

F. BERGERET

Y. CHANDON

C. LE GALL

De la statistique dans l'industrie : un exemple à Freescale

Journal de la société française de statistique, tome 145, n° 1 (2004),
p. 71-95

http://www.numdam.org/item?id=JSFS_2004__145_1_71_0

© Société française de statistique, 2004, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) 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/>

DE LA STATISTIQUE DANS L'INDUSTRIE : UN EXEMPLE À FREESCALE (anciennement MOTOROLA semiconducteurs)

F. BERGERET *, Y. CHANDON **, C. LE GALL ***

RÉSUMÉ

L'objectif de cet article est d'apporter une contribution sur la pratique de la statistique dans l'industrie et sur les rapports qui peuvent s'établir entre le monde industriel et la recherche statistique. Cette contribution se base sur une expérience de terrain dans une usine de développement et de production de composants électroniques. À partir de besoins, on décrit les techniques statistiques employées selon les « métiers ». Le système de formation et de management de la statistique est exposé. Les opportunités d'amélioration apportées par l'emploi des techniques statistiques sont mises en perspective.

ABSTRACT

The goal of this paper is to discuss applied statistics in the semiconductor industry, and also the links between statistical research and industry. This contribution is based on an actual experience in an integrated circuits fabrication unit. Starting with the needs, statistical methods are described, depending on the job. Training system and management of statistics are presented. We also detail improvements that can be obtained by statistical methods.

1. Introduction

Cet article vise à faire partager notre expérience de statisticien dans une usine de fabrication de circuits intégrés, Freescale semiconducteurs France, site de Toulouse, anciennement Motorola semiconducteurs. En particulier, nous décrivons comment notre travail s'insère dans les différentes activités de l'usine, quelles difficultés ont été rencontrées et comment certaines ont pu être résolues. Nous détaillons aussi nos relations avec la recherche statistique

* Freescale semiconducteurs France SAS, 134 avenue Eisenhower, BP72329, 31023 Toulouse cedex 1.

E-mail : francois.bergeret@freescale.com

** Freescale Crolles 2 Alliance, 850, rue Jean Monet 38926 Crolles Cedex.

E-mail : yves.chandon@freescale.com

*** Freescale semiconducteurs France SAS, 134 avenue Eisenhower, BP72329, 31023 Toulouse cedex 1.

E-mail : caroline.le.gall@freescale.com

et l'université avec tous les apports amenés à notre usine. La section 2 décrit l'activité de cette usine en production ; les principaux « métiers » et la façon dont ces « métiers » utilisent la statistique à tous les niveaux de l'entreprise sont détaillés dans la section 3. On verra dans la section 4 comment l'entreprise se structure pour tenir compte de ses besoins statistiques par la formation, les outils informatiques, la spécialisation d'ingénieurs, l'intégration de statisticiens dans les équipes et les collaborations universitaires. On réfléchira aux opportunités importantes qu'offrent des développements dans ce domaine et à la façon de surmonter les obstacles qui se présentent régulièrement. Enfin, nous consacrons la section 5 à une synthèse des réflexions de George Box sur la statistique, catalyseur de l'apprentissage scientifique, et à la façon dont elles sont illustrées dans la pratique de l'entreprise. Ces apports sont en effet très importants et devraient permettre de progresser encore dans l'utilisation de la statistique.

2. Description de l'activité et des « métiers » principaux

L'usine compte environ 2300 employés répartis en 40 % de cadres, 40 % d'opérateurs et 20 % de personnels techniques et intermédiaires. Les circuits intégrés et systèmes embarqués fabriqués sont utilisés dans l'automobile pour les freins ABS et les systèmes airbags, dans la téléphonie mobile, dans les imprimantes, etc. Ces circuits intégrés sont fabriqués en salle, dite blanche, où un système de filtration dépoussière l'air de façon à ce qu'il y ait moins de 300 particules par mètre cube contre un million dans l'air ambiant habituellement.

Les circuits sont fabriqués par lots de 24 plaquettes de silicium, chaque plaquette comptant entre 200 et 2000 circuits. Chaque circuit est fait de plusieurs milliers de composants élémentaires : transistors à effets de champs et bipolaires, diodes, capacités, résistances ... Une usine peut fabriquer par exemple 5 technologies différentes, à chaque technologie correspond une bibliothèque de composants élémentaires différente (par leur fonctions, leur taille minimum, les courants et les tensions limites qu'ils supportent). Pour chaque technologie, on fabrique de une à plusieurs dizaines de produits différents.

La fabrication consiste en plus de deux cents étapes élémentaires, qui sont des procédés physico-chimiques : dépôt cristallin par épitaxie, définition de motifs jusqu'à $0.8 \mu\text{m}$ par photolithographie, dopage des zones de silicium par implantation ionique et diffusion thermique, dépôt de couches isolantes conductrices. Il y a en tout 60 procédés dont certains reviennent à plusieurs étapes réparties sur plus de 400 équipements. Le temps de cycle pour fabriquer un lot de plaquettes varie entre 30 et 60 jours selon la complexité. Chaque équipement fait l'objet de procédures de maintenance préventive et de contrôles sur des paramètres qu'on appellera d'équipement ou d'entrée de procédé. Ces paramètres sont des flux de gaz, des températures, des pressions, des compositions chimiques, ...

Au cours de la fabrication de nombreuses mesures sont faites sur les lots à des fins de maîtrise statistique des procédés. On les appellera paramètres de

sortie du procédé. Ce sont des mesures de dimensions, d'épaisseurs de couches isolantes et conductrices, de nombre de particules, de résistivité, ...

En fin de fabrication, on mesure un échantillon de composants élémentaires de chaque lot sur quelques dizaines de paramètres principaux par composant ; on les appelle paramètres de technologie, ce sont des paramètres physiques caractéristiques des composants élémentaires utilisés, qui sont combinés ensemble pour réaliser le circuit intégré, ce sont des gains de transistors, des claquages de diodes, par exemple. Cela représente 300 paramètres mesurés pour l'ensemble des composants élémentaires.

Puis, chaque circuit est testé individuellement sur les 50 à 1000 paramètres qui le caractérisent, il s'agit des tests électriques. Tous les résultats de ces tests sont transférés dans une base de données techniques. Si un test est hors spécification, alors le circuit est déclaré mauvais et marqué d'une tâche d'encre rouge. Le pourcentage de puces bonnes après ce test est le rendement. C'est une donnée économique essentielle et un critère décisif de compétitivité de l'unité de fabrication. On cherche à introduire une technologie nouvelle aussi rapidement que possible, avec un rendement d'introduction aussi élevé que possible. Vu la complexité technologique, ce rendement n'est jamais optimum dès l'introduction de la technologie, on cherche à atteindre le plus rapidement possible le niveau de rendement de maturité.

Les plaquettes sont ensuite envoyées dans une usine d'assemblage où on les découpe par circuits qui sont assemblés dans un boîtier à l'intérieur duquel les accès électriques du circuit sont reliés par des fils aux pattes métalliques qui servent d'accès au boîtier, puis les produits finis subissent un test final qui doit garantir la conformité absolue du produit. Le rendement test final est également une donnée de coût importante.

On peut grossièrement résumer la structure d'enchevêtrement des paramètres. Un paramètre de sortie de procédé est fonction de plusieurs paramètres d'entrée de procédé, un paramètre de technologie est fonction de plusieurs paramètres de sortie de procédé, les tests électriques et donc le rendement sont fonction des paramètres de technologie et de facteurs de limitation de rendement, comme la contamination par les particules. On peut voir un schéma de ces relations sur la figure 1.

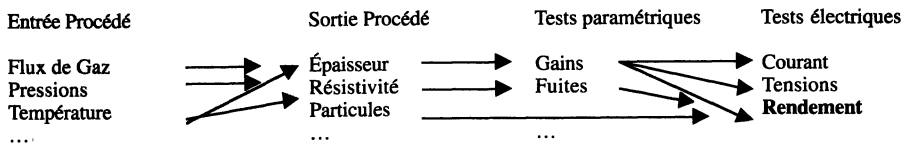


FIG 1. — La jungle des données dans l'industrie du semi-conducteur.

