

REVUE DE STATISTIQUE APPLIQUÉE

CAROLINE LE GALL

Comparaison de différentes analyses statistiques spatiales pour détecter une défaillance dans la fabrication de circuits intégrés

Revue de statistique appliquée, tome 52, n° 1 (2004), p. 5-37

http://www.numdam.org/item?id=RSA_2004__52_1_5_0

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

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

NUMDAM

*Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques*

<http://www.numdam.org/>

COMPARAISON DE DIFFÉRENTES ANALYSES STATISTIQUES SPATIALES POUR DÉTECTER UNE DÉFAILLANCE DANS LA FABRICATION DE CIRCUITS INTÉGRÉS

Caroline LE GALL⁽¹⁾

*Motorola Semiconducteurs S.A.S.
Avenue Eisenhower, 31023 Toulouse Cedex
Laboratoire de Statistique et Probabilités, UMR CNRS C5583
Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex*

⁽¹⁾ *Ce travail a été réalisé dans le cadre d'une convention CIFRE entre le Laboratoire de Statistique et Probabilités de l'Université Paul Sabatier et Motorola Semiconducteurs au Centre Électronique de Toulouse*

RÉSUMÉ

Cet article fait la synthèse de quatre techniques statistiques visant à détecter des sources de défaillance dans la fabrication de circuits intégrés. Leur objectif est de caractériser plus finement le rendement selon la répartition spatiale des puces défailtantes sur les plaquettes de silicium. Pour certaines approches, cela se traduit par une modélisation du rendement selon différentes composantes spatiales. Pour d'autres, cela consiste à décomposer les plaquettes et les rendements en deux parties : une partie représentant la dispersion uniforme des circuits défectueux sur la plaquette et une partie caractérisant plutôt leur proximité spatiale sur la plaquette. La comparaison des différentes approches, dans plusieurs configurations de défaillance, permet de dégager une combinaison de méthodes, stratégiquement optimale, pour répondre à la problématique posée.

Mots-clés : *Statistique spatiale, Champ markovien, Modèle de Poisson, Test d'indépendance spatiale, Arbre de classification, Contrôle qualité.*

ABSTRACT

This paper makes the synthesis of four statistical techniques aiming at detecting sources of failure in integrated circuit manufacturing. Their objective is to more finely characterize the yield, according to spatial parameters of the silicon wafers. For some approaches, it consists in modeling the yield according to various spatial components. For other approaches, it consists in dividing the wafers and the yields into two parts : a part for the defective integrated circuit dispersion on the wafer and a part rather for their spatial proximity. The comparison of the various approaches leads to an efficient combination of methods that makes it possible to identify the spatial signature of various wafers in several cases of failure.

Keywords : *Spatial statistics, Markov field, Poisson model, Test of spatial randomness, Classification tree, Quality control.*

1. Introduction

L'objet de cet article est de comparer l'efficacité et la pertinence de différentes techniques statistiques visant à détecter des sources de défaillance dans la fabrication de circuits intégrés de différents constructeurs (Baron *et al.* (2001) pour Texas Instruments Inc., Stapper *et al.* (1983) pour IBM Inc., Hansen *et al.* (1997) pour Lucent Inc. et Duvivier (1999) pour ST Microelectronics Inc.) Les circuits intégrés, ou puces électroniques, sont obtenus à partir d'un procédé de fabrication complexe nécessitant plusieurs centaines d'étapes physico-chimiques appliquées à des plaquettes de silicium. Selon la taille des puces, les plaquettes comprennent plusieurs centaines voire plusieurs milliers de puces (Figure 1). À l'issue du procédé de fabrication, les fonctionnalités de chaque puce sont testées et les puces défaillantes sont identifiées et rejetées (les puces noires sur la Figure 1). Le rendement d'une plaquette est le rapport du nombre de puces ayant passé tous les tests avec succès sur le nombre total de puces. L'objectif est évidemment d'optimiser ce rendement et donc de déceler le plus rapidement possible dans le procédé de fabrication, les causes originelles d'un faible rendement. Certaines de ces causes se caractérisent par la structure spatiale des défauts qu'elles engendrent. Par exemple, un mauvais nettoyage des plaquettes après certaines étapes peut générer des tâches de puces défaillantes à divers endroits de la plaquette. Un mauvais étalement de la résine aux étapes de photographie peut générer des puces défaillantes en bord de plaquette. Dans les deux cas, le point commun est qu'il existe une dépendance spatiale entre les puces défaillantes. On parle aussi d'un mode de défaillance systématique. Le dépôt de particules sur les plaquettes constitue une autre source de défaillance. Ce dépôt s'effectue de manière aléatoire sur la plaquette, il ne génère pas nécessairement des regroupements de puces défaillantes. Par opposition aux deux exemples précédents, on parle d'indépendance spatiale entre les puces défaillantes et d'une mode de défaillance aléatoire.

