Recherche d'une étape défaillante dans un procédé industriel : quand le désordre est source d'information

François Bergeret
Motorola Semiconducteurs S.A. — 31023 Toulouse cedex
Philippe Besse
Laboratoire de Statistique et Probabilités UMR CNRS 5583
Université Paul Sabatier — 31062 Toulouse cedex 4
Mél : Philippe.Besse@cict.fr

12 novembre 1998

Résumé Le but de cet article est de montrer comment des techniques statistiques classiques peuvent être utilisées avec un objectif original : détecter une étape défaillante dans un processus de fabrication pour laquelle le test d'adéquation du produit n'intervient qu'en fin de processus. Les seules informations disponibles sont les dates et heures de passage des lots à chacune des étapes de la fabrication. Des méthodes paramétriques et non paramétriques sont utilisées : la régression logistique et les lissages splines sont appliqués sur deux exemples réels qui proviennent de l'industrie du semi-conducteur. Pour une étape donnée, plus l'ajustement du paramètre final en fonction du temps est bon, plus la probabilité d'une défaillance de cette étape est forte. En effet, les lots se mélangeant lors de la fabrication, l'ajustement en fonction du temps sera moins bon pour les étapes éloignés de l'étape défaillante.

Mots-clefs : Contrôle de qualité, modélisation, lissage.

Abstract The purpose of this paper is to show how standard statistical methods can be used with a new objective: detecting in a manufacturing process the stage involved in a drift. The drift can only be observed after the fabrication and the only available information is the date of processing of the lots at the different process stages. Both parametric and non-parametric methods are used: logistic regression and spline smoothing are applied on two examples taken from the semiconductor industry. For a given stage, the goodness of fit of the final parameter as a function of the time indicates the likelihood to be involded in the drift. This works because the lots are mixed during the fabrication: the goodness of fit will be lower for the stages far from the guilty stage.

Key-words: Quality control, modelisation, smoothing.

1 Introduction

1.1 Contexte industriel

Dans une industrie aussi exigeante que la fabrication des circuits intégrés, la recherche de la qualité est une nécessité quotidienne. De nombreux outils existent, en particulier statistiques. On peut citer par exemple les cartes de contrôle qui servent à contrôler la conformité du procédé lors de la fabrication, les plans d'expérience pour développer une nouvelle technologie ou optimiser un procédé. Dans un souci d'amélioration continue, de nouveaux outils, basés sur l'information contenue dans une base de données documentée au cours de la production, sont sans cesse développés. Le cadre industriel de notre approche est présenté ci-dessous.

Les circuits intégrés sont fabriqués à partir de plaquettes de silicium qui subissent différents types de traitement physiques ou chimiques lors d'un cycle de fabrication qui dure plusieurs semaines. Les plaquettes sont regroupées en lots, chaque lot passant sur des machines lors d'une étape du procédé. Dans les unités de fabrication modernes, il y a plus d'une centaine d'étapes de procédé. Après la fabrication, il existe deux grands types de tests :

- le test de contrôle de procédé vérifie si celui-ci est conforme à la technologie. Ces tests sont effectués sur un échantillon de circuits pour chaque lot. Ils permettent de détecter des dérives du procédé mais ne servent généralement pas à rejeter des circuits.

- Le test sous pointes, ou test électrique, est effectué sur chaque circuit de chaque lot. Le test sous pointes consiste en un ensemble de mesures électriques (courants, fréquences, tensions) dont les spécifications ont été déterminées avec les clients. Tout circuit qui n'est pas dans la spécification pour un test est rejeté. La proportion de circuits ayant passé l'ensemble des tests est appelé rendement.