3. Les « métiers » et leurs principaux rapports avec la statistique

3.1. Les concepteurs

En amont de la fabrication, les concepteurs définissent le circuit à partir des besoins exprimés par le client pour remplir des fonctions. Ils choisissent la technologie qui permet de remplir ces fonctions au meilleur coût, puis assemblent virtuellement à l'aide de simulateurs les composants de cette technologie en sous-ensembles et ensembles pour remplir les fonctions électroniques. L'emploi de la statistique commence à se répandre dans ce secteur où l'approche utilisée classiquement consiste à prédire, en plus des valeurs cibles, les valeurs supposées « meilleures » et « pires ». Des applications prometteuses commencent à voir le jour. Cela consiste notamment à générer avec les simulateurs des distributions des paramètres du circuit à partir des distributions des paramètres de procédé. Ces distributions des paramètres de procédé sont obtenues par une méthode de rétro-propagation de la variance des paramètres liées à la technologie. Ce type de simulation permet d'estimer *a priori* ce que l'on appelle la capacité des paramètres du circuit : la capacité d'un paramètre ou d'un procédé indique la faculté du procédé de fabrication d'être toujours à l'intérieur des valeurs limites de spécifications. Pour cela une simulation Monte Carlo ou un échantillonnage stratifié de type hypercube latin qui offre un meilleur compromis entre nombre d'essais et qualité d'estimation est utilisé. Voir Liu (1998) pour des détails sur les hypercubes latins. Les simulations prennent du temps et exigent de la puissance de calcul, on peut les approximer, dans le domaine exploré, par un polynôme du second degré à l'aide d'un plan d'expérience de réponse de surface ou un hypercube latin (il n'y a pas ici d'erreur dans le modèle), cela permet d'avoir un compromis entre nombre d'essais et temps de simulation. Des exemples de ces applications ont récemment permis d'optimiser une conception, notamment en s'intéressant au compromis entre rendement et surface de silicium utilisée par circuit, compromis qui est évidemment lié au coût. L'intégration de la statistique à l'intérieur de l'environnement complexe des logiciels de conception rend ces outils plus accessibles et l'utilisation de ces méthodes se développe actuellement grâce à cela.

3.2. Les ingénieurs produit et de test ¹

La fonction d'ingénieur produit est simplement la responsabilité du produit. Cela comprend l'introduction du produit en production, l'écriture des programmes de test, faite par l'ingénieur de test, qui garantissent au client que son produit correspond bien aux spécifications. L'ingénieur produit est en

1. On utilise ici le mot ingénieur dans son sens le plus traditionnel pour les métiers proches des sciences physiques, et le mot statisticien pour désigner les personnes ayant fait des études en statistique même s'ils sont ingénieurs en titre et ont des fonctions qui relèvent de l'ingénierie au sens large, cette distinction ayant pour but de bien distinguer leur apport spécifique.

autre responsable de l'équilibre économique de son produit. Actuellement, le test est le principal domaine d'application statistique de cette responsabilité. Le test a trois caractéristiques principales :

- La couverture de test doit être parfaite. Cela consiste à soumettre le circuit à des stimuli équivalents à ceux qu'il recevra dans son application réelle et à mesurer des réponses électriques, principalement des tensions et des courants. Typiquement ce que l'on veut éviter c'est qu'un circuit passe le test et provoque une panne dans le dispositif dans lequel il est employé. Les conséquences peuvent être graves quand une telle panne touche un dispositif de sécurité, par exemple des systèmes ABS ou airbags sur les voitures. Les constructeurs automobiles sont d'ailleurs particulièrement exigeants sur les niveaux de qualité et demandent avec fermeté des plans pour n'avoir plus aucun défaut de fonctionnalité de circuit quelle que soit l'application. Notons tout de même que le taux de défaut est actuellement extrêmement faible, et que, dans ce domaine, l'industrie du semiconducteur est une des plus avancées en terme de qualité. Une méthode très simple consiste à tester les paramètres avec des limites de spécification statistiquement établies de façon à écarter les circuits atypiques. La façon classique était d'établir des limites fonctionnelles pour la garantie du circuit, mais on s'est aperçu qu'un circuit ayant une donnée atypique sur un paramètre pouvait être révélateur d'une faiblesse susceptible d'entraîner un défaut par la suite (exemple : une micro particule à un endroit du circuit qui ne l'empêche pas de fonctionner dans l'immédiat mais va causer par la suite une dégradation). On estime que cette méthode permet de diminuer de moitié les pannes client au prix d'une petite perte en rendement. Elle est appelée PAT pour Part Average Testing. Une application récente de cette technique simple sur un produit ayant des problèmes chez le client a permis d'atteindre zéro défaut depuis plusieurs mois. Le PAT est donc déployé sur l'ensemble des produits automobiles.
- La mesure du test elle-même doit être répétable et la variation de la mesure doit être négligeable par rapport à la largeur de spécification. Les circuits sont mesurés sur différents testeurs avec différentes cartes électroniques et différents socles. On fait des analyses de variation (utilisant les résultats d'analyse de variance) pour décomposer les contributions à la variance totale de ces différents éléments. Cela permet de détecter et d'améliorer certains systèmes de mesure. Quand on ne le peut pas, on protège le client en resserrant les limites au prix d'une perte potentielle de rendement.
- L'optimisation du nombre de tests. Comme déjà indiqué, les circuits intégrés sont testés au minimum deux fois, une première fois juste après la fabrication lors du test électrique aussi appelé test sous pointe. Le but de ce test est surtout économique : il s'agit de détecter les rejets avant la mise sous boîtier car ensuite le coût d'un rebut est plus élevé, incluant le coût du boîtier. Ensuite, il y a un second test, dit final, qui garantit au client que les circuits répondent à ses exigences. En effet, la mise en boîtier peut générer des rejets supplémentaires qu'il faut absolument détecter. En fait, pour nos applications les plus critiques type ABS ou airbag, le test final est

répété plusieurs fois pour simuler l'environnement de fonctionnement d'une voiture : test final à chaud, c'est-à-dire à plus de 100° C, et test final à froid c'est-à-dire à moins de 40° C. Cette multiplication de tests, pour garantir les exigences du client, conduit à un coût du test de plus en plus élevé. Pour réduire ce coût, une méthode appelée TOAST (Trimming Of Almost Similar Tests) a été développée dans le cadre d'une convention CIFRE (Convention Industrielle de Formation par la REcherche) entre Motorola Toulouse et le laboratoire de Statistique et Probabilités de l'Université Paul Sabatier entre 1992 et 1995. Sans entrer ici dans le détail de la méthode, TOAST fonctionne de la façon suivante : au niveau du test sous pointe, pour chaque circuit testé et pour chaque test, on enregistre l'information binaire : le test est bon ou mauvais pour ce circuit. En utilisant un échantillon de plusieurs milliers de circuits, il est théoriquement possible d'identifier un sous-ensemble de tests qui captent tous les rejets de l'échantillon. En pratique, TOAST ne cherche pas à capter tous les rejets, mais une très grande proportion d'entre eux. Les rejets non-détectés sont ensuite détectés au test final, générant un coût supplémentaire, le coût du boîtier principalement. Il s'agit donc d'optimiser une fonction de coût qui prend en compte les économies réalisées lors du test sous pointe (en supprimant des tests) et le coût des rejets supplémentaires au final. C'est un problème d'optimisation combinatoire et, vu le très grand nombre de tests (plusieurs centaines), le nombre de solutions est très grand, par exemple $3.27E150$ pour 500 tests. Différents algorithmes ont été évalués, des algorithmes pas à pas, décrits dans Bergeret et Chandon (1995) et des algorithmes stochastiques dont le recuit simulé et l'algorithme génétique décrits dans Bergeret et Besse (1997). Il est intéressant de noter que, sur un problème très concret comme celui de la réduction des temps de test, le recuit simulé s'est révélé nettement supérieur à l'algorithme génétique, alors que ce dernier était nettement plus présent dans la littérature sur l'optimisation combinatoire. Au-delà de cet aspect technique, la mise en oeuvre concrète de la réduction de temps de test n'a pu se mettre en place qu'en surmontant un certain nombre de réticences. De nombreux ingénieurs de test et ingénieurs produits, responsables devant le client, craignaient que la suppression de tests au niveau du test sous pointe ait un impact négatif sur la protection du client et la connaissance du produit. Sur un des premiers produits analysés, il nous a été répondu « il y a certains tests que je ne supprimerai jamais ». La réponse du statisticien et de sa hiérarchie consiste alors en un travail d'explication et de discussion pour convaincre qu'il n'y a aucun risque vis à vis du client en supprimant des tests car un ou plusieurs tests non supprimés couvrent systématiquement le test supprimé (ils rejettent à sa place le circuit mauvais). La vision graphique d'une simple table de contingence peut grandement aider à convaincre l'auditeur. Le test final, comme indiqué, est de plus un dernier rempart avant la livraison au client. Dans ce cas précis, l'immersion du statisticien dans le service impliqué par la réduction du temps de test a été un facteur très important de réussite. Après le passage en production des deux premiers produits avec réduction de temps de test, aucun problème n'ayant été décelé, la méthode

s'est généralisée et permet maintenant de diminuer de 55 % en moyenne le temps de test. Les économies réalisées à ce jour sont supérieures à neuf millions de dollars.