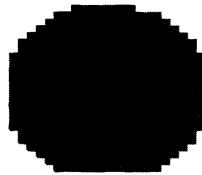


FIGURE 1
Cartographie des puces défaillantes

Les cartographies associées à ces trois exemples sont très différentes, pourtant leur rendement peuvent être très proches. En conséquence, la signature spatiale des puces défaillantes ne s'identifie pas à la valeur du rendement de la plaquette. En outre, une même plaquette peut être touchée simultanément par les deux modes de défaillance. Selon la contribution de chaque mode de défaillance dans les pertes totales, les actions correctives à mettre en place sont différentes. Si le mode de défaillance majoritaire est de type systématique, il témoigne d'un dysfonctionnement du procédé de fabrication ; si le mode de défaillance majoritaire est de type aléatoire, il témoigne d'une concentration en particules trop importante au sein de l'outil de production. La différenciation des deux modes de défaillance est donc une précieuse

source d'information pour accélérer le processus d'amélioration des rendements. Par ailleurs, lorsque le mode de défaillance majoritaire est de type systématique, il est souvent associé à une signature spatiale particulière. La caractérisation de cette signature permet aux ingénieurs de cibler plus rapidement les étapes du procédé qui peuvent en être à l'origine.

Quatre approches statistiques différentes proposées par quatre constructeurs de circuits intégrés essaient de répondre à cette problématique. La première approche (Baron *et al.*, 2001) basée sur une modélisation markovienne cherche à caractériser la signature spatiale à l'aide d'un ensemble de 10 paramètres. La deuxième approche (Stapper *et al.*, 1983) est basée sur une modélisation du rendement selon un modèle de Poisson. Elle conduit à la décomposition du rendement total en deux rendements : systématique et aléatoire. La troisième approche (Hansen *et al.*, 1997) basée sur l'analyse des corrélations spatiales vise à décomposer la plaquette en deux sous-plaquettes : une plaquette systématique et une plaquette aléatoire. La décomposition du rendement en parties systématique et aléatoire s'en déduit. Enfin, la quatrième approche (Duvivier, 1999) propose de caractériser la signature spatiale d'une plaquette à l'aide d'un arbre binaire de type CART (Breiman *et al.*, 1984).

La description de ces méthodes est l'objet de la section 2. Elles sont appliquées, section 3, sur un échantillon de plaquettes représentant différents modes de défaillance. La comparaison des résultats de chaque méthode permet d'identifier la méthode la plus efficace, voire de déterminer la combinaison de méthodes qui est stratégiquement la plus optimale pour répondre avec succès à la problématique posée. Les conclusions de cette comparaison sont l'objet de la discussion de la section 4.

2. Approches méthodologiques

Les différentes approches se basent toutes sur la modélisation du rendement. La première approche procède selon une modélisation markovienne. L'hypothèse est que l'état d'une puce (fonctionnelle ou défaillante) peut dépendre de l'état d'une puce voisine, mais elle ne dépend pas de l'état des autres puces de la plaquette. Cette approche cherche à identifier les caractéristiques de la signature spatiale selon un ensemble de paramètres. Ces paramètres peuvent servir ensuite de variables explicatives pour la classification des plaquettes selon leur cartographie. La deuxième approche est une approche plus classique de la décomposition du rendement en rendements aléatoire et systématique. Elle est basée sur la modélisation du nombre de défauts par puce selon un modèle de Poisson. La troisième approche conduit à la décomposition des plaquettes selon la Figure 2. Il en découle un calcul des rendements aléatoires et systématiques. Cette approche est basée sur l'analyse des corrélations spatiales et sur un test statistique dont l'indépendance spatiale des puces défaillantes est l'hypothèse nulle et la présence d'une dépendance spatiale est l'hypothèse alternative. La quatrième approche est aussi basée sur la décomposition du rendement en rendements aléatoire et systématique. Néanmoins, son principal attrait est la définition d'un arbre de classification dont l'objectif est de décrire la structure des dépendances spatiales.

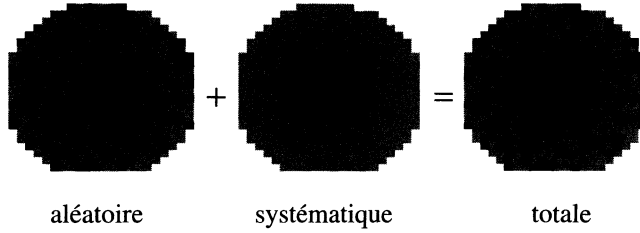


FIGURE 2

Décomposition de la plaquette en parties aléatoire et systématique

2.1. Modélisation markovienne

Un modèle probabiliste est défini pour décrire la distribution jointe des puces d'une plaquette. Il est construit sur la base d'une dépendance spatiale des puces selon un modèle markovien : l'état d'une puce (fonctionnelle ou défective) peut être influencé par ses voisins directs mais il est indépendant de l'état des puces qui sont en dehors de son voisinage. Selon Besag (1974) et Cressie (1991), toute mesure de probabilité dont les distributions conditionnelles définissent une structure de voisinage est un champ markovien aléatoire. Longtin *et al.* (1996) sont les premiers à avoir construit une distribution jointe des puces d'une plaquette selon un modèle markovien. Ils ont choisi une structure de voisinage définie par les puces à gauche, à droite, au-dessus et au-dessous de la puce courante. Dans ce cas, seuls les effets verticaux et horizontaux sont donc pris en compte. Baron *et al.* (2001) suggèrent d'améliorer ces résultats en tenant compte des interactions diagonales.