Les tests de contrôle de procédé, les tests électriques, et surtout le rendement sont des paramètres très importants dans l'industrie du semi-conducteur car ils déterminent en grande partie la rentabilité de la fabrication. Ils seront appelés, dans la suite de cet article, paramètres finaux des circuits. Ces paramètres sont suivis de façon quotidienne et la moindre dérive nécessite des analyses pour détecter la cause origine du problème. Ce dernier peut venir d'une erreur de fabrication, par exemple un temps d'implantation trop long sur une série de lots. Il peut aussi s'agir d'un équipement de production défectueux, qui est parfois difficile à identifier car il en existe plusieurs centaines de différents types dans une unité de fabrication de circuits intégrés. Une méthode a été mise en place pour détecter rapidement les équipements fautifs, le tableau de bord statistique (Bergeret et col. 1996) : il consiste à réaliser des analyses de variance avec les paramètres finaux comme réponses et les *équipements d'une étape* de production donnée comme facteur. Les lots sont en effet répartis de manière aléatoire entre les équipements équivalents réalisant une même étape. Ces analyses de variance, répétées pour toutes les étapes de production, permettent de vérifier si certains équipements posent problème. Cette méthode s'est révélée efficace puisqu'elle a permis d'améliorer les rendements de plusieurs pourcents et reste largement utilisée.

1.2 Problématique

Cette méthodologie basée sur une batterie d'analyses de variance ne permet toutefois pas d'appréhender des dérives liées à une *étape entière* du procédé c'est-à-dire simultanées à tous les équipements réalisant en parallèle cette étape. Ceci peut survenir s'il existe une marginalité de procédé : une variation dans la fabrication, même minime, qui entraîne ensuite des non conformités des paramètres finaux. Un autre type de dérive peut venir de facteurs environnementaux, par exemple une contamination de l'eau ou de l'air. Dans ce cas là, comme dans le cas précédent, il est important de noter qu'une étape entière va être affectée.

Dans les cas les plus favorables, une analyse physique permet d'identifier rapidement l'étape responsable, soit parce que le test qui pose problème a déjà été relié à une cause physique bien précise, soit parce que des analyses physico-chimiques permettent de le faire. Il arrive fréquemment que des dérives ne puissent pas être expliquées de façon simple, ce qui a motivé la recherche d'une méthode pour *identifier l'étape responsable.* Dans la section suivante, nous montrons comment le mélange des lots lors de la fabrication peut apporter une information précieuse pour atteindre un tel objectif.

Chaque semaine, plusieurs dizaines de lots sont lancés dans la chaîne de fabrication. À chaque étape du procédé, les lots doivent passer sur un certain

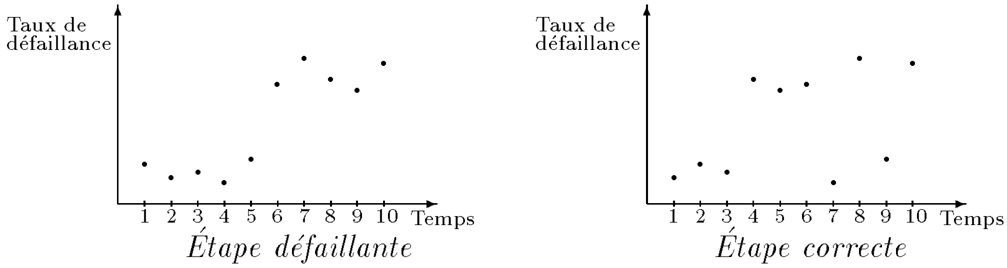

FIG. 1: *Exemple fictif : Taux de défaillance d'une série de lots en fonction de l'ordre de passage des lots dans une étape donnée.*