3.3. L'ingénieur de procédé

L'ingénieur de procédé est responsable de la mise au point et du bon fonctionnement de un ou plusieurs des soixante procédés élémentaires de fabrication. La performance du procédé doit se situer de façon stable bien à l'intérieur des limites de spécification qui ont été déterminées de façon à assurer le bon fonctionnement de la technologie et des produits, et cela doit être accompli à un coût minimum. L'outil principal d'optimisation est le plan d'expérience sous toutes ses formes. On utilise :

- Des plans d'écrantage ou de tamisage (de l'anglais screening design) de type Plackett-Burman ou Taguchi, qui permettent de sélectionner les paramètres d'entrée les plus influents parmi un certain nombre (choix des quelques paramètres parmi les nombreux potentiels selon le principe de Pareto qui veut que 80 % des effets soient dus à 20 % des causes). De par le nombre de paramètres, ces plans sont pratiquement plus compliqués à mettre en œuvre et sont les moins acceptés. Ils sont néanmoins très efficaces et cela a été un des principaux apports de le montrer dans l'industrie japonaise des années 1970. On peut faire des plans d'une dizaine de facteurs, 18 facteurs étant jusqu'à maintenant le «record» interne. Il faut que la personne qui conçoit le plan explique bien les objectifs de l'expérience aux personnels qui vont la réaliser et documente précisément. Ce dialogue permet d'ailleurs d'enrichir l'expérience.
- Les plans factoriels ou fractionnels classiques qui permettent de quantifier les effets et les interactions.
- Les plans de réponse de surface type Central Composite ou Box Benken qui permettent d'approximer des effets non linéaires.

On applique sur ces plans des techniques d'optimisation de plusieurs variables de type désirabilité maintenant présentes dans les logiciels et qui sont particulièrement adaptées à l'optimisation conjointe des critères de qualité et de coût. Ces techniques sont décrites dans Derringer et Suich (1980). Un des buts de l'utilisation des plans d'expérience est de réduire les variabilités des procédés en même temps qu'on vise une cible et qu'on cherche à minimiser les coûts. Cela a été l'objet de la première collaboration entre Motorola et l'Université Paul Sabatier en 1988. Un plan d'expérience avait été réalisé dans le but de réduire la variabilité d'un équipement de mesure de dimensions critiques, ce plan est décrit dans Chandon et Mathieu (1993). Il fallait donc traiter en variable de sortie l'écart type ou la variance qui ne présentent pas la propriété d'homoscédasticité. Il avait été imaginé de comparer des variances pondérées avec un test de Fisher ce qui était sensé mais non justifié théoriquement. Une solution fut alors proposée en utilisant le modèle linéaire généralisé. L'intérêt pour cette problématique alors encore peu répandue s'est depuis considérablement développé dans la littérature. L'impact d'une bonne

utilisation des plans d'expérience est très significatif. Un exemple d'utilisation de plan d'écrantage avec optimisation conjointe de l'uniformité sur plaquette, de la vitesse de dépôt et de la contamination particulière a par exemple permis d'améliorer ces caractéristiques tout en augmentant la charge de plaquettes dans la machine de 50 %. Cela permet l'économie d'investissement d'un équipement soit un million de dollars.

Un autre domaine important est la maîtrise des procédés. Au cours de la fabrication de nombreuses mesures sont faites sur les plaquettes à des fins de maîtrise statistique des procédés. Ce sont des mesures de dimensions, d'épaisseurs de couches isolantes et conductrices, de nombre de particules, de résistivité, ... Certains contrôles se rattachent à l'équipement et d'autres sont faits sur les lots. Les opérateurs utilisent les règles des cartes de contrôle de type Shewart afin d'intervenir sur un équipement ou un procédé quand la dérive est significative et avant qu'elle ne puisse avoir d'impact sur la qualité du produit. Ils disposent de procédures dites « plans d'actions en cas de hors contrôle », qui décrivent les vérifications et réglages élémentaires permettant de revenir en contrôle. Quand ces procédures sont insuffisantes, les équipements où procédés sont arrêtés et cela demande l'intervention des équipes de maintenance d'engineering procédé. Il y a plusieurs milliers de cartes de contrôles, une partie des données est transférée du logiciel de Maîtrise Statistique des Procédés (MSP) dans la base de données techniques. On regarde aussi les capacités qui doivent être élevées et l'indice d'instabilité (pourcentage de points hors contrôle) qui doit rester faible. Dans le cas contraire, les procédés font l'objet de plans d'amélioration. Les performances des procédés les plus critiques sont revues de façon systématique dans des réunions. On y regarde comment les hors contrôle sont résolus. On y discute et on y décide les actions pouvant amener des améliorations, changement de procédure, modification des équipements, plans d'expériences pour améliorer la performance. D'autre part, dans les unités de fabrication de circuits intégrés les plus récentes, on cherche à faire le contrôle directement au niveau de paramètres de type « entrée procédé » collectés directement dans les équipements par des capteurs. Cette approche est appelée APC (Advanced Process Control). Ces mesures peuvent être assez nombreuses et peuvent être décrite par une Analyse en Composantes Principales. Un fonctionnement anormal de l'équipement peut être détecté par un changement de signature sur les graphes de l'ACP (Fault Detection). On cherche également à lier ces variables d'entrée de procédé aux variables de sortie de procédé en utilisant la régression PLS. Ces approches ont été décrites par Pasqualini *et al.* (2002) et Ambrozic (2003).

3.4. L'ingénieur d'amélioration de rendement

Les ingénieurs d'amélioration de rendement sont appelés couramment ingénieurs « device ». Leur rôle principal est l'amélioration des rendements par produit ou famille de produit. Pour ce faire, ils disposent de leur connaissance analytique du circuit et de techniques d'analyse de défaillance qui permettent de remonter à l'origine des causes de défaillances. Ces techniques

sont puissantes mais leur complexité limite l'échantillonnage. L'analyse de la base de données fournit donc des renseignements précieux.

3.4.1. Un exemple simple : le tableau de bord

Une technique très couramment utilisée est appelée tableau de bord. Elle répond à la question suivante : pour une caractéristique donnée, comme le rendement, y-a-t-il pour chaque étape une différence significative entre les différents équipements utilisés à cette étape ? Pour y répondre on fait une série d'analyses de variance (incluant le test de Kruskal Wallis si les données sont fortement non normales). Les résultats sont mis à disposition sur une page intranet : étapes ayant une différence statistiquement significative, distributions par équipement, graphe d'évolution en fonction de l'ordre de passage à l'étape pour les équipements différents. L'utilisation des tableaux de bord a révélé des marginalités d'équipement ou de procédé. Souvent, ces outils révèlent des points que les ingénieurs n'auraient jamais soupçonnés autrement, ou même qui sont des surprises complètes pour eux. Les résultats sont présentés dans des réunions, les responsables des procédés suspectés vérifient les cartes de contrôle, il peut s'agir dans certains cas de fausse alarme. Certains équipements vont ressortir de façon récurrente. On va alors aller plus loin dans l'analyse et faire un plan d'expérience pour caractériser cette différence entre équipements. On peut alors résoudre le problème en optimisant le procédé, ou les procédures de maintenance préventives ou les cartes de contrôle (type de mesure, type de carte, recalcul des limites). Des améliorations nombreuses ont ainsi été réalisées, ce qui a permis d'atteindre rapidement le niveau de rendement de maturité d'une technologie complexe. Le programme très complexe du tableau de bord a été réalisé par plusieurs stagiaires statisticiens.

3.4.2. Un exemple complexe : la recherche d'une étape défaillante

Un autre exemple d'amélioration a été la mise en œuvre d'une méthode de recherche d'étape défaillante appelée SODA (Step Origin of a Drift Analysis) dans le procédé de fabrication. Le principe de SODA est présenté dans Bergeret et Besse (1999). D'une manière générale, le problème se pose de la façon suivante :

- Le tableau de bord n'a pas permis de détecter un équipement responsable d'une baisse de rendement.
- La présence de la défaillance n'est observée qu'en fin de cycle malgré le contrôle de procédé réalisé tout au long de la fabrication. En effet, sur certains problèmes de rendement, les paramètres de procédé ne donnent pas d'information pertinente pour résoudre la question.
- Les seules informations disponibles pour remonter à la défaillance sont alors les dates de passage des lots aux étapes.

La recherche d'une étape défaillante est basée exclusivement sur l'étude du mélange des lots aux étapes, le mélange étant décrit par les dates de passage. Durant la fabrication, le mélange est généré par les aléas de la production :

certains lots sont plus urgents que d'autres du fait des exigences de clients, si l'unité de production tourne à plein volume les arrêts de machines pour des maintenances retardent certains lots, etc. À l'étape défaillante, les lots bons et mauvais ne se mélangent pas : après apparition du dysfonctionnement à l'étape, les lots mauvais se succèdent, alors qu'ils se mélangent avec les lots bons aux autres étapes. En pratique, la séparation entre lots bons et mauvais n'est pas aussi nette, une simple inspection graphique n'est donc pas suffisante pour identifier l'étape défaillante. Toutefois, en faisant l'hypothèse que les lots bons et mauvais sont mélangés aux étapes non-défaillantes, il est raisonnable de considérer que le rendement observé en fonction du temps sera plus facilement modélisable à l'étape défaillante car elle est moins bruitée par le mélange. À partir de cette observation, un critère de comparaison des modélisations a été établi pour identifier l'étape défaillante.