Les puces d'une plaquette peuvent être identifiées à partir d'un repère cartésien. Le couple (i, j) identifie une puce. L'état des puces constitue une variable aléatoire de Bernoulli X_{ij} qui prend la valeur 1 si la puce est défective et 0 sinon. L'ensemble \mathcal{V}_{ij} définit le voisinage associé à la puce (i, j) . L'hypothèse de champ markovien se traduit par :

$$P(X_{ij} = 1 | X_{kl}, (k, l) \neq (i, j)) = P(X_{ij} = 1 | X_{kl}, (k, l) \in \mathcal{V}_{ij}).$$

Le voisinage préconisé par Baron *et al.* (2001) est défini par :

$$\mathcal{V}_{ij} = \{(k, l) \neq (i, j), |k - i| \leq 1 \text{ et } |l - j| \leq 1\}.$$

Soit $s = (i, j)$, une puce. Son voisinage est décrit par les 8 puces qui l'entourent (Figure 3), avec $t(s) = (i, j + 1)$, $b(s) = (i, j - 1)$, ...

lt(s)	t(s)	rt(s)
l(s)	s	r(s)
lb(s)	b(s)	rb(s)

FIGURE 3
Voisinage de la puce s

Un ensemble de puces dans lequel toutes les puces sont adjacentes définit une clique. Par exemple, $(r(s), s, t(s))$ est une clique contrairement à l'ensemble $(r(s), s, l(s))$ dans lequel $r(s)$ et $l(s)$ ne sont pas adjacentes. À partir de la structure de voisinage décrite Figure 3, 10 types de cliques peuvent être définis (Figure 4) :

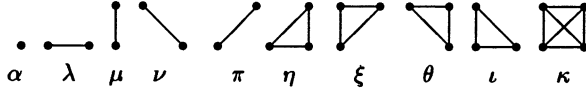


FIGURE 4
Cliques associées à la structure de voisinage choisie

À ces cliques sont associés 10 paramètres, identifiés par les mêmes symboles. Ces paramètres définissent les 10 paramètres canoniques de la distribution jointe. Pour chaque puce s et un sous-voisinage \mathcal{V}'_s de \mathcal{V}_s , des logits conditionnels sont définis :

$$\alpha_s(\mathcal{V}'_s) = \text{logit}(P(X_s = 1 | X_t = 1 \text{ pour } t \in \mathcal{V}'_s \text{ et } X_t = 0 \text{ pour } t \in \mathcal{V}_s \setminus \mathcal{V}'_s).$$

La plaquette est supposée stochastiquement homogène : $\alpha_s(\mathcal{V}'_s)$ dépend de la forme de \mathcal{V}'_s mais ne dépend pas de la localisation de la puce s . Sans cette hypothèse, le nombre de paramètres à estimer excéderait le nombre de puces, ce qui les rendrait inestimables. Sous la condition de positivité qui admet qu'une plaquette puisse avoir 100 % de puces fonctionnelles, *i.e.*

$$P(0) = P(X_s = 0, \forall s) > 0,$$

Baron *et al.* (2001) obtiennent les résultats suivants.

(a)

$$\begin{aligned}
\alpha &= \alpha(\emptyset) \\
\lambda &= \alpha(r) - \alpha = \alpha(l) - \alpha \\
\mu &= \alpha(t) - \alpha = \alpha(b) - \alpha \\
\nu &= \alpha(rb) - \alpha = \alpha(lt) - \alpha \\
\pi &= \alpha(rt) - \alpha = \alpha(lb) - \alpha \\
\eta &= \alpha(l, t) - \lambda - \mu - \alpha = \alpha(r, rt) - \lambda - \pi - \alpha = \alpha(b, lb) - \mu - \pi - \alpha \\
\xi &= \alpha(r, b) - \lambda - \mu - \alpha = \alpha(l, lb) - \lambda - \pi - \alpha = \alpha(t, rt) - \mu - \pi - \alpha \\
\theta &= \alpha(l, b) - \lambda - \mu - \alpha = \alpha(r, rb) - \lambda - \nu - \alpha = \alpha(t, lt) - \mu - \nu - \alpha \\
\iota &= \alpha(r, t) - \lambda - \mu - \alpha = \alpha(l, lt) - \lambda - \nu - \alpha = \alpha(b, rb) - \mu - \nu - \alpha \\
\kappa &= \alpha(r, t, rt) - \eta - \xi - \iota - \lambda - \mu - \pi - \alpha \\
&= \alpha(l, t, lt) - \eta - \theta - \iota - \lambda - \mu - \nu - \alpha \\
&= \alpha(r, b, rb) - \theta - \xi - \iota - \lambda - \mu - \nu - \alpha \\
&= \alpha(l, b, lb) - \eta - \xi - \theta - \lambda - \mu - \pi - \alpha
\end{aligned}$$