type d'équipement. Quand un lot arrive à l'étape, il est envoyé sur l'équipement le plus disponible. Si l'unité de fabrication tourne à pleine capacité, ce qui est généralement le cas, il y a des temps d'attente devant chaque équipement. Il est alors clair que, en fonction des aléas de la fabrication (équipement en maintenance, mesure hors contrôle qui entraîne des réglages ...), les lots vont circuler à des vitesses différentes et donc *ne vont pas passer dans le même ordre aux différentes étapes*. De ce mélange aléatoire, nous pouvons tirer une information, c'est ce qui est illustré par l'exemple fictif de la figure 1. Supposons qu'une étape génère 5 lots consécutifs corrects puis 5 lots présentant un défaut significatif, défaut qui sera détecté sur un paramètre final du circuit. Si l'on trace sur un graphe le taux de défaillance par lot au sens de ce paramètre final en fonction de l'ordre de passage des lots à l'étape (ce graphe est appelé une "time slide"), les 5 lots défectueux seront consécutifs. Si l'on trace le même type de graphe à une autre étape, les 5 ne seront généralement pas consécutifs, car mélangés par les aléas de la production. Quand on ne connaît pas l'étape défectueuse, l'idée consiste à tracer un graphique pour chaque étape du procédé afin de voir qu'elle est l'étape pour laquelle les lots défectueux sont le mieux discriminés des autres, cette étape sera alors fortement suspectée. Dans la pratique toutefois, compte tenu de la sensibilité des procédés de fabrication, il est rare qu'un ensemble de lots consécutifs soient tous défaillants. Il y a une alternance de lots bons et de lots mauvais, même à l'étape défectueuse. C'est pourquoi l'approche graphique citée ci-dessus ne se révèle généralement pas efficace, surtout lorsque une centaine d'étapes doivent être examinées, mais le principe de notre approche reste le même : les lots seront moins mélangés à l'étape défaillante qu'aux autres étapes. C'est cette information que nous allons exploiter dans les sections suivantes.

1.3 Démarche

Plus généralement, l'objectif recherché, qui n'est certainement pas spécifique à l'industrie électronique, est la détection d'une étape défaillante dans une chaîne de fabrication, chaîne dans laquelle le test d'adéquation du produit, son contrôle, n'intervient nécessairement qu'en fin de procédé. Ainsi, à moins que la dérive de certains paramètres physiques de fabrication propres à une étape ait été observée, tout le processus de fabrication peut apparaître comme une boîte noire dans laquelle il est difficile, par manque d'information, de déceler une étape défaillante. Les seules informations utilisables, sont disponibles et stockées dans une base de données, sont la date et l'heure de passage d'un lot dans une étape donnée.

Différents outils classiques et critères de sélection ont été utilisés et comparés sur deux exemples réels. Le degré de suspicion des différentes étapes est mesuré par l'adéquation d'une modélisation de la défaillance en fonction du temps ; meilleure est l'adéquation du modèle, par exemple en terme de déviance, et plus suspecte est l'étape concernée. Nous nous proposons donc de construire un classement par ajustement décroissant des trajectoires au modèle choisi. Certains outils font référence à des modèles paramétriques (régression logistique, polynômiale) d'autres à une approche non-paramétrique (lissage spline).

Une autre information intervient en deuxième lieu, c'est la liste chronologique des étapes dans le procédé de fabrication. En effet, il est facile de comparer les deux classements : le chronologique et celui basé sur la qualité d'ajustement. Ainsi, si l'éloignement à l'étape suspectée défaillante dans ce 2ème classement est compatible avec l'ordre chronologique, au moins dans les premières étapes, la présomption en est renforcée.

Les outils classiques, qui se sont révélés les plus efficaces pour les exemples considérés, sont brièvement décrits en section 2 et appliqués en section 3 sur deux jeux de données présentant des caractéristiques différentes.

2 Les outils

L'entreprise impose la contrainte, dans un premier temps, de se restreindre à un logiciel commercial classique (SAS). Ceci limite les méthodes pouvant être mises en œuvre facilement. Ainsi, différentes techniques comme la recherche d'une rupture ou discontinuité dans un modèle (Couallier et col. 1997) ou l'estimation d'un coefficient de "rugosité" comme la dimension de Hausdorff (Bardet 1998) pourront être testées dans une deuxième approche. D'autre part, une étude exploratoire préalable par une Analyse en Composantes Principales avec ou sans lissage (Besse et col. 1997) a été réalisée. C'est indispensable pour apprécier les éventuelles classes ou dispersions de trajectoires mais les résultats, sans influence sur la modélisation, ne sont pas présentés. Après différents essais "empiriques" [1], il semble que deux types d'outils sont à privilégier en fonction des circonstances et des types de données observées.

