

UNIVERSITE D'ALGER
ECOLE NATIONALE POLYTECHNIQUE
DEPARTEMENT ECONOMIE

5/73

157



PROJET de FIN D'ETUDES



Proposé par
H. AIT-OUYAHIA

Etudié par
A. KERCHOUNI

Promotion 1973

UNIVERSITE D'ALGER
ECOLE NATIONALE POLYTECHNIQUE
DEPARTEMENT ECONOMIE

PROJET de FIN D'ETUDES



Proposé par
H. AIT-OUYAHIA

Etudié par
A. KERCHOUNI

Je tiens à exprimer ma gratitude à M^r Hamid
ATT - OUYAHIA pour avoir suivi de près cette étude.

Je remercie M^r Rabah OUFERHAT pour l'aide précieuse
qu'il m'a apportée.

Je remercie également M^r MAMMERI à qui je dois la
troisième partie de cette étude.

Que les membres du Jury et l'ensemble des Professeurs
ayant contribué à ma formation trouvent ici l'expression de ma reconnaissance.

----- SOMMAIRE -----

GENERALITES

I- PARTIE

- I1 Les Moindres Carrés Ordinaires
- I2 Les Moindres Carrés Généralisés.
 - I2.1 Propriétés
 - I2.2 Le Problème de la Prévion
 - I2.3 Conséquences de la Correlation des aléas
 - I2.4 Tests d'Autocorrélation
 - I2.5 Estimation

II- PARTIE

Application au modèle de prévision à court terme

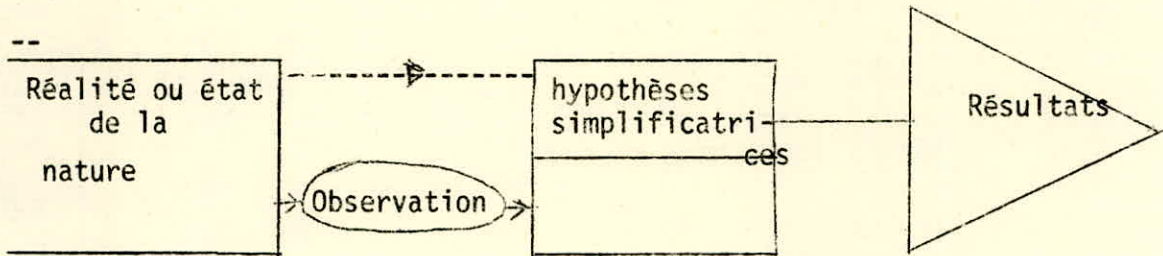
- II-1 Définition des variables - Notation
- II-2 Séparation de la tendance et des cycles
- II-3 Extrapolation de la tendance
- II-4 Analyse des cycles
- II-5 Indépendance des composantes
- II-6 Extrapolation des cycles
- II-7 La charge prévue et sa densité de probabilité
- II-8 Densité pour la pointe annuelle et la semaine où elle se présente.

III - PARTIE

PROGRAMMATION.

- III-1 Dossier des différentes routines
- III-2 Listing. en annexe.

AVANT PROPOS



Pour bien utiliser un modèle, il faut en connaître les possibilités, les limites.

Il convient pour cela :

- de spécifier l'objectif qu'on veut atteindre
- de distinguer :
 - la réalité ou l'état de la nature
 - le modèle simplificateur
 - l'observation.

La connaissance du phénomène réel, l'ensemble des concepts qui s'y rattachent, permettent de postuler le corps d'hypothèses du modèle et de poser la forme générale des relations entre les variables.

L'observation permettra d'estimer les paramètres des hypothèses et des relations fonctionnelles retenues.

GENERALITES

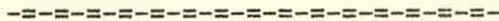
L'étude qui va suivre concerne la prévision d'une variable hebdomadaire présentant une tendance et une modulation annuelles.

Une variable exogène stationnaire, en l'occurrence, la température est utilisée pour expliquer une partie de la modulation annuelle.

Il s'agit d'isoler chacune des composantes de la variable (trend, cycles) et à les extrapoler séparément. Si ces composantes sont indépendantes, il est possible de recombinaer leurs prévisions respectives par le procédé de la convolution et d'obtenir ainsi une densité de probabilité pour la variable prévue.

Dans une première partie, on rappellera quelques résultats fondamentaux de la méthode économétrique, avant d'aborder en seconde partie l'application à la prévision de la demande d'électricité.

PREMIERE PARTIE



I.1.- MOINDRES CARRÉS ORDINAIRES

$$Y = Xb + U$$

$$\begin{array}{c}
 \text{(n,1)} \\
 \left| \begin{array}{c} \\ \\ \\ \end{array} \right. \\
 = \\
 \left[\begin{array}{c} X_{n1} \\ \\ \\ X_{np} \end{array} \right] \times \\
 \left[\begin{array}{c} \\ \\ \\ \end{array} \right] \\
 \text{(n,p)} \\
 \left| \begin{array}{c} \\ \\ \\ \end{array} \right. \\
 \left| \begin{array}{c} b \\ \\ \\ \end{array} \right. \\
 \text{(p,1)} \\
 + \\
 \left| \begin{array}{c} \\ \\ \\ \end{array} \right. \\
 \text{(n,1)}
 \end{array}$$

X_{ik} $i = 1, n$ est la $k^{\text{ième}}$ variable exogène

Y variable endogène

Pour inclure le terme constant du modèle dans la notation ci-dessus la $p^{\text{ième}}$ "variable" exogène est la variable particulière dont toutes les composantes sont égales à 1.

Problème :

1/ Quelles valeurs numériques convient-il d'attribuer aux paramètres $b_1 \dots b_p$?

2/ Quelles valeurs convient-il d'attribuer aux espérances mathématiques, variances et covariances des aléatoires u_j

1 - Principe de la méthode des moindres carrés ordinaires :

1/ Espace des variables

Les variables $Y_1 X_1 \dots X_p$ définissent un espace vectoriel à $p + 1$ dimensions; l'équation :

(1) $Y = b_1 x_1 + \dots + b_p x_p$ y est représentée par un hyperplan à p dimensions. D'autre part chacune des n observations $y_i, x_{i1} \dots x_{ip}$ est représentée par un point.

D'après le modèle :

$y_i = b_1 x_{i1} + \dots + b_p x_{ip} + u_i$ les observations forment un nuage de points dispersés autour du plan (1).

Il s'agit, connaissant les n observations seulement d'estimer la structure du modèle (2) qui les a engendrées.

Les points représentant les observations doivent être d'autant plus proches du plan (1) que le modèle est bien choisi.

La méthode des moindres carrés consiste à résoudre

$$\text{Min } Q = \sum_{i=1}^n (y_i - \sum_{j=1}^p b_j x_{ij})^2 \rightarrow \hat{b}_j \quad j = 1 \dots p$$

2/ Espace des observations

Dans l'espace des observations (\mathbb{R}^n) l'estimateur des moindres carrés a une interprétation géométrique commode.

L'échantillon $(y_1 \dots x_{i1} \dots x_{ip})$ est représenté par $p + 1$ points.

Y a pour coordonnées $y_1 \dots y_i \dots y_n$

X_1 a pour coordonnées $x_{11} \dots x_{i1} \dots x_{n1}$

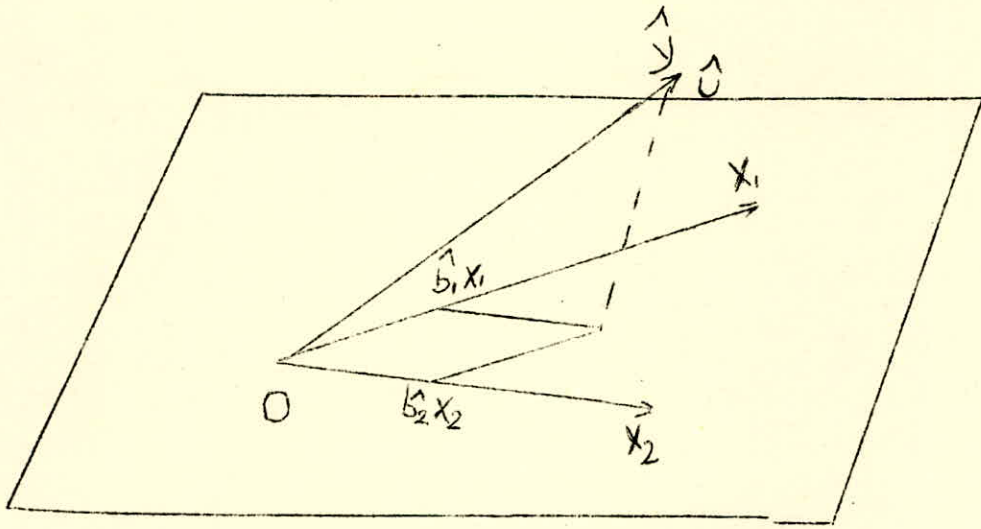
Si $n > p$, $x_1 \dots x_p$ engendrent un sous-espace vectoriel de dimension au plus égale à p, le modèle (2) s'écrit alors :

$$Y = b_1 X_1 + \dots + b_p X_p + u \quad (3)$$

Il s'agit de trouver le point du sous-espace vectoriel le plus proche de y; ce qui revient à trouver la projection orthogonale de Y sur ce sous-espace.

Il suffit pour cela d'annuler les p produits scalaires :

$$\left\langle Y - \sum_{j=1}^p b_j X_j, X_j \right\rangle = 0 \quad j = 1 \dots p$$



$$\langle X_j, \hat{u} \rangle = \sum_{i=1}^n x_{ij} \hat{u}_i = 0 \quad j = 1, p$$

Ce qui s'écrit sous forme matricielle

$$X'(Y - \hat{X}b) = 0 \quad \text{ou} \quad X'Y - X'\hat{X}b = 0$$

$$\text{D'où} \quad \hat{b} = (X'X)^{-1}X'Y$$

où

X	est de dimension	(N,p)
X'	"	(p,n)
(X'X)	"	(p,p)
(X'X) ⁻¹ X'	"	(p,n)
Y	"	(n,1)
\hat{b}	"	(p,1)

Autre manière de calculer \hat{b} :

$$\text{Min } Q = \sum_{i=1}^n (y_i - \sum_{j=1}^p b_j x_{ij})^2$$

$$Q = \sum_{i=1}^n \hat{u}_i^2 = \hat{U}' \hat{U}$$

$$Q = (Y - \hat{X}b)' (Y - \hat{X}b)$$

$$Q = (Y' - \hat{b}'X') (Y - \hat{X}b) = Y'Y - Y'\hat{X}b - \hat{b}'X'Y + \hat{b}'X'\hat{X}b$$

$$= X'Y - 2\hat{b}'X'Y + \hat{b}'X'\hat{X}b \quad (\text{car } \hat{b}'X'Y \text{ est un scalaire})$$

$$\text{Min } Q \rightarrow \begin{cases} \partial Q / \partial \hat{b} = 0 \\ \partial^2 Q / \partial \hat{b}^2 > 0 \end{cases} \Rightarrow \hat{b} = (X'X)^{-1} X'Y$$

PROPRIETES STATISTIQUES DES ESTIMATEURS DES MOINDRES CARRES

Hypothèses :

Y et X_j : variable certaine mesurée sans erreur

$$E(U) = 0 \quad (H_1)$$

$$V(U) = \sigma^2 I \quad (H_2)$$

$$U \rightsquigarrow N(0, \sigma^2 I) \quad (H_3)$$

1/ \hat{b} est un estimateur sans biais

$$\hat{b} = (X'X)^{-1} X'Y$$

$$= (X'X)^{-1} X' (Xb + U)$$

$$= (X'X)^{-1} X'Xb + (X'X)^{-1} X'U$$

$$= b + (X'X)^{-1} X'U$$

$$E(\hat{b}) = E(b + (X'X)^{-1} X'U)$$

$$= b + E((X'X)^{-1} X'U)$$

$$= b + (X'X)^{-1} E(U)$$

$$E(\hat{b}) = b \quad (\text{d'après } H : E(U) = 0)$$

$$2/ V(\hat{b}) = \sigma^2 (X'X)^{-1}$$

$$V(\hat{b}) = E \left\{ (b - E(\hat{b})) (b - E(\hat{b}))' \right\}$$

$$= E \left\{ (\hat{b} - b) (b - b)' \right\}$$

$$= E \left\{ (X'X)^{-1} X'U ((X'X)^{-1} X'U)' \right\}$$

$$= E \left\{ (X'X)^{-1} X'U U' X (X'X)^{-1} \right\}$$

$$= (X'X)^{-1} X' E(UU') X (X'X)^{-1}$$

$$= (X'X)^{-1} X' \sigma^2 I X (X'X)^{-1}$$

$$V(\hat{b}) = \sigma^2 (X'X)^{-1}$$

Théorème de Gauss-Markov :

\hat{b} est de variance minimum dans la classe des estimateurs linéaires sans biais.

Soit \hat{b} un estimateur linéaire quelconque de b

$\hat{b} = C'Y$ où C' est une matrice (p,n) fonction de l'échantillon (X)

$$\hat{b} = C' (Xb + U)$$

$$= C'Xb + C'U$$

$$E(\hat{b}) = E(C'Xb + C'U)$$

$$= C'Xb + C' E(U) \quad (\text{d'après } H_1)$$

Comme \hat{b} doit être sans biais on aura $E(\hat{b}) = b \implies C'X = I$
où I est de dimension (p,p)

On a donc :

$$\hat{b} = b + C'U$$

$$V(\hat{b}) = E \left\{ (\hat{b} - E(\hat{b})) (\hat{b} - E(\hat{b}))' \right\}$$

$$= E \left[(C'UU'C) \right] = C'E(UU')C$$

$$= \sigma^2 C'C$$

Posons $C' = (X'X)^{-1}X' + D'$ où D' est une matrice (p,n)

On exprime sur D' la condition $C'X = I$

On a alors :

$$((X'X)^{-1}X' + D')X = I \text{ ou } D'X = 0$$

on développe l'expression de $V(\hat{b})$

$$V(\hat{b}) = \sigma^2 ((X'X)^{-1}X' + D') ((X'X)^{-1}X' + D)'$$

$$= \sigma^2 ((X'X)^{-1}X'X(X'X)^{-1} + (X'X)^{-1}X'D + D'X(X'X)^{-1} + D'D)$$

$$= \sigma^2 (X'X)^{-1} + D'D > V(\hat{b})$$

\hat{b} est le BLUE (best linear unbiased estimator) de b .

Un estimateur sans biais de σ^2 est :

- 7 -

$$s^2 = \frac{E(\hat{U}'\hat{U})}{n-p}$$

La variance σ^2 des aléatoires U_i étant inconnue à priori, on peut chercher à l'estimer à partir de la variance empirique des écarts U_i -

$$\sum \hat{U}_i^2 = \hat{U}'\hat{U}$$

$$\begin{aligned}\hat{U} &= (Y - \hat{X}b) = Y - X(X'X)^{-1}X'Y \\ &= (I - X(X'X)^{-1}X')U\end{aligned}$$

$$\text{On pose } M = I - X(X'X)^{-1}X'$$

M est symétrique et idempotente de sorte que :

$$U'U = U'M'MU = U'M^2 = U'MU$$

$$E(U'U) = E(U'MU)$$

Utilisons l'opérateur trace :

$U'MU$ étant un scalaire il est identique à sa trace

$$E(U'U) = E(\text{Tr}(U'MU))$$

$$= E(\text{Tr}(MUU'))$$

$$= \text{Tr}ME(UU')$$

$$= \text{Tr}M\sigma^2I \quad (\text{d'après } H_1 \text{ et } H_2)$$

$$= \sigma^2\text{Tr}M$$

$$= \sigma^2\text{Tr}(I_n - X(X'X)^{-1}X')$$

$$= \sigma^2(\text{Tr}I_n - \text{Tr}(X(X'X)^{-1}X'))$$

$$= \sigma^2(\text{Tr}I_n - \text{Tr}X'X(X'X)^{-1}) = \sigma^2(\text{Tr}I_n - \text{Tr}I_p)$$

$$E(\hat{U}'\hat{U}) = \sigma^2(n - p)$$

5/ \hat{b} est l'estimateur de Maxi de vraisemblance

La vraisemblance de l'échantillon observé est

$$d(x_i) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-U_i^2/2\sigma^2}$$

d'après (H_2 et H_3)

$$\text{Max } d(x_i) \equiv \text{Min } \sum U_i^2$$

Soit le modèle $y = Xb + u$ (1)

avec : $E(u) = 0$

$E(uu') = \sigma^2 W$ où σ^2 est inconnu.

et W une matrice symétrique définie positive d'ordre n .

W peut donc s'écrire :

$W = PP'$ où P est une matrice non singulière

$$P^{-1} W P^{-1} = I$$

$$P^{-1} \cdot P^{-1} = W^{-1}$$

Pré-multiplions (1) par P^{-1}

$$P^{-1} y = P^{-1} Xb + P^{-1} u$$

$$y^* = X^* b + u^*$$

$$E(u^* u^{*\prime}) = E(P^{-1} u u' P^{-1}) = P^{-1} E(u u') P^{-1}$$

$$= P^{-1} \sigma^2 W P^{-1} = \sigma^2 I$$

Si u est gaussien, u^* l'est aussi

$$\hat{b} = (X^{*\prime} X^*)^{-1} X^{*\prime} Y^*$$

$$\hat{b} = (X' W^{-1} X)^{-1} X' W^{-1} Y$$

I.2.1 - Propriétés

* \hat{b} est encore le BLUE de b

$$* V(b) = (X' W^{-1} X)^{-1}$$

* Un estimateur sans biais de σ^2 est $\frac{1}{n-p} \hat{u}' W^{-1} \hat{u}$

où \hat{u} est le vecteur résidu, noté dans la suite e .

* \hat{b} est encore l'estimateur du maximum de vraisemblance.

Remarque :

Si on appliquait les O.L.S l'estimateur obtenu serait quand même sans biais, mais il ne serait plus de variance minimum dans la classe des estimateurs linéaires sans biais.