Le paramètre α correspond au logit de la probabilité qu'une puce soit défective alors que ses voisines sont opérationnelles. Le paramètre λ correspond au logit additionnel issu de la défaillance d'une puce placée à gauche ou à droite de la puce courante... Le paramètre κ correspond au logit additionnel issu de la défaillance simultanée de 3 puces voisines appartenant à la clique κ . En appliquant la terminologie du modèle linéaire généralisé, les paramètres λ , μ , ν et π correspondent aux effets principaux, η , ξ , θ et ι correspondent aux interactions doubles et κ correspond à l'interaction triple. Le choix du voisinage contraint à négliger les interactions d'ordre supérieur. À ces 10 paramètres sont associées 10 statistiques.

$$\begin{aligned}
N_\alpha &= \sum X_{ij} \\
N_\lambda &= \sum X_{ij} X_{i+1,j} \\
N_\mu &= \sum X_{ij} X_{i,j+1} \\
N_\nu &= \sum X_{ij} X_{i+1,j-1} \\
N_\pi &= \sum X_{ij} X_{i+1,j+1} \\
N_\eta &= \sum X_{ij} X_{i,j+1} X_{i-1,j} \\
N_\xi &= \sum X_{ij} X_{i,j-1} X_{i+1,j} \\
N_\theta &= \sum X_{ij} X_{i,j-1} X_{i-1,j}
\end{aligned}$$

$$N_l = \sum X_{ij} X_{i,j+1} X_{i+1,j}$$

$$N_\kappa = \sum X_{ij} X_{i,j+1} X_{i+1,j} X_{i+1,j+1}$$

Les sommations sont réalisées à partir de l'ensemble des puces qui ont un voisinage entièrement compris dans la plaquette. Elles comptabilisent le nombre de cliques qui ne sont composées que de puces défectives.

- (b) Les 10 statistiques précédentes sont suffisantes pour l'estimation des paramètres $(\alpha, \lambda, \mu, \nu, \pi, \eta, \xi, \theta, \iota, \kappa)$.
- (c) La distribution jointe de toutes les puces X_s d'une plaquette est de la forme :

$$P(X) = P(0) \exp\{\alpha N_\alpha + \lambda N_\lambda + \mu N_\mu + \nu N_\nu + \pi N_\pi + \eta N_\eta + \xi N_\xi + \theta N_\theta + \iota N_\iota + \kappa N_\kappa\}$$

- (d) La distribution conditionnelle de la puce X_s est de la forme :

$$P(X_s | \mathcal{V}_s) = \frac{1}{1 + \exp\left\{\sum_C \mathcal{C} \Delta_C(s, X_s, \mathcal{V}_s)\right\}}$$

avec $\mathcal{C} \in \{\alpha, \lambda, \mu, \nu, \pi, \eta, \xi, \theta, \iota, \kappa\}$. La statistique Δ_C est définie par

$$\Delta_C(s, X_s, \mathcal{V}_s) = N_C(s, 1 - X_s, \mathcal{V}_s) - N_C(s, X_s, \mathcal{V}_s)$$

avec $N_C(s, X_s, \mathcal{V}_s)$ qui correspond à la statistique N_C appliquée à la puce s et son voisinage \mathcal{V}_s .

Pour une puce s défective, $\Delta_C(s, X_s, \mathcal{V}_s)$ représente le nombre négatif de cliques de type \mathcal{C} qui impliquent s et les puces défectives de son voisinage. Pour une puce s opérationnelle, $\Delta_C(s, X_s, \mathcal{V}_s)$ représente le nombre de cliques défectives de type \mathcal{C} qui serait obtenu si la puce s était défective.

Une démonstration de ces résultats est proposée par Baron *et al.* (2001).

La distribution jointe s'exprime à partir de la probabilité $P(0)$ d'avoir 100% de puces opérationnelles sur la plaquette. Cette probabilité est définie par

$$P(0) = \left(\sum_W \exp\{\alpha N_\alpha + \lambda N_\lambda + \dots + \kappa N_\kappa\} \right)^{-1}.$$

La sommation s'effectue sur l'ensemble de toutes les combinaisons possibles de puces opérationnelles sur une plaquette représentée par l'ensemble W , soit $2^{|W|}$

combinaisons. Les paramètres $(\alpha, \lambda, \mu, \nu, \pi, \eta, \xi, \theta, \iota, \kappa)$ qui caractérisent la signature spatiale des puces défailtantes sont donc plutôt estimés à partir des distributions conditionnelles. Ainsi, le vecteur $\Theta = (\alpha, \lambda, \mu, \nu, \pi, \eta, \xi, \theta, \iota, \kappa)$ est estimé à partir de la maximisation de la pseudo-vraisemblance associée aux distributions conditionnelles :

$$\begin{aligned}\hat{\Theta} &= \arg \max_{\Theta} \prod_{s \in W_o} P(X_s | \mathcal{V}_s) \\ &= \arg \min_{\Theta} \prod_{s \in W_o} (1 + \exp\{\sum_c \mathcal{C} \Delta_c(s, X_s, \mathcal{V}_s)\}).\end{aligned}$$

L'ensemble W_o représente un ensemble de puces conditionnellement indépendantes les unes des autres, compte tenu de l'ensemble des puces restantes $W \setminus W_o$. Cette indépendance justifie la multiplication des distributions conditionnelles dans le calcul de la pseudo-vraisemblance.

