

REVUE DE STATISTIQUE APPLIQUÉE

J. C. LIGERON

R. GOARIN

Modélisation multidimensionnelle des taux de défaillance de composants électroniques

Revue de statistique appliquée, tome 22, n° 3 (1974), p. 43-67

http://www.numdam.org/item?id=RSA_1974__22_3_43_0

© Société française de statistique, 1974, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

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

MODÉLISATION MULTIDIMENSIONNELLE DES TAUX DE DÉFAILLANCE DE COMPOSANTS ÉLECTRONIQUES ⁽¹⁾

J. C. LIGERON **

Ingénieur à la Section Fiabilité des Engins Matra

R. GOARIN

Ingénieur au C.N.E.T.

I – INTRODUCTION [37] [38] [39]

Dans le cadre du développement de la banque de données de fiabilité, parallèlement au Recueil d'informations de plus en plus précises, il a fallu mettre au point des méthodes d'analyse de ces données. Notre souci a été de caractériser un résultat de fiabilité sur un composant en y attachant toutes les caractéristiques de ce composant qui puissent avoir une influence sur la fiabilité. Cela a entraîné la gestion d'un volume important d'informations et la recherche d'outils mathématiques permettant d'extraire de cette masse de données, des renseignements objectifs et concis.

La notion de taux de défaillance pour les composants électroniques est très largement admise par la majorité des spécialistes en Fiabilité. Cette notion a déjà donné lieu à de très nombreuses publications.

On admet, de façon courante, que le taux de défaillance est fonction de la température et des contraintes électriques subies par le composant. Dans ce domaine, les recherches et les expérimentations ont été très nombreuses. Elles ont donné naissance à un grand nombre de modèles dont certains sont très populaires.

Néanmoins, ces modèles ont été établis à l'aide de données sur des résultats d'essais, antérieurs à 1964, ce qui les rend relativement périmés.

Ces modèles ont été établis, soit par des fabricants de composants sur un produit bien déterminé, avec plus ou moins de succès, (l'idée est généralement dans ce cas, de rechercher une description physique de la dégradation des composants) à partir d'un volume plus ou moins grand de données.

- - - - -

(1) Article remis le 21/3/73. Révisé le 28/9/73.

** Ce travail fait partie d'une thèse de Doctorat d'Université effectué sous la direction de Monsieur le Professeur Dugué.

L'idée est donc de décrire statistiquement, le comportement d'une population de composants.

On ignore dans ce cas, si les données ont un caractère de généralité suffisant et de plus la méthode utilisée a été parfois uniquement empirique.

L'ingénieur de Fiabilité s'occupant d'un équipement, peut toujours s'interroger sur la valeur du modèle proposé par les recueils ou par un fabricant de composants.

Nous nous sommes proposé dans ce travail, de bâtir des méthodes permettant de trouver des modèles mathématiques de taux de défaillance, et à l'aide de ces méthodes, vérifier la validité de ceux existant.

II – ETUDES ET SELECTION DES DONNEES

Les données de fiabilité relatives aux défaillances catalectiques seront caractérisés par le taux de panne.

La sélection et la critique des données sont une partie très importante du travail de construction du modèle. La précision de celui-ci dépend du soin apporté à analyser les données ; pour cela, rien ne doit être laissé au hasard, l'échantillon retenu pour bâtir les modèles devra avoir une parfaite homogénéité, de même, l'on devra connaître l'ensemble des corrélations existant entre les différentes variables semblant influencer sur un taux de défaillance.

Variables influentes ou variables primaires –

Nous pouvons dresser une liste de variables qui influent directement sur le taux de défaillance d'un composant électronique, variables d'ailleurs fournies dans la plupart des résultats d'essais.

Critères de défauts –

Ce sont les valeurs critiques de certains paramètres accessibles sur un composant au-delà desquelles le composant est considéré comme défaillant.

On peut distinguer deux types de défaillances :

- a) Les défaillances catalectiques
- b) Les défaillances par dérive.

Contraintes électriques –

Suivant le type de composant étudié, on peut distinguer 2 types de contraintes électriques :

– Contrainte en tension :

$$\frac{V_2}{V_1} = \frac{\text{Tension appliquée}}{\text{Tension nominale}}$$

– Contrainte en puissance :

$$\frac{P_2}{P_1} = \frac{\text{Puissance dissipée}}{\text{Puissance nominale}}$$

– Contraintes thermiques

La température considérée dépend du type de composant étudié.

– Les composants passifs sont caractérisés par la température ambiante en °C, la température de couche ou la température de boîtier.

– Les composants actifs sont caractérisés par 2 types de température :

1 – la température de jonction en °C

2 – la température de jonction normalisée.

Variables qualitatives et variables secondaires –

On peut dresser une liste non exhaustive des variables qualitatives et secondaires, c'est-à-dire, semblant influencer, d'une manière moins directe, le taux de défaillance, ce sont :

- | | |
|-------------------------|----------------------------------|
| – le fabricant | – l'année de fabrication |
| – la technologie | – le nombre de composants testés |
| – le type d'utilisation | – la durée de l'essai |

III – CHOIX DES DONNEES

– Critique des données –

Cette étude doit être entreprise avant tout traitement, elle a pour but, de déceler les "vices cachés" dont voici quelques exemples :

– un domaine de contrainte peut ne pas être couvert entièrement. Il est alors dangereux d'extrapoler.

– des critères de défaillance peuvent ne pas être homogènes, pour un fabricant, ce qui élimine l'effet fabricant en le remplaçant par une variable critère de rejet.

– des regroupements peuvent se produire pour une même technologie, il est bon de noter cette dépendance qui dans certains cas, peut provenir d'autres variables "cachées" par les durées d'essais normés pour cette technologie.

– critère de défaillance inexistant : ce genre de données est à rejeter car on ne possède pas de point de comparaison.

IV – ETUDE DE L'INFLUENCE DES VARIABLES SELECTIONNEES SUR LE TAUX DE DEFAILLANCE –

But –

Nous avons pu voir dans les chapitres précédents, que la sélection des données revêtait une grande importance.

Le problème étant en fait, d'avoir des échantillons par fabricant, qui soient homogènes, c'est-à-dire, que les mêmes causes aient les mêmes effets.

En réalité, il est très difficile, étant donné le nombre de variables influant sur le "*taux de défaillance*", de déterminer à priori leur influence.

Sans intercorrélations entre les variables, le problème serait simple à résoudre. Les coefficients de corrélation simple nous renseigneraient sur les variables influentes :

Aussi, il nous faut trouver une méthode nous séparant les données au mieux en perdant le moins d'informations possible.

Une première analyse des données devrait permettre d'entreprendre une étude de modèle si les données sont homogènes ou de l'arrêter à ce stade.

Il faut noter dès à présent que du fait d'une certaine imprécision lors de l'établissement du taux de panne avec peu de défaillances (et que dire du résultat ayant conduit à aucune défaillance) les conclusions de l'étude ne sauraient être plus précises que les informations de départ.

Les paramètres décrivant l'essai de fiabilité seront la valeur du taux de panne catalectique estimée avec un niveau de confiance de 60%.

Pour étudier les caractères influant sur la fiabilité, nous pouvons distinguer la technologie des composants qui bien souvent est reconnue en tant que paramètre influant, mais il est intéressant de savoir à quel point elle agit.

La liste des autres paramètres pouvant avoir de l'influence est très large ; mode d'encapsulation, caractéristiques thermiques ou électriques, dimensions, conditions de fonctionnement, ambiance, fabricant, date de fabrication . . . etc

Il est à noter aussi que bien souvent nous ne disposons pas de tous les éléments influant sur la fiabilité comme les fluctuations des approvisionnements de matières premières chez les fabricants, ce qui explique le manque de précision des conclusions. La répartition non uniforme des essais effectués sur une fabrication par exemple ou dans une gamme de température ou de puissance est aussi une limitation à la qualité des conclusions.