I.2.2- Prévision

$$Y = Xb + u \quad \text{avec } E(u) = 0$$

$$E(uu') = V$$

Le problème de la prévision est le suivant :
 connaissant un vecteur de prévisions sur les variables exogènes, quelle est la prévision Y_0 sur la variable expliquée ?

$$\text{Soit } Y_0 = X_0b + u_0$$

où u_0 est la vraie, mais inconnue, valeur prévue pour l'erreur.

Il est naturel de supposer :

$$E(u_0) = 0$$

$$E(u_0^2) = \sigma_0^2$$

$$E(u_0 u) = \begin{pmatrix} E(u_1 u_0) \\ E(u_2 u_0) \\ \vdots \\ E(u_n u_0) \end{pmatrix} = \omega$$

ω représente le vecteur des covariances de la prévision u_0 de l'erreur théorique avec les résidus.

On définit une "projection" linéaire :

$$p = C'Y \quad \text{où } C \text{ est un vecteur à } n \text{ composantes constantes}$$

Pour que p soit la meilleure prévision sans biais on doit choisir C tel que :

$$\text{Min } \sigma_p^2 = E((p - y_0)^2)$$

$$E(p - y_0) = 0$$

$$p - y_0 = (C'X - X_0')b + C'u - u_0$$

La condition d'absence de biais implique :

$$C'X - X_0' = 0$$

L'erreur de prévision est alors :

$$p - y_0 = C'u - u_0 ; \text{ sa variance est :}$$

$$\begin{aligned} \hat{\sigma}_p^2 &= E \left\{ (p - y_0)^2 \right\} \\ &= E \left\{ (p - y_0) (p - y_0)' \right\} \\ &= E \left\{ C' u u' C + u_0^2 - 2C' u u_0 \right\} \\ &= C' V C + \sigma_0^2 - 2C' \omega \end{aligned}$$

$$\begin{cases} \text{Min} \left\{ C' V C + \sigma_0^2 - 2C' \omega \right\} \\ C' X - X_0 = 0 \end{cases}$$

$$\equiv \text{Min } F = C' V C - 2C' \omega - 2(C' X - X_0) \lambda$$

où λ est un vecteur de multiplicateurs de la contrainte.

En annulant les dérivées partielles par rapport à C et à λ on obtient :

$$\begin{bmatrix} V & X \\ X' & 0 \end{bmatrix} \begin{bmatrix} C \\ -\lambda \end{bmatrix} = \begin{bmatrix} \omega \\ X'_0 \end{bmatrix} \quad \text{ou}$$

$$\begin{pmatrix} \hat{C} \\ \hat{\lambda} \end{pmatrix} = \begin{bmatrix} V & X \\ X' & 0 \end{bmatrix}^{-1} \begin{pmatrix} \omega \\ X'_0 \end{pmatrix}$$

Tous calculs faits on arrive à :

$$\hat{p} = X_0' \hat{b} + \omega' V^{-1} \hat{u}$$

où \hat{u} est le vecteur des résidus G.L.S.

Cas où les résidus sont autocorrélés.

Supposons qu'on ait un échantillon d'observations pour les périodes 1 à n auxquelles nous voulons ajuster le modèle $y = Xb + u$

avec $V = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & & & \\ \vdots & & \ddots & & \\ \rho^{n-1} & & & & 1 \end{bmatrix}$

On suppose en outre qu'on connaît les valeurs que prendront les variables explicatives durant la période $n + 1$; on se donne donc le vecteur ligne X_{1+1} dans la matrice X .

On a vu précédemment que la prévision sans biais de variance minimale est :

$$\hat{Y}_{n+1} = X_{n+1} \hat{b} + \bar{\omega}' V^{-1} e$$

$$\text{où } e = Y - X \hat{b}$$

$$\bar{\omega} = \begin{pmatrix} E(u_1 \ u_{n+1}) \\ E(u_2 \ u_{n+1}) \\ \vdots \\ E(u_n \ u_{n+1}) \end{pmatrix}$$

Si les erreurs sont générées par un processus autorégressif du 1er ordre, on verra que :

$$E(u_t \cdot u_{t-s}) = \rho^s \sigma_u^2 \quad (\text{cf. II-3})$$

$$\text{d'où } \bar{\omega} = \rho \sigma_u^2 \begin{pmatrix} \rho^{n-1} \\ \rho^{n-2} \\ \vdots \\ 1 \end{pmatrix}$$

$\bar{\omega} = \rho$ fois la dernière colonne de V

$\bar{\omega} V^{-1}$ est alors égale à ρ fois la dernière ligne I
d'où $\bar{\omega} V^{-1} e = \rho e_n$

$$Y_{n+1} = X_{n+1} \hat{b} + \rho e_n \quad (2)$$

(1) est utilisée lorsque V est connue (i.e ρ est connu)

(2) est plus immédiate quand on ne fait qu'estimer ρ , ce qui est souvent le cas.

Supposons ρ positif et les observations sur la variable exogène générées par le même processus que les aléas; alors pour n très grand l'expression entre parenthèses vaut approximativement : $\frac{1 + \rho^2}{1 - \rho^2}$

Ainsi si $\rho = 0,8$ $\frac{1 + \rho^2}{1 - \rho^2} = 4,5$ la variance réelle représente plus que 4 fois la variance obtenue par application des O.L.S.

En pratique cette sous-estimation est encore plus exagérée : en effet, ne connaissant généralement pas ρ on l'estime par :

$$\frac{E(e'e)}{n-1} = \sigma_u^2 \quad (\text{cf. O.L.S.})$$

Alors que l'estimateur réel (compte tenu de l'autocorrelation est :

$$\frac{E(e'e)}{n - \frac{1 + \rho^2}{1 - \rho^2}}$$

$$e'e = u'u - u'X(X'X)^{-1}X'u$$

$$E(e'e) = E(u'u) - E \left\{ u'X (X'X)^{-1} X'u \right\}$$

$$E \left\{ u'X(X'X)^{-1}X'u \right\} = E \text{Tr} \left\{ u'X(X'X)^{-1} X'u \right\}$$

L'argument étant scalaire on a :

$$E \left\{ u'X(X'X)^{-1}X'u \right\} = \text{Tr} (X(X'X)^{-1} X'u)$$

$$E(u'u) = n\sigma_u^2$$

Si on revient au cas d'une seule variable exogène alors :

$$E(e'e) \approx \sigma_u^2 \left(n - \frac{1 + \rho^2}{1 - \rho^2} \right)$$

Si $\rho = 0$ on retrouve bien l'estimateur de σ_u^2 fourni par les O.L.S

Pour $n = 20$ et $\rho = 0,5$ la sous-estimation est de l'ordre de 4 %

Pour $n = 20$ et $\rho = 0,8$, elle est de 19 %

Plus ρ est proche de 1 plus grand est le biais

Plus n est grand, moins important est le biais.

I.2.4- Test d'autocorrélation

Le fait d'ignorer l'autocorrélation entraînant d'importantes conséquences, il est nécessaire de tester cette autocorrélation dès résidus dans une série. Pour les grands échantillon on peut considérer les écarts $e = y - Xb$ comme étant des observations sur les aléas u et appliquer des tests non paramétriques (tests basés sur des considérations de classement); cependant ces tests ne sont rigoureux que lorsque les résidus e sont effectivement des observations sur les erreurs.

1/ Ratio de Von Neumann

$$v = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2 / (n-1)}{\sum_{t=1}^n (e_t - \bar{e})^2 / n}$$

Pour n très grand $v \rightsquigarrow N \left(\frac{2}{n-1}, \frac{4n^2 (n-2)}{(n+1)(n-1)^3} \right)$

v ne suit en réalité cette loi que lorsque les e_t sont indépendamment distribués, et ceci n'est pas le cas pour les résidus des O.L.S., même lorsque les aléas u ont des distributions indépendantes. (cf. I.2.42)

2/ Test de Durbin et Watson

$$d = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2} \quad \text{on suppose } \bar{e} = 0$$

Il est intuitif que pour une corrélation positive les lères différences auront tendance à être faibles, en valeur absolue, par rapport aux valeurs absolues de e , tandis que pour une autocorrélation négative, elles seront plus grandes. Durbin et Watson ont tabulé 2 limites. d_u et d_L qui permettent de tester l'hypothèse de la corrélation nulle contre celle de l'autocorrélation positive.

$d < d_L$: rejeter l'hypothèse de l'autocorrélation en faveur de la contre hypothèse.

$d > d_u$: ne pas rejeter l'hypothèse nulle

$d_L < d < d_u$: on ne peut pas conclure.

C'est cette dernière éventualité qui fait perdre au test son intérêt.

Remarque : Si $d > 2$; on comparera $d' = (4-d) d_u$ et d_L

alors si $d' < d_L$: Autocorrélation négative du 1er ordre.

Un autre test permettant de lever le doute de la 3ème éventualité a été proposé par les mêmes auteurs; mais il nécessite beaucoup de calcul :

d peut s'écrire :

$$d = e' A e / e' e$$

$$\text{où } A(n,n) = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & -1 & 2 & -1 & \\ & & & -1 & 1 \end{pmatrix}$$

Les valeurs propres de A sont

$$\lambda_j = 2 \left(1 - \cos \frac{j\pi}{n} \right) \quad j = 0, 1, \dots, n-1$$

Les vecteurs propres correspondantes sont :

$$l_j = \begin{pmatrix} \cos \frac{j\pi}{2n} \\ \cos \frac{3j\pi}{2n} \\ \vdots \\ \cos \frac{(2n-1)j\pi}{2n} \end{pmatrix} \quad j = 0, 1, \dots, n-1$$

* On prend les K' plus petites valeurs propres autres que $=0$ avec $K' = K - 1$ (nombre de variables explicatives).

* On joint les K' vecteurs propres correspondantes pour former la matrice L :

$$L = (l_1, l_2, \dots, l_{K'})$$

Soit $X_{j*} = X_j$ observée - moyenne des observations sur X_j

* On applique les O.L.S. au modèle

$$y = a \cdot 1 + X b_1 + L b_2$$

* Déterminer les matrices triangulaires inférieures

P_1 et P_2 telles que :

$$P_1 P_1' (X_L' X_L)^{-1} \text{ et } P_2 P_2' = (L_X' L_X)^{-1}$$

où $X_L = X - L (L'L)^{-1} L'X$

$$L_X = L - X (X'X)^{-1} X'L$$

* Calculer

$$C = P_1 P_2^{-1} b_2 \quad b_2 \text{ étant issu des O.L.S.}$$

$$* Z = y - a \cdot 1 - X b_1 - L b_2 + X_L C$$

$$d' = \frac{\sum_{t=2}^n (z_t - z_{t-1})^2}{\sum_{t=1}^n z_t^2}$$

d' suit la même loi que d_u , ainsi on peut comparer d' à d_u tabulé

$d > d_u$: accepter l'hypothèse nulle.

2/ La procédure BLUS

Cette procédure est due à THEIL; l'idée est la suivante : Au lieu de rechercher des tests de plus en plus compliqués, il est peut-être préférable d'envisager un autre estimateur des aléas, lequel aura des propriétés plus simples que le vecteur des résidus des moindres carrés.

La difficulté, avec les résidus des moindres carrés, est le fait que ceux-ci sont autocorrélés même si les aléas ne le sont pas. En effet :

$$e = Mu$$

avec $M = I - X(X'X)^{-1}X'$ symétrique, idempotente.

$$\text{Ainsi } E(ee') = \sigma_u^2 M$$

On voit que les résidus sont autocorrélés quand même les aléas ne le sont pas.

L'idée de THEIL est de rechercher un estimateur e de u qui ait les propriétés suivantes :

- linéaire
- sans
- $V(\hat{e} - u)$ scalaire et minimum, si bien que le test du ratio de Von Neuman peut être appliqué .

D'où le nom de BLUS (Best linéar Unbiased Scalar)

La procédure est la suivante :

- * Partitionner le vecteur des résidus O.L.S.

$$e = \begin{pmatrix} e_0 \\ e_1 \end{pmatrix}$$

e_0 a k composantes

e_1 a $n-k$ composantes.

$$\hat{e}_1 = e_1 - (X_1 X_0)^{-1} \left[\sum_{i=1}^k \frac{d_i}{1+d_i} q_i q_i' \right] e_0$$

On ne peut de cette manière qu'estimer (n-k) aléas.

d_i^2 $i = 1, k$ représentent les valeurs propres de $X_0 (X'X)^{-1} X_0'$

d_i = racine carrée positive de d_i^2

La sommation dans la formule précédente ne s'étend qu'aux valeurs propres inférieures à l'unité.

Un problème se pose encore : le choix des k aléas non estimés.

Pour tester une autocorrélation nulle contre l'autocorrélation positive, on procède comme suit :

* Prendre pour e_0 les m^{ier} et les (k-m) derniers résidus

$$(0 \leq m \leq k)$$

Il y a ainsi k + 1 partitions possibles.

* Pour chaque valeur de m, on partitionne X de telle sorte que les m et les (k-m) lignes de X constituent X_0 et les (n-k) lignes restantes constitueront X_1

* Le choix entre les k + 1 partitions se fait comme suit :

Pour tout m_i $i = 0, k$:

- Calculer $X_0 (X'X)^{-1} X_0'$ et les valeurs propres d_1^2, \dots, d_k^2

Si $d_1^2 = 0$ éliminer la/les partitions P_i

- Calculer $d_1 + \dots + d_k$

Max $d_1 + \dots + d_k \rightarrow$ partition optimale, celle qui donne le minimum de variance de l'erreur.

- Calculer le vecteur propre de

$$X_0 (X'X)^{-1} X_0' \text{ correspondant à } d_i^2 < 1$$

* Calculer \hat{e}_1

* Appliquer aux résidus ainsi obtenus le test du ration de Von Numam

Conclusion :

Les tests d'autocorrélation conduisent à des temps de calculs trop longs (ils nécessitent plusieurs sous-programmes d'analyse numérique). Aussi on pratique se contente-t-on de supposer l'autocorrélation. De toute façon il est plus risqué de conclure à l'autocorrélation nulle quand elle existe en réalité, plutôt que d'admettre à priori l'autocorrélation alors qu'elle n'a pas lieu, puisque en faisant $\rho = 0$ dans les équations du I 2 3 on retombe sur le cas des O.L.S.

En pratique donc, on préfère supposer l'autocorrélation et utiliser une procédure d'estimation de ρ et revenir par transformation des variables à un modèle O.L.S., c'est l'objet du paragraphe sui suit.

I.2.5- Estimation

Soit le modèle :

$$Y_t = bX_t + U_t \quad t = 1 \dots n$$

$$U_t = \rho U_{t-1} + \varepsilon_t \quad |\rho| < 1$$

$$E(\varepsilon_t) = 0$$

$$E(\varepsilon_t \cdot \varepsilon_{t-s}) = \begin{cases} \sigma_\varepsilon^2 & \text{pour } s = 0 \quad \forall t \\ 0 & \text{pour } s \neq 0 \quad \forall t \end{cases}$$

En général, ni l'ordre d'autocorrélation, ni les valeurs des paramètres ne sont connus, si bien que les estimateurs des G.L.S. ne peuvent pas être calculés directement.

$$b = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

où $V^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1-\rho & & & & \\ -\rho & 1+\rho^2-\rho & & & 0 \\ & -\rho & & & \\ 0 & & 1+\rho^2-\rho & & \\ & & -\rho & & \\ & & & & 1-\rho \end{pmatrix}$ V^{-1} est tridigonale

Le coefficient ρ est inconnu.

Cependant les estimateurs des G.L.S. peuvent être approximatés par la méthode suivante :

Si on applique aux variables Y_t et X_t la transformation suivante :

$$Y_t - \rho Y_{t-1} = b(X_t - \rho X_{t-1}) + \varepsilon_t$$

On obtient un nouveau modèle où les aléas ne sont pas corrélés.

Notons \hat{b}, r les estimateurs respectifs de b et ρ
la somme des carrés des résidus est donnée par :

$$\sum_{t=1}^n e_t^2 = \sum_{t=1}^n [(y_t - r y_{t-1}) - \hat{b} (x_t - r x_{t-1})]^2$$

La sommation s'étendant de 1 à n si y_0 et x_0 sont donnés, sinon on sommerait de 2 à n.

La minimisation de la somme des carrés des résidus par rapport à \hat{b} et r ne donnent pas d'expressions analytiques pour \hat{b} et r si bien que ces estimateurs ne peuvent pas être directement obtenus à partir de la minimisation de $\sum_{t=1}^n e_t^2$

Un processus itératif d'approximation des valeurs de \hat{b} et r dû à Cochrane-Orcutt est décrit ci-dessous :

- 1 - On fixe arbitrairement une valeur de r soit r_1
- 2 - $\text{Min}_{\hat{b}} \sum e_t^2 \rightarrow \hat{b}_1$
- 3 - $\hat{b} = \hat{b}_1$ puis $\text{Min}_r \sum e_t^2 \rightarrow r_2$
- 4 - $r = r_2$ aller en 2, arrêter les itérations quand 2 valeurs consécutives de r par exemple sont suffisamment proches.

En pratique on prend $r_1 = 0$, la somme des carrés des résidus est alors :

$$\sum_{t=1}^n (y_t - b x_t)^2$$

1/ $\text{Min}_b \sum_{t=1}^n e_t^2$ b estimateur fourni par les O.L.S. appliqués au modèle initial.