L'estimation $\hat{\Theta}$ décrit numériquement la signature spatiale des puces défailtantes. L'estimation $\hat{\alpha}$ donne le taux de défaillance «nette», *i.e.* la probabilité qu'une puce soit défailtante alors que les puces de son voisinage ne le sont pas. Cette probabilité est estimée par $\frac{e^{\hat{\alpha}}}{1+e^{\hat{\alpha}}}$. Elle traduit l'hypothèse d'indépendance spatiale entre les puces défailtantes. Les estimations des autres paramètres décrivent plutôt la dépendance spatiale entre les puces défailtantes. Des valeurs élevées des 4 estimations $\hat{\lambda}$, $\hat{\mu}$, $\hat{\nu}$ et $\hat{\pi}$ témoignent d'une forte interaction entre les puces adjacentes résultant probablement de la présence d'un regroupement massif de puces défailtantes sur la plaquette. En revanche, des valeurs faibles de ces estimations suggèrent plutôt une dispersion des puces défailtantes. La comparaison de ces estimations donnent aussi des indications sur la forme d'éventuels regroupements. Si $\hat{\lambda} \approx \hat{\mu}$ et $\hat{\nu} \approx \hat{\pi}$, la forme du regroupement est probablement circulaire. Si $\hat{\lambda} > \hat{\mu}$, les effets horizontaux sont plus influents que les effets verticaux, etc.

La maximisation de la pseudo-vraisemblance ne conduit pas toujours à des estimations uniques et finies. Les différents cas à considérer sont explicités par Baron *et al.* (2001).

Les 10 paramètres estimés permettent de caractériser la cartographie des plaquettes, mais ils permettent aussi de comparer les plaquettes entre elles et donc de les classer.

Il est à noter que cette méthode présente quelques limites dans son application. Le procédé de fabrication génère naturellement des effets de bord : les puces de bord ont une plus forte probabilité d'être défailtantes que les autres puces de la plaquette. En conséquence, l'homogénéité du modèle de distribution peut être remise en cause. Selon Baron *et al.* (2001), le modèle a pour objet l'identification de regroupements massifs de puces défailtantes et ces effets dominent les autres effets susceptibles d'être présents sur la plaquette. Ils préconisent donc d'éliminer de l'analyse, les puces qui n'ont pas un voisinage inclus complètement sur la plaquette. Lorsque la taille des puces est élevée et qu'elles sont peu nombreuses sur la plaquette, cette suppression peut représenter plus de 20% des puces contenues sur la plaquette. Dans ce cas, la présence d'une couronne de puces défailtantes en bord de plaquette risque d'être ignorée. Par ailleurs, le calcul de la pseudo-vraisemblance ne peut se faire qu'à partir

d'un échantillon de puces conditionnellement indépendantes l'une de l'autre. Pour répondre à cette contrainte, Baron *et al.* (2001) préconisent d'utiliser l'ensemble le plus large qui vérifie cette propriété : $W_o = \{(i, j) \text{ tel que } (i/2, j/2) \in W\}$. Dans le cas de plaquettes avec un nombre restreint de puces, cet ensemble W_o représente un faible nombre de puces, et finalement, l'ensemble complet W de puces doit être utilisé malgré la perte d'indépendance. Toutefois, Comets (1992) a prouvé la consistance de l'estimation $\hat{\Theta}$ dans ce contexte.

2.2. Modélisation du rendement selon un modèle de Poisson

Stapper *et al.* (1983) décrivent plusieurs approches pour modéliser le rendement d'une plaquette. Ils proposent notamment d'utiliser un modèle de Poisson appliqué au nombre de défauts observés sur une puce. Associé à des résultats de Paz *et al.* (1977), ce choix de modèle permet de décomposer facilement le rendement total en des rendements systématique et aléatoire. Hansen *et al.* (1997b) présentent de manière simplifiée cette approche.

Soit D_0 , le nombre moyen de défauts par unité d'aire de puce et A l'aire active des puces. Dans le cadre d'un modèle de Poisson, la probabilité qu'une puce ait k défauts indépendamment de l'état des autres puces est définie par :