2.1 Lissage spline

Même en situation "normale", lorsqu'un procédé est sous contrôle, des défaillances sont toujours présentes, dues essentiellement, dans le cas de la fabrication de circuits intégrés, au dépôt aléatoire de particules étrangères microscopiques. Ceci engendre un bruit résiduel incompressible sur les mesures de test dont on peut supposer qu'il est de variance constante. Un lissage élémentaire (spline ou noyau) permet alors de gommer ce bruit résiduel afin de révéler les seules ruptures qui décèlent soit un problème de fabrication dans l'étape qui nous intéresse ou des permutations dans l'ordre des lots pour les étapes qui en sont chronologiquement éloignées.

Malgré leurs excellentes propriétés, les méthodes non-paramétriques ne sont pas encore largement répandues dans les logiciels commerciaux. Elles apparaissent dans le module interactif SAS/INSIGHT (1995), très pratique pour une approche exploratoire, mais inefficace lorsque il s'agit d'automatiser une démarche sur des données volumineuses. Ceci conduit à laisser de côté la régression non-paramétrique par la méthode du noyau et à privilégier le lissage spline mis en œuvre dans la procédure `transreg` du module SAS/STAT (1989).

Sous l'hypothèse que le test est l'observation bruitée d'un phénomène régulier en fonction du temps t, c'est-à-dire supposé dérivable jusqu'à un certain ordre :

$$y_j = f(t_j) + \varepsilon_j,$$

l'estimation par spline cubique de f est obtenue en recherchant, dans le bon espace fonctionnel (espace de Sobolev), la solution de

$$\min_f \left[\sum_{j=1}^{p} (f(t_j) - y_j)^2 + \lambda \int (f''(t))^2 dt \right].$$

Le paramètre de lissage λ contrôle l'équilibre entre fidélité aux données et régularité de la solution. Il est fixé par l'utilisateur, qui cherche à approcher

1. Les auteurs remercient les étudiants du Magistère d'Économiste Statisticien et de la Maîtrise de Mathématiques de l'Université Paul Sabatier pour leur concours.

une valeur “optimale” dépendant de la variance inconnue du bruit à éliminer. La stratégie la plus répandue pour pour estimer cette valeur est la validation croisée généralisée (Wahba 1990) qui est implémentée dans la plupart des logiciels (Splus, SAS/INSIGHT).

Dans notre situation, le choix de la valeur de λ est déterminé sur une période où le procédé est sous contrôle de façon à éliminer le bruit jugé usuel et laisser apparaître les problèmes réels. Ainsi, une fois les courbes lissées avec cette même valeur de paramètre de lissage, il reste à comparer les qualités d'ajustement des rendements de chaque étape en considérant les erreurs quadratiques moyennes :

$$\mathrm{MSE} = \frac{1}{p}\sum_{j=1}^{p}(y_j - \widehat{f}_\lambda(t_j))^2,$$

$\widehat{f}_\lambda(t_j)$ étant la valeur estimée de $f(t_j)$, associée au paramètre λ.

La démarche qui vient d'être décrite avec les outils disponibles dans SAS/INSIGHT peut être automatisée par l'utilisation de la procédure `transreg`. Celle-ci estime un modèle additif conjointement avec la recherche de transformations optimales des variables mais, restreinte à une variable explicative, elle exécute une régression spline sur une base de B-splines (de Boor 1978). Formellement, ce n'est pas exactement la solution du problème d'approximation ci-dessus ; néanmoins les résultats sont en pratique très proches et conduisent aux mêmes diagnostics dans notre problème. Le paramètre de lissage est alors contrôlé par le nombre de nœuds déclarés dans le modèle.