Soit n le nombre de lots de l'échantillon, $Y = (y_i, i = 1, \dots, n)$ la variable représentant le rendement des lots i et $X = (x_i, i = 1, \dots, n)$, la variable représentant l'ordre de passage des lots aux étapes. Pour chaque étape, il s'agit donc de modéliser Y par X selon une fonction f telle que $y_i = f(x_i) + \varepsilon_i$, $i = 1, \dots, n$ avec $\varepsilon = (\varepsilon_i, i = 1, \dots, n)$ représentant l'erreur aléatoire d'espérance nulle. Différentes études ont permis de tester différentes modélisations : des modélisations paramétriques avec des régressions polynomiales ou logistiques et des modélisations non-paramétriques avec des lissages splines ou par noyau. Les modélisations non-paramétriques à l'aide de lissages donnent de très bons résultats, contrairement aux approches paramétriques. Toutefois, la recherche d'une fonction f relativement régulière ne semble pas *a priori* très adaptée pour modéliser des ruptures dans le fonctionnement d'une étape; de plus, en pratique les méthodes de lissage nécessitent le choix d'un paramètre (taille de fenêtre pour les noyaux et paramètre du lissage spline) dont l'influence sur les résultats a incité à tester d'autres approches. Deux méthodes destinées à la détection de ruptures ont donc été testées : l'une, décrite dans Green (1996), basée sur un algorithme MCMC (Markov Chain Monte Carlo) et l'autre, basée sur la segmentation optimale d'un processus aléatoire, décrite dans Lavielle (2002).

Dans le premier cas, il s'agit de construire une chaîne de Markov dont la distribution stationnaire unique représente la distribution *a posteriori* du nombre de ruptures et de leur position. À partir du nombre de ruptures et de leur position, une fonction étagée peut être estimée par moindres carrés. Lorsque la distribution stationnaire est atteinte, le théorème d'ergodicité est appliqué et la moyenne ergodique des fonctions étagées est calculée pour finalement modéliser la fonction f . Pour construire la chaîne de Markov, il est nécessaire d'avoir un *a priori* sur le nombre moyen de ruptures (*i.e.* nombre moyen d'occurrence de la défaillance). À partir d'exemples concrets, on montre que les résultats obtenus sont peu sensibles au choix de l'*a priori*, ce qui constitue un avantage certain sur les lissages traditionnels (splines et noyaux).

La deuxième approche est basée sur la segmentation optimale d'un processus aléatoire. Dans ce cas, $Y = (y_i, i = 1, \dots, n)$, avec n le nombre de lots, est

supposé être un processus réel stationnaire par morceaux. La segmentation optimale de Y s'obtient à partir de $R = (R_1, \dots, R_n)$, le vecteur aléatoire tel que R_i prend la valeur 1 si i correspond à la position d'une rupture et 0 sinon. La distribution *a priori* π de R est supposée être une suite de variables de Bernoulli indépendantes telle que

$$P(R = r) = \pi(r) = \lambda^{\sum_{i=1}^n r_i} (1 - \lambda)^{n - \sum_{i=1}^n r_i}$$

avec λ représentant la probabilité *a priori* qu'une rupture ait lieu à la position i . L'estimateur de l'*a posteriori* maximum de R représente la réalisation la plus probable de R étant donné Y . Elle est donc obtenue en maximisant la distribution *a posteriori* de R . Cela revient à optimiser une log-vraisemblance (ou de manière plus générale, une fonction contraste) pénalisée par le nombre de ruptures. Cette méthode permet de détecter des ruptures dans la moyenne et/ou la variance de Y . L'approche basée sur la modélisation markovienne et par le lissage spline sont apparues les plus efficaces pour répondre à la problématique de recherche d'une étape défaillante. Des applications sont présentées dans Bergeret et Le Gall (2003). Toutefois, du fait de la complexité des mécanismes de défaillance, ces approches ne sont pas toujours adaptées et peuvent donc générer des fausses alarmes sur des étapes. Sachant que le temps de cycle de fabrication est de plusieurs semaines, le moindre retard dans l'identification de la défaillance peut être fortement préjudiciable en terme de rendement. L'estimation de la fiabilité des résultats obtenus à l'issue de la modélisation ou même l'estimation de la confiance qui peut être attribuée aux résultats s'est donc avérée indispensable. On cherche alors à apprécier la qualité du mélange des lots à l'étape jugée défaillante par la modélisation. En procédant à des simulations du procédé de fabrication, on peut comparer le mélange des lots durant le procédé avec le mélange d'un jeu de cartes. Plus précisément, à partir des bornes théoriques des vitesses de mélange d'un jeu de cartes développées dans Diaconis (1988), il s'agit d'évaluer si le mélange des lots bons et mauvais à l'étape jugée défaillante est éloigné d'un mélange uniforme des lots. Si c'est le cas, les résultats seront diffusés avec confiance. Toutefois, certains mécanismes de défaillance sont tellement complexes que cette approche ne permet pas toujours de conclure entre fausse alarme et étape réellement défaillante. Pour cette raison, un test d'hypothèse a été construit. Il s'agit de tester l'hypothèse nulle d'absence d'une étape défaillante dans le procédé contre l'hypothèse alternative de présence d'une étape défaillante. L'hypothèse nulle suggère qu'à l'issue de la modélisation n'importe quelle étape pourrait être jugée défaillante ce qui reviendrait à choisir aléatoirement l'étape défaillante selon le tirage d'une loi uniforme. En s'inspirant du concept de bagging de Breiman (1996) et du théorème central limite, on montre alors que tester l'hypothèse nulle revient à tester l'hypothèse de normalité du critère retenu pour évaluer la présence d'une défaillance. Plus de détails sur ces approches sont donnés dans Besse et Le Gall (2004).

Présentons maintenant un exemple. Alors que l'usine était presque en situation de rupture de livraison auprès d'un client en raison d'une augmentation

du nombre de circuits non conformes aux exigences électriques, il était indispensable d'identifier la cause à l'origine des non-conformités des puces. À partir d'un mois d'historique sur les données électriques, une recherche d'étape défaillante a été réalisée. Cette méthode a nettement pointé des étapes de fin de fabrication parmi les centaines d'étapes du procédé de fabrication. Une analyse physico-électronique a confirmé la suspicion sur une étape en fin de fabrication. La Figure 2 est une illustration de l'étape suspectée. Le rendement électrique de chaque lot est représenté en fonction de son ordre de passage à l'étape. La courbe du rendement ainsi obtenue est ajustée par un lissage. L'étape qui regroupe le mieux les lots avec un faible rendement, ce qui est la définition « théorique » d'une étape dite défaillante, est celle qui aura le lissage le mieux ajusté. Avec l'ingénieur responsable de l'étape, nous avons fait des analyses complémentaires qui ont fortement confirmé la suspicion sur l'étape et nous avons envisagé un mécanisme de défaillance. Lors de la présentation des résultats, les solutions d'améliorations proposées n'ont pas été retenues, notamment parce que le procédé de fabrication faisait l'objet de profondes modifications qui avaient pour but d'éliminer les rejets aux tests électriques. Toutefois, l'application du nouveau procédé n'étant pas immédiate, le mois suivant, on a continué à observer des rejets importants. Une nouvelle recherche d'étape défaillante a été réalisée. Une nouvelle fois l'étape préalablement suspectée a été pointée par l'analyse. En raison de cette redondance et des analyses complémentaires déjà réalisées, notre hypothèse sur le mécanisme de défaillance a été retenue et nous avons pu ainsi diminuer nos rejets et surtout livrer notre client sans retard. Cet exemple souligne plusieurs aspects de l'utilisation de la statistique dans notre industrie :

- La nécessité d'un travail en étroite collaboration avec les ingénieurs. Quand l'analyse pointe une ou des étapes, la validation physique est évidemment indispensable.
- La nécessité d'évaluer le taux de confiance qui peut être donné aux résultats. Dans notre industrie, les ingénieurs sont familiers de la notion de test d'hypothèse et des probabilités de rejet d'hypothèses qui en découlent. La notion du risque de première espèce (α) est bien maîtrisée. Il a donc été nécessaire de développer cette notion à la recherche d'une étape défaillante. En effet, la puissance avec laquelle la méthode peut identifier une étape défaillante pas toujours soupçonnée par les ingénieurs, sans connaissance *a priori* du problème, doit être accompagnée d'une mesure de confiance permettant d'être crédible auprès des ingénieurs.
- La nécessité d'une grande persévérance. Il est déjà arrivé que la méthode pointe une étape à laquelle aucun ingénieur ne croyait et qui s'est pourtant avérée à l'origine des rejets aux tests électriques. Les ingénieurs peuvent être difficiles à convaincre, il ne faut donc pas hésiter à rééditer des analyses afin de confirmer les suspicions sur des échantillons différents.