2/ Comparer 2 valeurs consécutives de r :

$$r_i - r_{i+1} < \hat{\eta} \quad \text{arrêter les itérations sinon continuer}$$

3/ $\text{Min}_r \sum_{t=1}^n [(Y_t - bX_t) - r(Y_{t-1} - bX_{t-1})]^2 \rightarrow r_2$

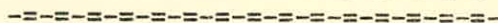
r_2 est l'estimateur des D.L.S. appliqués à $u_t = \rho u_{t-1} + \varepsilon_t$

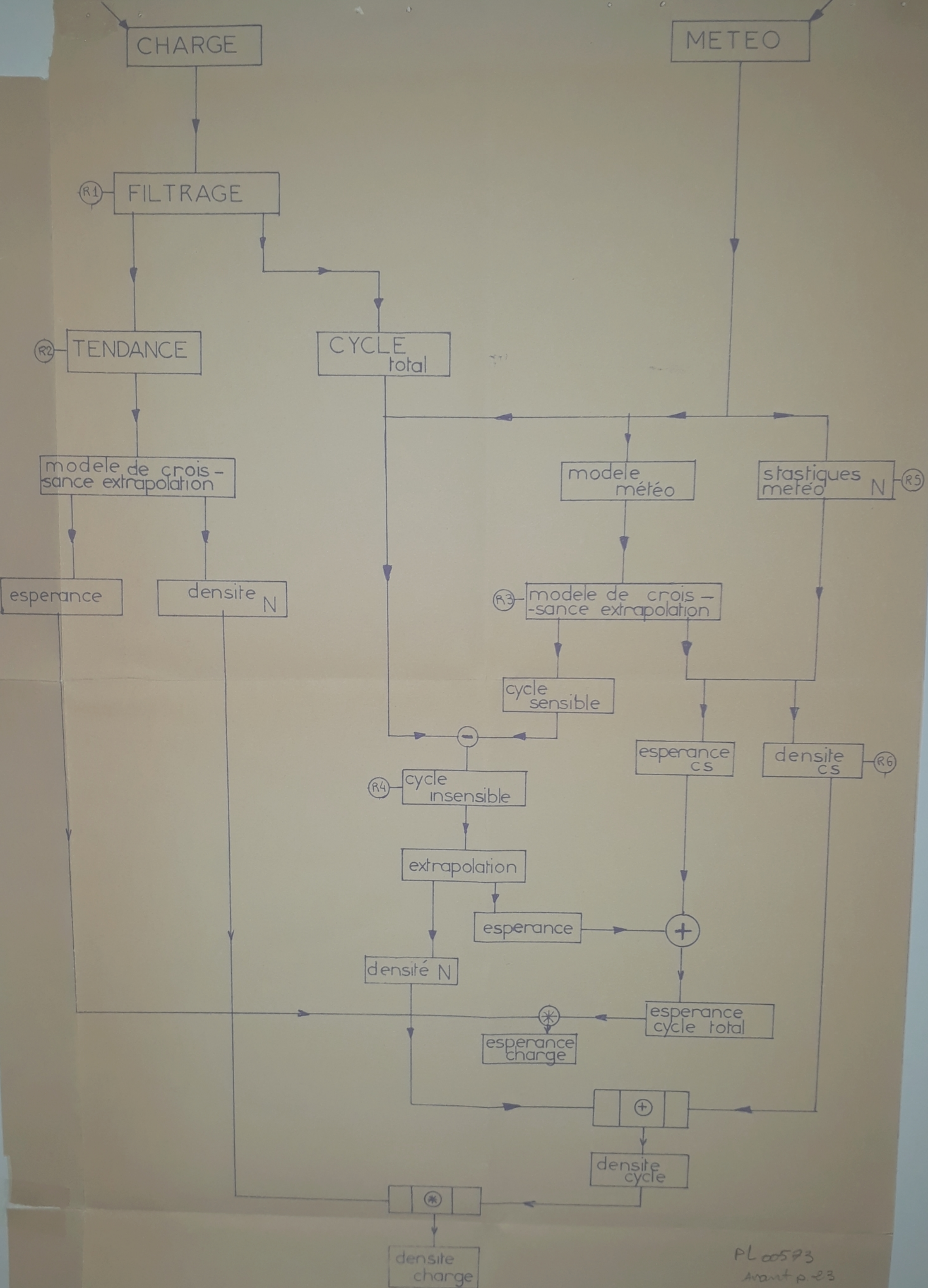
4/ Appliquer les O.L.S. aux variables transformées

$$(y_t - r_2 y_{t-1}), (x_t - r_2 x_{t-1}) ; \text{ aller en 2}$$

C'est cette procédure qui a été appliquée dans le cas qui va suivre, où l'autocorrélation a été admise à priori du 1er ordre.

DEUXIEME PARTIE





PL 00593
 Avant p. 23

R i est le renvoi aux routines principales utilisées pour chaque partie.

R1 : CYCL , CRLN.
R2 : MCSP, MCGE , RØS2
R3 : CFTE , EXTR , FNCS , CYNM , CYSB .
R4 : CYSB
R5 : CSNY
R6 : PRGS , CSNY , FNCS , PROB , FNTP .

II-1.- DEFINITION DES VARIABLES - NOTATIONS

D_t = demande observée la semaine t

T_t = ligne de tendance pour la semaine t

$C_{i,j}$ = Valeur du coefficient multiplicatif du cycle pour la semaine j de l'année i (semaine t)

$W_t = W_{i,j}$ = variable météo observée la semaine t

$CS_{i,j}$ = valeur du coefficient sensible pour la semaine j de l'année i

$CI_{i,j}$ = Valeur du coefficient du cycle insensible pour la semaine j de l'année i

$[TW, TS]$ = intervalle d'insensibilité.

Equations

$$D_t = T_t + C_{i,j}$$

$$T_t = a + b t + c t^2 + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t \text{ avec } \begin{cases} E(\varepsilon_t) = 0 \\ \sigma^2(\varepsilon_t) = \\ \text{cov}(\varepsilon_t, \varepsilon_{t-\theta}) = 0 \end{cases}$$

$$C_{i,j} = CS_{i,j} + CI_{i,j}$$

$$CS_{i,j} = \begin{cases} KW_i (W_t - TW) & \text{pour } W_t < TW \\ 0 & \text{pour } TW \leq W_t \leq TS \\ KS_i \times (W_t - TS) & \text{pour } W_t > TS \end{cases}$$

$$KW_i = \alpha_w + \beta_w * i$$

$$KS_i = \alpha_s + \beta_s * i$$

$$CI_{i,j} = \gamma_j + \eta_{i,j}$$

KW_i et KS_i sont des coefficients inconnus mais certains; ils expriment la sensibilité à la température respectivement en hiver et en été.

$$E(\eta_{i,j}) = 0$$

$$V(\eta_{i,j}) = I \sigma_{\eta_j}^2 \quad \text{homoscédasticité des erreurs}$$

semaine par semaine.

La variable météo imprévisible à long terme est représentée dans l'avenir par sa moyenne et sa variance :

$$W_{i,j} = \bar{w}_j + \theta_{i,j}$$

$$E(\theta_{i,j}) = 0$$

$$V(\theta_{i,j}) = I \cdot \sigma_{w_j}^2$$

Hypothèse :

ε, η et θ suivent des lois de Laplace-Gauss

Problème :

Estimer les paramètres :

1/ $a, b, c,$ et σ_{ε}^2 pour la ligne de tendance

2/ $\alpha_w, \beta_w, \alpha_s, \beta_s, \bar{w}_j, \sigma_{w_j}^2$ pour le cycle sensible

3/ $\gamma_j, \sigma_{\eta_j}^2$ pour le cycle insensible.

II.2- SEPARATION DE LA TENDANCE ET DE CYCLES. (routine CYCL)

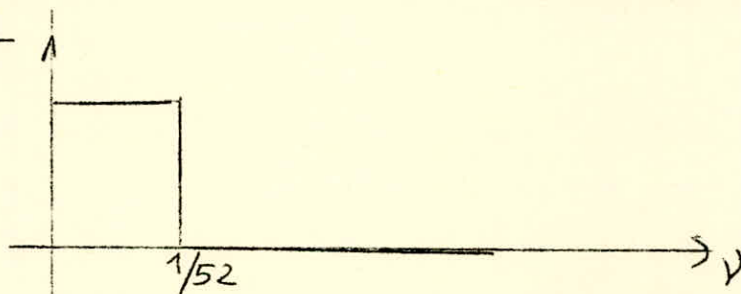
On entend par tendance, aussi bien le trend de développement que le cycle conjoncturel. Les cycles longs ne nous intéressent pas, puisque nous envisageons uniquement le court terme.

Le cycle est la partie de la charge dont la fréquence est supérieure à $\frac{1}{52}$

Pour séparer ces deux composantes on utilisera le filtrage numérique.

1/ Le choix du filtre

Le filtre parfait serait tel que le carré du module de sa fonction de transfert soit égal à 1 pour les fréquences inférieures ou égale à $\frac{1}{52}$



On considère 2 moyennes mobiles de longueur 52.

- moyenne mobile quasi-centrée

$$T(t) = \frac{1}{52} \sum_{j=t-26}^{25} D_j$$

- moyenne mobile décentrée de 26 semaines en avant

$$T(t) = \frac{1}{52} \sum_{j=t-51}^t D_j$$

Les fonctions de transfert respectives de ces deux filtres sont :

$$F_1(\omega) = \frac{1}{52} \left((1 + \cos 26\omega + 2 \sum_{k=1}^{25} \cos \omega k) - 2j \sum_{k=1}^{25} \sin \omega k - j \sin 26\omega \right)$$

$$F_2(\omega) = \frac{1}{52} \left(1 + \sum_{k=1}^{51} \cos \omega k - j \sum_{k=1}^{51} \sin \omega k \right)$$

Le filtre F_2 introduit un déphasage important par rapport à celui de F_1 , mais ces deux filtres sont équivalents du point de vue de la séparation fréquentielle, en particulier ils ont la même fréquence d'atténuation maximum pour la fondamentale $\frac{1}{52}$ que l'on veut extraire.

L'application d'un filtre numérique à une série stationnaire X dont la densité spectrale de puissance est $f_X(\omega)$, donne une nouvelle série Y telle que $f_Y(\omega) = |f(\omega)|^2 \cdot f_X(\omega)$ où $f(\)$ est la fonction de transfert du filtre

Plus le carré du module de la fonction de transfert est proche de 1 mieux est le filtrage.

La routine SEPAR propose F_1 et F_2 en option. On pourrait être tenté, pour simplifier le programme, de retenir seulement F_2 : en effet en appliquant F_2 la série lissée comporterait 51 valeurs en moins au début, ce qui à priori ne gênerait pas pour la mesure et l'extrapolation des cycles. Par contre si on appliquait F_1 , on perdrait 26 valeurs au début et 25 à la fin de la série représentant la tendance; ces dernières sont nécessaires à l'extrapolation des cycles.

Mais cette simplification introduirait un trop grand biais dans l'extrapolation des cycles en perdant l'information concernant 51 semaines. Il est donc nécessaire de reconstituer les valeurs perdues.

Reconstitution des valeurs perdues

$$T_1, T_2, \dots, T_{t-1}$$

Soit à reconstituer T_t

Si on suppose que T suit une loi parabolique alors

$$T_t - T_{t-3} = 3 (T_{t-1} - T_{t-2})$$

$T_t = T_{t-3} + 3(T_{t-1} - T_{t-2})$ pour les valeurs à reconstituer en fin de série.

$$T_{t-3} = T_t - 3(T_{t-1} - T_{t-2})$$

Dans le cas de la moyenne mobile quasi centrée, 26 valeurs sont à reconstituer au début de la série et 25 valeurs à la fin de la série.

Dans le cas de la moyenne mobile décentrée, 51 valeurs sont à reconstituer au début de la série.

En pratique on préfère utiliser l'extrapolation exponentielle.
La transformation logarithmique.

On pourrait, dans le cas d'une série présentant un trend plus que linéaire et un cycle d'amplitude variable, envisager la transformation logarithmique avant le filtrage par moyenne mobile; cette transformation est utile dans le cas d'une évolution exponentielle de la charge.

$$D_t = Ae^{bt+c} \cos \omega t$$

en passant aux logarithmes on obtient :

$$\log D_t = \log A + bt + c \cos t$$

L'application d'une moyenne mobile de longueur

$$l = \frac{2\pi}{\omega} \text{ réalisera la séparation :}$$

$$\log T(t) = a + bt$$

$$\log C_t = C \cos \omega t$$

$$T(t) = e^{a+bt}$$

$$C_t = e^C \cos \omega t$$

$$D(t) = T(t) * C(t)$$

II.3- EXTRAPOLATION DE LA TENDANCE

1/ Spécification du modèle

On retient le modèle parabolique pour sa souplesse : 3 paramètres à estimer; en outre la parabole peut-être considérée comme le développement de Taylor limité à l'ordre 2 pour n'importe quelle courbe; les variables explicatives $1, t, t^2$ sont additives alors qu'avec un modèle exponentiel par exemple il faudrait calculer le logarithme des variables et utiliser des densités Log-Normales.

$$Y = a + bt + Ct^2 + U_t$$

Il s'agit d'estimer a, b et c . Les O.L.S. (Ordinary Least Squares) ne peuvent être appliqués du fait de l'interdépendance des erreurs. On utilisera donc les G.L.S. (Generalized Least Squares).

2/ Estimation

De façon générale le modèle s'écrit

$$Y = X\beta + U$$

Supposons que les erreurs sont générés par un processus auto-régressif du 1er ordre

$$U_t = \rho U_{t-1} + \varepsilon_t$$

où $|\rho| < 1$

$$E(\varepsilon_t) = 0$$

$$E(\varepsilon_t \cdot \varepsilon_{t+s}) \begin{cases} = \sigma_\varepsilon^2 & \text{si } s = 0 \\ = 0 & \text{si } s \neq 0 \end{cases} \quad \left. \vphantom{E(\varepsilon_t \cdot \varepsilon_{t+s})} \right\} V_t$$

$$\varepsilon \rightsquigarrow N(0, \sigma_\varepsilon^2)$$

Nous avons alors :

$$\begin{aligned}
u_t &= \rho u_{t-1} + \varepsilon_t \\
&= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \dots \\
&= \varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots + \rho^j \varepsilon_{t-j}
\end{aligned}$$

$$\left. \begin{aligned}
u_t &= \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j} \\
E(\varepsilon_t) &= 0
\end{aligned} \right\} \Rightarrow E(u_t) = 0 \quad \forall t$$

$$E(u_t^2) = E(\varepsilon_t^2) + \rho^2 E(\varepsilon_{t-1}^2) + \rho^4 E(\varepsilon_{t-2}^2) + \dots$$

$$E(u_t^2) = \sigma_{\varepsilon}^2 (1 + \rho^2 + \rho^4 + \dots)$$

$$\sigma_{u_t}^2 = \frac{\sigma_{\varepsilon}^2}{1 - \rho^2} \quad \forall t$$

$$\begin{aligned}
E(u_t \cdot u_{t-1}) &= E[(\varepsilon_t + \rho \varepsilon_{t-1} + \rho^2 \varepsilon_{t-2} + \dots)(\varepsilon_{t-1} + \rho \varepsilon_{t-2} + \rho^2 \varepsilon_{t-3} + \dots)] \\
&= E\{[\varepsilon_t + \rho(\varepsilon_{t-1} + \rho \varepsilon_{t-2} + \dots)](\varepsilon_{t-1} + \rho \varepsilon_{t-2} + \dots)\} \\
&= \rho E[(\varepsilon_{t-1} + \rho \varepsilon_{t-2} + \dots)^2]
\end{aligned}$$

$$E(u_t \cdot u_{t-1}) = \rho \sigma_u^2$$

De même on aura :

$$E(u_t \cdot u_{t-2}) = \rho^2 \sigma_u^2$$

$$E(u_t \cdot u_{t-5}) = \rho^5 \sigma_u^2$$

Remarque : La relation précédente peut encore s'écrire

- 31 -

$$\frac{E(U_t \cdot U_{t-s})}{\sigma_u^2} = \rho^s$$

Cette nouvelle expression calculée pour toutes les valeurs de s représente le corrélogramme de la série U_t .

Le coefficient d'ordre 0 est égal à 1 pour n'importe quelle série. Pour une série purement aléatoire tous les coefficients d'ordre supérieur sont nuls.

Revenons à la matrice des variances covariances des erreurs U_t - En regroupant les expressions précédentes on obtient :

$$E(UU') = V = \frac{\sigma_e^2}{1-\rho^2} \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & & & \rho^{n-2} \\ \vdots & & \ddots & & \vdots \\ \rho^{n-1} & \rho^{n-2} & \dots & \rho & 1 \end{pmatrix}$$

Si ρ est connu, on a immédiatement la solution BLUE :

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y$$

$$\sigma^2(\hat{\beta}) = (X'V^{-1}X)^{-1}$$

avec

$$V^{-1} = \frac{1}{\sigma_e^2} \begin{pmatrix} 1 & -\rho & 0 & \dots \\ -\rho & 1+\rho^2 & -\rho & \\ & -\rho & 1+\rho^2 & -\rho \\ 0 & & & \dots & -\rho & 1 \end{pmatrix}$$

Mais malheureusement ρ n'est pas connu; on utilisera alors une procédure itérative qui estimera à la fois ρ et $\hat{\beta}$

- 1/ Une régression O.L.S. (Routine **MESP**) sur les variables initiales fournira une suite de résidus U_t .
- 2/ Une régression O.L.S. de U_t sur U_{t-1} donne une première estimation de ρ : $U_t = U_{t-1} + \varepsilon_t$
- 3/ Cette première estimation de ρ permet de calculer les variables transformées \tilde{Y}_1, \tilde{X}_1 et \tilde{Y}_t, \tilde{X}_t
- 4/ Une régression O.L.S. de \tilde{Y}_t sur \tilde{X}_t (le modèle $\tilde{Y}_t = \tilde{X}_t b + \varepsilon_t$ vérifiant les hypothèses O.L.S.) $\rightarrow \hat{b}$
- 5/ $\hat{b} \rightarrow$ nouvelle suite de résidus aller en 2.

$\tilde{Y}_t = \tilde{X}_t b + \varepsilon_t$ doit vérifier les hypothèses O.L.S.
 quelle transformation appliquer à Y et X pour obtenir \tilde{Y}_t et \tilde{X}_t vérifiant O.L.S. ?