Méthode d'analyse des correspondances –

Une étude a été entreprise à partir de données sur des condensateurs "tante"

Nous avons défini 7 classes de taux de défaillances catalectique par le tableau suivant :

λ en $10^{-6}/h$

0,5 1 2,5 5 10 15

Symbôle	LA1	LA2	LA3	LA4	LA5	LA6	LA7
---------	-----	-----	-----	-----	-----	-----	-----

Les paramètres considérés sont les contraintes électriques S avec trois classes selon que la contrainte est inférieure S-1, égale S = 1 ou supérieure S + 1 à la valeur maximale du fabricant. Les températures sont divisées en quatre classes regroupées autour des valeurs 65°C, 85°C, 130°C, 170°C. Les fabricants sont repérés par les lettres de A à G et trois technologies d'encapsulation sont repérées par les symboles HER hermétiques, TRE trempé, MOU moulé. Nous avons ainsi obtenu un tableau de répartition dans chaque classe.

HE	LA1	LA2	LA3	LA4	LA5	LA6	LA7
S - 1	.	1.	1.	0.	3.	0.	1.
S = 1	.	9.	7.	18.	7.	13.	6.
S + 1	.	0.	0.	1.	2.	0.	1.
T 65	.	0.	2.	0.	0.	0.	0.
T 85	.	8.	5.	14.	8.	10.	3.
T 13	.	2.	1.	5.	3.	2.	3.
T 17	.	0.	0.	0.	1.	1.	2.
A	.	3.	0.	3.	0.	0.	0.
B	.	0.	3.	4.	8.	10.	7.
C	.	4.	0.	0.	1.	0.	0.
D	.	0.	0.	3.	1.	1.	0.
E	.	0.	1.	2.	0.	1.	0.
F	.	2.	0.	1.	2.	0.	1.
G	.	1.	4.	6.	0.	0.	0.
HER	.	10	8.	16.	10.	9.	7.
TRE	.	0.	0.	2.	0.	2.	0.
MOU	.	0.	0.	1.	2.	2.	1.

CONDENSATEURS "TANTALE"

Figure 1

La projection des points dans le plan des axes 1 et 2 (figure 2) fait apparaître sur l'axe 1 un taux d'inertie de 40 % qui met en opposition les taux de pannes LA1 et LA7 les plus forts et les plus faibles, ce qui permet de classer les technologies, les condensateurs hermétiques HER étant meilleurs que les moulés MOU ou les trempés TRE d'où l'influence importante de la technologie les fabricants A, C et F se placent du côté des bons taux de panne, en effet, cela est normal, ils fabriquent des condensateurs hermétiques.

Le fabricant D fabrique des condensateurs trempés et le fabricant B des moulés ce qui explique aussi leur position. L'effet des contraintes électriques et thermiques n'est pas évident car les points correspondants sont répartis selon l'axe 2 qui est décorrélé du premier. Sur cet axe 2, on peut malgré tout, faire apparaître les taux de panne meilleurs pour des températures faibles T 65 que pour les températures élevées T 17 du côté de LA6.

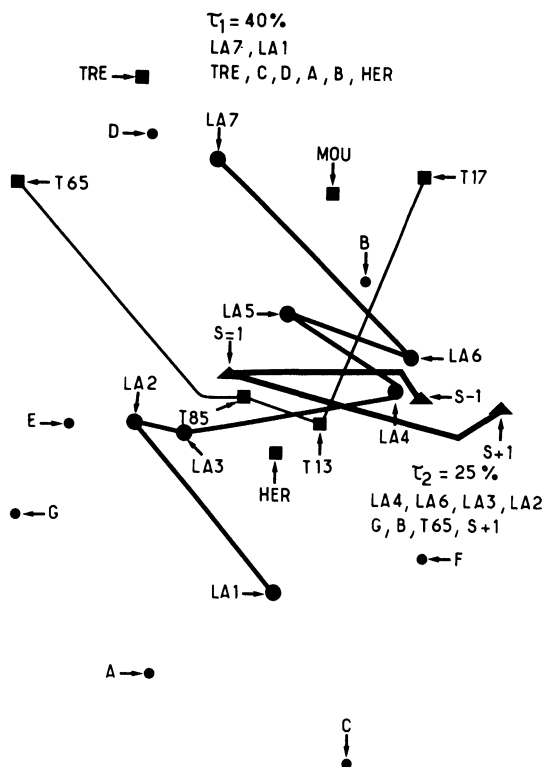


figure 2

La position du point S-1 semble anormale, elle voudrait signifier qu'une contrainte électrique faible a entraîné des taux de panne plus élevés que S=1, en fait une analyse plus fine des données indique que les essais effectués à contrainte électrique faible S - 1 l'on été à température élevée donc la position de ce point est une conséquence plus qu'une cause. Ceci prouve aussi que des conclusions trop hâtives ne peuvent être données sans une analyse détaillée des données. D'autre part, les effectifs ayant servi à établir cette étude sont relativement faibles, il faut donc donner à ces résultats une importance toute relative dont l'objectivité est en rapport avec la quantité des données et leur répartition.

Toutefois, pour avoir la confirmation de la signification de l'axe 2, nous avons considéré le plan des axes 1 et 3 (figure 3), le taux d'inertie correspondant à l'axe 3, n'est pas négligeable (19 %) cet axe mettant une nuance aux conclusions de l'étude précédente.

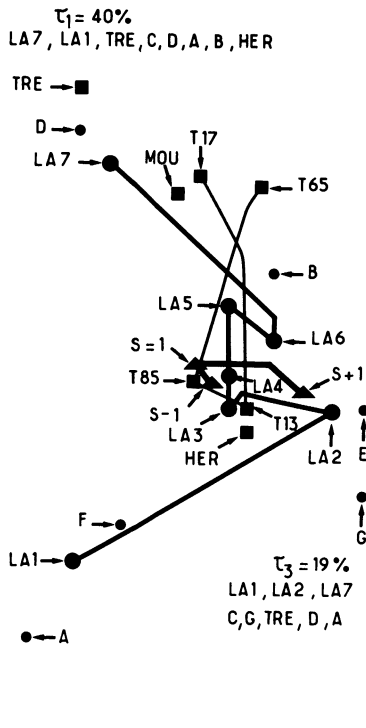


figure 3

On remarque malgré tout que les températures et les contraintes électriques ont une projection regroupée, cela est normal car l'axe 3 est perpendiculaire à l'axe 2 de même les taux de panne autres que LA 1 et LA 7 sont répartis de manière aléatoire. Cet axe 3 correspond aux fabricants qui modulent les conclusions tirées à partir des axes 1 et 2.

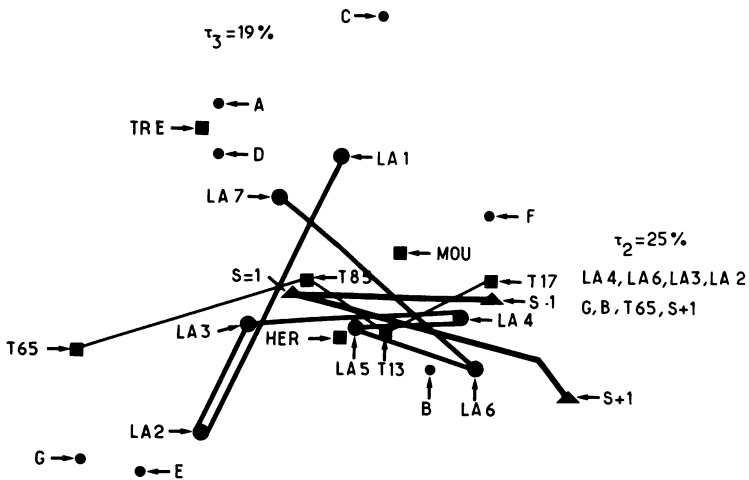


figure 4

Le plan des axes 2 et 3 correspond à un pourcentage d'inertie de 44 % et on voit que l'opposition entre LA1 et LA7 n'existe plus mais les conclusions tirées sur l'axe 2 à partir de l'observation du plan des axes 1 et 2 se retrouvent.