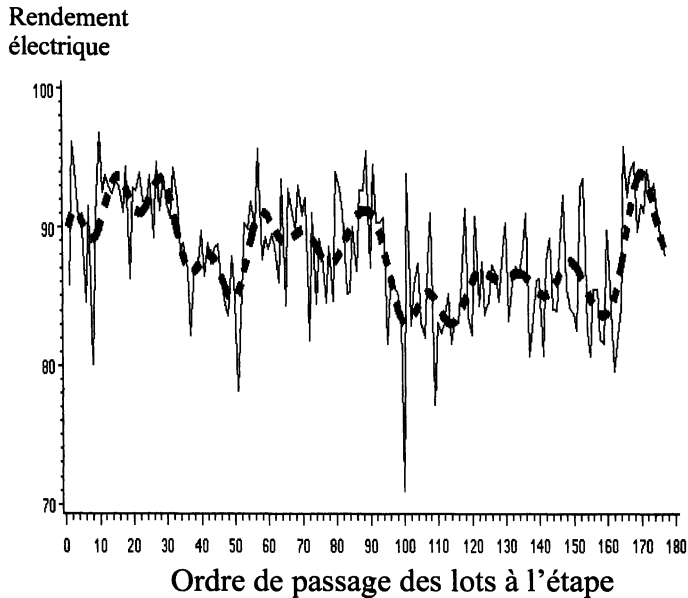


FIG 2. — Graphique de l'étape défaillante identifiée par l'application d'un lissage des données (en continue : données réelles, en pointillé : lissage).

4. Le système Six Sigma : les personnes, les outils, la formation, la structure

Notre expérience personnelle se basant sur le Six Sigma, nous concentrons ce paragraphe sur cette méthodologie. Il est entendu que d'autres méthodes d'amélioration par la statistique existent mais elles ne seront pas détaillées dans cet article.

4.1. Du développement de l'utilisation de la statistique dans l'industrie au « Six Sigma »

L'intérêt pour la statistique dans l'industrie remonte à loin mais il connaît un développement significatif pendant l'effort de guerre aux États-Unis et en Grande Bretagne où il commence à se montrer particulièrement efficace. À la fin de la seconde guerre mondiale, les qualitatifs et statisticiens qui ont participé à cet effort voient l'intérêt pour leurs méthodes baisser considérablement. Une grande partie du monde est en ruine et l'industrie américaine n'a pas de concurrence. Certains vont alors au Japon où tout est à reconstruire et où ils vont rencontrer une grande attention. Le plus influent d'entre eux fut sans doute William Edward Deming qui montra tout l'intérêt de la statistique dans l'industrie entre autres contributions. À la fois physicien et statisticien, Deming a joué un grand rôle dans le développement de la Statistique et nous essayons toujours d'appliquer ses concepts dans notre pratique

quotidienne. Nous ne pouvons pas ici résumer la pensée de Deming tant elle est profonde, on lira avec profit Deming (1994) pour une première approche, puis Deming (1982) qui est l'ouvrage de référence. Nous allons dans cette section tenter de montrer, à l'aide d'exemples rencontrés dans notre usine, en quoi il est important de garder à l'esprit les recommandations de Deming, tout à fait d'actualité alors que nos clients, en particulier automobile, sont de plus en plus exigeants. Les références citées ci-dessous proviennent toutes de « Out of the Crisis » (Deming, 1982).

- L'optimisation du test. Deming insiste (page 28) sur le fait que la qualité n'est jamais garantie par l'inspection. Notre expérience nous a montré qu'il est possible de réduire le coût du test sans effet sur la qualité, et surtout, il ne faut pas perdre de vue que c'est en améliorant les procédés et les rendements que la qualité chez le client sera meilleure. Même s'il est vrai que, dans l'industrie du semiconducteur, il y aura toujours un test exhaustif (en ce sens que l'on teste toutes les pièces avant livraison au client) à cause de la complexité et de l'évolution rapide des technologies, il ne faut jamais perdre de vue ces deux objectifs complémentaires : améliorer les procédés et réduire les tests.
- L'analyse statistique des données existantes. Deming (page 391) insiste sur le fait qu'il n'est pas toujours nécessaire de commencer par l'expérimentation pour résoudre un problème. En effet, les procédés varient de façon naturelle et, si l'on conserve cette information, il est souvent possible d'en tirer de précieuses informations. Dans notre industrie où il y a des centaines d'étapes de procédé et des milliers de paramètres, il est préférable de toujours analyser les données existantes qui vont fournir des pistes pour une expérimentation ultérieure. En fait, comme disait Deming, expérimentation et analyse des données existantes sont complémentaires.
- La formation. Deming cite un exemple coûteux de confusion entre limites de contrôle et de spécification (en page 356). Nous avons eu récemment un exemple similaire pour lequel cette confusion a eu un impact sur l'activité. Nous pensons que la formation à Motorola/Freescale est un point fort de l'approche qualité. Même si le rôle des statisticiens est important, la formation permet à chaque employé de comprendre les concepts statistiques et donc d'améliorer la qualité dans son travail quotidien.

Deming eut des disciples, dont Genichi Taguchi, le premier à insister sur l'importance des plans d'expérience d'écrantage. Le Japon connut la croissance que l'on sait à cette époque et la statistique revint au début des années 80 dans le monde occidental. En 1984, quatre fabricants japonais prirent les quatre premières places dans la fabrication des semiconducteurs et l'intérêt pour la statistique dans notre industrie vint par l'introduction des plans d'expérience factoriels puis de type Taguchi. Les constructeurs automobiles américains, dont Ford, poussèrent tous leurs fournisseurs à l'utilisation des cartes de contrôle. En 1986, Bob Galvin, alors PDG du groupe Motorola, adopta l'idée de Bill Smith, ingénieur du groupe. Jusqu'à lors, il était assez communément admis d'avoir des limites de spécification à 3 sigma de la moyenne. En pratique,

on a quand même 0.27 % de points hors spécification. De plus, quand la moyenne varie, ce chiffre se dégrade rapidement. Enfin, quand on a 200 étapes élémentaires qui interagissent entre elles, viser le 3 sigma pour chacune conduit à un risque global non négligeable. L'idée du Six Sigma consiste à garder les distributions toujours très à l'intérieur des spécifications. Cela permet de garder un taux de rejet à 3.4 pièces par million même dans le cas où la moyenne dériverait de 1.5 sigma. La qualité de Motorola évolua rapidement de la médiocrité vers un très bon niveau et Motorola reçut en 1987 le Malcolm Baldrige Award qui est le prix qualité le plus prestigieux aux États-Unis. En 1992 Motorola fonda avec quelques autres entreprises le Six Sigma Institute. On introduisit la notion de Six Sigma Black Belt, qui est attribuée à des ingénieurs spécialisés ayant reçu des formations statistiques leur permettant de former eux-même leurs collègues et de les aider dans leurs analyses. En 1998, d'autres compagnies adoptèrent cette approche. General Electric est la plus connue d'entre elle. Le rôle de Black Belt évolua pour devenir un expert à plein temps en résolution de problèmes et amélioration de processus. Le Six Sigma allie maintenant l'utilisation des techniques statistiques avec la gestion de projet et d'autres techniques d'amélioration (analyse des modes de défaillance et de leur criticité, *brainstorming*, techniques de créativité) le tout articulé dans une méthode séquentielle d'amélioration appelée DMAIC (Define, Measure, Analyse, Improve, Control). Motorola, qui a été à l'origine du Six Sigma, a à son tour adopté cette évolution du Six Sigma. Aujourd'hui Freescale, forte de son expérience Motorola, s'inscrit pleinement dans cette démarche qualité. Nous allons maintenant voir comment cela fonctionne en pratique.

4.2. Les outils

Les techniques statistiques sont maintenant très souvent disponibles sous forme de logiciel. Pour nos applications principales, maîtrise statistique des procédés, analyse de donnée et expérimentation, on a des besoins différents :

- Maîtrise statistique des procédés. Cet outil est utilisé par les opérateurs, on lui demandera d'être d'utilisation conviviale, interconnecté avec les équipements de mesure pour que les mesures lui soient directement transférées ce qui évite les erreurs de saisie, connecté avec le système de gestion de production et de maintenance afin de pouvoir programmer les contrôles. Il doit aussi être interconnecté avec la base de données technique. Les opérateurs doivent avoir accès aussi directement que possible aux instructions et actions à réaliser en cas de hors contrôle.
- Analyse de données pour l'amélioration de rendement. Dans tout cet article, le terme analyse de données veut dire analyse statistique sur des données non collectées par expérimentation, mais disponibles dans une base de données. La base de données doit rassembler l'ensemble des données décrites plus haut, car il faut pouvoir corrélérer les données de rendement avec les données de paramètres technologiques, les données de sortie et d'entrée de procédé et de fabrication (dates de passage et équipements employés). Il faut un système d'extraction de ces données. Il est essentiel

que les clients utilisateurs, en particulier les statisticiens, définissent le contenu de la base de données. On voit croître parallèlement les possibilités de stockage et les données candidates. C'est aux utilisateurs de choisir les priorités des données à stocker en fonction des analyses. On peut voir des bases de données où certains types de données sont sur-représentés et d'autres faiblement parce que la communication entre informaticiens et utilisateurs ne s'est pas faite correctement. Un logiciel statistique complet et programmable, avec possibilité d'automatisation, doit être disponible pour faire ces analyses. Les ingénieurs « d'amélioration de rendement » sont les utilisateurs principaux de ces outils, mais les ingénieurs de procédé peuvent aussi les utiliser. Il doit y avoir une personne responsable de ces outils, dans le but de faire les analyses les plus sophistiquées, de former les ingénieurs utilisateurs, d'assurer l'interface avec les informaticiens responsables de la base donnée, et enfin de développer les outils d'analyses comme les tableaux de bord ou SODA. Ce poste est parfois appelé EDS pour Engineering Data Specialist. Un ingénieur formé et spécialisé en statistique peut remplir ce poste. Dans notre usine, il a été choisi de le faire remplir par un statisticien notamment pour renforcer la partie développement de méthodes. Il est absolument nécessaire qu'un statisticien à ce poste soit très à l'écoute de ses clients, les ingénieurs, si on veut que ceux-ci utilisent les méthodes et acceptent les rigueurs de l'outil statistique.