2.2 Régression logistique

Un simple histogramme permet de visualiser la distribution d'une variable test. Si celle-ci apparaît nettement bimodale et permet ainsi de discriminer les lots jugés défectueux des autres, il est alors facile de recoder la variable test en une variable binaire (0,1). On élimine ainsi radicalement le bruit résiduel de fabrication. Dans cette situation, une régression logistique vise à expliquer le logit de la probabilité de défaillance π comme une fonction affine ou polynômiale du temps t :

$$\ln(\frac{\pi}{1-\pi}) = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4$$

ce qui est équivalent à modéliser π comme une fonction sigmoïdale du prédicteur linéaire :

$$\pi = \frac{\exp(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4)}{1 + \exp(\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4)}.$$

La fonction affine est utilisée pour distinguer deux périodes d'une étape, la première correcte de la deuxième présentant des défaillances systématiques. Un polynôme du 4ème degré est nécessaire afin de modéliser une période de défaillance isolée entre deux périodes normales de fabrication.

On cherche alors les étapes dont le modèle est susceptible de discriminer au mieux bons et mauvais lots autour d'une date inconnue mais estimable (β_0) ; ce sont les étapes présentant des déviances minimales pour leur modèle respectif. Les étapes sont donc ordonnées au sens de ce critère. Le lecteur peut se reporter à Antoniadis et col. (1992) pour plus de détails sur la régression logistique. Tous les logiciels standards proposant une telle modélisation, l'automatisation de cette démarche sur un grand nombre d'étapes ne pose aucun problème.

3 Application

Parmi les cas de défaillances rencontrés à l'usine de Motorola Toulouse, nous avons sélectionné deux cas représentatifs des types de tests, et donc de distributions de ces tests, faisant partie du contrôle final des circuits. Un travail préalable de gestion des données a permis de rapprocher, pour chaque cas étudié, une extraction de la base de données archivant les dates de passage de chaque lot à chaque étape avec le fichier des résultats des tests finaux de chaque lot. Nous avons ainsi constitué deux fichiers donnant, pour chaque étape, la date de passage ainsi que la valeur moyenne au résultat d'un paramètre final de chacun des lots.

Nous n'avons pas jugé nécessaire de décrire les procédés physico-chimiques mis en œuvre dans chacune des 150 étapes de la fabrication d'un circuit intégré. Précisons simplement que les codes se terminant par DIF ou REC désignent des étapes de diffusion ou recuit au cours desquelles les plaquettes passent dans des fours à haute température pour diffuser des dopants à l'intérieur du silicium. Si le code se finit par OXC, OXP ou REO, ce sont des étapes de propagation d'un oxyde. Enfin, une étape désignée par GRA grave des motifs dessinés lors d'une étape précédente de masquage (MAS).

3.1 Test de contrôle de procédé

Le premier jeu de données est relatif à une brusque hausse d'un paramètre de contrôle de procédé. Cette augmentation génère ensuite un nombre de rejets important lors du test électrique, mais c'est le test de contrôle de procédé, une résistance mesurée en ohms, qui sera le paramètre final étudié car il est plus représentatif du problème. Dans ce premier cas, l'étape respon-

Étape	Rang	MSE
DHDIF	12	33.27
DHREO	13	36.86
GHGRA	11	38.90

TAB. 1: *Régressions splines avec 30 noeuds de la moyenne par lot de la valeur du test final en fonction de l'ordre de passage des lots. Classement des meilleurs ajustements par ordre croissant de l'erreur quadratique moyenne.*