Soit P cette transformation.

$$PY = PX + PU$$

On doit avoir :

$$E(PUU'P') = I$$

Considérons la matrice P_1 (n X n)

$$P_1 = \begin{pmatrix} \sqrt{1-\rho^2} & 0 & \dots & 0 \\ -\rho & 1 & & \\ 0 & -\rho & 1 & \\ & & & \ddots \\ 0 & & & -\rho & 1 \end{pmatrix}$$

$$E(P_1 U U' P_1') = \sigma^2 I \text{ si bien que :}$$

$$\tilde{Y}_t = P_1 Y = \tilde{X}_t \beta + \varepsilon_t \rightsquigarrow \text{O.L.S}$$

Remarques:

1/ Lorsque l'estimation de ρ fourni par regression sur les résidus est supérieure à 1, on la remplace par la 1ere valeur du corrélogramme des U_t corrigée du biais. Si cette valeur est encore supérieure à 1 on supprime la correction de biais (ROUTINE ROS2). Le facteur de correction du biais est $(1 + \frac{3}{N})$. D'une façon générale le facteur de correction du biais est $\frac{K}{N}$ où K est le nombre de paramètres du modèle et N le nombre de résidus utilisés pour estimer la 1ere valeur du corrélogramme.

$$\sum_{t=2}^N \hat{U}_t \cdot \hat{U}_{t-1} / N-1$$

soit $r = \frac{\sum_{t=2}^N \hat{U}_t \cdot \hat{U}_{t-1}}{N-1}$

$$\sum_{t=2}^N \hat{U}_t^2 / N$$

2/ On peut estimer de 2 façons différentes σ_{ε}^2

$$\sigma_{\varepsilon}^2 = \frac{1}{N-1} \sum_{i=1}^N (U_t - \rho U_{t-1})^2$$

d'autre part σ_{ε}^2 est estimé sans biais par

$$S = \frac{1}{N-3} \sum_{i=1}^N (T_i - FM)^2$$

où T_i = ligne de tendance transformée ; FM = parabole ajustée

La comparaison de ces 2 valeurs permettra d'apprécier le modèle retenu.

En appliquant le processus itératif décrit plus haut on obtient les estimations :

$$\hat{a}, \hat{b}, \text{ et } \hat{c} \text{ ainsi que les variances :}$$

$$\sigma_{\hat{\epsilon}}^2, \sigma_u^2 = \sigma_{\epsilon}^2 / (1 - \hat{\rho}^2)$$

et la matrice de variance-covariance des coefficients $\hat{a}, \hat{b}, \hat{c}$.

Matrice de variance-covariance des coefficients :

$$\begin{bmatrix} \sigma_{\hat{a}}^2 & \text{COV}_{\hat{a}, \hat{b}} & \text{COV}_{\hat{a}, \hat{c}} \\ \text{COV}_{\hat{a}, \hat{b}} & \sigma_{\hat{b}}^2 & \text{COV}_{\hat{b}, \hat{c}} \\ \text{COV}_{\hat{a}, \hat{c}} & \text{COV}_{\hat{b}, \hat{c}} & \sigma_{\hat{c}}^2 \end{bmatrix}$$

La prévision pour U_t est : $E(\hat{U}_t / U_0) = \rho^t \hat{U}_0$

La prévision pour Y_t est alors :

$$T_t = \hat{a} + \hat{b}t + \hat{c}t^2 + \rho^t \hat{U}_0$$

Cette expression montre que le dernier écart U_0 entre la ligne de tendance et la parabole est progressivement amorti : la prévision se raccorde exponentiellement à la parabole estimée par les G.L.S.

Variance de l'erreur de prévision

$$Y_t = a + bt + ct^2 + \rho^t U_0 + \sum_{j=0}^{t-1} \rho^j \epsilon_{t-j}$$

$$Y_t = \hat{a} + \hat{b}t + \hat{c}t^2 + \rho^t \hat{U}_0$$

$$(Y_t - \hat{Y}_t) = (a - \hat{a}) + (b - \hat{b})t + (c - \hat{c})t^2 + \rho^t (U_0 - \hat{U}_0) + \sum_{j=0}^{t-1} \rho^j \epsilon_{t-j}$$

Or on a : $U_0 = Y_0 - a$

$$\frac{U_0 = Y_0 - \hat{a}}{U_0 - \hat{U}_0 = \hat{a} - a}$$

$$Y_t - \hat{Y}_t = (a - \hat{a}) (1 - \rho^t) + (b - \hat{b}) t + (c - \hat{c}) t^2 + \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}$$

$$Y_t - \hat{Y}_t = A + B$$

En supposant que le terme $B = \sum_{j=0}^{t-1} \rho^j \varepsilon_{t-j}$ est

indépendant de A, on peut calculer la variance de l'erreur de prévision en ajoutant les variances de A et B :

$$E (Y - \hat{Y})^2 = \sigma_A^2 + \sigma_B^2$$

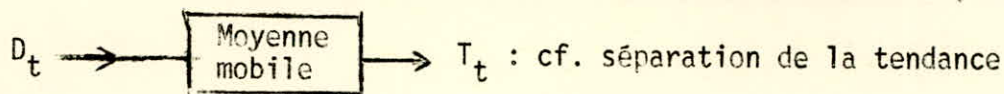
$$\sigma_A^2 = \begin{pmatrix} (1 - \rho^t \ t \ t^2) \begin{pmatrix} \sigma_{\hat{a}}^2 & \text{cov}_{\hat{a}, \hat{b}} & \text{cov}_{\hat{a}, \hat{c}} \\ \text{cov}_{\hat{a}, \hat{b}} & \sigma_{\hat{b}}^2 & \text{cov}_{\hat{b}, \hat{c}} \\ \text{cov}_{\hat{a}, \hat{c}} & \text{cov}_{\hat{b}, \hat{c}} & \sigma_{\hat{c}}^2 \end{pmatrix} \begin{pmatrix} 1 - \rho^t \\ t \\ t^2 \end{pmatrix} \end{pmatrix}$$

$$\sigma_B^2 = \sum_{j=0}^{t-1} \rho^{2j} \sigma_{\varepsilon}^2 = \sigma_{\varepsilon}^2 \sum_{j=0}^{t-1} \rho^{2j}$$

$$\sigma_B^2 = \sigma_{\varepsilon}^2 \frac{1 - \rho^{2t}}{1 - \rho^2} = \sigma_u^2 (1 - \rho^{2t})$$

σ_B^2 Croît exponentiellement vers une asymptote égale à σ_u^2

1 - Détermination des coefficients hebdomadaires (*routine CFTE*)



$$C_{i,j} = \frac{D_t}{T_t}$$

Les cycles sont exprimés en facteurs multiplicatifs par rapport à la ligne de tendance.

Un cycle est représenté par une suite de 52 coefficients multiplicatifs.

Le cycle annuel se décompose en 2 cycles :

- L'un sensible à la température, la température est alors introduite comme variable exogène pour expliquer une partie des variations observées au cours de l'année.
- L'autre insensible à la température.

$$C = C_S + C_I$$

La liaison entre le cycle et la météorologie est approximée par le modèle :

$$C_{S_{i,j}} = \begin{cases} KW_i (W_{i,j} - TW) & \text{Si } W_{i,j} < TW \\ 0 & \text{Si } TW \leq W_{i,j} \leq TS \\ KS_i (W_{i,j} - TS) & \text{Si } W_{i,j} > TS \end{cases}$$

Les coefficients KW_i et KS_i représentent la sensibilité du cycle à la température en hiver et en été

$$KW_i = \frac{\sum_j (W_{i,j} - TW) (C_{i,j} - X^M)^2}{\sum_j (W_{i,j} - TW)^2} \quad \text{Si } W_{i,j} < TW$$

$$KS_i = \frac{\sum_j (W_{i,j} - TS) (C_{i,j} - XM)}{(W_{i,j} - TS)^2} \quad \text{Si } W_{i,j} > TS \quad - 37 -$$

où XM = moyenne des valeurs du cycle observées pour $W \in [TW, TS]$ si $TW \neq TS$

XM = moyenne des valeurs du cycle observées pour $W \in [TW-2, TS+2]$ si $TW = TS$

On peut vérifier par l'intermédiaire de ce modèle que la sensibilité du cycle à la température évolue d'année en année. Cette évolution peut-être extrapolée pour les années futures par la moyenne, la droite ou la parabole de régression ajustée sur la suite des valeurs KW_i et KS_i

On suppose KW_i et KS_i certains, on ne tiendra donc pas de leur variance d'estimation et d'extrapolation; mais en fait cette variance est reportée, sur celle du cycle insensible, car ce dernier contiendra les erreurs d'estimation sur le cycle sensible et les aléas qui lui sont propres.

On a en effet :

$$CI_{i,j} = C_{i,j} - CS_{i,j}$$

On extrapolera le cycle insensible par sa moyenne hebdomadaire

$$CIBAR_j = \frac{1}{N} \sum_{i=1}^N CI_{i,j} \quad j = 1 \dots 52$$

Variance :

$$CIVAR_j = \frac{1}{N-1} \sum_{i=1}^N (CI_{i,j} - CIBAR_j)^2$$

(Cf. Routine CYSB)

II.5. - INDEPENDANCE ENTRE LA TENDANCE ET LE CYCLE

- 38 -

(Subroutine CRLN)

Il est nécessaire de vérifier l'indépendance entre la tendance et le cycle total d'une part, et le cycle sensible et le cycle insensible d'autre part, ceci afin d'extrapoler séparément chacune de ces composantes et de recombinaison ces dernières pour la prévision de la charge.

1/ Correlation entre la tendance et le cycle total

T : tendance non stationnaire

C : cycle total stationnaire

Soit $R_t = T_t / FM_t$ où t est l'indice des semaines et FM_t le point d'abscisse t , puis sur la parabole ajustée à la tendance.

$R (R_1, R_2, \dots, R_t \dots R_N)$

$C (C_1, C_2, \dots, C_t \dots C_N)$

Le coefficient de corrélation r entre R et C est donné par ..

$$r_{(R,C)} = \frac{\text{COV}(\vec{R} - \vec{m}_R, \vec{C} - \vec{m}_C)}{\sqrt{V(R) \cdot V(C)}}$$

avec $\vec{m}_R (\sum R_t/N, \sum R_t/N \dots \dots \dots \sum R_t/N)$

$m_C (\sum C_t/N, \sum C_t/N \dots \dots \dots \sum C_t/N)$

On pose $m_R = \sum R_t/N$

$m_C = \sum C_t/N$

r devient :

$$r(\vec{R}, \vec{C}) = \frac{\vec{R} \cdot \vec{C} - Nm_R \cdot m_C}{\sqrt{(R^2 - Nm_R^2)(C^2 - Nm_C^2)}}$$

2 - Correlation entre CI et CS

CI (CI₁ , CI₂ CI_N)

CS (CS₁ , CS₂ CS_N)

On pose $m_1 = \sum CI/N$

$$m_2 = \sum CS/N$$

$$r(\vec{CI}, \vec{CS}) = \frac{\vec{CI} \cdot \vec{CS} - Nm_1 \cdot m_2}{\sqrt{(CI^2 - Nm_1^2)(CS^2 - Nm_2^2)}}$$

II.6 - EXTRAPOLATION DES CYCLES

1/ Extrapolation des paramètres KW, KS (routine **EXTR**)

On extrapole les sensibilités KW, et KS de façon à tenir compte de l'évolution de l'influence de la météorologie sur la charge.

3 types d'extrapolation

- extrapolation par la moyenne observée sur les années successives (la longueur de cette moyenne sera de plus en plus grande) :
- Extrapolation linéaire
- Extrapolation parabolique

Donc on a en sortie un couple de coefficients pour chaque année future.

2/ Extrapolation de la composante sensible du cycle (routine **EXCS**)

On a supposé, dans l'analyse des cycles qu'il y a une relation déterministe entre le cycle sensible et la température : les coefficients KW et KS sont inconnus mais certains; ainsi la variance du cycle sensible provient uniquement de la variance de la température.

Cette dernière est supposée normale, elle est représentée dans l'avenir par 52 moyennes (WBAR) et 52 variances (WVAR) correspondants aux 52 semaines de l'année (Routine **CYSE**).

La densité du cycle sensible, n'est pas gaussienne (elle résulte d'une transformation non linéaire d'une variable gaussienne).

L'espérance $CS(\bar{x})$ et la variance $VS(\bar{x})$ de cette composante sont obtenues en calculant sa densité PCS de la manière suivante :

(Routines *PROB*, *CSNX*, *PRCS*, *FNCS*)

- * Calcul des valeurs extrêmes de la température :
 TMIN et TMAX : températures ayant une probabilité 0,01 d'être dépassées. Ceci pour chaque semaine à prévoir.
- * Calcul des valeurs extrêmes de la composante sensible SMIN, SMAX qui dépendent de TMIN, TMAX, des sensibilités KW et KS et de la position de TMIN, TMAX par rapport à TW et TS
- * Pour toutes les valeurs So telles que $SMIN \leq So \leq SMAX$ on calcule $Prob(S < So)$. Cette probabilité dépend de KW et KS (du signe de ces coefficients surtout) et des températures :

$$T_1 = S_1^{-1} (S_0) \quad \text{pour } T \leq TW$$

$$T_2 = S_2 (S_0) \quad \text{pour } T \geq TS$$

S(T) désignant la liaison cycle - température (voir schémas ci-contre).

L'intervalle (SMIN, SMAX) est divisé en 20 intervalles; la suite des probabilités représente la fonction de répartition du cycle sensible.

La densité de probabilité PCS est obtenue en mesurant l'accroissement de la fonction de répartition en chacun des 20 points de l'intervalle (SMIN, SMAX).

La moyenne CSBAR et la variance CSVAR sont alors :

$$CSBAR = \sum_{k=1}^{20} \frac{CS_K + CS_{K+1}}{2} * PCS_K$$

$$CSVAR = \sum_{K=1}^{20} \left(\frac{CS_K + CS_{K+1}}{2} - CSBAR \right)^2 * PCS_K$$