- Les outils informatiques. Il est essentiel d'avoir un outil qui possède un grand nombre de techniques statistiques de base et qui soit convivial d'utilisation. Un point très important est que cet outil dispose de bons graphes, car l'association d'un bon graphe à l'œil augmente nettement les facultés d'analyse de l'homme. On lira avec profit à ce sujet Valois (2000). Un autre point important est que ce logiciel doit être interactif pour sélectionner rapidement des données d'intérêt, refaire et comparer des analyses. Ce logiciel doit disposer d'un langage de programmation statistique pour automatiser certaines analyses, notamment dans le cas où il y a beaucoup de variables. On pourrait rêver d'un logiciel qui posséderait toutes les qualités requises pour les trois applications. Il n'est pas impossible de le faire avec SAS[®] ou Statistica[®]. Mais ce qui peut être très bon pour l'analyse des données peut être difficilement utilisable par les ingénieurs faisant des cartes de contrôles et des plans d'expérience. En pratique on utilise des logiciels différents, SAS[®], pour l'analyse de données, un logiciel maison pour les cartes de contrôle, et JMP[®] ou Minitab[®] pour les analyses statistiques ponctuelles. Il est alors nécessaire que les données puissent circuler facilement entre les trois applications.

4.3. Les moyens humains

Comme on a pu le voir la statistique est utilisée par différents personnels :

- Les opérateurs suivent des règles de Shewhart pour la Maîtrise Statistique des Procédés sur la ligne de production.

- Les ingénieurs et techniciens sont amenés de façon soit ponctuelle, soit régulière à utiliser certaines techniques : exploration, comparaison d'échantillons, analyse de variabilité des procédés ou des systèmes de mesure, plans d'expérience.
- Les Black Belts et Green Belts sont des ingénieurs avec des degrés différents de spécialisation dans les méthodes d'amélioration continue, y compris statistiques.
- Les statisticiens sont présents pour utiliser des méthodes un peu complexes ou en développer de nouvelles.
- Des personnels à compétence extérieure auxquels on a recours en cas de besoin : stagiaires souvent utilisés pour développer des macros d'analyse systématique. Quand il y a des méthodes très sophistiquées à développer, on travaille contractuellement avec une université dans le cadre d'une thèse avec convention CIFRE par exemple.

4.4. La formation

C'est un aspect indispensable car il faut donner les moyens aux personnes de se servir des outils pour les différents niveaux d'utilisation cités. La formation est donnée soit par des sociétés extérieures pour lesquelles les formateurs ont été certifiés par Motorola/Freescale, soit en interne par des ingénieurs, des statisticiens ou des Black Belts.

Pour les opérateurs, on dispense une formation de 4 heures dans laquelle on présente les bases élémentaires de la statistique, indicateurs de tendance centrale, de dispersion, capabilité et logiciels de MSP. Pour les ingénieurs il existe un cursus progressif pour les former aux techniques de base. Ce cours est fait pour moitié de théorie et pour moitié d'exercices pratiques et d'application, ces deux parties étant imbriquées. Il comprend six cours que les ingénieurs ne suivent pas forcément en totalité :

- deux jours d'introduction aux distributions, capabilité, statistique univariée et maniement de logiciel,
- trois jours sur les méthodes statistiques de comparaisons d'échantillons et sur les tests d'hypothèse,
- deux jours sur les sources de variabilité et les capabilité d'instrument de mesure,
- deux jours sur les cartes de contrôle,
- trois jours sur les plans d'expérience factoriels, fractionnels factoriels et d'écrantage,
- deux jours sur les plans de réponse de surface et d'optimisation.

Les ingénieurs de type Green Belt ou Black Belt sont spécialisés dans les méthodes d'amélioration continue. Ils reçoivent une formation de base sur la méthodologie systématique d'amélioration DMAIC pour Définir, Mesurer, Analyser, Améliorer, Contrôler. Cette formation de base allie des notions d'économie, de méthodes qualité comme l'analyse des modes de défaillance

et de leur criticité, de travail de groupe, avec les notions statistiques, qui sont complétées selon les besoins par des formations spécifiques (données catégorielles, régression multiple, dépouillement d'enquête, ...). Les méthodes statistiques sont ici intégrées aux autres et agissent en complémentarité.

Des formations sont également mises en place pour les personnels dirigeants. Sans entrer dans les détails techniques, elles donnent les éléments de base afin d'aider ceux-ci dans la sélection des projets selon leur impact et l'appui des groupes de travail en mettant à disposition les moyens nécessaires.

4.5. La structure

Les formations et le personnel compétent ne sont pas suffisants en soi. Il est nécessaire d'avoir une structure qui intègre l'utilisation de la statistique dans le travail de groupe. En effet, l'utilisation de ces outils reste un changement radical dans la façon habituelle de travailler. La façon habituelle de travailler des ingénieurs en groupe où individuellement consiste le plus souvent à procéder analytiquement par raisonnement à partir de leurs études et de leur expérience. C'est un peu le déterminisme, on parle en anglais de « best educated guess » qu'on peut traduire « meilleure devinette ». L'observation et l'expérimentation sont souvent conçues comme des moyens de confirmer les déductions en faisant varier une variable à la fois. L'utilisation de la statistique et du Six Sigma modifie quelque peu ces habitudes. Les cartes de contrôle demandent de recentrer une valeur parce qu'on est statistiquement différent de la moyenne alors que, du point de vue de l'utilisateur, on a des valeurs correctes puisqu'en spécification et même parfois à l'intérieur des limites de contrôle. Les analyses de données apportent des quantités de renseignements, par exemple à une étape l'identification d'un équipement sur lequel une majorité de lots avec des rendements faibles sont passés.

Nous avons vécu de nombreux exemples où l'analyse de données oriente la résolution de problèmes vers des domaines ignorés des experts, ce qui peut les faire réagir et parfois même vivement. Mais c'est justement là toute la puissance de la statistique dans ce domaine que de faire surgir des espaces d'investigation nouveaux ce qui peut être particulièrement utile dans des cas où le travail d'amélioration piétine plus ou moins. Les plans d'expériences vont être décidés pour confirmer les résultats d'une analyse de données. La simultanéité du changement des niveaux des facteurs dans un plan d'expérience n'est pas non plus toujours acceptée facilement. L'idée que l'on voit mieux les évolutions en faisant évoluer un seul facteur à la fois recule mais garde un certain ancrage. La formation, l'expérience et la gestion de projets permettent d'aider à l'évolution des esprits sur le sujet. La structure doit donc permettre de guider les groupes, mais aussi d'être à l'écoute des ingénieurs clients. Une relation de confiance est indispensable à l'adoption des méthodes par les ingénieurs. La structure comprend donc des groupes de travail réguliers avec des réunions hebdomadaires ou bimensuelles comprenant des exposés de résultats, des revues des plans d'actions et faisant l'objet de comptes rendus. Dans les réunions consacrées au contrôle de procédé, on voit l'évolution des mesures de capacité, on regarde les cartes de contrôles des équipements les

moins performants ainsi que les meilleurs. On vérifie comment sont redémarrés les équipements après une situation de hors contrôle. On remet en question le contrôle, la fréquence, l'échantillonnage, on recalcule les limites. D'autres réunions sont consacrées aux résultats des analyses de données. On regarde les résultats, on vérifie comment les lots passés à l'étape se situent dans le temps, on vérifie les cartes de contrôle des équipements en question à la période et si rien n'a été vu on peut remettre en question ce contrôle de procédé. On recherche la confirmation et l'explication des différences par des plans d'expériences. Dans ces réunions est présent l'ensemble des personnes concernées y compris des membres de l'encadrement de haut niveau. Le but est en effet de prendre des décisions. L'avantage des méthodes statistiques est qu'elles apportent le maximum d'éléments pour prendre ces décisions. Il peut s'agir de choisir le meilleur compromis. On peut décider d'un investissement dans un équipement si on montre qu'il y a un gain significatif en rendement ou au contraire décider de supprimer un équipement ou utiliser un équipement moins coûteux si on a montré qu'il n'y pas d'impact négatif. L'implication de cet encadrement de haut niveau est essentiel car c'est lui qui maîtrise tout le système et lui donne les moyens de fonctionner en terme de financement des logiciels et bases de données, formations, embauches, affectation de ressources. L'enjeu est très important, car de la façon dont sont utilisées ces techniques dépendent en partie les résultats, donc la compétitivité de l'entreprise et de ses unités de production, avec des conséquences sur l'emploi dans un environnement économique de plus en plus dur.