Le programme permet de mettre certains paramètres en éléments supplémentaires c'est-à-dire qu'on attribue à ces points une masse nulle et de ce fait n'interviennent pas dans la détermination des axes d'inertie. La figure 5 représente dans le nouveau plan des axes 1 et 2 la projection des points sauf la technologie.

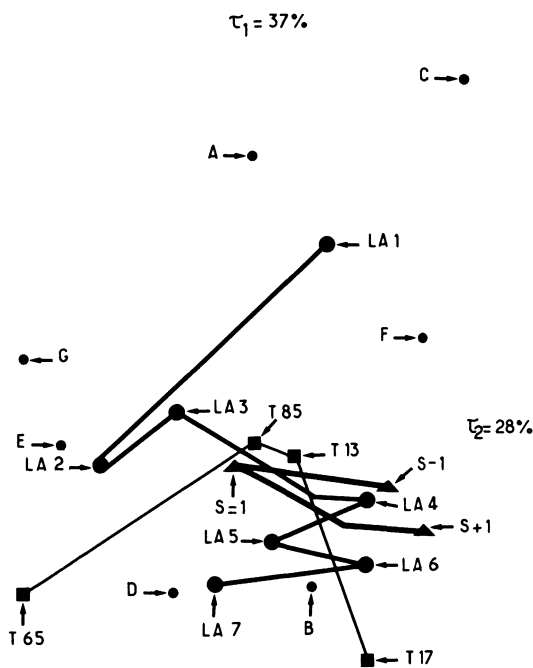


figure 5

Ce paramètre mis à part, on constate que le point LA7 s'est rapproché du centre de gravité et il n'y a que le point LA1 qui se dégage bien sur l'axe 1, c'est donc que les mauvais taux de pannes sont dus davantage à la technologie qu'aux fabricants qui interviennent dans la détermination des axes. On constate ainsi que l'inertie totale a certes diminuée mais est restée importante (165 à 135), ce qui prouve que les fabricants ont une grande influence sur le résultat.

Une étude identique a été menée en considérant les fabricants en paramètres supplémentaires, on constate que le point LA1 s'est rapproché du centre de gravité alors que le point LA7 reste éloigné, c'est donc la technologie qui explique les mauvais taux de panne.

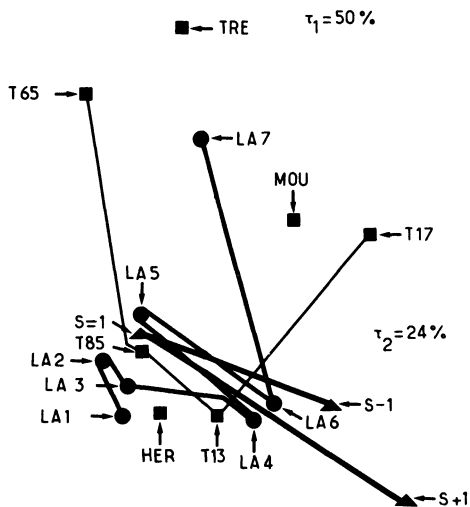


figure 6

V – CHOIX DU MODELE

Les modèles usuels sont généralement à base physicochimique.

Nous avons essayé plusieurs modèles :

- Les modèles non linéaires, non linéarisables
- Les modèles du type RADC TR 67.108 (Rome Air Development Center Reliability notebook) qui se présentent par exemple pour les condensateurs sous la forme :

$$\lambda = \lambda b \left[\left(\frac{S}{NS} \right)^H + 1 \right] \cdot e^{\left(\frac{T}{NT} \right)^G}$$

- Les modèles non linéaires linéarisables à partir de transformations en Log.

- Le modèle de Bazovsky :

$$\lambda = \lambda b \left(\frac{V_2}{V_1} \right)^\alpha \cdot K^{(\theta_2 - \theta_1)}$$

- Le modèle exponentiel ou de Tatar :

$$\lambda = \lambda b \cdot e^{b_1 T + b_2 S + b_3 S^2 + b_4 T^2 + b_5 ST}$$

Dans ces modèles :

$\frac{V_2}{V_1}$ ou S représente la contrainte électrique $\frac{\text{tension appliquée}}{\text{tension nominale}}$

$(\theta_2 - \theta_1)$ où T représente la contrainte thermique.

Le choix d'un modèle est généralement assez subjectif. Il doit éviter d'introduire des autocorrélations entre variables, et doit être d'une forme relativement simple de manière à éviter les temps de calcul trop importants.

VI – CHOIX DE LA METHODE APPLICABLE –

A partir des N essais retenus nous pouvons former un tableau :

$$\lambda_i = f(S_i, T_i)$$

λ_i étant les taux de défaillance.

S_i les contraintes électriques

T_i les contraintes thermiques.

On a, d'autre part, une expression algébrique théorique donnée par exemple, par le modèle de Tatar.

$$\lambda = \lambda_0 \cdot e^{b_1 T + b_2 S + b_3 S^2 + b_4 T^2 + b_5 ST}$$

Le problème qui se pose est le suivant :

Peut-on ajuster à l'échantillon le modèle indiqué, les paramètres $\lambda_0, b_1, b_2, b_3, b_4, b_5$ ayant des valeurs fixes. Si oui, quelles sont les valeurs à donner aux paramètres ?

Posé sous cette forme, le problème est celui de l'ajustement d'une courbe à des points expérimentaux.

Deux méthodes sont alors utilisables :

- la régression linéaire dans le cas où le modèle est linéaire ou linéarisable ;
- une méthode d'ajustement du type méthode de gradient ou méthode de Fletcher Powell dans le cas où le modèle n'est pas linéarisable : par exemple, les modèles RADC.

Ces méthodes consistent à minimiser la somme des carrés des écarts entre valeurs estimées λ_i et les valeurs ajustées λ'_i soit la somme :

$$Q = \sum_{i=1}^n (\lambda_i - \lambda'_i)^2$$

Nous avons appliqué ces deux méthodes.

– La régression sous une forme modifiée permettant d'inclure les variables en fonction de leur importance [2] [26] [30] [34] [35] [36].

Cette méthode a permis l'étude des modèles de Bazovsky et de Tatar.

– Les méthodes de gradient sous une forme développée par Fletcher et Powell permettant d'étudier des modèles du RADC [10] [11] [12].

Néanmoins, d'importants problèmes de pondération des données se sont posés ; en effet, l'estimation du taux de défaillance est d'autant plus précise que le nombre de défauts observés en cours d'essai est grand ; il est naturel d'accorder plus de crédit aux résultats d'essais précis qu'aux résultats imprécis, par

exemple, 0 défaut ; de même, les hypothèses de base de la régression impliquent une stabilité de la variance sur les valeurs observées, Aussi, avons-nous dû utiliser une méthode de stabilisation consistant à pondérer les données par $\sqrt{C+1}$ fois le nombre de défauts [29]. (C étant le nombre de défauts)

De plus le traitement des variables qualitatives a nécessité une étude spéciale permettant de traiter en une seule fois l'ensemble des variables, la distinction entre elles se faisant par variation du terme constant du modèle [28].

Il est bien évident qu'à chaque étude de modèle par régression est rattachée l'analyse de variance sous jacente permettant de vérifier :

– si l'hypothèse de linéarité peut être acceptée (Test du F) ; [1] [33] [31] [32]

– si les coefficients de régression sont significativement différents de 0 (Test du T). [1] [33] [31] [32].

Forme des résultats –

Traités par régression.

On peut voir un exemple de résultat obtenu par régression linéaire sur les transistors Silicium NPN, sur la figure 7.

