et le découpant comme il vient de l'être précisé, nous avons obtenu des coefficients propres à chacune des périodes.

Connaissant la prévision annuelle, on peut donc établir une prévision au niveau de chacune des périodes par application d'un pourcentage de répartition lissé exponentiellement et rapporté à un total 100.

Soit :

$$x_t^I = \alpha \bar{x}^I + (1 - \alpha)x_{t-1}^I$$

Avec $\alpha = 0.9$

x_t^I = Prévision du pourcentage de l'année t pour la période I .

\bar{x}^I = Pourcentage réel de la période I dans l'année $t - 1$.

x_{t-1}^I = Dernière prévision connue du pourcentage de l'année $t - 1$ pour la période I .

CHAPITRE II

ETABLISSEMENT DU MODELE POUR LA TROISIEME PERIODE

La troisième période est la plus longue ; c'est celle qui est la plus instable dans le cours des années.

Les courbes représentant, par exemple, les réceptions de bobines, appellent les observations suivantes.

Nous trouvons :

- un certain mouvement de fond ;
- un certain nombre d'impulsions s'amortissant rapidement et qui correspondent aux fêtes ;
- d'autre part, l'ensemble de ces quantités semblent s'appuyer de plus sur un mouvement exponentiel ; c'est ce que nous baptiserons « le profil Produit X » caractéristique du phénomène suivant : la tendance générale pour ce produit étant une réception de plus en plus forte du mois de février aux mois d'été ;
- enfin, toute valeur à l'instant t est fonction également de tout ce qui s'est produit aux instants :

$$\begin{aligned}t - 1 \\t - 2 \\t - 3, \text{ etc...}\end{aligned}$$

C'est ainsi que nous sommes amenés à représenter une réception d'une semaine par la formule suivante :

$$B_t = [A_t + W_k \cdot f(t)] \cdot e^{\alpha t}$$

Avec les notations suivantes :

B_t : c'est la quantité de produits X reçus au cours de la semaine t .

A_t : c'est une quantité représentative de la valeur du mouvement de fond à l'instant t ,

$f(t)$: valeur du coefficient d'amortissement à l'instant t relatif au phénomène k ,

W_k : c'est la valeur de l'impulsion relative au phénomène k . Elle sera représentative d'une fête.

$e^{\alpha t}$: c'est le support exponentiel de coefficient α à l'instant t .

Nous poserons :

$$\begin{aligned} C_t &= B_t e^{-\alpha t} \\ C_t &= [A_t + W_k \cdot f(t)] \end{aligned}$$

Notons que $f(t)$ et A_t obéissent à des équations aux différences finies du 1^{er} ordre ou du 2^e ordre.

Soit pour le 1^{er} ordre :

$$\begin{aligned} A_{t+1} - aA_t &= u(t) \\ f(t+1) - bf(t) &= v(t) \end{aligned}$$

avec :

- $E[u(t)] = E[v(t)] = 0$
- Les variances de u et v indépendantes du temps.

Pour le 2^e ordre, nous avons :

$$f(t+1) - b_1f(t) - b_2f(t-1) = v(t)$$

avec :

- $E[u(t)] = E[v(t)] = 0$
- Les variances u et v indépendantes du temps.

1. Détermination de la valeur de α

Nous avons remarqué sur les réceptions annuelles, une courbure non négligeable dont la concavité est toujours tournée dans le même sens.

Nous avons donc effectué un ajustement logarithme-linéaire portant sur B_t .

$$\text{Soit : } B_t = C_t e^{\alpha t}$$

$$\text{Log } B_t = \alpha t + \text{Log } C_t$$

Les calculs sont classiques par les moindres carrés, et nous avons obtenu une estimation des α permettant l'élimination du trend.

2. Calcul des C_t

Ayant α , nous pouvons nous ramener des variables B_t aux variables C_t .

Soit le système simplifié :

$$\begin{aligned} C_t &= A_t + W_k \cdot f(t) \\ v_t &= f(t+1) - b \cdot f(t) \\ u_t &= A_{t+1} - a \cdot A_t \end{aligned}$$

avec :

- $E[u(t)] = E[v(t)] = 0$
- Les variances de u et v indépendantes du temps.