On achèvera cette partie en faisant remarquer qu'on s'est cantonné ici à décrire l'utilisation de la statistique dans les métiers techniques parce que c'est dans cette branche que les avancées ont été les plus nombreuses et les plus exemplaires. Les expériences récentes dans le domaine du Six Sigma montrent qu'il y a des possibilités insoupçonnées dans les autres domaines de l'entreprise qu'on appelle parfois « transactionnels » : les services, le marketing, les ventes, la finance, les ressources humaines, la logistique, l'informatique, etc. Certaines références (www.isixsigma.com) indiquent que les applications les plus rentables pourraient se trouver là.

On peut résumer ainsi : la statistique est pour un praticien industriel la façon rationnelle de regarder des données. L'utiliser correctement et au bon niveau est la garantie d'un avantage compétitif sur les résultats.

5. Les rapports entre la pratique et la recherche en statistique

On pourra lire avec intérêt Tassi et Dreesbeke (1997) sur l'histoire de la statistique. Aussi Roberts (1990), Banks (1993) et Woodall (2000) amènent une contribution très intéressante à ce débat.

5.1. L'apport de Box : la statistique, catalyseur de l'apprentissage par la méthode scientifique

On peut rattacher cette expérience à un cadre plus général des relations entre la statistique et la pratique. On se basera sur le numéro du « Journal of Quality Technology » consacré à ce sujet à partir d'une remarquable contribution de Box et Liu (1999). Cette réflexion ouvre des champs de progression à la fois pour les praticiens et pour les chercheurs en Statistique. Il est d'abord indispensable de se livrer au difficile exercice de résumer au moins les idées principales de la contribution de Box. Le premier article de Box et Liu, « Statistics as a catalyst to learning by scientific method – an example », décrit les innovations industrielles importantes comme le résultat d'une investigation, un processus d'apprentissage dynamique et non stationnaire, adaptatif et pas « d'un bloc ». Dans ce cadre, le chemin emprunté pour atteindre l'objectif se dévoile au fur et à mesure de la progression et chaque résultat sert à déterminer la phase d'expérimentation suivante. Box préconise alors pour catalyser l'innovation que le statisticien fasse partie de l'équipe et ne soit pas seulement celui qui analyse les données des autres. Il est remarquable de noter que différents chemins peuvent être empruntés pour arriver au même point car l'itération scientifique tend à être auto-correctrice, tout comme l'itération mathématique. Certains chemins sont cependant plus longs que d'autres et cela peut faire une différence importante. Un exemple d'optimisation de la conception et de la fabrication d'un hélicoptère en papier avec pour objectif la maximisation du temps de vol est ensuite présenté en utilisant des plans d'écrantage, factoriels et de surface de réponse, chemin de plus grande pente, ridge regression.

Dans la seconde partie de l'article, Box décrit son trajet : il vint à apprendre la statistique en participant à la recherche sur la protection contre les armes chimiques dans l'armée britannique pendant la seconde guerre mondiale. Par la suite, lors de son premier emploi dans la société chimique ICI, il vécut l'expérience de proposer à ses collègues l'utilisation des plans d'expérience et de la voir rejeter au premier abord avec force moqueries. C'est en travaillant dans les équipes pour aider au montage des expériences que Box développa la méthodologie des surfaces de réponse.

Le développement d'un nouveau produit ou procédé comprend une phase souvent longue de mise au point suivie d'une phase de test. On peut comparer ce processus à la résolution d'une affaire criminelle comprenant l'enquête et finissant par un procès. L'enquête policière et judiciaire est une procédure séquentielle où l'on va s'efforcer de générer des hypothèses à partir d'observations, tout comme dans une analyse de données. À ce stade, on commence par interroger des témoins, observer des indices et trouver des suspects. Mais suspecter n'est pas condamner et fixer un niveau de test d'hypothèse à 0.05 peut conduire à éliminer des pistes capitales. Ainsi, une analyse de variance sur des données non structurées (issues d'une base de données) peut conduire à ne pas rejeter l'hypothèse nulle de non signification d'un paramètre avec une probabilité de 0.15, pourtant l'expérience des ingénieurs va les conduire à poursuivre les investigations en utilisant, par exemple, des plans

d'expériences. En effet, il est fréquent qu'après un plan d'expérience, avec un test plus puissant, une taille d'échantillon plus importante ou des mesures moins « bruitées », on arrive à confirmer l'effet pointé par l'analyse de variance avec un très bon niveau de confiance. En poussant le raisonnement, Box questionne : « si on sait qu'il y a une chance sur deux de trouver une pépite d'or derrière un arbre, ne doit-on pas quand même aller voir ? ». Quant au procès, c'est une procédure « d'un bloc », l'hypothèse nulle de l'innocence doit être rejetée avec un bon niveau de confiance pour prononcer une condamnation. C'est équivalent aux conclusions d'un plan d'expérience confirmées par des essais en situation. En pratique, ce qu'il faut prendre en compte c'est la notion de compromis entre les fausses alarmes qui peuvent être générées et les pistes capitales qui ne peuvent être négligées.

Le paradigme de l'apprentissage scientifique compare une déduction d'un modèle avec des données observées. De l'inadéquation entre les deux on induit une hypothèse et bâtit un nouveau modèle. L'expertise technique sur le sujet est bien sûr capitale et nourrit la façon de collecter les données ou de monter une expérience en accélérant l'efficacité. Le paradigme mathématique est lui « d'un bloc » : « À partir d'un certain nombre d'hypothèses, une proposition suit logiquement, j'en fournis alors la preuve ou la démonstration ». Traduit en statistique cela donne : « À partir d'une hypothèse à tester, de son hypothèse alternative, et de tout ce que l'on suppose d'autre à propos du modèle, je fournis une procédure de décision optimale » ou pour les plans d'expériences : « À partir des variables importantes, je détermine la région d'intérêt dans l'espace des facteurs ». L'amélioration continue est un principe central dans la qualité moderne. Cette notion rejoint l'optimisation mathématique dans son objectif mais en diffère en ce que cette dernière se réalise à modèle fixe tandis que, dans l'amélioration continue, ni la forme du modèle, ni l'identité des facteurs, ni même celle des réponses ne sont complètement fixées, grâce à quoi l'optimum lui-même a la possibilité d'évoluer.

Il n'y a pas de théorème mathématique sur l'investigation scientifique, et on ne peut prédire les différents chemins que prendront des investigateurs différents. On peut alors comprendre que des statisticiens peuvent se tenir à l'écart de ces problématiques pour se concentrer sur le développement de procédures mathématiques respectables « d'un bloc » pour lesquelles le travail de groupe n'est pas indispensable. Le malaise qui affecte l'application statistique a son origine dans cette confusion entre paradigmes mathématiques et scientifiques. Dans l'investigation scientifique le sujet est l'investigation elle-même et l'objectif, son succès.

Malgré les possibilités de générer des plans optimaux, Box défend l'utilisation des plans factoriels pour leur pouvoir d'induction car ils guident vers la suite de l'investigation. L'examen d'un graphe cube représentant un plan factoriel 2^3 permet de figurer l'évolution le long des arêtes et les interactions d'ordre 2 et 3. Par ailleurs les expérimentateurs hésitent souvent à utiliser des plans d'écrantage, en leur reprochant de ne pas voir les interactions, or les plans factoriels fractionnels et les plans d'écrantage type Plackett-Burmann ont de remarquables propriétés de projectivité. Ainsi, si l'on expérimente 8 facteurs

avec 16 essais, on peut voir les interactions de tout groupe de 3 facteurs. On peut ainsi voir toutes les interactions de 3 facteurs sur 11 dans un plan Plackett-Burman à 12 essais, et toutes les interactions doubles de 5 facteurs sur 11 avec un plan Plackett-Burman à 24 essais. Bien sûr, dans l'analyse de ces plans, il faut rester prudent sur les conclusions et tenir compte des confusions entre les différents effets étudiés.

Les analyses doivent être robustes. La robustesse dépend de l'écart par rapport aux hypothèses et de la sensibilité par rapport à cet écart. Deming a bien décrit comment les résultats d'une expérience sont très rarement mis en œuvre dans des conditions similaires aux circonstances de cette expérience. L'utilisation de la statistique pour concevoir un plan qui fonctionnera correctement dans le monde réel remonte au moins aux travaux de Gosset (dit Student) aux brasseries Guinness pour trouver une variété d'orge telle que les qualités de la bière soient insensibles aux différentes conditions de temps, de sol et de techniques agricoles dans les différentes régions de l'Irlande. Taguchi divisa les objectifs des plans robustes en deux catégories : minimiser la variation dans la performance du système transmis par ses composants, puis minimiser les effets sur les systèmes des variables d'environnement dans l'utilisation.