$$P(k) = \frac{(D_0 A)^k \exp(-D_0 A)}{k!}.$$

La probabilité qu'une puce soit fonctionnelle parce qu'elle ne compte aucun défaut s'écrit alors $\pi = \exp(-D_0 A)$. Il est raisonnable de penser que cette probabilité varie d'une région à une autre de la plaquette. Schématiquement, la plaquette peut être décomposée en deux parties : une partie où la probabilité qu'une puce soit défectueuse est faible, elle correspond à la partie aléatoire de la plaquette ; une partie où cette probabilité est élevée, elle correspond à la partie systématique de la plaquette avec une forte concentration de puces défectueuses. Chaque partie peut être composée de l'union de plusieurs sous-parties. Par exemple, la partie systématique peut être composée d'une couronne en bord de plaquette et d'une tâche au centre de la plaquette. La partie aléatoire est alors définie par le reste de la plaquette. Soit \mathcal{W} , l'ensemble des puces de la plaquette, \mathcal{C}_1 l'ensemble des puces de la partie aléatoire et \mathcal{C}_2 l'ensemble des puces de la partie systématique. On vérifie que $\mathcal{W} = \mathcal{C}_1 \cup \mathcal{C}_2$ et $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$. Aux ensembles \mathcal{C}_1 et \mathcal{C}_2 sont associées respectivement les probabilités π_1 et π_2 que les puces n'aient aucun défaut. Par définition de l'ensemble \mathcal{C}_2 , la probabilité π_2 est très proche de 0.

Soit c une puce de la plaquette, la probabilité qu'une puce soit fonctionnelle s'écrit alors :

$$\begin{aligned} P(c \text{ fonctionnelle}) &= P(c \in \mathcal{C}_1)P(c \text{ fonctionnelle} | c \in \mathcal{C}_1) \\ &\quad + P(c \in \mathcal{C}_2)P(c \text{ fonctionnelle} | c \in \mathcal{C}_2) \\ &= P(c \in \mathcal{C}_1)\pi_1 + P(c \in \mathcal{C}_2)\pi_2. \end{aligned}$$

Soit W le nombre de puces sur la plaquette. Le rendement total de la plaquette Y s'écrit :

$$Y = \frac{1}{W} \sum_{c \in \mathcal{W}} I_{\{c \text{ fonctionnelle}\}}$$

avec $I_{\{c \text{ fonctionnelle}\}}$ égal à 1 si la puce c est fonctionnelle et 0 sinon. L'espérance de la variable Y s'écrit alors :

$$\mathbb{E}(Y) = \frac{1}{W} \sum_{c \in \mathcal{W}} P(c \text{ fonctionnelle}).$$

D'après l'expression de $P(c \text{ fonctionnelle})$ donnée précédemment, cette espérance peut s'écrire aussi :

$$\begin{aligned} \mathbb{E}(Y) &= \frac{1}{W} \sum_{c \in \mathcal{W}} P(c \in \mathcal{C}_1) \pi_1 + \frac{1}{W} \sum_{c \in \mathcal{W}} P(c \in \mathcal{C}_2) \pi_2 \\ &= \frac{\pi_1 \mathbb{E}(C_1)}{W} + \frac{\pi_2 \mathbb{E}(C_2)}{W}. \end{aligned}$$

Les variables C_1 et C_2 représentent le nombre puces dans chaque ensemble \mathcal{C}_1 et \mathcal{C}_2 , soit

$$C_i = \sum_{c \in \mathcal{W}} I_{\{c \in \mathcal{C}_i\}}$$

avec $i = 1, 2$ et $I_{\{c \in \mathcal{C}_i\}}$ égal à 1 si c appartient à l'ensemble \mathcal{C}_i et 0 sinon. Par hypothèse, π_2 est proche de 0, l'espérance du rendement s'écrit alors :

$$\mathbb{E}(Y) = \frac{\pi_1 \mathbb{E}(C_1)}{W} = \pi_1 \left(1 - \frac{\mathbb{E}(C_2)}{W}\right).$$

L'expression $\left(1 - \frac{\mathbb{E}(C_2)}{W}\right)$ représente le rendement associé à la partie systématique de la plaquette, alors que la probabilité π_1 représente le rendement associé à la partie aléatoire. Cela revient à poser : $Y_r = \pi_1$ et $Y_s = \left(1 - \frac{\mathbb{E}(C_2)}{W}\right)$. Par définition de π_1 , le rendement aléatoire s'écrit $Y_r = \exp(-D_0 A)$. Ainsi,

$$\mathbb{E}(Y) = \exp(-D_0 A) Y_s.$$

En prenant le logarithme de cette expression, on obtient une relation linéaire entre l'espérance du rendement et la taille de la puce :

$$\ln(\mathbb{E}(Y)) = -D_0 A + \ln Y_s.$$

En procédant à une estimation par Moindres Carrés, il est possible d'estimer la pente D_0 et l'ordonnée à l'origine $\ln Y_s$ de la relation linéaire. On en déduit

l'estimation de Y_r et Y_s . Cette estimation requiert l'observation d'un échantillon bivarié (A_i, Y_i) , c'est-à-dire un échantillon qui donne pour différentes tailles de puce A_i , le rendement total Y_i associé. Pour obtenir un tel échantillon à partir d'une même plaquette, des super-puces sont créées artificiellement. Ainsi, à partir d'une plaquette avec des puces de taille A et un rendement Y , on construit des super-puces de taille $k \times A$. Autrement dit, si A représente une puce, les super-puces représentent k puces. Une super-puce est dite défaillante si l'une au moins des puces qui la composent est défaillante, elle est dite fonctionnelle si toutes les puces qui la composent sont fonctionnelles. On obtient ainsi un nouveau rendement de la plaquette qui est le ratio du nombre de super-puces défaillantes sur le nombre de super-puces contenues sur la plaquette. En itérant ce calcul pour plusieurs valeurs de k , on obtient l'échantillon bivarié souhaité.