3. Régression multiple d'ordre 2

Si nous résolvons le système d'équations, pour le système d'équations aux différences finies d'ordre 1, nous avons les équations de bases sui-

vantes. La résolution doit nous permettre de déterminer les valeurs de a et b

Soit

$$\begin{aligned} C_{t+1} &= A_{t+1} + W_k \cdot f(t+1) \\ C_t &= A_t + W_k \cdot f(t) \\ C_{t+1} &= A_{t+1} + W_k \cdot f(t-1) \\ A_{t+1} - a \cdot A_t &= u(t) \end{aligned}$$

avec :

- $E[u(t)] = E[v(t)] = 0$
 - Les variances u et v indépendantes du temps.
- $$f_{t+1} - bf(t) = v(t)$$

a) *Lissage par C_t*

Partant de la formule suivante :

$$C_{t+1} - bC_t = A_{t+1} + W_k \cdot f(t+1) - bA_t - bW_k \cdot f(t)$$

Nous aboutissons par des transformations simples à :

$$C_{t+1} - bC_t - aC_t = -bA_t - a \cdot W_k \cdot f(t) + u(t) + v(t) \quad (1)$$

b) *Lissage par C_{t+1}*

Par des transformations analogues, nous obtenons :

$$C_{t+1} - bC_t - a \cdot C_t + ab \cdot C_{t-1} = -[b \cdot u(t) + a \cdot v(t)] + u(t) + v(t) \quad (2)$$

Comme :

- $E[u(t)] = E[v(t)] = 0$
- Les variances de u et v indépendantes du temps.

On peut poser :

$$R(t) = -[b \cdot u(t) + a \cdot v(t)] + u(t) + v(t) = \text{Résidus}$$

Dès lors :

- $E[R(t)] = 0$
- Variance de R indépendante du temps.

c) *Équation finale*

Nous arrivons finalement à l'équation :

$$C_{t+1} - (a + b)C_t + (ab)C_{t-1} = R(t)$$

avec :

- $E[R(t)] = 0$ (3)
- Variance de R indépendante du temps.

Pour la suite du calcul nous poserons :

$$\begin{aligned} a + b &= \gamma \\ a \cdot b &= \beta \end{aligned}$$

L'équation (3) se ramènera à :

$$C_{t+1} - \gamma \cdot C_t + \beta \cdot C_{t-1} = 0$$

d) *Recherche de γ et β*

Désignons par C_3 le vecteur de $n + 1$ composantes non nulles.

$$\begin{array}{ccc} - & C_2 & - & n & - & - \\ - & C_1 & - & n-1 & - & - \end{array}$$

n étant le nombre de semaines de la période.

L'équation (4) se ramène à :

$$C_3 - \gamma \cdot C_2 + \beta \cdot C_1 = R \quad (5)$$

On minimise par les moindres carrés le carré des écarts. Ce qui revient vectoriellement à minimiser l'expression :

$$\text{minimum} = \frac{1}{2} R \cdot R'$$

Nous obtenons alors le système suivant :

$$C_2' C_3 = \gamma C_2' C_2 - \beta C_2' C_1$$

$$C_1' C_3 = \gamma C_1' C_2 - \beta C_1' C_1$$

Nous résolvons ce système de deux équations à deux inconnues β et γ et comme :

$$ab = \beta$$

$$a + b = \gamma$$

on en tire facilement les valeurs de a et b

Dans l'exemple traité :

$$a = + 0,9830$$

$$b = - 0,1520$$

Ce qui s'interprète physiquement de la façon suivante :

pour a : le mouvement de fond est proche d'une valeur horizontale tout au long de la période ;

pour b : l'amortissement des fêtes est très rapide.

e) *Recherche des valeurs des impulsions dues aux fêtes*

Nous avons dit précédemment que nous envoyions une impulsion au moment des fêtes et autres phénomènes pouvant se produire.

Notre modèle devient dès lors :

$$C_t = A \cdot a^t + [W_{k1} \cdot b^{t-k1} + W_{k2} \cdot b^{t-k2} + \dots + W_{ki} \cdot b^{t-ki}] + \varepsilon(t)$$

avec W_{ki} = valeur de l'impulsion due à la fête i située à l'instant ki .

$W_{ki} \cdot b^{t-ki}$ = valeur amortie de l'impulsion W_{ki} à l'instant t .

Soit X la matrice des paramètres à estimer.

Soit C la matrice des coefficients C_t

Soit M la matrice des coefficients a et b

Le modèle peut s'écrire :

$$M \cdot X - \varepsilon = C \quad (6)$$

L'équation (6) est résolue par la méthode de Lagrange, c'est-à-dire que nous formons l'expression avec le paramètre μ

$$L = \frac{1}{2} \varepsilon \cdot \varepsilon' + \mu'(M \cdot X - \varepsilon - C) = 0$$

et l'on aboutit à :

$$M' \cdot \varepsilon = 0 \quad (7)$$

Or :

$$M \cdot X - C = \varepsilon$$

D'où :

$$M' \cdot \varepsilon = M'(M \cdot X - C)$$

$$X = \frac{M' C}{M' M}$$

f) *Calculs*

Les calculs, réalisés sur ordinateur, qui ont porté sur un historique de cinq années, ont donné une estimation satisfaisante sauf pour la pointe de Pâques qui est toujours sous-estimée (voir Annexe II).