De par leurs propriétés d'équilibre et d'orthogonalité, les plans factoriels ont des caractéristiques de robustesse supérieures aux plans optimaux divers, ces derniers restant utiles dans certains cas particuliers. On peut tirer parti des puissances de calcul informatique :

- en utilisant les nombreuses possibilités graphiques qui permettent de regarder les résultats d'un point de vue différent,
- pour tester rapidement les différentes possibilités de combinaisons d'interaction dans des plans qui se projettent (exemple : 969 possibilités de combinaisons de 3 facteurs sur 19 d'un Plackett-Burman à 20 essais). Cette possibilité est particulièrement intéressante, la façon de trouver les interactions a été étudiée par Samset (1999). Nous avons pour notre part identifié un bon modèle avec une interaction triple en partant d'un Plackett-Burman à 24 essais avec 10 facteurs au départ, en utilisant les arbres de régression pour trouver les variables actives et en testant les interactions possibles.

Après discussion, Box conclut en citant quelques uns des nombreux exemples où des développements théoriques ont été réalisés à partir de problématiques pratiques quand les techniques ne donnaient pas de résultat satisfaisant. Il oppose cette saine pratique au recyclage à l'infini des mêmes jeux de données qu'il qualifie de « tentative de réanimation de cadavres par des pilleurs de tombes ».

5.2. L'expérience de Motorola/Freescale

Les thèmes développés par Georges Box recourent en grande partie notre expérience sur la pratique de la statistique en entreprise et cela ouvre des perspectives tout à fait intéressantes. En effet, on remarque une similitude frappante entre le paradigme mathématique écrit par Box et un comportement fréquent chez les ingénieurs, plutôt déterministes et orientés sur le développement de mécanismes d'explications où l'expérience a plus un rôle de vérification que d'exploration. Cela n'est pas surprenant dans la mesure où la culture d'origine de ces deux communautés, les ingénieurs et les mathématiciens, est identique. De même, l'idée que le monde industriel ne recèle que des problématiques statistiques relativement simples semble partagée autant dans le milieu de la statistique que dans l'industrie. Ainsi, même dans les milieux industriels où l'on est persuadé que l'utilisation de la statistique alliée aux autres techniques dans le cadre du Six Sigma est fondamentale pour faire progresser les sociétés, l'idée est très répandue que la qualité principale requise est d'avoir des « change-agents », c'est-à-dire des gestionnaires de projets, et que des techniques statistiques de bases suffisent comme si la force de changement était incompatible avec la sophistication statistique. Si cette idée est répandue, c'est qu'elle est basée sur l'expérience et le fait qu'on a vu beaucoup plus d'améliorations reposant sur une bonne dynamique de groupe et des techniques simples. Pourtant notre expérience indique que, si des comparaisons simples d'échantillons et autres techniques élémentaires peuvent apporter beaucoup de renseignements, certains progrès ne peuvent être réalisés qu'avec des techniques beaucoup plus sophistiquées : recuit simulé, MCMC, bagging, etc. Encore notre expérience reste-t-elle très limitée. La modélisation statistique complète d'une unité de fabrication de circuits intégrés reste, par exemple, un défi à réaliser. Mais les statisticiens placés au cœur d'équipes opérationnelles ne sont pas si nombreux que cela et une partie des statisticiens industriels sont employés dans des services support et non intégrés dans des équipes opérationnelles. La perception d'un statisticien en dehors de la réalité des problèmes, qui serait là uniquement comme garant de l'orthodoxie, peut contribuer à alimenter l'idée d'une utilité très limitée de ce rôle. À l'inverse on a constaté que, quand des statisticiens sont bien intégrés dans des équipes d'ingénieurs qui sont eux-mêmes formés, voire autonomes, sur une partie des analyses dans un système encouragé par l'encadrement supérieur, la pratique commune et l'accumulation des expériences permettent de faire évoluer les esprits et les pratiques des uns et des autres dans un sens très positif. Alors cette progression amène à l'utilisation de méthodes sophistiquées. Il faut noter que, même dans ce cas, la communication de ces méthodes à l'encadrement supérieur doit être particulièrement soignée pour être efficace.

6. Conclusion

La recherche et l'industrie font face à des problématiques de plus en plus compliquées, les données sont de plus en plus nombreuses, les possibilités

de calcul et de stockage croissent, les logiciels offrent de plus en plus de possibilités. Il y a un réel besoin d'utilisation et de développement des méthodes statistiques. La confrontation de notre expérience avec le Six Sigma, les idées développées par Deming, puis par Box, nous indiquent d'intéressantes voies pour progresser.

La capacité mathématique reste nécessaire mais la formation statistique devrait inclure l'expérimentation et la pratique. De même, les thèses ne devraient pas être évaluées uniquement en fonction de la quantité de mathématique contenue, qui devrait dépendre des besoins du sujet, mais devraient aussi tenir compte de la capacité de l'étudiant à catalyser l'apprentissage et à développer de nouvelles méthodes dans ce but.

Remerciements

Nous tenons particulièrement à remercier Jean-René Mathieu et Philippe Besse. Sans eux, le fonctionnement de notre usine ne serait peut être pas exactement ce qu'il est aujourd'hui.

Références

- AMBROZIC C. (2003), *Automation for the 300mm fab : A Strategy for Integrating FDC Using Multivariate Analysis*, AEC/APC 15.
- BANKS D. (1993), Is industrial statistics out of control?, *Statistical Science*, 8, 356-409.
- BERGERET F., BESSE P. (1997), Simulated Annealing, Weighted Simulated Annealing and Genetic Algorithm at work, *Computational Statistics*, vol. 12 (4), pp. 447-465.
- BERGERET F., BESSE P. (1999), Recherche d'une étape défaillante dans un procédé industriel : quand le désordre est source d'information, *Revue de Statistique Appliquée*, vol. XLVII (3), pp. 27-37.
- BERGERET F., CHANDON Y. (1995), Recherche d'une séquence de tests optimale en contrôle de fabrication, *Revue de Statistique Appliquée*, XLIII (3), pp. 21-33.
- BERGERET F., LE GALL C. (2003), Yield Improvement using Statistical Analysis of Process Dates, *IEEE transactions on Semiconductor manufacturing*, vol. 16 (3), pp. 535-542.
- BESSE P., LE GALL C. (2004), Application and Reliability of Change-Point Analyses for Detecting a Defective Stage in Integrated Circuit Manufacturing. Prépublication à soumettre.
- BOX G.E.P., LIU P. (1999), Statistics as a Catalyst to Learning by a Scientific Method, *Journal of Quality Technology*, vol. 31 (1), 1-15.
- BREIMAN L. (1996), *Bagging Predictors*, *Machine Learning*, vol. 24 (2), pp. 123-140.
- CHANDON Y., MATHIEU J.R. (1993), *Design of experiment to test sigma reduction in a critical dimension measurement apparatus*, Characterization and Optimization, Motorola University Press.
- DEMING W.E. (1982), *Out of the Crisis*, MIT center for advanced engineering study.

- DEMING W.E. (1994), *The New Economics*, MIT center for advanced educational services.
- DERRINGER D., SUICH R. (1980), Simultaneous Optimization of Several Responses Variables, *Journal of Quality Technology*, vol.12 (4), pp. 214-219.
- DIACONIS P. (1988), *Groups Representation in Probabilities and Statistics*, IMS, Hayward, CA.
- GREEN P.J. (1996), Reversible jump Markov chain computation and bayesian model determination, *Biometrika*, vol. 82, pp. 711-732.
- LAVIELLE M. (2002), On the use of penalized contrasts for solving inverse problems. Application to DDC (Detection of Diverse Changes) problem. Soumis.
- PASQUALINI F., BAUDIER M., LAFAYE de MICHEAUX D., LEMAIRE F. (2002), Using Global Process Control in Semiconductor fabs to achieve APC, *Micromagazine*. <http://www.micromagazine.com/archive/02/01/Pasqualini.html>
- ROBERTS H.V. (1990), Applications in Business and Economic Statistics, *Statistical Science*, 4, 372-402.
- SAMSET O. (1999), *Identifying active factors from non-geometric Plackett-Burman designs and their half-fractions*, technical report 4, department of Mathematical Sciences, The Norwegian University of Science and Technology, Norway.
- TASSI P., DROESBEKE, J.-J. (1997), Histoire de la statistique, Que sais-je?, *Presse Universitaire de France*.
- VALOIS J.P. (2000), Sur l'approche graphique en analyse des données (avec discussion), *Journal de la Société Statistique de France*, 141, 4.
- WOODALL W.H. (2000), Controversies and Contradictions in Statistical Process Control, *Journal of Quality Technology*, vol. 32 (4), pp. 341-350.
- YE K.Q. (1998), Orthogonal Column Latin Hypercube and their application in computer experiments, *Journal of the American Statistical Association*, vol. 93, No 444, pp. 1430-1439.