$i=1$

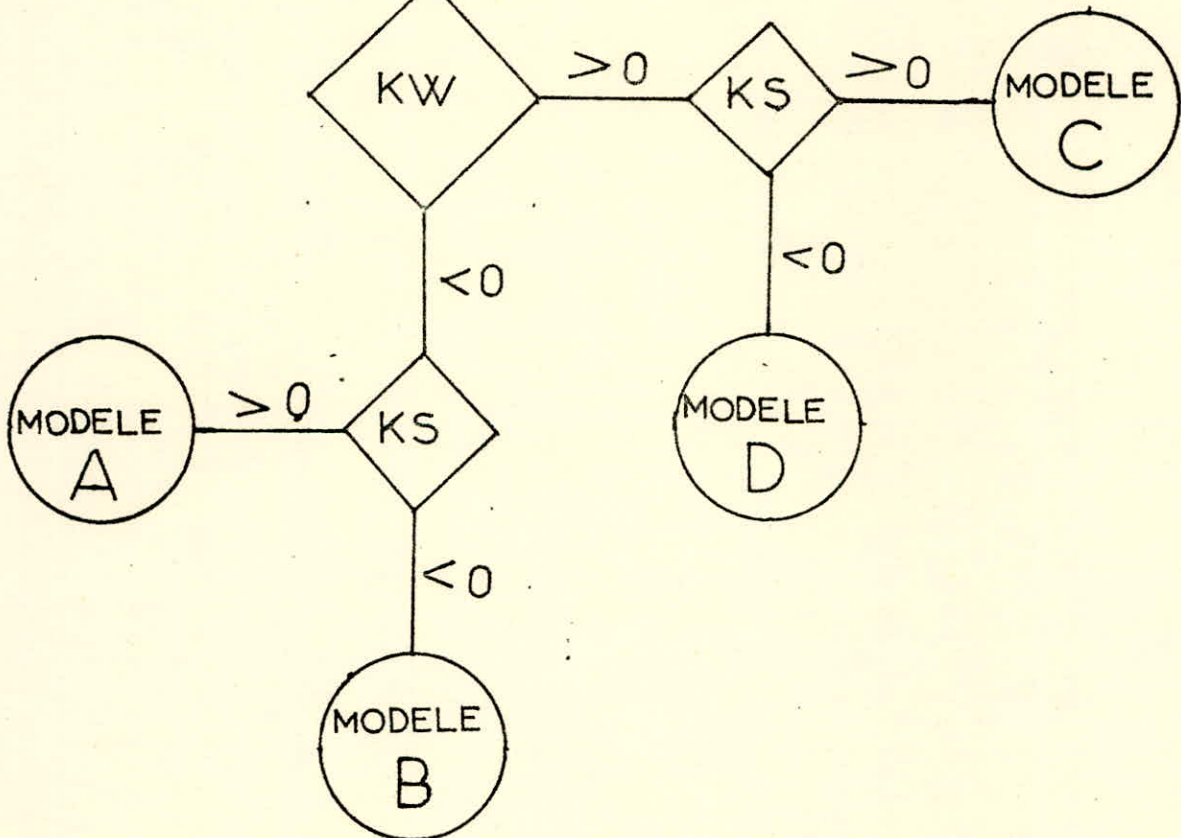
1^{ère} année à prévoir

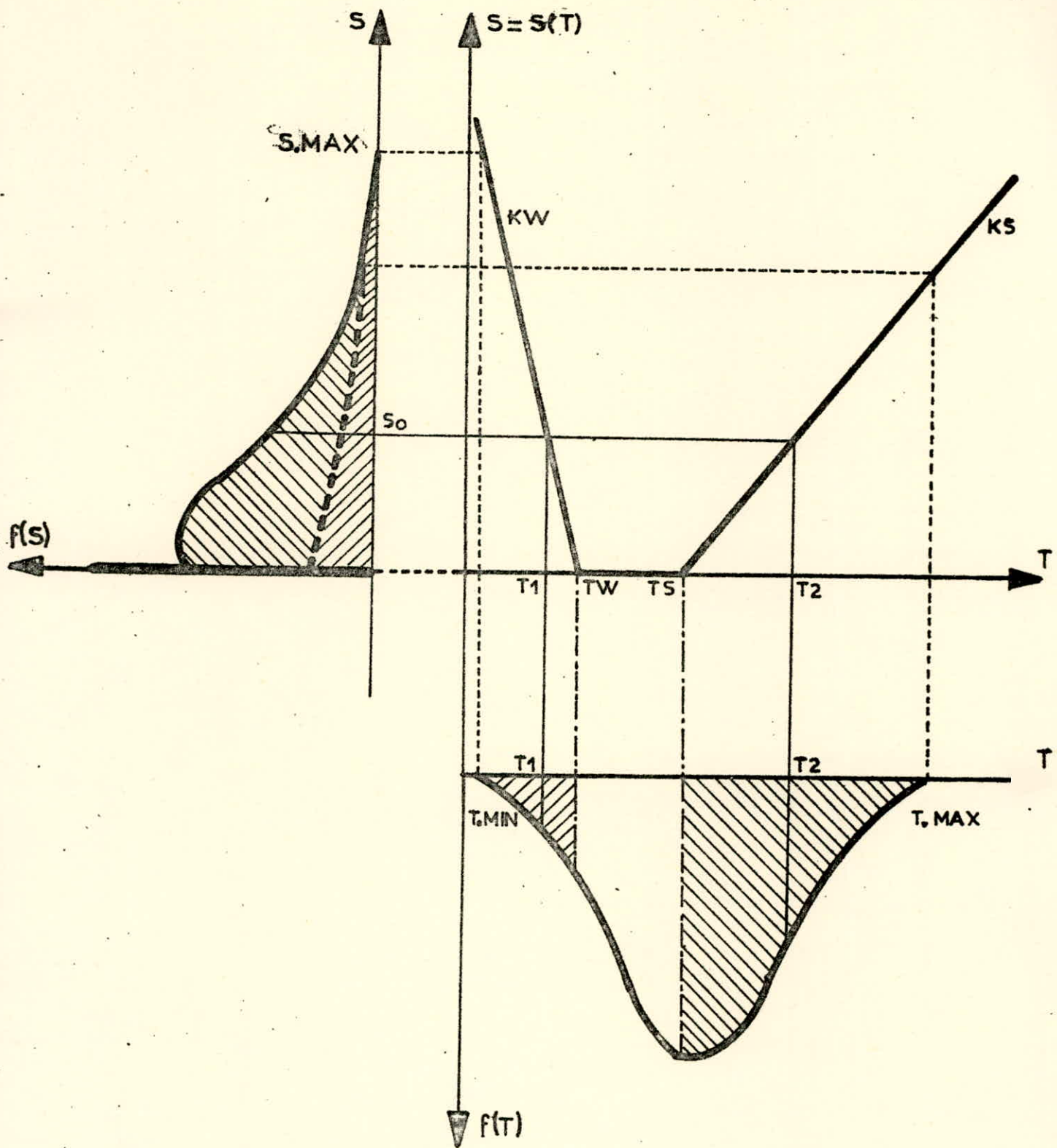
CALCULER
KS(i) et KW(i)
routine EXTR

J=1

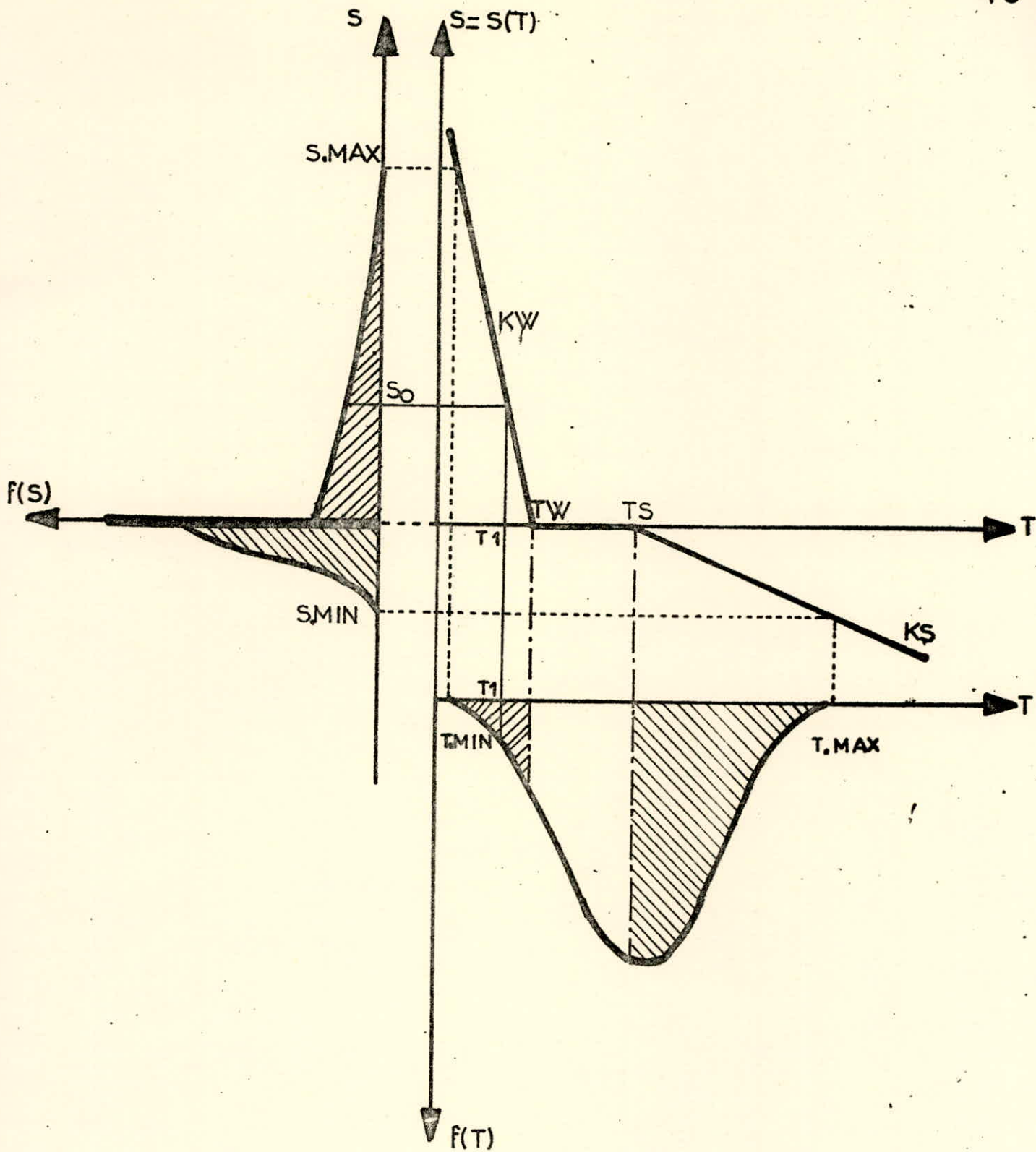
1^{ère} semaine de la l^{ème} année

CALCULER
TMIN(i)
TMAX(i)
cf routine CSNX



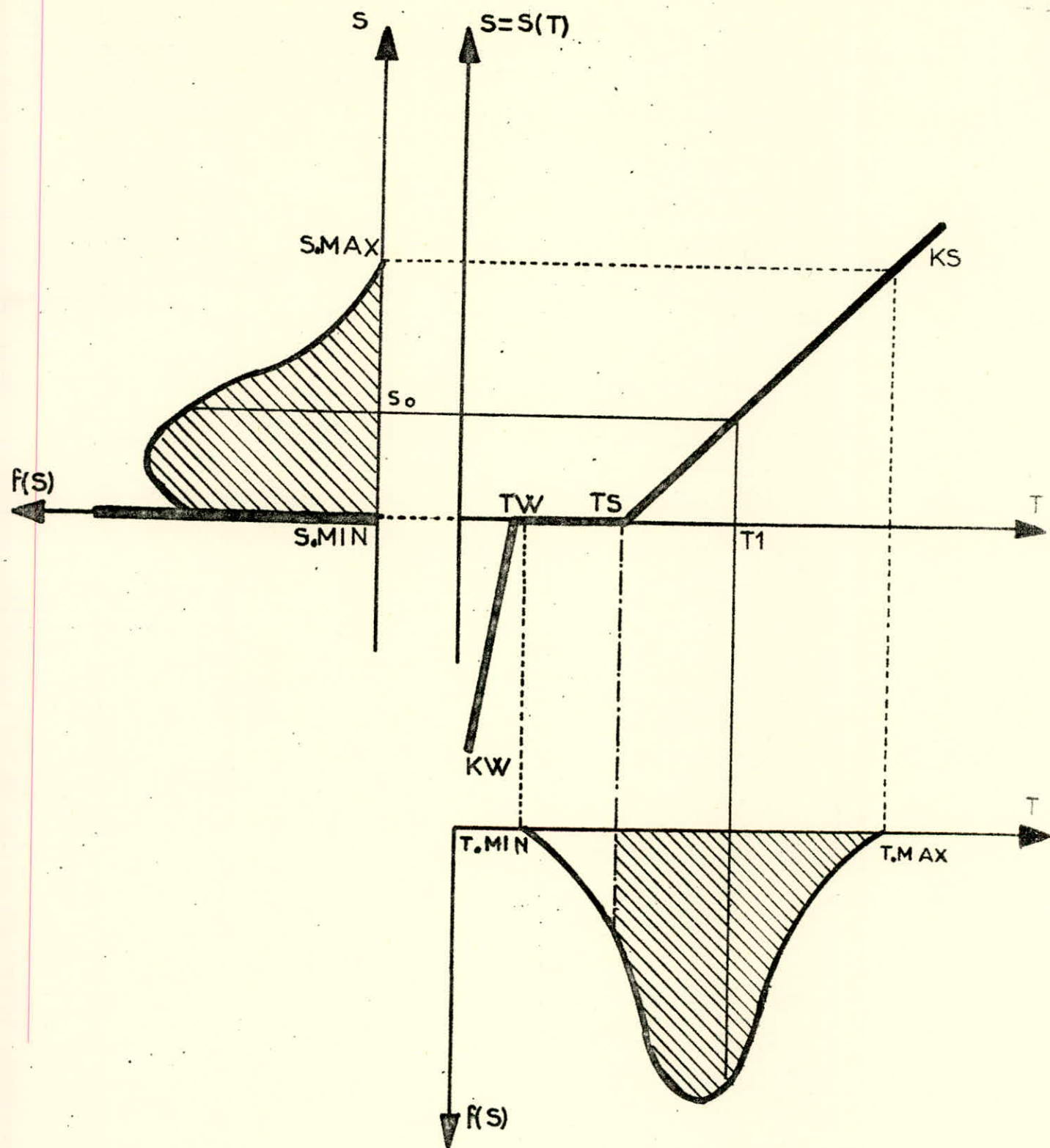


MODELE A : $K_W < 0$ $K_S > 0$
 $Prob(S < S_0) = Prob(T_1 < T < T_2)$

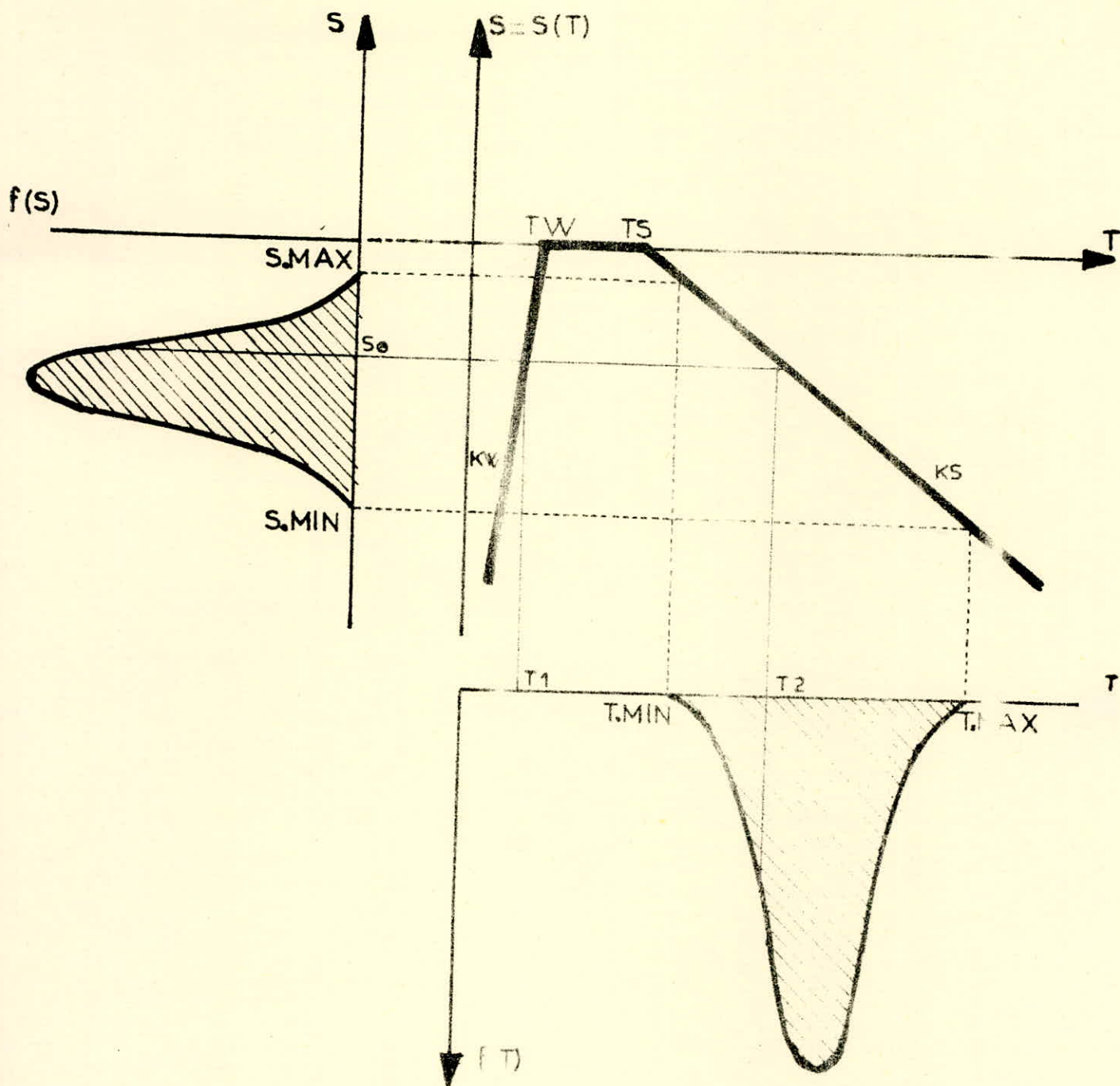


MODELE B : $KW < 0$ $KS < 0$

$Prob(S < S_0) = Prob(T > T_1)$



MODELE C : $KW > 0$ $KS > 0$
 $\text{Prob}(S < S_0) = \text{Prob}(T < T1)$



MODELE D: $K_W > 0$ $K_S < 0$

$$\text{Prob}(S < S_0) = \text{Prob}(T < T_1) + \text{Prob}(T > T_2)$$

PCS_K est supposée concentrée au milieu de l'intervalle $(K, K+1)$. Plus K est grand, mieux est l'approximation. (voir organigramme général ci-contre).

3/ Le cycle insensible

Ce cycle n'a pas de modification importante d'une année à l'autre.

On l'extrapole par un cycle insensible moyen représenté par 52 moyennes $CIBAR_j$ et 52 variances $CIVAR$ ($j = 1 \dots 52$)

$$CIBAR_j = \frac{1}{N} \sum_{i=1}^N CIBAR_{ij} \quad N \text{ étant le nombre d'années d'analyse.}$$

$$CIVAR_j = \frac{1}{N-1} \sum_{i=1}^N (CI_i - CIBAR_j)^2$$

Les valeurs $PCI_K = \text{Prob} (CI_K \leq CI + < CI_{K+1})$

sont tirées d'une table de la loi Normale.

$$C = CS + CI$$

Si les densités de 2 variables aléatoires indépendantes sont connues analytiquement, la densité de leur somme est donnée par :

Soient X_1 de densité $f_1(x)$

X_2 de densité $f_2(x)$

Alors la densité de $Z = Y_1 + Y_2$ est $f_Z(y) =$

$$\int_{-\infty}^{+\infty} f_1(x) f_2(y-x) dx$$

Dans notre cas les densités de eS et CI ne sont pas analytiques.

Elles se présentent sous la forme d'une suite de probabilités élémentaires.

Composantes sensible :

$$PCS_i = \text{Prob} (CS_i \leq CS * \leq CS_{j+1})$$

Composante insensible :

$$PCI_j = \text{Prob} (CI_j \leq CI * CS_{i+1})$$

On réalisera la convolution de ces 2 densités par la méthode suivante :

1/ Associer les probabilités PCS_i et PCI_j aux centres des intervalles correspondants

$$PCS_i = \text{Prob} (CS * = \frac{CS_{i+1} + CS_{i+1}}{2})$$

$$PCI_j = \text{Prob} (CI * = \frac{CI_j + CI_{j+1}}{2})$$

2/ Calculer toutes les valeurs possibles de la somme

$$Z_{ij} = \frac{CS_i + CS_{i+1}}{2} + \frac{CI_j + CI_{j+1}}{2}$$

Les probabilités associées sont :

$$P_{Z_{ij}} = PCS_i * PCI_j$$

Si la densité de CS est définie en NS points et celle de CI en NI points, alors celle de $Z = CS + CI$ sera calculée en $NZ = NS \times NI$ points.

3/ Condenser les valeurs calculées pour Z et PZ en divisant l'intervalle des valeurs possibles (Z_{MIN}, Z_{MAX}) en N sous-intervalles :

$$Z_1 \ Z_2 \ \dots \ Z_K \ \dots \ Z_{N+1}$$

La probabilité élémentaire pour chacun de ces intervalles sera calculée en additionnant les PZ_{ij} associés aux Z_{ij} appartenant à chaque intervalle on aura ainsi :

$$P_1 \ P_2 \ \dots \ P_K \ \dots \ P_N$$

avec

$$P_K = \text{Prob} (Z \in [Z_K, Z_{K+1}])$$

$$P_K = \sum P_{Z_{ij}} \quad \forall Z_{ij} \in [Z_K, Z_{K+1}]$$

on réalise ces opérations en adoptant $NS = NI = 21$,

$$NZ = 441, N = 20$$

Comme résultat on aura donc la densité du cycle total :

$$PC_K = \text{Prob} (C_K \leq C < C_{K+1})$$

Alors

$$\bar{C} = \sum_{K=1}^{20} \frac{C_K + C_{K+1}}{2} PC_K$$

$$\sigma_C^2 = \sum_{K=1}^{20} \left(\frac{C_K + C_{K+1}}{2} - \bar{C} \right)^2 \cdot PC_K$$

CS CI	CS ₁	CS ₂		CS _j		CS ₂₁
CS ₁	Z ₁₁	Z ₁₂				Z ₁₄
CS ₂						
CS _e				Z _{ek}		
CI ₂₁						Z ₂₁₂₁

$$ZMIN = \min_{e,k} (Z_{ek})$$

$$ZMAX = \max_{e,k} (Z_{ek})$$

à ce tableau correspond un autre tableau, celui des probabilités $P_{Z_{1k}} = PCS_i * PCI_j$

(ZMIN, ZMAX) est divisé en 20 intervalles

$$Z_1 \quad Z_2 \quad \dots \quad Z_K \quad \dots \quad Z_{21}$$

$$Z_K = ZMIN + \frac{ZMAX - ZMIN}{20} \times K$$

$$PC_K = \text{Prob} [Z \in (Z_K, Z_{K+1})]$$

$$PC_K = \sum P_{Z_{ek}} \quad \forall Z_{e,k} \in [Z_K, Z_{K+1}]$$

Prévision de la tendance

$$y_t = \hat{a} + \hat{b}t + \hat{c}t^2 + \rho^t u_0$$

Les espérances des réalisations futures de la tendance s'obtiennent en extrapolant la parabole ajustée sur la ligne de tendance, puis en la corrigeant de l'amortissement du dernier résidu.