4. Régression multiple d'ordre 3

Nous avons résolu le problème précédent avec une régression multiple d'ordre 2, ce qui correspondait physiquement à un amortissement rapide du phénomène de pointe.

N'était-il pas intéressant de regarder la réaction du système vis-à-vis d'un amortissement beaucoup plus faible ?

Ceci nous a conduit à la régression multiple d'ordre 3. En effet, nous avons désormais :

$$A_{t+1} - aA_t = u(t)$$

avec :

$$- E[u(t)] = E[v(t)] = 0$$

— Les variances de u et v indépendantes du temps

$$f(t+1) - b_1 f(t) - b_2 f(t-1) = v(t)$$

Nos équations de base sont devenues les suivantes :

$$C_{t+1} = A_{t+1} = W_k \cdot f(t+1)$$

$$C_t = A_t + W_k \cdot f(t)$$

$$C_{t-1} = A_{t-1} + W_k \cdot f(t-1)$$

$$C_{t-2} = A_{t-2} + W_k f(t-2)$$

Nous avons abouti à la forme généralisée suivante :

$$C_{t+1} - C_t(a + b_1) + C_{t-1}(ab_1 - b_2) + C_{t-2}(ab_2) = R'_t$$

avec :

$$\begin{aligned} a + b_1 &= \alpha \\ ab_1 - b_2 &= \beta \\ ab_2 &= \gamma \end{aligned}$$

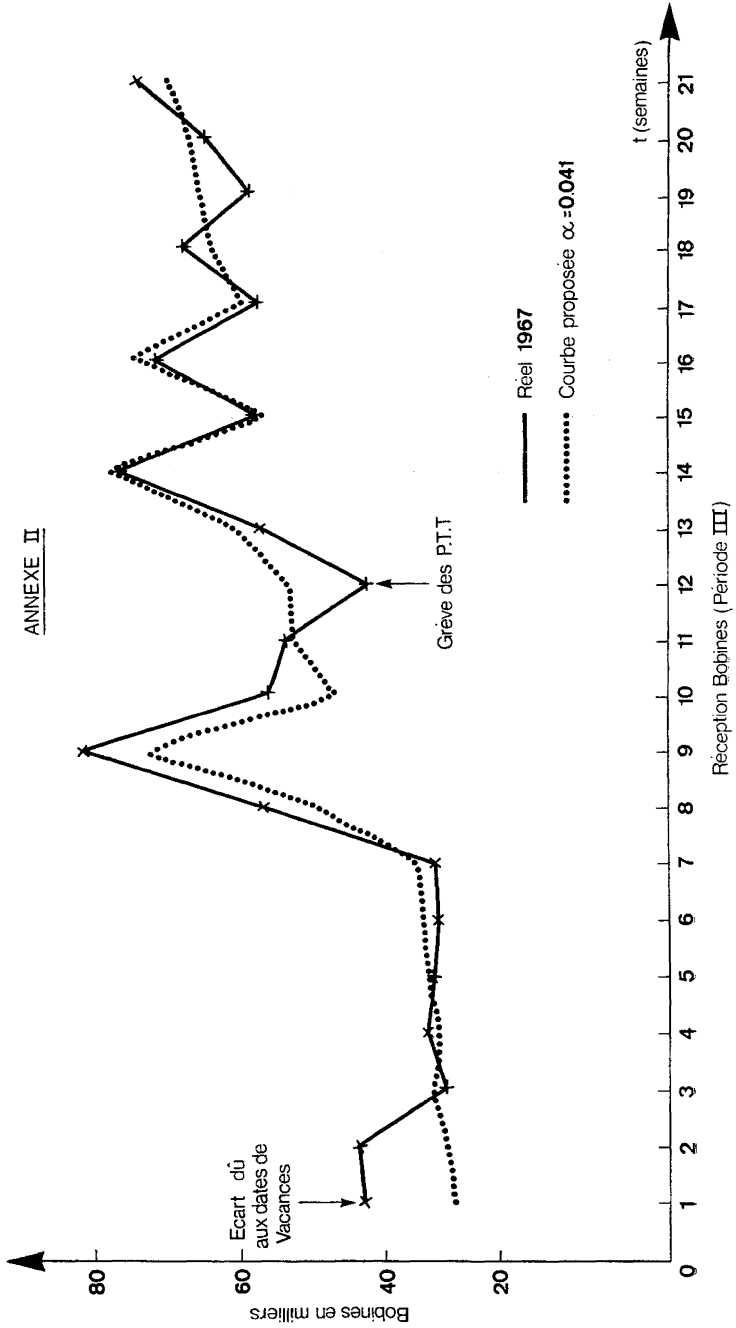
Nous obtenons :

$$C_3 - \alpha C_2 + \beta C_1 + \gamma C_0 = R'_t$$

avec :

- $E[R'(t)] = 0$
- Variance de R' indépendante du temps.