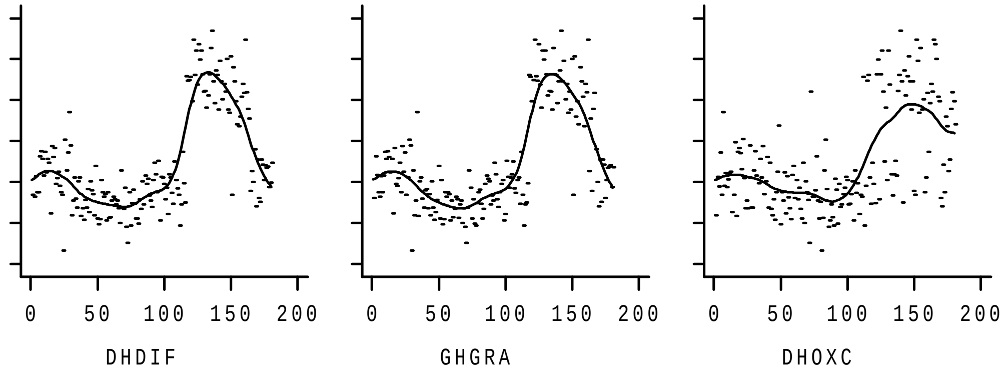

FIG. 2: *Valeur moyenne par lot du test final en fonction de l'ordre de passage de chaque lot pour l'étape défaillante, la suivante et une étape chronologiquement plus éloignée.*

sable de la dérive est connue car il a été possible, grâce à une étude clinique fastidieuse des graphiques associés à chacune des étapes, de remonter jusqu'à un fournisseur de matière première qui a livré certains lots défectueux. Cela explique la forme des courbes (cf. figure 2), sur lesquelles la brusque hausse du paramètre final se maintient un certain temps, puis les valeurs de ce paramètre redescendent à un niveau acceptable. Il faut également noter que ce problème, quoique bien réel, correspond à un cas d'école. Dans la pratique il est rare que la dérive soit si nette et que la séparation entre une bonne et une mauvaise période soit si visible, même à l'étape origine de la dérive. Ce cas d'école va donc nous permettre de tester l'outil de diagnostic. Dans l'autre jeu de données, la distinction en deux périodes n'est pas aussi nette.

L'exemple considéré se prête bien à un lissage spline car l'hypothèse, selon laquelle la variance du bruit est constante en situation de défaillance ou non, semble raisonnable à la vue des graphes (figure 2). Une fois le paramètre de lissage fixé, ou, dans le cas des B-splines, le nombre de nœuds, les estimations pour chaque étape conduisent à la construction du tableau 1 fournissant les critères des meilleurs ajustements. On a noté que ces résultats sont peu sensibles (± 5) au choix du nombre de nœuds et ils sont cohérents avec l'intuition. L'étape la plus suspecte est directement suivie, dans l'ordre décroissant d'ajustement, des étapes qui l'encadrent chronologiquement. C'est bien celle qui avait été identifiée comme défaillante par les responsables du procédé et des rendements. On remarque, à la vue des graphes de la figure 2, que la recherche visuelle est loin d'être évidente à cause du nombre très restreint de modifications qui peuvent apparaître dans l'ordre des lots d'une étape à la suivante. Il a fallu beaucoup de perspicacité et de connaissance physique du procédé avant d'arriver à la solution alors que le critère issu du lissage fournit, dans ce cas, directement la solution.

Une autre stratégie a été entreprise sur ce même jeu de données. Une analyse en composantes principales, suivie d'un examen visuel de chaque graphe, a permis d'éliminer les étapes qui, de toute évidence, sont les plus bruitées et donc pas suspectes. La distribution des valeurs de test observées pour les lots des étapes restantes apparaissant clairement bimodale, il a été facile de discriminer bons et mauvais lots par un recodage (0, 1). Un modèle de régression logistique polynômiale, de degré 4 afin de suivre la croissance puis la décroissance des tests moyens, a ensuite été ajusté pour chaque étape. L'étude des déviances à ces modèles a strictement confirmé les résultats obtenus par lissage. Cette démarche, plus manuelle et donc moins automatisable n'est pas développée pour ce jeu de données. En revanche, l'illustration d'une utilisation de la régression logistique est l'objet de la section suivante.

3.2 Nombre moyen de défaillances

Le second fichier correspond à un problème dont la cause origine n'a pas encore été trouvée à ce jour. Un changement de procédé a toutefois permis de contenir rapidement le problème. Il s'agit d'une augmentation brutale du nombre de rejets à un test électrique, mais, à la différence du cas précédent, il y a des lots bons et des lots mauvais même pendant la mauvaise période, ce qui rend la recherche de l'étape suspecte plus difficile. Ce type de phénomène est plus représentatif des dérives rencontrées, dérives dont la cause est généralement difficile à trouver, et pour lesquelles une méthode d'analyse systématique et "objective" peut fournir des pistes précieuses.