Transistor silicium NPN					
Caractéristiques	N° du Step	Variables	% Explique	Valeurs calculées	λ_b
Défauts totaux Tj	1	FAB 2	23.6	b1 = 1.01	0.711
Référence : basse puis.	2	S ²	12.5	b2 = 1.06	
Epitaxial Planar	3	T	5.8	b3 = 0.013	
Fournisseur +	4	Grown	3.7	b4 = 1.14	
Type +	5	Puissance	4.3	b5 = 0.947	
Technologie	6	FAB 3	2.4	b6 = 0.594	
	7	ST	1.8	b7 = 0.005	
	8	Mesa	1.5	b8 = 1.57	
			Total		
			55.6		

figure 7

L'influence des variables qualitatives est assez marquée et permet de chiffrer les différences entre technologies, fabricants et type d'utilisation ; dans ce cas bien précis, la technologie référence était Epitaxial planar et le type était basse-puissance.

Ce type de résultat a été trouvé pour les composants permettant de rôder la procédure, cette méthode a permis en outre, de trouver les modèles de Bazovsky équivalents à ceux du RADC et ayant une forme beaucoup plus simple.

Modèle exponentiel ou de Tatar

$$\lambda = 0,711 e^{0,013T + 1,06S^2 + 0,005ST + 1.01 \text{ (si fabricant 2)} + 1,14 \text{ (si technologie grown)} + 0,947 \text{ (si transistor de puissance)} + 1,57 \text{ (si technologie mesa)}}$$

Traités par méthode d'optimisation –

Le RADC 67 108 utilise des modèles différents pour chaque famille de composants.

Pour la famille des condensateurs, le modèle général est :

$$\lambda = \lambda b \left[\left(\frac{S}{NS} \right)^H + 1 \right] e^{\left(\frac{T}{NT} \right)^G}$$

avec $S = \frac{\text{tension utilisation}}{\text{tension nominale}}$

T = température ambiante en °K

$\lambda b, NS, H, NT, F$ Constantes Caractéristiques du type de composants.

Pour les condensateurs au tantale à électrolyte solide et les condensateurs au polyester les valeurs numériques sont rappelées au tableau 1.

Pour la famille des semi-conducteurs, le modèle général est de la forme :

$$\lambda = \lambda b \exp\left(\frac{NT}{TJ}\right) \times \exp\left[\left(\frac{TJ}{To}\right)^P \times \left(\frac{S}{NS}\right)^H\right]$$

S est le rapport de la puissance réellement dissipée sur la puissance maximale permise à 25°C.

TJ est la température de jonction en °K.

$\lambda b, NT, To, P, NS, H$ sont les constantes caractéristiques du type de semi-conducteur.

Pour les transistors silicium NPN et les diodes ZENER, les valeurs sont reportées dans le tableau 1.

Type de composant	λb	NS	H	NT	G	
Condensateurs au tantale	$1,3 \times 10^{-8}$	0.52	3	358	8.8	
Condensateur au polyester,	$3,2 \times 10^{-9}$	0.73	5	350	17.4	
Type de composant	λb	NT	To	NS	H	P
Transistor silicium NPN	$2,7 \times 10^{-7}$	-1052	448	1	1	10.1
Diodes ZENER	$2,74 \times 10^{-6}$	-2242	448	1	1	10.7

TABLEAU 1

D'autres résultats obtenus sont repérés dans le tableau 2. On trouvera Ref [37] un critère d'efficacité de ces modèles.

Type de composants	λb	NS	H	NT	G	% expliqué	
Condensateurs au tantale	$0,114 \times 10^{-7}$	0,557	4,41	278	4,551	59 %	
Condensateurs au polyester	$0,165 \times 10^{-10}$	0,511	3,946	308	8,25	77 %	
	λb	NT	TO	P	NS	H	%
Diodes ZENER	1261×10^{-6}	– 3586	325,7	36,47	0,001028	0,0692	56,2 %
Transistors silicium NPN	162×10^{-6}	– 1930	432,5	3,39	1,45	1,83	62,06 %

TABLEAU 2

Analyse des résultats obtenus à l'aide des différentes méthodes :

Si l'on compare les résultats obtenus sur les deux modèles étudiés TATAR et BAZOVSKY, on s'aperçoit donc : que les pourcentages expliqués sont en général plus faibles pour le modèle de BAZOVSKY, du fait de la forme du modèle de TATAR qui est plus sophistiquée. Mais, on constate que les variables qualitatives sont prépondérantes par rapport aux variables quantitatives, dans certains cas par exemple : les défauts par dérive de transistors silicium NPN, les résultats obtenus sont absolument identiques dans les deux cas.

D'une façon générale, on peut dire que lorsque les données permettent un ajustement correct avec un modèle, le deuxième modèle donne aussi de bons résultats et inversement. Il n'y a pas de différences significatives entre la validité des deux modèles.

Ceci nous conduit à analyser de plus près les données. L'existence d'un grand nombre d'essais n'ayant donné lieu à aucun défaut observé, peut être une des causes de l'influence prépondérante des variables qualitatives.

En effet, supposons qu'on sache qu'un composant a un taux de panne égal à 10^{-6} /heures, et que nous effectuons des essais cumulant respectivement 10 pièces heures, 100 pièces heures, 10000 pièces heures, aucun essai n'ayant donné lieu à observation de défauts. Selon la méthode classique de calcul du meilleur estimateur du taux de panne, on évaluerait $\lambda_1 = X_1$, $\lambda_2 = X_2$, $\lambda_3 = X_3$.

Supposons de plus, que les fabricants aient des durées standard d'essais de 10, 100, 10000 heures respectivement ; par la méthode utilisée jusqu'alors, on trouvera que λ_1 est corrélé à la variable fabricant 1, etc. . . les variables S et T n'intervenant manifestement pas dans le comportement de ces curieux "Lambda".

Les essais qui cumulent un nombre de pièces heures faible devant la MTBF approximative du composant conduisent à des résultats de toute évidence erronés.

La méthode de pondération des essais en $\sqrt{C+1}$ utilisée jusqu'alors avait pour but d'essayer de tenir compte de cette différence de validité des données. Ceci, est très certainement valable lorsqu'une faible proportion de résultats est manifestement biaisée, ce n'est certainement pas valable lorsque la majorité des essais le sont. Pour notre cas, nous allons essayer de tenir compte de ces erreurs.

Méthodes Bayes [3] [4] [5] [6] [7] [8]

– Tenir compte de la connaissance à priori que l'on a de l'ordre de grandeur des taux de panne et intégrer les résultats d'essais observés à cette connaissance à priori par la méthode de Bayes. On obtiendra ainsi, des distributions de taux de panne à postériori (étant connus les résultats d'essais) qui pourront :

. soit être utilisées dans une régression avec pondération en $\sqrt{C+1}$ comme nous l'avons fait jusqu'à présent.

. soit être utilisées pour fabriquer des essais systématiques et chercher les courbes enveloppes du modèle.

VII – UTILISATION DES METHODES BAYESIENNES –

Définition –

Soit une variable aléatoire X, dont la distribution de probabilité dépend d'un ensemble de paramètres.

$$U = U_1, U_2, U_v$$

Les valeurs exactes des paramètres n'étant pas connues avec certitude le raisonnement Bayésien attribue une distribution de probabilité aux diverses valeurs possibles de ces paramètres qui sont alors considérés comme des variables aléatoires.

$$U = U_1, U_2, U_v$$

Dans le cas présent, nous nous intéressons à la distribution du taux de défaillance, la méthode Bayésienne consiste à fabriquer une fonction de distribution à priori du taux de défaillance, dépendante du nombre de résultats d'essais ou de l'expérience que nous possédons à un instant donné. La forme de la distribution est modifiée en fonction des résultats de nouveaux essais.

Rappels sur le théorème de Bayes –

Le théorème de Bayes s'exprime généralement sous la forme suivante :

$$P(A/B) = P(A) \times \frac{P(B/A)}{P(B)}$$

où P(A) est la probabilité à priori de l'évènement A.

P(A/B) celle de A, étant donné que l'évènement B s'est produit, ou probabilité à postériori.