Conclusion : Nous ne nous étendrons pas sur ce point ; les résultats obtenus ont été relativement médiocres. L'amortissement « traînait » trop en longueur pour ce type de série chronologique.



CHAPITRE III

ETABLISSEMENT DES COEFFICIENTS POUR LES QUATRE AUTRES PERIODES

1. Description

Les périodes I, II, IV et V étant plus stables que la troisième, la méthode pour déterminer les coefficients hebdomadaires est plus classique. Nous avons utilisé la méthode dite du lissage.

Cette fois-ci, le lissage n'est pas de type exponentiel, mais du type dit « I.N.S.E.E. ».

L'unité de temps est représentée par la semaine à l'intérieur de la période, et le travail a été effectuée sur des pourcentages.

Soit le modèle :

$$x_t = f_t(1 + s_t) + z_t$$

Avec :

x_t = donnée brute de la semaine t étant considéré comme le numéro de la semaine dans la séquence

f_t = le mouvement extra-saisonnier (c'est-à-dire : tendance générale + cycle)

s_t = le mouvement saisonnier

z_t = les fluctuations résiduelles ou accidentelles.

REMARQUES :

a) Nous avons fait une hypothèse sur la composante saisonnière : nous l'avons prise multiplicative.

b) Nous avons admis la rigidité du mouvement saisonnier.

c) L'extra-saisonnier f est estimé par la moyenne mobile.

d) On fera l'hypothèse $E[z(t)] = 0$.

2. Première étape

On effectue pour chaque période une moyenne mobile m basée sur son nombre de semaines.

Or :
$$x_t = (1 + s_t)f(t) + z_t$$

Soit :

$$\frac{x_t}{f_t} = (1 + s_t) + \frac{z_t}{f_t}$$

Les rapports saisonniers hebdomadaires que nous calculons sont donc peu différents de la somme du coefficient saisonnier et d'un élément aléatoire de moyenne nulle par hypothèse.

Si nous prenons pour une semaine donnée, la moyenne des rapports obtenus pour n années consécutives, nous éliminons cette composante aléatoire et ne laisserons subsister que le coefficient saisonnier cherché.

3. Deuxième Etape

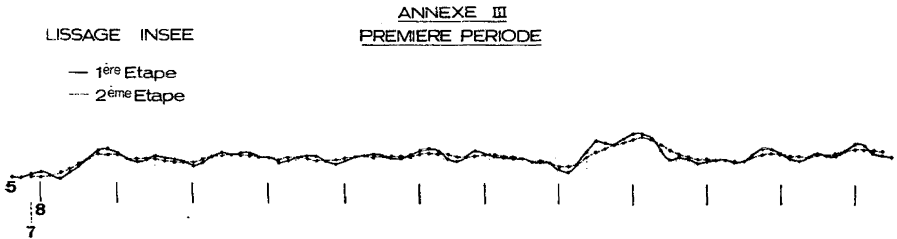
a) Elle consiste à utiliser une méthode dite améliorée.

C'est-à-dire qu'après avoir effectué une première désaisonnalisation au cours de la première étape, on en effectue une seconde au cours de la deuxième.

Cette deuxième désaisonnalisation n'a été effectuée que sur la première période et la quatrième période, les autres étant trop réduites.

b) *Calculs*

Les résultats des deux étapes figurent en Annexe III. On verra que la deuxième méthode améliore le lissage.



CONCLUSION

Ce modèle tel qu'il vient d'être décrit, fonctionne depuis trois ans.

Avant d'être mis en application, il a été testé sur le « Réel » d'une année et ensuite généralisé à d'autres produits.

Les résultats qu'il donne sont fournis au Responsable de la prévision qui apporte quelques modifications pour tenir compte de phénomènes tels que :

— modification dans la structure des dates de vacances scolaires (par exemple : passage du système à une zone au système à deux zones décalées) ;

— existence cette année-là de « ponts » (1^{er} Mai, 14 Juillet, 15 Août, ...)
dont on étudie actuellement la prise en compte automatique.

Les ajustements étant effectués, la prévision devient alors officielle.

Les méthodes manuelles anciennes donnaient un écart absolu moyen hebdomadaire de 12 à 15 %. Cet écart est tombé entre 3 et 5 %, à l'exception de la pointe de Pâques qui n'a jamais pu être bien prévue ; le système de vacances scolaires, 1^{re} zone et 2^e zone, ayant perturbé l'historique.

Ces prévisions ont deux types d'utilisation en production :

— la mise en place de moyens en personnels et matériel (plan d'embauche, de formation et d'entretien du matériel) d'une part ;

— après correction à court terme, elles sont éclatées à la journée, puis aux types de réceptions pour l'ordonnancement et le suivi de la production d'autre part.