Cette méthode de calcul présente certaines limites. Quand la taille des puces est élevée et que les plaquettes comptent juste quelques centaines de puces, la création de super-puces est délicate. En effet, compte tenu du faible nombre initial de puces sur la plaquette, le nombre de super-puces est fortement restreint rendant le calcul du rendement associé approximatif. Dès lors, la qualité de l'échantillon peut être remise en cause pour une bonne estimation de D_0 et de $\ln Y_s$. Par ailleurs, les puces de bord nécessitent un traitement spécifique, les super-puces qui leur sont associées n'étant pas toujours entièrement comprises dans la plaquette. Il est donc nécessaire de corriger le calcul du rendement pour ne pas sur-estimer les défaillances en bord de plaquette.

2.3. Modélisation par un Processus Spatial

La méthode proposée par Hansen *et al.* (1997) consiste à décomposer les plaquettes en deux sous-plaquettes comme décrit Figure 2. À partir de cette décomposition, sont déduits le rendement aléatoire issu de la dispersion des puces défaillantes, et le rendement systématique issu des regroupements non-aléatoires de puces défaillantes. Cette décomposition permet d'évaluer la part de chaque mode de défaillance dans la valeur du rendement total. En outre, elle permet d'éliminer le bruit issu du dépôt de particules pour se concentrer sur l'analyse des regroupements de puces défaillantes ainsi identifiés.

Pour la suite, on a besoin d'un ensemble de notations/définitions :

- \mathcal{W} est l'ensemble des puces d'une plaquette,
- W est le nombre de puces sur une plaquette,
- i est une puce de la plaquette, $i \in \mathcal{W}$
- \mathcal{N}_1 (resp. \mathcal{N}_0) est l'ensemble des positions des puces défaillantes (resp. fonctionnelles),
- N_1 (resp. N_0) est le nombre de puces défaillantes (resp. fonctionnelles),
- $I_{\mathcal{N}_1}(i) = 1$ si $i \in \mathcal{N}_1$, i.e. si i est défaillante et 0 sinon,
- $I_{\mathcal{N}_0}(i) = 1$ si $i \in \mathcal{N}_0$, i.e. si i est fonctionnelle et 0 sinon,
- $p_i = \mathbb{E}(I_{\mathcal{N}_1}(i))$ est la probabilité que la puce i soit défaillante,
- $q_i = 1 - p_i$ est la probabilité que la puce i soit fonctionnelle.

Par définition, $\mathcal{W} = \mathcal{N}_1 \cup \mathcal{N}_0$ avec $\mathcal{N}_1 \cap \mathcal{N}_0 = \emptyset$ et $W = N_1 + N_0$. Il vient aussi que $I_{\mathcal{N}_0}(i) = 1 - I_{\mathcal{N}_1}(i)$ pour tout i .

La modélisation est basée sur deux hypothèses. L'hypothèse nulle suppose que les puces sont aléatoirement dispersées sur la plaquette. On parle d'indépendance spatiale. Sous cette hypothèse, les variables $I_{\mathcal{N}_1}(i)$ sont donc indépendantes et identiquement distribuées selon des variables de Bernoulli de paramètre p , avec $p = p_i$ pour toutes les puces i de \mathcal{W} . Autrement dit, les puces de la plaquette ont toutes la même probabilité d'être défailante. Cela revient à considérer un processus de Bernoulli spatialement homogène sur l'ensemble des puces de la plaquette. La probabilité p peut être estimée globalement par $\hat{p} = \frac{N_1}{W}$ ou localement par $\hat{p}_i = \sum_{j \in \mathcal{W}} w_i(j) I_{\mathcal{N}_1}(j)$. Pour chaque puce i de la plaquette, l'ensemble $\{w_i(j)\}$ est un ensemble de poids non négatifs et normalisés tel que $\sum_{j \in \mathcal{W}} w_i(j) = 1$. La structure des poids est identique pour chaque puce : ils sont constants et égaux pour les puces qui sont dans le voisinage et nuls en dehors. Pour les puces en bord de plaquette, le calcul des poids est adapté au nombre de puces présentes dans le voisinage. Le voisinage le plus couramment utilisé est un voisinage 3×3 , identique à celui de la méthode précédente (Figure 3). Ce choix n'est pas rigide, de nombreux autres voisinages peuvent être privilégiés. L'hypothèse alternative suppose la présence de regroupements de puces défailtantes. Implicitement, cela suggère qu'il existe des zones sur la plaquette où la probabilité qu'une puce soit défailtante est plus importante. Dans ce cas, des écarts importants entre l'estimation globale \hat{p} et les estimations locales \hat{p}_i de la probabilité qu'une puce soit défailtante sont observés. Pour cette raison, l'hypothèse nulle d'indépendance spatiale est testée à partir de statistiques de la forme