Problèmes d'utilisation –

L'utilisation du théorème de Bayes peut être récursive, c'est-à-dire, que si l'on part d'une distribution à priori P(A) et que l'on a une succession d'essais ayant donné les résultats B₁, B₂, B_n. On a successivement :

$$P(A/B_1) = P(A) \frac{P(B_1/A)}{P(B_1)}$$

Puis :

$$P(A/B_1, B_2) = P(A) \frac{P(B_1/A)}{P(B_1)} \frac{P(B_2/A)}{P(B_2)}$$

Ainsi de suite :

$$P(A/B_1, B_2 \dots B_n) = P(A/B_1, B_2 \dots B_{n-1}) \frac{P(B_n/A)}{P(B_n)}$$

Au fur et à mesure que des résultats d'essais sont connus, on utilise la distribution à postériori déjà obtenue comme l'actuelle distribution à priori et on recommence la séquence des opérations. On applique ainsi ce qui est appelé la méthode Bayésienne.

L'ancienne distribution à priori $P(A)$ tend à être périmée, et modifiée au fur et à mesure que de nouveaux résultats sont pris en compte.

La vitesse avec laquelle une distribution à priori est périmée, détermine la "force" de cette distribution.

Si une distribution à priori est "forte", elle a un faible écart type, et la connaissance à priori est très peu dispersée autour de la valeur moyenne, on ne pourra donc, que difficilement modifier la distribution à priori par les apports successifs.

A l'opposé, une distribution à priori "faible", représente une connaissance à priori très imprécise et les apports successifs, s'ils confirment la tendance centrale de la connaissance à priori transformeront l'à priori "faible" en une à postériori "forte" ; sinon les essais viennent modifier la tendance centrale de l'à priori. Dans ce dernier cas, la quantité d'information contenue dans l'essai ne vient plus en addition de celle contenue dans l'à priori mais en différence, et ne pourront l'emporter sur la connaissance à priori que si elle est supérieure à celle-ci.

On se trouve alors, dans une situation pire que si l'on avait aucune connaissance à "priori".

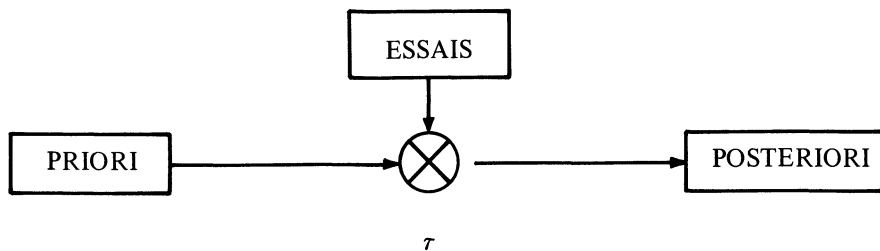
On voit donc que l'utilisation de la méthode de Bayes nécessite un choix minutieux de la distribution à priori, en particulier, il faut vérifier que les essais ne mettent pas globalement en évidence un taux de panne significativement différent de la valeur moyenne du taux de panne de la distribution à "priori".

Par la suite, si les composants successivement mis en essais sont de même nature et, partant du même ordre de grandeur de fiabilité, il n'est plus nécessaire de faire des vérifications.

Dans la pratique, on utilisera une "à priori" faible, du fait que l'on ne sait pas grand chose sur la fiabilité, et que l'on veut pouvoir détecter les écarts et modifications introduits par les essais, pour tester après coup la validité du choix "à priori".

Distribution "à priori" choisie —

On peut considérer l'utilisation du théorème de Bayes comme l'application d'une transformation τ représentée ci-dessous :



Dans l'étude de toute transformation on recherche les invariants. Dans le cas du théorème de Bayes, il existe des distributions à priori qui restent invariantes, aux valeurs de paramètres près, lorsqu'on applique le théorème de Bayes.

Si l'on fait des essais binomiaux par tout ou rien, la distribution invariante dans l'application du théorème de Bayes est la distribution *BETA*,

Dans le cas des essais de mesure de λ , l'invariant bayésien est la distribution *gamma*.

Il se trouve que la distribution gamma (Γ) recouvre la distribution du χ^2 pour certaines valeurs des paramètres.

L'artifice des invariants Bayésiens est très commode sur le plan numérique, mais il impose que la distribution à postérieure ait une forme donnée, ce qui peut être une contrainte. De plus, il réduit, du moins en apparence, l'utilisation de la méthode de Bayes à une série d'additions banales.

La distribution gamma est donnée par la relation suivante :

$$P(\lambda) d\lambda = \frac{\tau(e^{-\lambda\tau}) (\lambda\tau)^{(\Phi-1)}}{\Gamma(\Phi-1)} d\lambda$$

τ : est un temps correspondant à un essai qui produirait la même distribution de λ que $P(\lambda)$.

Φ : est un nombre équivalent de défauts dans le même essai fictif que ci-dessus :

$$\Gamma(x) = \int_0^{\infty} e^{-z} z^{x-1} dz$$

Si $x = n$, nombre entier $\Gamma(n) = (n-1)!$;

$$\text{Moyenne : } \mu = \frac{\Phi}{\tau} = \lambda$$

$$\text{Variance : } \sigma^2 = \frac{\Phi}{\tau^2} = \frac{\lambda}{\tau}$$

Pour un essai ayant donné le résultat B : f défauts, t temps cumulé :

$$P(B/A) = \frac{e^{-\lambda t} (\lambda t)^f}{f!} \text{ Loi de Poisson}$$

Dans le cas continu le théorème de Bayes s'exprime sous la forme :

$$P(\lambda/B) = \frac{P(\lambda) P(B)/\lambda}{\int_0^\lambda P(B/\lambda) P(\lambda) d\lambda}$$

On obtient :

$$P(\lambda/B) = \frac{(t+\tau) \exp. [(-\lambda(t+\tau))] \cdot [(t+\tau) \lambda]^{(\Phi+f-1)}}{(\Phi+f-1)!}$$

Ainsi :

On voit que la distribution Γ de paramètres : ϕ et τ est transformée lors d'essais poissonniens en une distribution Γ de paramètres :

$$\Phi + f \text{ et } \tau + t.$$

On dit que les distributions γ et de Poisson sont conjuguées Bayésiennes pour exprimer le fait que la distribution γ reste invariante par la transformation de Bayes lors d'un essai poissonnien.

Les moyennes et variances sont indiquées dans le tableau ci-dessous :

	à priori	à postérieur
Moyenne	Φ/τ	$(\Phi + f)/(\tau + t)$
Variance	Φ/τ^2	$(\Phi + f)/(\tau + t)^2$

Application

Etablissement de la distribution à priori –

Pour établir la distribution "à priori" nous utiliserons le RADC TR 67 108 ou sa version française ' recueil de données de fiabilité du CNET.

Nous admettons que ce recueil de données est exhaustif des cas qui se sont présentés dans le passé et reflète bien le comportement moyen des composants.

Les résultats d'essais disponibles permettent alors de déterminer les distributions "à postérieur".

Si pour un composant déterminé nous avons les renseignements suivants :

S = contrainte électrique de l'essai

T = contrainte thermique de l'essai

tcum = durée cumulée de l'essai

c = nombre de défauts observés lors de l'essai

Et si nous portons sur un graphique $\lambda=f(S,T)$ paramètre par S il est possible de :

– déterminer un certain nombre de strates, c'est-à-dire des valeurs d'essais à S et T fixes pour lesquelles nous possédons des résultats d'essais (voir figure 8).

– construire sur ces strates des distributions à priori indépendamment du fabricant et de la technologie.