A ces espérances futures sont associées des variances résultant de l'incertitude d'estimation des coefficients de la parabole, de l'incertitude d'estimation du dernier résidu.

On suppose que la prévision ainsi constituée suit une loi normale. On peut construire sa densité à partir des percentiles de la loi normale.

$$PT_K = \text{Prob} (T_K \leq T^* < T_{K+1})$$

Prévision de la charge globale

$$D = T + C$$

On possède les densités de T et C, il est possible de calculer la densité de D.

On applique une convolution analogue à celle utilisée pour les cycles :

$$Z_{ij} = \frac{T_i + T_{i+1}}{2} \times \frac{C_j + C_{j+1}}{2}$$

$$PZ_{ij} = PT_i \times PC_j \quad T \text{ et } C \text{ étant supposés indépendants.}$$

C	C ₁		C _K	C ₂₁
T				
T ₁				
T _l			Z _{l,k}	
T ₂₁				

$$D_{MIN} = \min_{e,k} (Z_{e,k})$$

$$D_{MAX} = \max_{1,n} (Z_{e,k})$$

$$D_1 \quad D_2 \quad \dots \quad D_K \quad \dots \quad D_{21}$$

avec

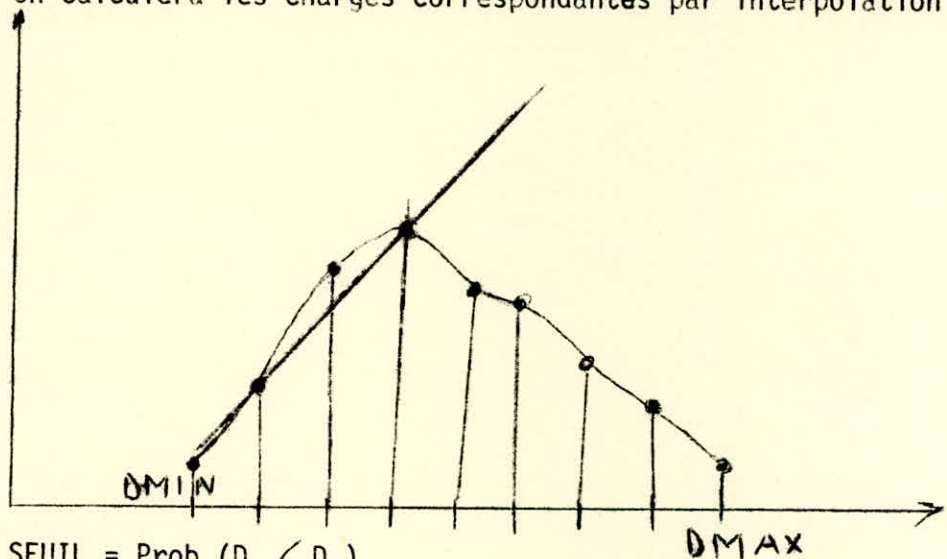
$$D_K = k \cdot \frac{D_{MAX} - D_{MIN}}{20} + D_{MIN}$$

$$PD_K = \sum P_{Z_{e,k}} \quad \forall Z_{e,k} \in [D_K, D_{K+1}]$$

Maintenant qu'on a la densité de la charge totale il serait intéressant d'avoir la fonction de répartition correspondant à certains seuils particuliers :

$$0,01 - 0,10 - 0,50 - 0,90 - 0,95 - 0,99$$

On calculera les charges correspondantes par interpolation linéaire.



$$SEUIL = \text{Prob} (D < D_L)$$

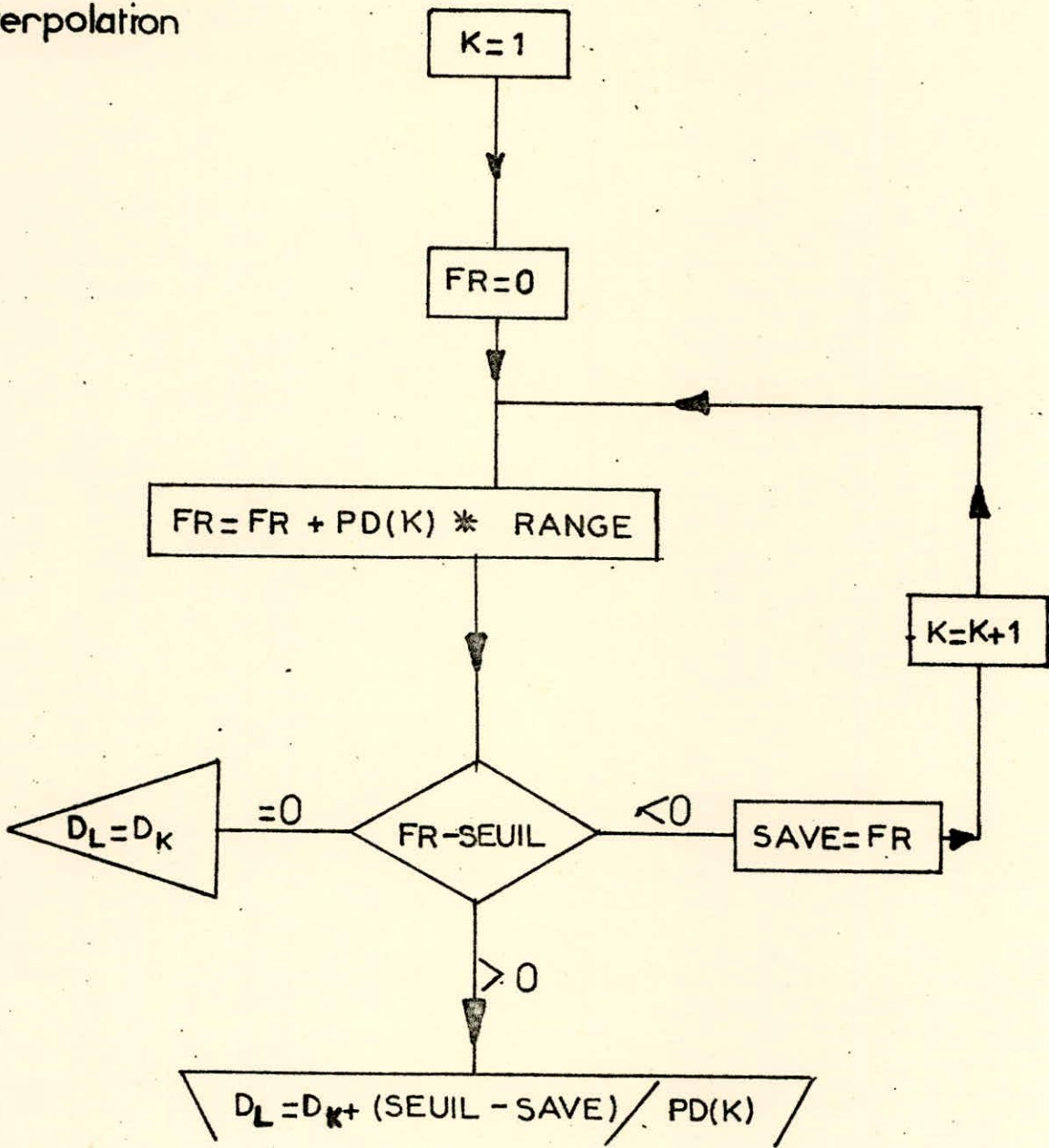
Il s'agit de trouver D_L connaissant le seuil.

Pour cela on calcule la somme cumulée

$$SOM(k) = \sum PD_K \cdot \frac{D_{MAX} - D_{MIN}}{20} \quad \text{qu'on comparera}$$

à chaque pas au seuil.

interpolation



$$SOM(K) = \sum_k PD_k \cdot RANGE$$

- 52 -

Si $SOM(K) = SEUIL$ alors $D_L = D_K$

Si $SOM(K) < SEUIL$ Continuer à Sommer

Si $SOM(K) > SEUIL$ procéder à l'interpolation

VARIANCE DE LA CHARGE

$D = T * C$ où T et C sont supposées indépendantes.

D'une manière générale la variance d'une variable aléatoire $Z = X.Y$ avec X et Y sont indépendantes est :

$$E((X.Y - \bar{Z})^2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x.y - \bar{Z})^2 f_X(x) \cdot f_Y(y) dx dy$$

$$E((X.Y - \bar{Z})^2) = \int_{-\infty}^{+\infty} x^2 f_X(x) dx \int_{-\infty}^{+\infty} y^2 f_Y(y) dy - 2\bar{Z} \int_{-\infty}^{+\infty} x f_X(x) dx \int_{-\infty}^{+\infty} y f_Y(y) dy + \bar{Z}^2$$

$$E((X.Y - \bar{Z})^2) = [E(X^2) + \bar{X}^2] [E(Y^2) + \bar{Y}^2] - \bar{Z}^2$$

$$= (\sigma_X^2 + \bar{X}^2) (\sigma_Y^2 + \bar{Y}^2) - \bar{Z}^2$$

$$= \sigma_X^2 \sigma_Y^2 + \bar{X}^2 \sigma_Y^2 + \bar{Y}^2 \sigma_X^2$$

Dans le cas de la charge on aura donc

$$\sigma_D^2 = \sigma_T^2 \cdot \sigma_C^2 + \bar{T}^2 \sigma_C^2 + \bar{C}^2 \sigma_T^2$$

Le terme central est le plus important.

Toutes ces opérations étant faites pour chaque semaine à prévoir on mémorise par semaine les éléments suivants :

$$\begin{aligned}
 CY &= \text{espérance du cycle insensible} \\
 CS &= \text{espérance du cycle sensible} = \sum_{K=1}^{20} CS_K \cdot PB_K \\
 CP &= \text{espérance du cycle total} = CY + \dots \\
 FM &= \text{espérance de la tendance} \\
 DP &= \text{espérance de la charge globale} = FM \times CP \\
 VY &= \text{variance du cycle insensible} \\
 VS &= \text{variance du cycle sensible} = \sum_{K=1}^{20} (CS_K - CS)^2 \cdot PB_K \\
 VC &= \text{Variance du cycle total} = VY + VS \\
 VT &= \text{VARIANCE DE LA TENDANCE} \\
 VD &= \text{Variance de la charge globale} \\
 &= VT \cdot VC + \overline{FM}^2 \cdot VC + \overline{CP}^2 \cdot VT
 \end{aligned}$$

NIVC (1 → 7) = 7 points de la fonction de répartition de D

0.01, 0.05, 0,10 0,50 0,90, 0,95, 0,99

Nota : Les noms des variables utilisées ci-dessus ne sont pas toujours les mêmes que ceux utilisés dans le programme.

II.8.- DENSITE JOINTE POUR LA POINTE ANNUELLE ET LA SEMAINE OU ELLE SE PRESENTE. - 54 -

Les opérations du II-6 et II-7 sont faites pour chaque semaine à prévoir.

La densité de la charge est mémorisée pour toutes les semaines d'hiver (semaine 49, 50, 51, 52, 1, 2, 3, 4; décembre et janvier) où la pointe annuelle peut se présenter.

Il est alors possible de calculer la densité jointe pour la pointe annuelle et la semaine où elle se présente.

Soient $K = 1 \dots 8$ les semaines considérées

X_K = charge de la semaine K

$$f_K(x) = \text{prob} (X_K \in (x, x + \Delta x))$$

$$F_K(x) = \text{Prob} (X_K < x)$$

$\emptyset(K, x)$ = probabilité que la pointe annuelle se produise à la $k^{\text{ième}}$ semaine et soit comprise entre x et $x + \Delta x$.

La probabilité que la charge soit comprise entre x et $x + \Delta x$ à la $k^{\text{ième}}$ semaine est $f_K(x)$

La probabilité que la charge soit inférieure à x pour toutes les autres semaines vaut

$$\prod_{i \neq K} F_i(x)$$

Si on suppose ces 2 événements indépendants

alors

$$\emptyset(K, x) = f_K(x) \cdot \prod_{i \neq K} F_i(x)$$

$$\emptyset(K) = \int_0^{\infty} \emptyset(K, x) dx$$

= probabilité que la pointe annuelle se produise à la $k^{\text{ième}}$ semaine indépendamment de sa valeur.

$$\Psi(x) = \int_1^8 \emptyset(K, x) dK$$

= probabilité que la pointe annuelle soit comprise entre x et $x + \Delta x$ quelque soit la semaine où elle se présente.

Dans notre cas on aura donc :

Semaines : 49 50 51 52 1 2 3 4

- 55 -

Densités $PD^{(h)}$:

Associées

On sait que la densité $PD^{(h)}$ pour chaque semaine est définie en 20 intervalles - $(PD_k^{(h)} = \text{Prob}(D_k < D < D_{k+1}))$.

Soit $PPIC(k, h) = \text{Probabilité pour que la pointe annuelle soit comprise dans l'intervalle } (k, k+1) \text{ et se produise pendant la semaine } h$

$$\equiv \text{Prob} \left(E_1 \cap E_2 \right) \text{ avec}$$

$E_1 =$ La charge est située entre D_k et D_{k+1} pendant la semaine K

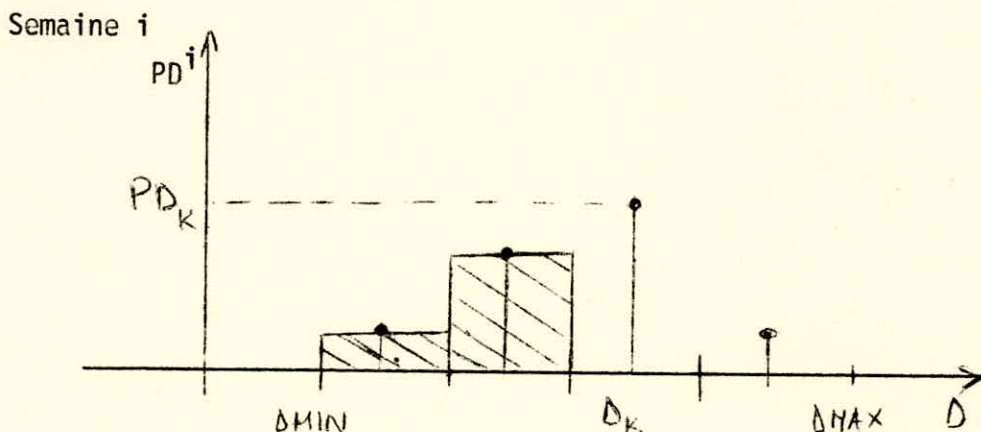
$E_2 =$ La charge est inférieure à D_k pour toutes les autres semaines.

$$PPIC(k, h) = PD_k^{(h)} \cdot \prod_{i \neq h} FR^{(i)}(DK)$$

$h =$ indice des semaines

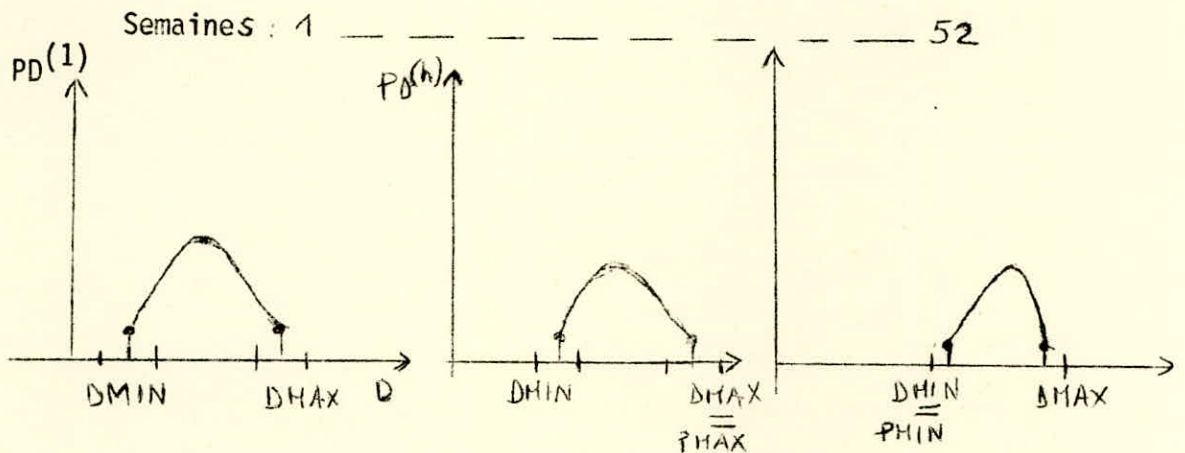
$k =$ indice de séparation des intervalles

$$\begin{aligned} FR^{(i)}(DK) &= \text{Prob}(D < D_k) \text{ pendant la semaine } i \\ &= \sum_{j=1}^k PD_{j-1}^i \end{aligned}$$



Exemple :

$$PPIC(k,50) = PD_k^{50} \cdot FR^{49}(DK) \cdot FR^{51}(DK) \cdot FR^{52}(DK) \cdot FR^{1}(DK) \cdot FR^2(DK) \cdot FR^3(DK) \cdot FR^4(DK)$$



Le graphe de la densité hebdomadaire se déforme et s'éloigne de l'origine selon les semaines.