Le test électrique considéré donne cette fois un résultat binaire pour chaque circuit présent dans un lot. La variable étudiée est donc le nombre moyen de circuits défectueux par lot et les processus étudiés présentent des caractéristiques de processus de Poisson. Néanmoins, même après transformation des valeurs (racine, logarithme), les graphiques de la figure 3 montrent clairement que la variance n'est pas constante mais croît avec le logarithme du nombre moyen de défaillance par lot. La technique de lissage élémentaire

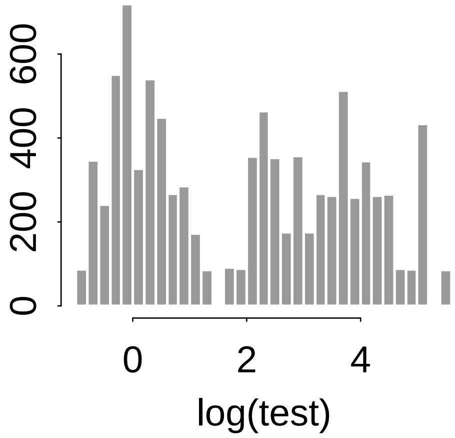

FIG. 3: *Distribution du logarithme du nombre moyen de puces défaillantes par lot.*

Étape	Rang	Déviance
DMOXP	52	100.655
PMMAS	53	100.785
DMREC	55	103.170

TAB. 2: *Régressions logistiques de la probabilité de défaillance en fonction du temps. Classement par ordre croissant de la déviance des meilleurs modèles.*

adaptée à l'exemple précédent n'est donc plus appropriée.

En revanche, l'histogramme de la figure 3 montre qu'il est facile de discriminer les deux populations de bons et mauvais lots qui sont donc recodées (0, 1) et des modèles de régression logistique sont ajustés en considérant une fonction affine du temps.

Les résultats du tableau 2 font ressortir deux étapes cohérentes avec l'ordre chronologique. L'examen visuel des graphes des étapes incriminées (figure 4) est encore plus délicate que dans l'exemple précédent. Seul un critère quantitatif peut apporter une aide effective aux ingénieurs responsables des rendements. Des investigations sont en cours afin de confirmer ou infirmer les indications fournies.

Comme pour l'exemple précédent, d'autres modèles ont été expérimentés. Ainsi, des résultats identiques sont obtenus en modélisant le logarithme du nombre moyen de défaillances par une fonction polynômiale du temps de degré 4. Seul l'ordre des 4ème et 5ème étape est inversé par rapport à celui donné par la régression logistique.

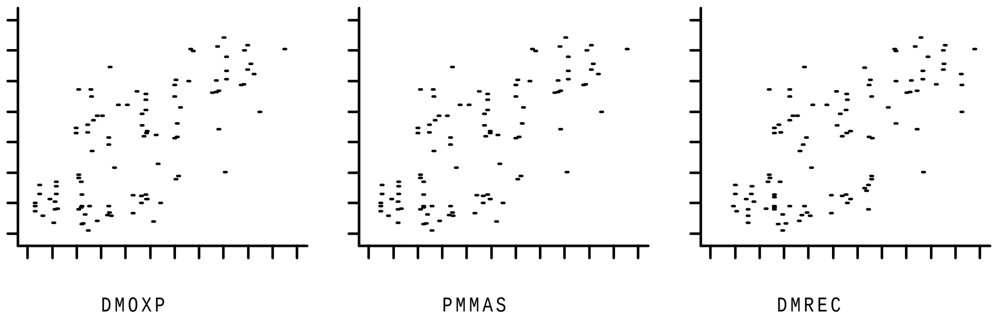

FIG. 4: *Logarithme du nombre moyen de défaillances par lot, en fonction de la date de passage de chaque lot, pour les étapes les plus suspectes.*