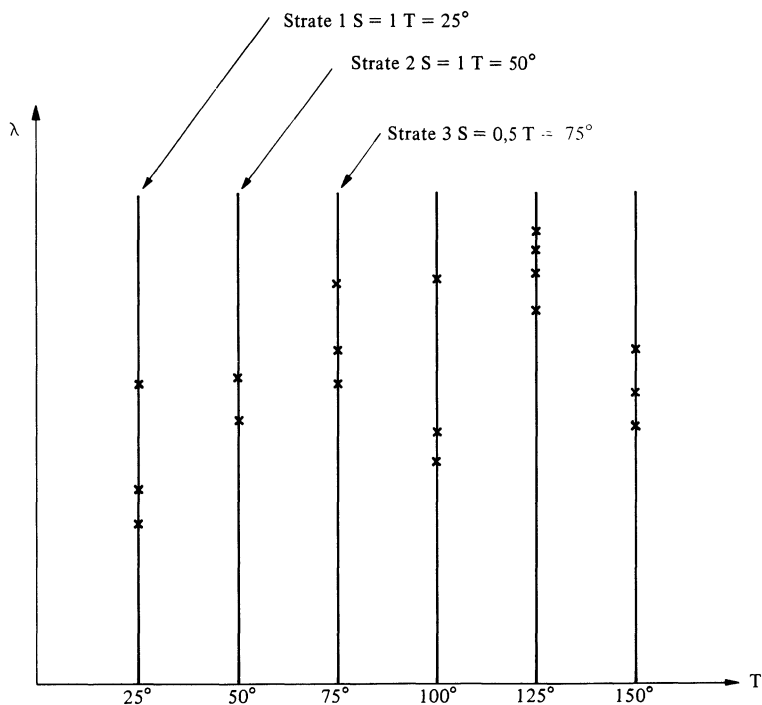


figure 8

La méthode est certes discutable, aussi serait-il nécessaire de travailler normalement sur un seul type de composant. Mais le but de ces modèles n'est pas de trouver un modèle correspondant à un transistor particulier, mais d'estimer une valeur moyenne du modèle pour un fabricant déterminé, voire pour l'ensemble des fabricants.

Choix des variables de la distribution gamma à priori —

Considérons le recueil de données de fiabilité, il donne l'expression du taux de défaillance en fonction de S et T dans un terme appelé taux de base λb .

Le taux de défaillance s'exprime sous la forme générale :

$$\lambda = \lambda b \times \pi E + \Sigma E$$

πE et ΣE sont des paramètres exprimant à la fois le type d'environnement et un niveau de qualité.

Il est possible de calculer un taux de défaillance en fonction du paramètre de qualité. Nous avons :

$$\lambda_{\text{inf}} = \lambda b \times \pi E + \Sigma E_{\text{inf}}$$

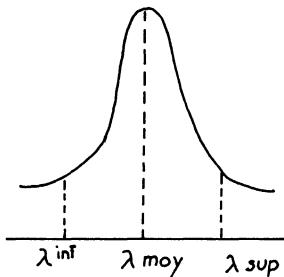
$$\lambda_{\text{sup}} = \lambda b \times \pi E + \Sigma E_{\text{sup}}$$

$$\lambda_{\text{moyen}} = \frac{\lambda_{\text{sup.}} + \lambda_{\text{inf}}}{2}$$

Le but de ces calculs est de déterminer une valeur équivalente du temps cumulé et des défaillances, compte-tenu de λ moyen, λ sup et λ inf.

Supposons le taux de défaillance en fonction de la qualité distribué suivant une loi à paramètres :

$$\lambda \text{ sup} - \lambda \text{ inf} = 6 \sigma$$



$$\text{Posons } \frac{(\lambda \text{ sup} - \lambda \text{ inf})^2}{36} = \sigma^2$$

La moyenne de la loi gamma étant égale à :

$$\frac{\Phi}{\tau} = \frac{\text{nombre défaillants}}{\text{temps cumulé}}$$

$$\text{Variance} = \frac{\Phi}{\tau} \times \frac{1}{\tau} = \lambda_{\text{moy}} \times \frac{1}{\tau} = \frac{\lambda \text{ moy}}{\tau}$$

$$\tau \text{ eq. temps} = \frac{\lambda \text{ moy}}{\text{Variance}}$$

Nombre de défauts équivalents : $\phi \text{ eq.} = \lambda \text{ moy} \times \tau$

Nous avons donc une loi gamma à priori de paramètres ($\phi \text{ eq.}$ $\tau \text{ eq.}$).

Mise en oeuvre de la méthode —

1) On sélectionne :

- les fabricants
- les strates, c'est-à-dire l'ensemble des couples S et T sur lesquels ont été effectués les essais.

2) Pour les strates sélectionnées on calcule le nombre de défauts et les temps équivalents à partir du RADC (C eq. T eq.).

3) Pour chaque fabricant, et pour chaque strate ; on calcule les temps et le nombre de défauts cumulés pour l'ensemble des essais correspondants. (C essai, T essai).

4) Pour chaque fabricant et chaque strate on calcule $c = c \text{ eq.} + c \text{ essai}$. $T = T \text{ eq.} + T \text{ essai}$. Ces résultats permettent de calculer les paramètres de la loi gamma correspondant à la distribution "à postériori".

Néanmoins, il convient d'ajouter des essais homogènes à la distribution à priori, c'est-à-dire, non significativement différents pour une strate donnée.

Il faut noter que, si dans chaque strate il y a au moins un fabricant pour lequel on dispose de résultat, il peut se produire que tous les fabricants n'aient pas fait d'essai dans toutes les strates.

Ceci entraîne que pour les fabricants qui n'ont pas fait d'essai dans une strate on crée un essai "fictif" caractérisé par $C_{\text{essai}} = 0$, $T_{\text{essai}} = 0$, on retrouve alors intégralement la distribution à priori caractérisée par C eq. et T eq.

5) On trouve un taux de défaillance par tirage au hasard (méthode de *Monte-Carlo*), dans la distribution à postériori par strate, par fabricant et l'on forme un tableau $\lambda = f(S, T)$.

6) On effectue une régression linéaire sur $\lambda = f(S, T)$

7) On répète la procédure décrite de 5 à 6, un grand nombre de fois et l'on obtient alors la distribution des coefficients de régression dont il est prudent de vérifier la normalité [27].

Néanmoins, il est nécessaire de vérifier que la distribution à priori n'a pas trop de force dans les résultats en pratiquant une étude de sensibilité.

La figure 9 indique la procédure utilisée dans la construction du programme.

Résultats obtenus par la méthode Bayésienne –

Choix des modèles –

Les deux modèles sélectionnés sont :

- le modèle de Tatar
- le modèle de Bazovsky

On trouvera sur le tableau 3 un exemple de résultats obtenus par le modèle de Bazovsky.

Modèle de Bazovsky

Variable	Moyenne	Ecart type	Pourcentage expliqué	Modèle
Polyester T Intercept.	1,07 $0,951 \cdot 10^{-3}$	$0,571 \cdot 10^{-2}$ $0,279 \cdot 10^{-3}$	80 %	$\lambda = 0,951 \cdot 10^{-3} + 1,07^{(\theta_2 - \theta_1)}$
Tantale T Intercept.	1.118 $0,24 \cdot 10^{-1}$	0,258 $0,107 \cdot 10^{-1}$	60 %	$\lambda = 0,24 \cdot 10^{-1} + 1,118^{(\theta_2 - \theta_1)}$

TABLEAU 3

On peut voir que les pourcentages expliqués sont relativement élevés.