Pour standardiser les probabilités joints PPIC(k,h) on calcule : les maximum et minimum possible de la pointe.

$$P_{MAX} = \max_h \left[\max_k D(h,k) \right]$$

$$P_{MIN} = \max_h \left[\min_k D(h,k) \right]$$

On divise le range $P_{MAX} - P_{MIN}$ en 20 intervalles $PTE(j) = P_{MIN} + j (P_{MAX} - P_{MIN})/20$ $j = 0 \dots 20$
 $PTE_0, PTE_1 \dots PTE_{20}$

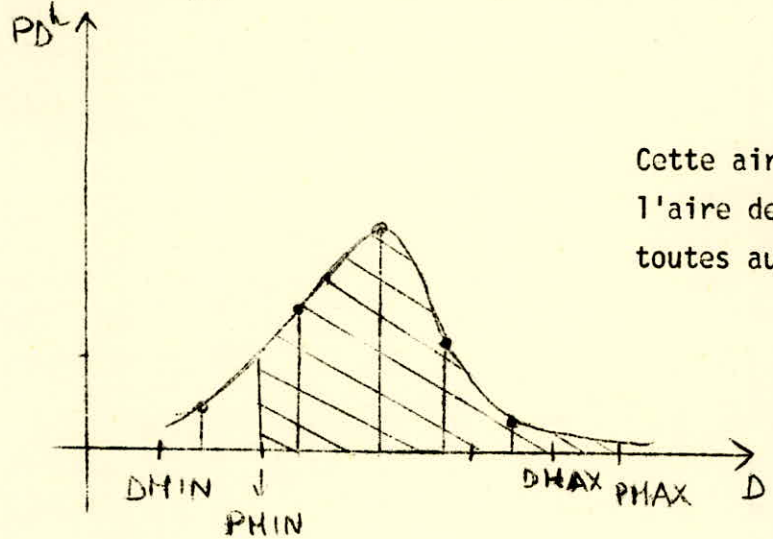
Les densités associées à ces intervalles seront :

$$PPTE^h(j) = \text{Prob} \left(PTE(j) \leq PTE < PTE(j+1) \text{ pendant la semaine } k \right)$$

$$PPTE^h(j) = \sum PPIC(h,k) \quad \forall k \in (j, j+1) -$$

$PPTE^h = \text{Prob (que la pointe se produise durant la semaine h independamment de sa valeur).}$

$$PPTE^h = \sum PPIC (h,k) \quad \forall D_k \in (P_{MIN}, P_{MAX})$$



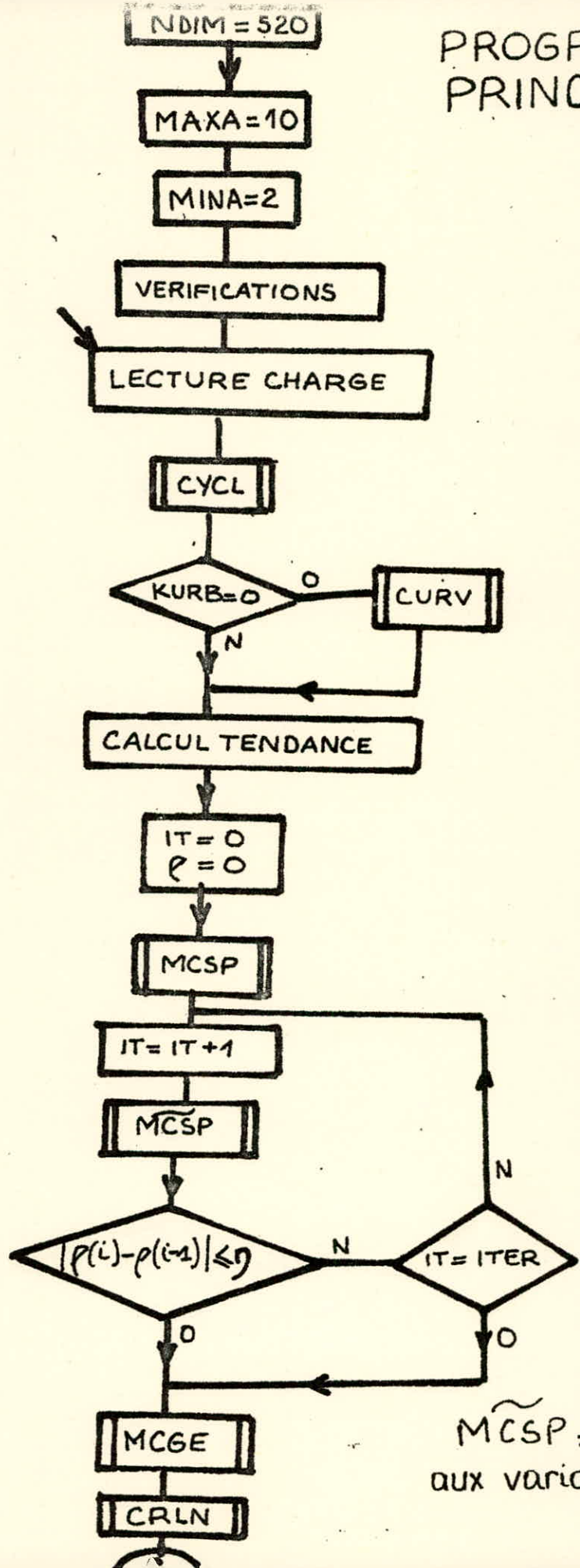
Cette aire est encore multipliée par l'aire des queues de courbe pour toutes autres semaine.

$PPTE(j) = \text{Prob [PTE(j) \leq PTE(j+1) independamment de la semaine.}$

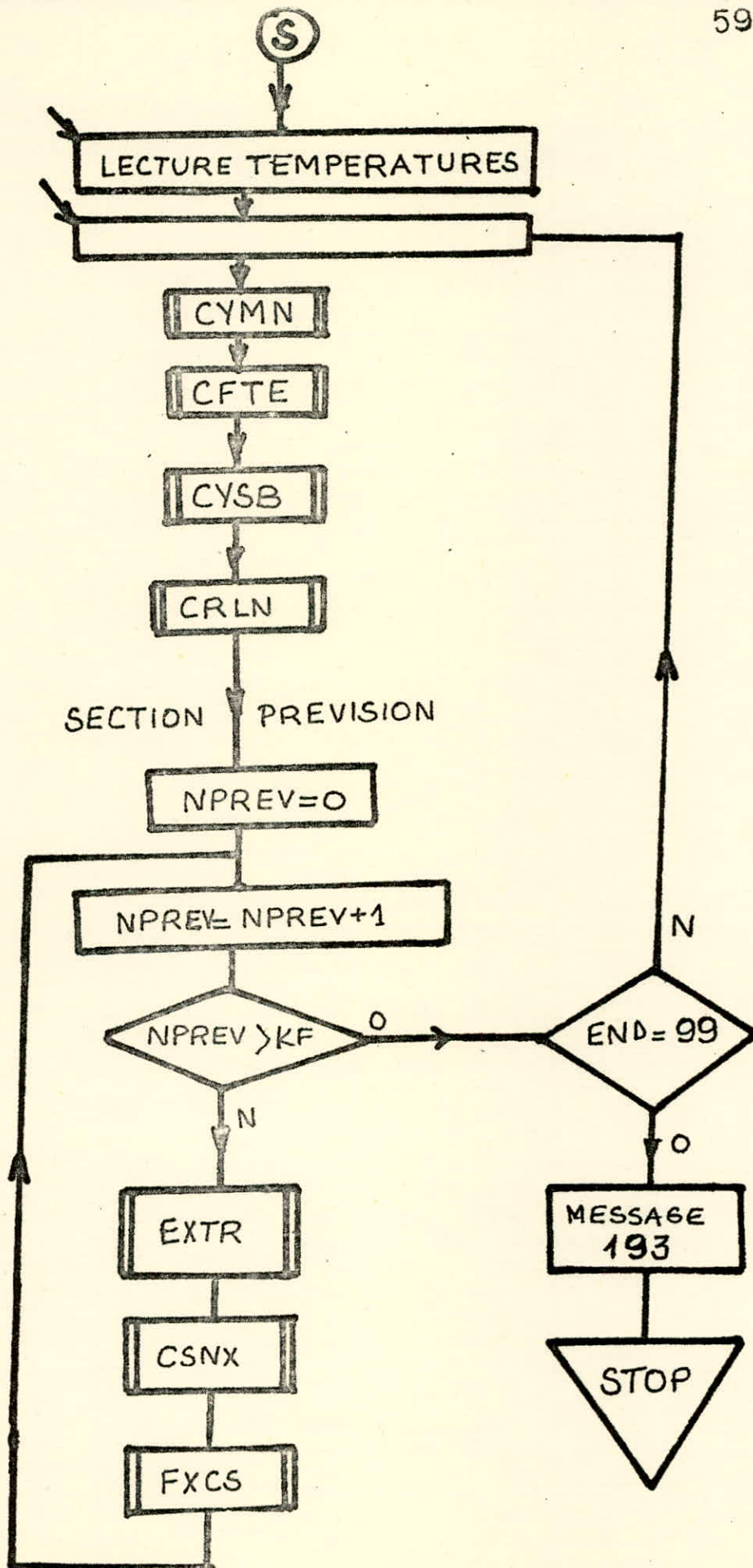
$$PPTE(j) = \sum_{h=1}^8 PPIC (h,k) \quad \forall k \in (j+1, j+1)$$

TROISIEME PARTIE

PROGRAMME PRINCIPAL



\widetilde{MCSP} = O.L.S appliqués aux variables transformées



PROGRAMME PRINCIPAL DECLARATIONS.

Integer Y,Z :

Y : numéro logique du lecteur de cartes

Z : numéro logique de l'imprimante

Donc il ne faut pas utiliser ces variables dans le programme principal.

External Kann :

La fonction Kann(i) qui calcule l'année de la i^{eme} semaine (voir plus loin) est appelée à partir d'un sous-programme.

Common N :

Le nombre total de semaines

Si Kp = nombre d'années

$N = 52 Kp.$

$Kp \leq 10 \Rightarrow n \leq 520$

Common /SENS/

Kp = nombre d'années

$Ktps$ = nombre de périodes par année = 52.

Common :/MØY/

T (520) au départ les vecteurs de la charge
après call cycl, celle de la tendance.

Common /CYC/

CT 5(520) après cycl le cycle total
après cysb le cycle sensible

Common /PAR/

a,b;c Coefficient de la parabole ajustée

Common /STA/

$$R \phi = \rho \quad (U_t = \rho U_{t-1} + \epsilon_t)$$

$$RC = (1 - \rho^2)^{1/2}$$

$$SU = \sigma_u$$

$$SE = \sigma_\epsilon$$

UZ = U_n le dernier residu servant de U₀ dans la prévision.

Common /CRL/

CR : Coefficient de corrélation entre les variables V(i) et C T(i)
au premier CALL CRLN, on a

$$V_i = t_i/\text{parabole}$$

$$C_{ti} = \text{cycle total}$$

au deuxième CALL CRLN, on a :

$V_i =$ cycle insensible

$C T_i =$ cycle sensible.

Common /LØ C/:

V(520) : tableau bidon, sert de tableau intermédiaire.

Common/MET/:

W (520) : valeurs de la températures.

WL, We : températures d'insensibilité d'hiver et d'été.

CM : cycle moyen entre WH et WE si WH WE
" " " WH - 2 et WF + 2 si WH = WE

Remarque: dans la 2^{eme} partie :

WH = TW

WF = TS

CM = XM

Common /CFT/:

FH (10), FE (10) coefficients (WH et WE) des droites du cycle sensible.

Remarque : Dans la 2^{eme} partie

FH = KW

FE = KS

Common /SYS/:

m = 3 Ordre du système linéaire auquel on est conduit dans la méthode des moindres carrés.

$$(X'X) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = X'Y.$$

D'une façon générale on utilise la notation suivante:

$$\begin{bmatrix} aa & ab & ac \\ ab & bb & bc \\ ac & bc & cc \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} ta \\ tb \\ tc \end{bmatrix}$$

Ce système est ensuite résolu par la subroutine RESL.

Common /CIB/:

CY (52) : moyennes hebdomadaires du cycle insensible.

VY (52) Variantes hebdomadaires du cycle insensible.

Common /ERR/:

KRIR Indicateur du déroulement de certains calculs (cf subroutine RØS2)

PROGRAMME PRINCIPAL DONNEES.

1 - Données introduites par instructions d'affectations

NDIM = 520

MAXA = 10

MINA = 2 il faut au moins 2 années d'analyse
(division par KP-1 dans CYSB).

Y = 1

Z = 3

m = 3 Ordre du système issu des moindres carrés.

1 - Données à introduire sur cartes :

read (100) KP , KTPS : nombre d'années d'analyse
et nombre de périodes.

read (101) ITER , ETA:

ITER : nombre maximum d'itérations tolérés pour une précision absolue
ETA surp si^e on atteint ITER il y a exécution de la méthode G.L.S
(avec la dernière valeur de ϵ).

read (100) KART : Nombre de valzurs de la charge perforées sur une carte.

KART = 13 4 cartes par année.

La lecture se fait dans les boucles 18 et 19.

Read (100) IDON , ICYC , KURB , IANA , METEO

Indicateurs d'impression des résultats

$\neq 0$ Si on veut le résultat

= 0 Si non

IDON : données : charge et météo

ICYC : toutes les valeurs des cycles.

KURB : courbes de charge et tendance

IANA : coefficients de la parabole, ,

METEO : influence de la température.

FUNCTION KANN (I)

Calcul de l'année correspondant à la i^{eme} semaine.

Common /SENS/ :

$K = I/KTPS$

$K.KTPS - I$ $\left\{ \begin{array}{l} = 0 \quad KANN = K \\ \\ = \quad K = K+1 \quad KANN = K. \end{array} \right.$

SUBROUTINE CYCL

Moyenne mobile ponderée

Entrée : Les valeurs de la charge dans /MØY/

Sortie : La moyenne dans /MOY/
Le cycle total dans /CYC/

Boucle 1 : * Calcul de la moyenne entre les indices 27 et n - 26.
Les deux 1/2 années extrêmes sont à reconstituer
* Cycle total.

Boucle 3 : Extrapolation exponentielle pour les deux 1/2 années extrêmes.
Calcul du rapport moyen t_i / t_{i-1} sur 4 périodes consécutives.

Remarque : Les indices KMT , KSEN , KSTE servent à reconstituer les valeurs des de 2 1/2 années extrêmes par une même séquence.

SUBROUTINE MCSP

Calcul des éléments du système issu de la méthode O.L.S, système qui va être résolu par la méthode de GAUSS (RESL).

Entrée ≠ La tendance dans /MØY/
La valeur de précédente (au départ = 0).

Sortie : Les variables transformées $E_i = t_i - t_{i-1}$
(avec le précédent) dans /CØC/
La matrice du système (3.3) et le second meubre.

O. L. S.

Les variables transformées \rightsquigarrow O.L.S

$V_i = t_i - t_{i-1}$ pour $\rho = 0$ on obtient les variables initiales.

Soient $rc = \text{sqrt}(1 - \rho^2)$

$xzz = (rc, 1 - \rho, \dots, 1 - \rho)$

$xu = (rc, 2 - \rho, \dots, i - \rho(i-1) \dots n - \rho(n-1)$

$xd = (rc, 4 - \rho, \dots, n^2 - \rho(n-1)^2)$

La formule des O. L. S. appliquée au système :

$$\begin{cases} t_1 = a \cdot xz_1 + b \cdot xu_1 + cxd_1 \\ \dots \\ t_n = a \cdot xz_n + b \cdot xu_n + cxd_n \end{cases}$$

donne :

$$A \begin{bmatrix} a \\ b \\ c \end{bmatrix} = B$$

Avec :

$$A = \begin{bmatrix} xz_1 & \dots & xz_n \\ xu_1 & \dots & xu_n \\ xd_1 & \dots & xd_n \end{bmatrix} \begin{bmatrix} xz_1 & xu_1 & xd_1 \\ \vdots & \vdots & \vdots \\ xz_n & xu_n & xd_n \end{bmatrix} = \begin{bmatrix} aa & ab & ac \\ ab & bb & bc \\ ac & bc & cc \end{bmatrix}$$

$$B = \begin{bmatrix} xz_1 & \dots & xz_n \\ xu_1 & \dots & xu_n \\ xd_1 & \dots & xd_n \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} ta \\ tb \\ tc \end{bmatrix}$$

$$\begin{cases} aa = rc^2 + (n-1) (1-)^2 \\ ab = rc^2 + xz_i \cdot xu_i \\ ac = rc^2 + xz_i \cdot xd_i \\ bb = rc^2 + (xu_i)^2 \\ bc = rc^2 + xu_i \cdot xd_i \\ cc = rc^2 + (xd_i)^2 \end{cases}$$

$$\begin{cases} ta = V_1 rc + V_i xz_i \\ tb = V_1 rc + V_i xu_i \\ tc = V_1 rc + V_i xd_i \end{cases}$$

partout ou somme de 2 a n.

SUBROUTINE MCGE

éléments du système de la méthode G.L.S.
(rendu dans RELS)

Entrée : . La tendance , dans /moy/
. Le ρ précédent, dans /sta/

Sortie : . la matrice 3.3. et le second membre.

Jusqu'à D \neq 1 : calcul des termes extrêmes de tous les produits scalaires.
(car expression différente du terme courant calculé dans
boucle 1).