4 Conclusion

Des outils statistiques élémentaires, limités à ceux immédiatement disponibles dans l'entreprise, se révèlent donc très utiles pour la détection d'une étape défectueuse dans un processus aussi sensible que la fabrication des circuits intégrés. Ils peuvent se montrer précieux et efficaces car les dérives sont fréquentes et la détection de leur origine peut être très longue et très coûteuse. Si les données sont correctement stockées dans une base de données, le temps nécessaire à une analyse utilisant une des méthodes présentées dans cet article est très court et peut permettre de réagir rapidement.

Sur le premier exemple traité dans cet article, nous avons pu montrer que la méthode est fonctionnelle car l'étape défectueuse, connue a priori, a été détectée. Sur le second exemple, la solution n'est pas connue avec certitude mais il est rassurant de constater que les étapes identifiées comme suspectes sont chronologiquement proches les unes des autres, ce qui est cohérent car le mélange des lots est très faible pour des étapes consécutives. De plus, d'autres analyses sont venues confirmer cette hypothèse.

Les deux exemples présentés correspondent à deux types de test réalisés en fin de procédé. Le premier, qui mesure une mesure électrique conduit à des trajectoires assimilables à des bruits blancs décentrés. La défaillance provocant un saut de la valeur moyenne mais laissant la variance relativement inchangée. Dans cette situation, même si le procédé présente des fluctuations alternant des périodes de fonctionnement correct et des périodes de défaillance, une procédure de lissage élémentaire semble tout à fait efficace. Dans le deuxième exemple, la modélisation la plus naturelle serait de considérer un processus de Poisson. Dans le cas traité, même après transformation (log), la variance de la variable test croit avec la moyenne. Aussi, le lissage élémentaire est mis en défaut. Néanmoins, dans le cas où il est facile de discriminer les lots défectueux des autres et donc de recoder (0, 1) ces lots, la régression logistique semble très adaptée. Dans le cas contraire plus défavorable : processus de Poisson et discrimination difficile, d'autres approches moins élémentaires seront sans doute nécessaires. Depuis cette étude préliminaire, la méthodologie présentée a été utilisée avec succès sur deux autres cas. Neuf autres sont en cours d'investigation afin de confirmer les soupçons révélés par un lissage spline. Certains ont montré que le choix du paramètre de lissage pouvait s'avérer délicat. Des développements sont en cours pour aider à ce choix. En conclusion, ce travail montre l'intérêt que peut présenter un certain désordre dans un procédé aussi complexe et méticuleux que la fabrication des circuits intégrés. Paradoxalement, il faudrait sans doute augmenter volontairement la composante de ce désordre afin d'amplifier les capacités de diagnostic de la stratégie proposée.

References

Antoniadis, A., Berruyer, J., and Carmona, R. (1992). *Régression non linéaire et applications*. Economica.

Bardet, J. (1998). *Test d'autosimilarité des processus gaussiens. Dimension fractale et dimension de corrélation*. PhD thesis, Université de Paris sud, France.

Bergeret, F. and Chandon, Y. (1996). Data analysis the third pillar of spc. *Soumis*.

Besse, P., Cardot, H., and Ferraty, F. (1997). Simultaneous non-parametric regressions of unbalanced longitudinal data. *Computational Statistics & Data Analysis*, 24:255–270.

Couallier, V., Sarda, P., and Vieu, P. (1997). Estimation non-paramétrique de discontinuités d'une fonction d'intensité. *Revue de Statistique Appliquée*, 45:89–106.

De Boor, C. (1978). *A Practical Guide to Splines*. Springer Verlag.

SAS (1989). *SAS/STAT User's Guide*, fourth edition. version 6.

SAS (1995). *SAS/INSIGHT User's Guide*, third edition. version 6.

Wahba, G. (1990). *Spline Models for Observational Data*. SIAM.