On peut observer pour un type de composant, les courbes enveloppes de variation des modèles, elles sont telles que : (voir figure 10)

$$\lambda_{\text{sup}} = (\lambda b + \sigma \lambda b) \times (a + \sigma A)^{(\theta_2 - \theta_1)}$$

$$\lambda_{\text{inf}} = (\lambda b + \sigma \lambda \cdot b) \cdot (a - \sigma A)^{(\theta_2 - \theta_1)}$$

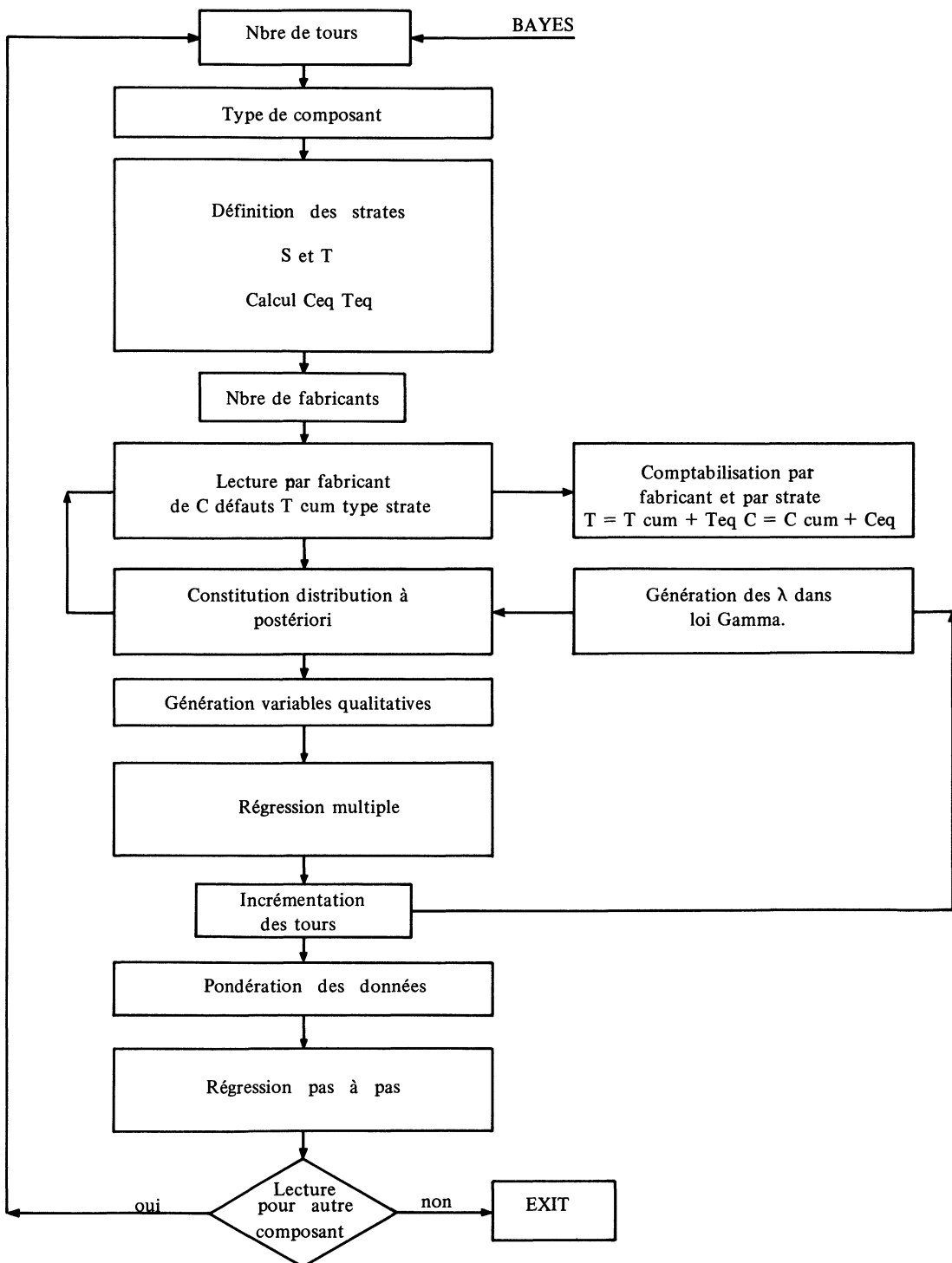


figure 9

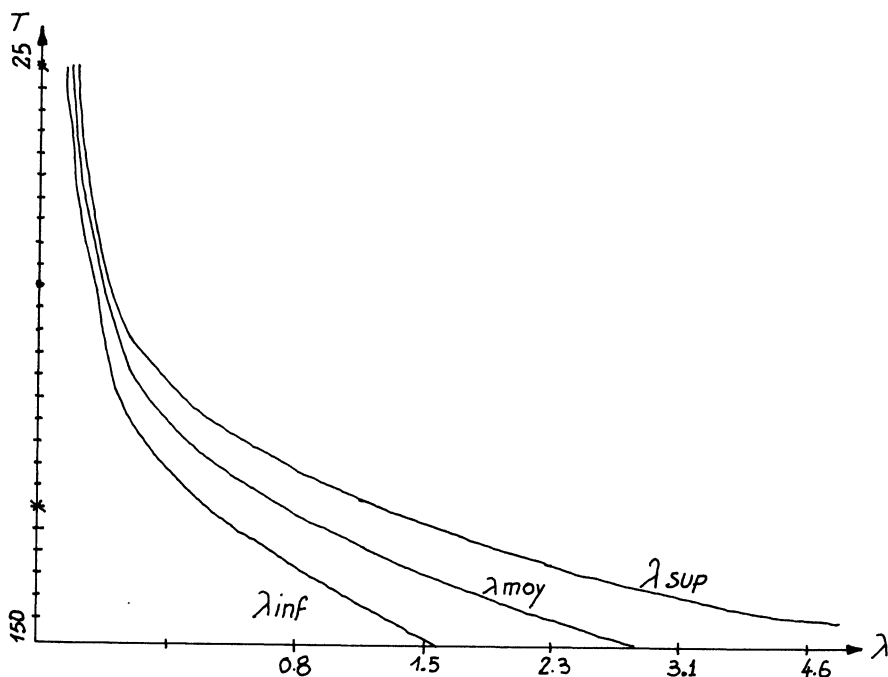


figure 10

IV – CONCLUSIONS GENERALES

Cette étude aura permis, d'une part de chiffrer pour 4 types de composants l'influence des divers facteurs agissant sur le taux de défaillance, d'autre part ce qui est fondamental, d'établir une procédure et des "outils" permettant de bâtir de tels modèles mathématiques. On s'aperçoit d'ailleurs que les techniques de statistiques les plus évoluées, notamment l'analyse des correspondances et la théorie Bayésienne des probabilités ont été nécessaires, les techniques classiques n'étant pas en effet assez puissantes et ayant des hypothèses contraignantes.

Sur la figure 11, on peut voir le déroulement de la méthodologie employée pour bâtir ces modèles.

Ces méthodes recouvrent un ensemble de programmes écrits en Fortran IV et implantés sur machine IBM.

De même nous avons pu voir que pour l'Ingénieur Concepteur de circuits, les influences de la contrainte et de la température étaient secondaires. (Il n'est pas en effet souhaitable d'utiliser des composants à haute température et à contrainte électrique élevée). Les variables du type fournisseur, technologie étaient prépondérantes dans le cas d'une utilisation normale.

Il est bien évident que la précision des modèles trouvés est fonction de la quantité de données disponibles ; dans notre cas, elles étaient relativement faibles ; souhaitons que dans un proche avenir une collaboration entre les différentes banques de données de fiabilité nous permette de bâtir des modèles de plus en plus précis.