Moindres carrés généralisés

on pose $xz = (1, 1, \dots 1)$
 $xu = (1, 2, i \dots n)$
 $xd = (1, 4 \dots i^2, \dots n^2)$

Soit x l'un de ces vecteurs, il faut résoudre :

$$(x'v^{-1}) \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (x'v^{-1} T)$$

où $T = (t_1, t_2 \dots t_n)$

et où v^{-1} a une expression connue, c'est une matrice tridiagonale :

$$\begin{bmatrix} 1 & -\rho & 0 \\ -\rho & 1+\rho^2 & -\rho \\ & & \dots \\ 0 & \dots & -\rho & 1 \end{bmatrix}$$

Posons $\alpha = (\alpha_1, \dots, \alpha_n) = xz \cdot v^{-1}$

$$\beta = (\beta_1, \dots, \beta_n) = xu \cdot v^{-1}$$

$$\gamma = (\gamma_1, \dots, \gamma_n) = xd \cdot v^{-1}$$

On a, pour α par exemple :

$$\begin{cases} \alpha_1 = x_1 - \rho x_2 \\ \alpha_2 = -\rho x_1 + (1 + \rho^2) x_2 - \rho x_3 \\ \alpha_i = -\rho x_{i-1} + (1 + \rho^2) x_i - \rho x_{i+1} \\ \alpha_{n-1} = \dots \\ \alpha_n = -\rho x_{n-1} + x_n \end{cases}$$

Pour les :

$$\begin{aligned} \alpha'_i : x_i &= xz_i \\ \beta_i : x_i &= xu_i \\ \gamma_i : x_i &= xd_i \end{aligned}$$

Le système à résoudre s'écrit alors

$$A \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = B$$

$$A = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \\ \gamma_1 & \dots & \gamma_n \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & n & n^2 \end{bmatrix} = \begin{bmatrix} aa & ab & ac \\ ab & bb & bc \\ ac & bc & cc \end{bmatrix}$$

$$B = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \\ \gamma_1 & \dots & \gamma_n \end{bmatrix} \begin{bmatrix} t_1 \\ \cdot \\ t_n \end{bmatrix} = \begin{bmatrix} ta \\ tb \\ tc \end{bmatrix}$$

$$aa = \sum \alpha_i xz_i$$

$$ab = \sum \alpha_i xu_i$$

$$ac = \sum \alpha_i xd_i$$

$$bb = \sum \beta_i xu_i$$

$$bc = \sum \beta_i xd_i$$

$$cc = \sum \gamma_i xd_i$$

$$ta = \sum \alpha_i t_i$$

$$tb = \sum \beta_i t_i$$

$$tc = \sum \gamma_i t_i$$

Dans le calcul des sommes précédentes, les termes extrêmes ont une expression particulière :

on a :

$$\begin{array}{ll} \alpha_1 = 1 - \rho & \alpha_n = 1 - \rho \\ \beta_1 = 1 - 2\rho & \beta_n = n - \rho(n-1) \\ \gamma_1 = 1 - 4\rho & \gamma_n = n^2 - \rho(n-1)^2 \end{array}$$

et pour $i \in [2, n-1]$

$$\alpha_i = -\rho + (1 + \rho^2) - \rho = (1 - \rho)^2 = \text{cste}$$

$$\begin{aligned} \beta_i &= -\rho(i-1) + (1 + \rho^2)i - \rho(i+1) \\ &= (1 - \rho)^2 \cdot i \\ &= i \cdot \alpha_i \end{aligned}$$

$$\begin{aligned} \gamma_i &= -\rho(i-1)^2 + (1 + \rho^2)i^2 - \rho(i+1)^2 \\ &= (1 + \rho^2)i^2 - 2\rho(1 + i^2) \end{aligned}$$

SUBROUTINE RØS2

Calcul de ρ , σ_u , σ_ε , .

Entrée . La tendance ou les **variables** transformées dans / ρ /.
(fait automatiquement dans MCSP et MCGE)
. la parabole dans /par/
. le ρ précédent ($\rho = 0$ initialement).

Il est également nécessaire de faire $\rho = 0$ après MCGE et avant RØS2 pour que dans V_i on ait la tendance et non les **variables** transformées).

Sortie . ρ , $\sqrt{1 - \rho^2} = RC$, $\sigma_u = SU$, $\sigma_\varepsilon = SE$, $U_n = UZ$

Remarque: KRIK donne la formule utilisée dans le calcul de ρ .

krik = 3 calcul normal (formules page ci-dessous): on a trouvé

$$|\rho| < 1.$$

krik = 2 calcul par la première valeur du urrelogramme corrigée du biais (biais = $1 + 3/n$.)

krik = 1 idem , non corrigée du biais.

krik = 0 calcul impossible (on a toujours $|\rho| > 1.$)

Il y a erreur quelque part.

le programme s'arrête.

calcul de ρ

on a

$$\begin{aligned}
 U_2 &= \rho \cdot u_1 + \varepsilon \\
 U_3 &= \rho \cdot u_2 + \varepsilon \\
 &\dots \\
 U_n &= \rho \cdot u_{n-1} + \varepsilon
 \end{aligned}$$

la résolution au sens des ϕ . L. S. s'écrit

où

$$A \begin{bmatrix} \rho \\ \varepsilon \end{bmatrix} = B$$

$$A = \begin{bmatrix} u_1 & \dots & u_{n-1} \\ 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} U_1 & 1 \\ \vdots & \vdots \\ U_{n-1} & 1 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \beta & \delta \end{bmatrix}$$

avec

$$\begin{aligned}
 \alpha &= \sum (U_i)^2 \\
 \beta &= \sum U_i \\
 \delta &= \sum_{i=1}^{n-1} 1 \quad \text{toujours}
 \end{aligned}$$

et où

$$B = \begin{bmatrix} \theta \\ \mu \end{bmatrix} \quad \text{avec} \quad \begin{aligned} \theta &= \sum u_i u_{i+1} \\ \mu &= \sum u_{i+1} \end{aligned}$$

La résolution donne:

$$\begin{aligned}
 \varepsilon &= (\mu - \beta\theta/\delta) / (\delta - \beta^2/\alpha) \\
 \text{et } \rho &= (\theta - \beta \cdot \varepsilon) / \alpha
 \end{aligned}$$

le calcul donne :

$$\boxed{\rho = \frac{\delta \cdot \theta - \beta \cdot \mu}{\alpha \delta - \beta^2}}$$

On écrit plutôt

$$\rho = (\theta - \beta \cdot \mu / S) / (\alpha - \beta^2 / S)$$

i.e.

$$\rho = \frac{\sum u_i u_{i+1} - \sum u_i \cdot \sum u_{i+1} / (n-1)}{\sum (u_i)^2 - (\sum u_i)^2 / (n-1)}$$

Remarque 1 :

$$\sigma_u^2 = \frac{1}{n-1} \left[\sum_1^n (u_i)^2 - \frac{1}{n} \left(\sum_1^n u_i \right)^2 \right]$$

Remarque 2 :

Lorsque la valeur de ρ ci dessus est > 1 , on la remplace par la valeur du corrélogramme obtenu en considérant $(n-1) - 1$ termes.

le dénominateur est le même.

Le numérateur s'écrit

$$\sum_{i=1}^{n-2} U_i \cdot U_{i+1} - \frac{1}{n-2} \sum_1^{n-2} U_i \sum_1^{n-2} U_{i+1}$$

Remarque 3 :

$$\rho = 1 + \frac{\Delta}{\text{den}}$$

$$\Delta = \sum U_i U_{i+1} - \sum (U_i)^2 + \sum U_i \{ U_i - U_{i+1} \} / (n-1)$$

$$= \sum U_i (U_{i+1} - U_i) + (U_1 - U_n) \cdot \sum U_i / (n-1) .$$

Calcul de m la première valeur du corrélogramme

Soit n le nombre de termes U_i .

On a
$$r_1 = \frac{\text{num}}{\text{den}}$$

Numérateur :

$$\sum_{i=1}^{n-1} U_i U_{i+1} - \bar{Y} \sum_{i=2}^n U_i$$

où
$$\bar{Y} = \frac{1}{n-1} \sum_{i=1}^{n-1} U_i$$

Dénominateur ;

$$\sum_{i=1}^n u_i^2 - \frac{1}{n} \left(\sum_{i=1}^n U_i \right)^2$$

Facteur de correction du biais

$$1 + \frac{3}{n-1}$$

SUBROUTINE RESL

Résolution par Gauss du système 3 - 3 issu de MCSP ou MCGE.

Entrée : aa , bb , cc , ab ; bc ,
 ta , tb , tc.

Sortie : a , b , c coeff. de la parabole.

N.B. toujours vérifier que on a bien écrit
 m = 3 quelque par avant dans le p.p.

SUBROUTINE CYMN

Calcul de la moyenne du cycle total aux températures comprises entre les bornes d'insensibilité wh, we.

Entrée : Le cycle ~~total~~, dans /cyc/ .

 Les températures et les bornes dans /met/ .

Sortie : le cycle moyen, CM , dans /met/ .

SUBROUTINE CFTE

Calcule les coefficients k_{wi} , k_{si} des droites représentatives du cycle sensible.

Entrée :

- . Le cycle total , dans /cyc/
- . Les températures , dans :/met/

Sortie : les coefficients k_{w_i} , k_{s_i} dans /cft/ .

On fait appel à la fonction KANN (i) de calcul de l'année.

La variable LFIN sert à repérer la fin des données de températures (car le passage d'une année à l'autre se fait par détection du changement du numéro de l'année, donc impossible à la fin).

SUBROUTINE CYSB

Calcul des cycles sensible et insensible
de la moyenne hebdomadaire du cycle insensible.

Entrée:

- . Le nombre d'années d'analyse : /sens/.
- . les températures : :/met/
- . les coefficients des droites du cycle sensible dans /cft/.

Sortie:

- . les cycles sensibles , dans /cyc/.
- . les cycles insensibles, dans /l ϕ e/.
- . les cycles insensibles hebd. moyens et les variances hebd. dans /cib/.

Boucle 4 : calcul des cycles sensibles et insensible

Boucle 7 : calcul des moyennes insensibles.

Remarque: division par $k_p - 1$ (ISN 32)

faut $k_p \gg \text{MINA} = 2$.

SUBROUTINE CURV

Tracé des courbes de charge et de tendance.

Entrée : { . la charge , dans / lØc / (fait dans cycl)
 { . la tendance , dans /mØy/ .

La courbe est tracée sur 101 ~~clames~~ *colonnes*
et cadrée par rapport aux valeurs extrêmes de la charge.

Charge : signes +

tendance: signes *

Lorsque (charge = tendance) à la précision près de l'échelle imposée,
il y a impression d'un point.

SUBROUTINE CRLN

calcul du coefficient de corrélation entre les variables V_i et ct_i .

Sortie { . le coefficient de corrélation cr , dans /crl/
 { . la variable d'erreur krik
 (= 0 si tentative d'extraire la racine carrée d'un nombre
 négatif).

Remarque: Si krik = 0 , on a dans cr le nombre négatif en question.

Entrée : le S/P est appelé 2 fois :

au premier appel

{ . $R_i = T_i$ / parabole_i dans /lØc/
 { . cycle total dans /cyc/

au Deuxième appel

{ . cycle insensible dans /lØc/
 { . cycle sensible dans /cyc/.

Fonction de répartition du cycle sensible :

Prob ($S < S_0$) : cf. II.6.2.

Un GO TO calculé permet un aiguillage sur l'un des modèles A, B, C, D suivant la valeur de la variable KCAS. Cette dernière peut prendre les valeurs 1, 2, 3, 4 selon les signes de CE et de CH (coefficients de sensibilité extrapolés) : cf. SUBROUTINE EXTR.

Les paramètres formels DM, DV seront remplacés à l'appel respectivement par TY (j) et VB(j), moyenne et variance hebdomadaire de la variable météo; elles sont calculées dans la subroutine CSMX.

DM et DV serviront à centrer et réduire la température de manière à pouvoir lui appliquer la fonction PROB.

FUNCTION FNCS (TP)

Calcul le cycle sensible pour une température TP

Entrée : La température dans /MET/

Sortie : La valeur du cycle sensible correspondant.

$$CS = \begin{cases} CH (TP - WH) & \text{Si } TP < WH \\ 0 & \text{Si } WH \leq TP \leq WE \\ CE (TP - WE) & \text{Si } TP > WE \end{cases}$$

FUNCTION FNTP (KSZ, S)

- 79 -

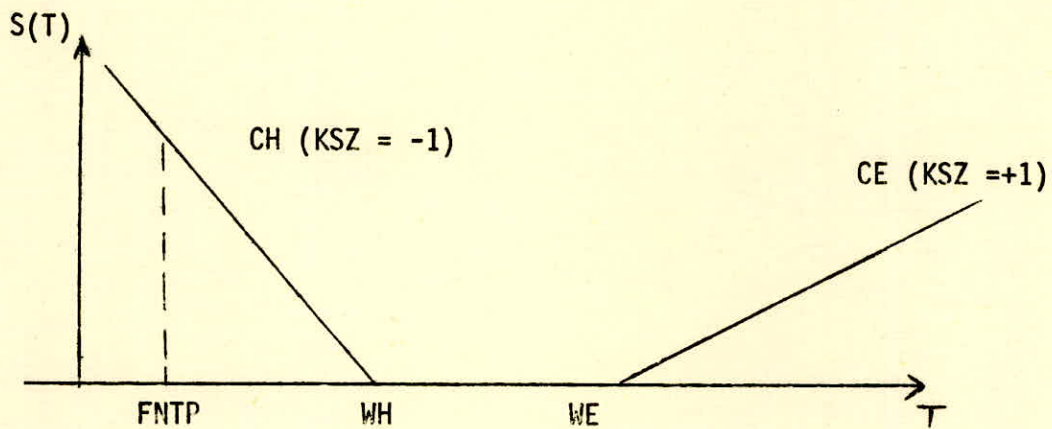
C'est la fonction $S^{-1}(S_0)$, inverse du cycle KSZ

KSZ prend les valeurs : $\begin{matrix} +1 & \text{été} \\ -1 & \text{hiver} \end{matrix}$

Normalement on a toujours :

$(FNTP - BE) \cdot KSZ > 0$ (cf : graphique A, B, C, D du II.6.2)

Sinon il y a erreur quelque part Krik = 0



$$FNTP - WH < 0$$

$$(FNTP - WH) \cdot KSZ > 0$$

SUBROUTINE EXTR

Extrapolation de KW et KS

Entrée : * Les coefficients KW(i) et KS(i) des Kp années d'analyse dans /CFT/

* KP, NPREV dans /SEN/

Sortie : * CH et CE dans /CFT/

après appel à RESL

Calcul de CH et CE

Il s'agit d'estimer la relation :

$$KW_i = \alpha + \beta i \quad (\text{voir II.1})$$

$$KW_1 = \alpha + \beta$$

$$KW_2 = \alpha + 2\beta$$

$$KW_3 = \alpha + 3\beta$$

$$KW_i = \alpha + i\beta$$

$$KW_{KP} = \alpha + KP \cdot \beta$$

La résolution au sens des O.L.S. s'écrit

$$A \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = B$$

$$A = \begin{pmatrix} aa & ab \\ ab & bb \end{pmatrix} \quad B = \begin{pmatrix} ta \\ tb \end{pmatrix}$$

où :

$$A = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & KP \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \cdot & \cdot \\ \cdot & \cdot \\ 1 & KP \end{bmatrix} = \begin{bmatrix} KP & P \cdot (P+1)/2 \\ P \cdot (P+1)/2, (2 P+1) \\ P(P+1)/6 \end{bmatrix}$$

$$B = \begin{bmatrix} \sum_{i=1}^{KP} KW_i \\ \sum_{i=1}^{KP} i KH_i \end{bmatrix}$$

- 81 -

Après CALL RES on a dans (PAR) A et B ce qui nous permet de calculer $CH = A + B * PP$ où $PP = KP * NPREV$

On refait la même séquence en considérant KS

Boucle 1 : Calcul de CH

Boucle 2 : Calcul de CE

SUBROUTINE CSNX

- 82 -

Calcul de :

- * Caractéristiques METEO :
Moyennes et variances de la température dans
les vecteurs respectifs TY(j) et VB(j)

- * SMIN, SMAX

Entrée :

- * Températures dans /MET/
- * KP, KPTS dans /SEM/

Sortie :

- * Statistiques météo dans /MST/
- * SMIN, SMAX dans /CNX/

Remarques * La boucle 2 permet de calculer la somme des températures de son carré pour les semaines de même nom à travers toutes les années.

$$* X_K = 2.326667 = N^{-1} \begin{matrix} (0,1) \\ 0,01 \end{matrix}$$

Compte tenu du fait que X_K est centrée réduite, on applique FNCS.

Extrapolation du cycle sensible.

On calcule :

- * La densité du cycle sensible en 20 points
- * La moyenne CS (j)
- * La variance VS(j)

Entrée :

- * Les caractéristiques météo dans /MST/
- * KPTS dans /SEM/

Sortie :

- * Le cycle moyen hebdomadaire CS(j) dans /CSB/
- * La variance hebdomadaire VS(j) dans /CSB/

Remarques :

- * La boucle 3 sert à mesurer l'accroissement de la fonction de répartition du cycle sensible sur 1 intervalle, ce qui donne la densité approximative PB(K) du cycle sensible. Cette dernière permet de calculer les moyenne et variance du cycle sensible.
- * La boucle 5 calcule précisément cette variance
- * La boucle 1 permet de réaliser les opérations ci-dessus semaine par semaine.

Conclusion:

Des perfectionnements du modèle peuvent être envisagés :

- 1/ Le filtrage peut être amélioré en recherchant un filtre dont le gain voisine l'unité, si bien qu'il ne subsistera quasiment pas de corrélation entre les composantes, ce qui améliorera la variance de la charge globale.
- 2/ Dans l'analyse de la corrélation entre les composantes, on n'avait considéré que la liaison entre des valeurs concomitantes.
On peut pousser plus loin l'analyse en utilisant la théorie spectrale.
- 3/ Une tendance exponentielle pourrait paraître plus vraisemblable, dans ce cas l'hypothèse de la nullité de l'espérance des aléas a plus de chance d'être vérifiée, ce qui garantit l'absence de biais des estimateurs.

Références bibliographiques

- Programme BUFFLE - Description
Direction des ETUDES ECONOMIQUES DE LA SONELGAZ.

- J. JOHNSTON :
Econometric Methods - Mac Graw-Hill. 1972

- E. MALINVAUD :
Méthodes Statistiques de l'économétrie - Dunod 1969

- RODIER et MAZODIER :
MODELES ECONOMETRIQUES

- J.C. LALOIRE
METHODES DU TRAITEMENT DES CHRONIQUES
Dunod - 1972

- C.W.J. GRANGER
Analyse spectrale des séries temporelles en économie Dunod. 1969.