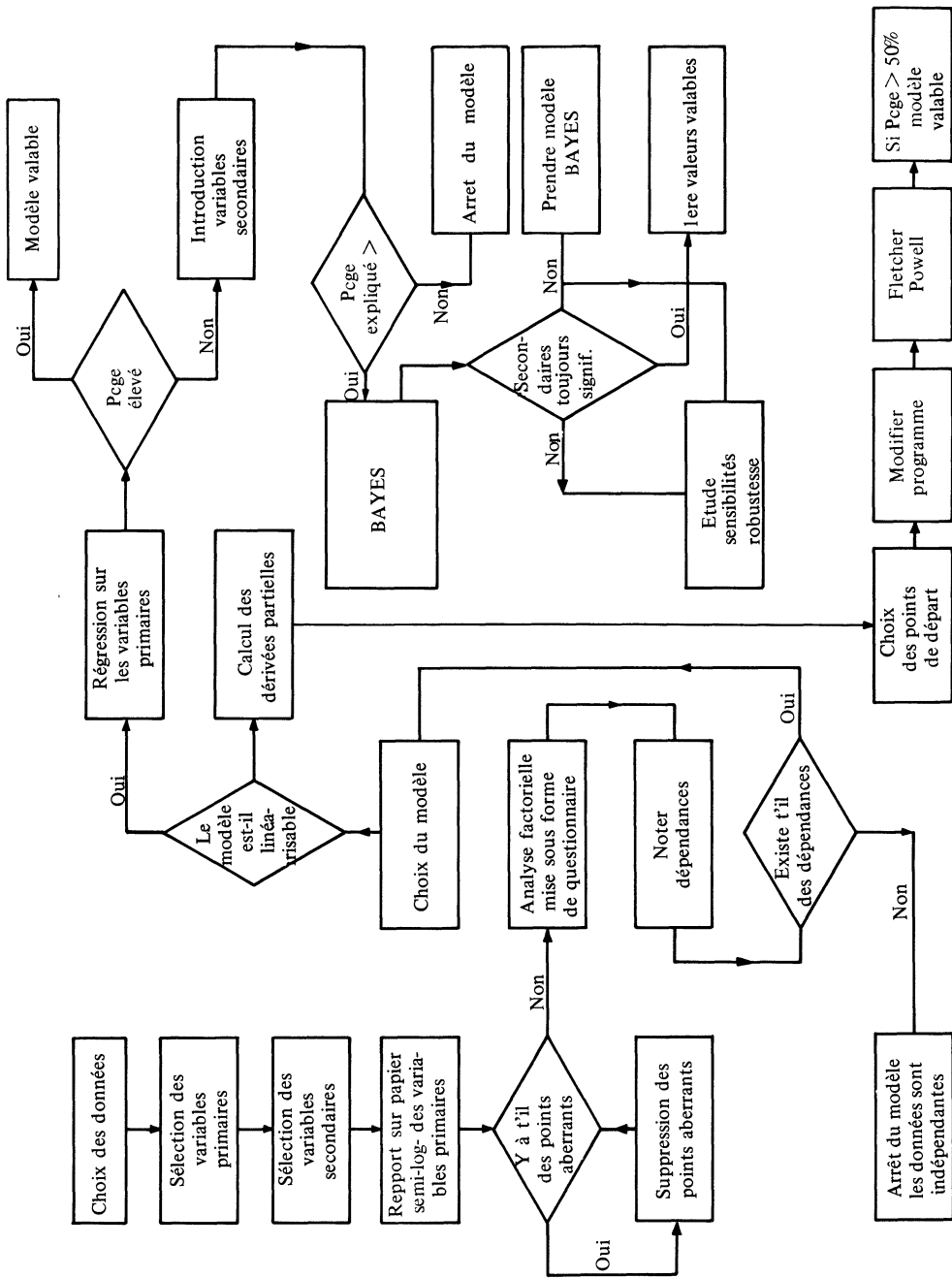


figure 11

BIBLIOGRAPHIE

- [1] DUGUE D. et GIRAULT M. – Analyse de Variance et Plan d'Expérience Dunod – 1959.
- [2] N.R. DRAPER H. SMITH – Applied Regression Analysis.
- [3] W.J. MAC FARLAND – Use of Bayes Theoreme in its discrete formulation for reliability estimation purposes. *Annal of assurance sciences* 1968.
- [4] W. Thomas WEIR – Bayesian Reliability Evaluation ? *Annal of Assurance Sciences* 1968.
- [5] J.J. DEELY and W.J. ZIMMER – Some comparisions of Bayesian and classical confidence Intervals in the exponential case.
- [6] John E. OLLSSON – Implementation of a Bayesian reliability measurement program.
- [7] A.M. BREIPHOL – A consideration of the Bayesian Approach in reliability evaluation.
- [8] G. HANSEL D. GROUCHKO – Prevision sequentielle par la méthode de Bayes. *Revue de statistiques appliquée* 1965 vol. XIII.
- [9] W. COOLEY P.R. LOHNES – Multivariate Procedures for the behavioral sciences.
- [10] R. FLETCHER, M.L.D. POWELL – A rapidly convergent descent method for minimization. *Brit. computer J. Vol 6–1963–pp 163–168*
- [11] M.J.D. POWELL – An efficient method for finding the minimum of a function of several variables without using derivatives. *Brit. Computer J. Vol. 7 1964 – pp 155 – 162.*
- [12] A.M. VERCOUSTRE – Etude comparative des méthodes de minimisation de FLETCHER et POWELL et DAVIDON E.D.F. – *Bulletin de la Direction des études et recherches – série C mathématique – informatique n° 1970 pp 57 – 76.*
- [13] L. LEBART et J.P. FENELON – Statistique et Informatique Appliquées. Dunod 1971.
- [14] J. GUIGOU – Expériences numériques comparatives concernant la méthode du gradient conjugué. *E.D.F. – Bulletin de la Direction des Etudes et Recherches – serie C mathématiques – informatique n°2 1968 pp. 79 – 92.*
- [15] F. TONNELIER – Deux méthodes de minimisation : la méthode de FLETCHER-POWELL, la méthode de FLETCHER-REEVES, Séminaire Y. Cherruault du 19.11.1969.
- [16] E. POLAK – G. RIBIERE – Note sur la convergence de méthodes direction conjuguées. *RIRO* 3ème année n° 16 1969 pp. 35–43.
- [17] R. FLETCHER et C.M. REEVES – Function minimisation by conjugate gradients, *The computer journal* pp. 149 – 154 (1964).
- [18] J.F. KENNEY, E.S. KEEPING – Mathematics of Statistics Part two chap. X.XI – Van Nostrand Company Fev. 1951.

- [19] L.S. LASDON, A.D. WAREN – Mathematical programming for optimal design *Electro Technology* – Nov. 1967 pp. 55–70.
- [20] R.J. HADER, AH GRANDAGE – Simple and multiple regression analyses *Experimental Designs in Industry*. John Wiley and Sons Inc. New-York 1958.
- [21] H.O. HARTLEY – The modified Gauss-Newton method of the fitting of non linear regression functions by least squares.
- [22] W.G. IRESON – Editor *Reliability Handbook* 1966 pp. 4.62–4.65.
- [23] A. RHALSTON, H.S. WILF – Méthodes mathématiques pour calculateur arithmétiques, Dunod Editeur.
- [24] BRICE, CARNAHAN, H.S. LUTHER, J.O. WILKES – Applied numerical methods, Wiley – Editor.
- [25] J.P. BENZECRI – Leçons sur l'analyse des données multidimensionnelles (Ronéo FAC).
- [26] D.F. MORISSON – Multivariate Statistical methods Mc Gram Hill
- [27] P. FERIGNAC – Test de Kolmogorov-Smirnov sur la validité d'une fonction de distribution.
- [28] D.B. SUITS – Use of Dummy variables in regression *Journal of Americal statistical association* 52 Dec. 1957 pp. 548.
- [29] KENNEY, KEEPING – Mathematics of statistics part two – chap X, XI Van Nostrand company inc. Princeton (N.J)
- [30] A. BENNETT L. FRANKLIN – Edit. WILEY and Sons, statistical Analysis in chemistry industry.
- [31] KENDALL and STUART – The advanced theory and statistics tome II
- [32] MORICE and CHARTIER – Methode statistique
- [33] Tables statistiques – ISUP
- [34] N.R. DRAPER and D.M. STONEMAN
Testing for the inclusion of variables in linear regression by a randomisation technique, *Technometrics* Vol. 8 n° 4, p. 695.
- [35] R.J. FREUND – A warning of round off erros in regression, the american statistician December 1965. p. 13.
- [36] J.C. LIGERON J.M. GRANGE J. DORLEANS – Modèle mathématiques de taux de défaillances de composants électroniques en fonction des contraintes électriques et thermiques. *Onde Electrique*.
- [37] R. GOARIN – Réalisation d'une banque de données de fiabilité (Journées d'Etudes sur les données de Fiabilité LANNION Juin 71)
- [38] J.C. LIGERON – Modélisation multidimensionnelle de taux de défaillance (Congrès National de Fiabilité Perros Guirec Sept. 72).
- [39] R. GOARIN – Application de la méthode d'analyse des correspondances à l'étude de la fiabilité des composants électroniques (Congrès National de Fiabilité Perros-Guirec Sept. 72).