$$T = \sum_{i \in \mathcal{N}_1} \phi_{\mathcal{N}_1}(\hat{p}_i, \hat{p}) + \sum_{i \in \mathcal{N}_0} \phi_{\mathcal{N}_0}(\hat{q}_i, \hat{q})$$

où $\phi_{\mathcal{N}_1}(\hat{p}_i, \hat{p})$ et $\phi_{\mathcal{N}_0}(\hat{q}_i, \hat{q})$ mesurent l'écart entre les estimations globales et les estimations locales. Il y a différents choix possibles pour les fonctions ϕ . Hansen *et al.* (1997) proposent de construire deux statistiques : $S_{\mathcal{N}_1}$ pour les puces défailtantes et $S_{\mathcal{N}_0}$ pour les puces fonctionnelles. Dans le premier cas, ils posent que $\phi_{\mathcal{N}_1}(x, y) = W^{-1}(x - y)$ et $\phi_{\mathcal{N}_0}(x, y) = 0$. Dans le deuxième cas, ils posent que $\phi_{\mathcal{N}_1}(x, y) = 0$ et $\phi_{\mathcal{N}_0}(x, y) = W^{-1}(x - y)$. On obtient alors les statistiques $S_{\mathcal{N}_1}$ et $S_{\mathcal{N}_0}$ suivantes :

$$S_{\mathcal{N}_1} = W^{-1} \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{W}} w_i(j) I_{\mathcal{N}_1}(i) I_{\mathcal{N}_1}(j) - p^2,$$

$$S_{\mathcal{N}_0} = W^{-1} \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{W}} w_i(j) I_{\mathcal{N}_0}(i) I_{\mathcal{N}_0}(j) - q^2.$$

Il s'agit de versions généralisées des statistiques proposées par Moran (1948) pour tester l'hypothèse de dépendance spatiale. Selon la structure de poids choisie (constants dans le voisinage et nuls au-delà), les statistiques $S_{\mathcal{N}_1}$ et $S_{\mathcal{N}_0}$ peuvent être interprétées comme des estimateurs du premier ordre de la covariance spatiale. Ainsi, $S_{\mathcal{N}_1}$ est un estimateur de $\mathbb{E}(I_{\mathcal{N}_1}(i) I_{\mathcal{N}_1}(j)) - \mathbb{E}(I_{\mathcal{N}_1}(i)) \mathbb{E}(I_{\mathcal{N}_1}(j))$ où j appartient au voisinage de i . Sous l'hypothèse nulle d'indépendance spatiale, cette covariance est

proche de 0. Pour la suite, on s'intéressera à une version simplifiée des statistiques S_{N_1} et S_{N_0} :

$$T_{N_1} = W^{-1} \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{W}} w_i(j) I_{N_1}(i) I_{N_1}(j),$$

$$T_{N_0} = W^{-1} \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{W}} w_i(j) I_{N_0}(i) I_{N_0}(j).$$

En conditionnant les valeurs observées, par l'estimation globale de la probabilité de défaillance ($\hat{p} = \frac{N_1}{W}$), ces statistiques sont équivalentes à S_{N_1} et S_{N_0} . Cliff et Ord (1981) ont étudié différentes statistiques pour tester l'hypothèse d'indépendance spatiale parmi lesquelles les statistiques T_{N_1} et T_{N_0} . En particulier, ils se sont intéressés à leur distribution et à leurs propriétés asymptotiques. Ainsi, ils ont montré que les deux premiers moments de la statistique T_{N_1} s'écrivent :

$$\mathbb{E}(T_{N_1} | \hat{p} = \frac{N_1}{W}) = \frac{N_1}{W} \times \frac{N_1 - 1}{W - 1},$$

$$\mathbb{E}(T_{N_1}^2 | \hat{p} = \frac{N_1}{W}) = \alpha_1 \frac{N_1}{W} + \alpha_2 \frac{N_1}{W} \times \frac{N_1 - 1}{W - 1} + \alpha_3 \frac{N_1}{W} \times \frac{N_1 - 1}{W - 1} \times \frac{N_1 - 2}{W - 2}$$

$$+ \alpha_4 \frac{N_1}{W} \times \frac{N_1 - 1}{W - 1} \times \frac{N_1 - 2}{W - 2} \times \frac{N_1 - 3}{W - 3}.$$

Les moments de la statistique T_{N_0} se déduisent de ceux de T_{N_1} en remplaçant N_1 par $(W - N_1)$. Ces égalités sont obtenues en supposant que pour toutes puces j , $w_j(j) = 0$. Dans le cas contraire, un facteur correctif est introduit pour prendre en compte le fait que $I_{N_1}(j)^2 = I_{N_1}(j)$. Les paramètres α_k sont définis par

$$\alpha_k = W^{-2} \sum_{i \in \mathcal{W}} \sum_{j \in \mathcal{W}} \sum_{i' \in \mathcal{W}} \sum_{j' \in \mathcal{W}} w_i(j) w_{i'}(j') I_k(i, j, i', j')$$

avec I_k égal à 1 s'il y a exactement k éléments distincts parmi les 4 arguments et 0 sinon (par exemple, $I_1(i, i, i, i) = 1$ et $I_1(i, j, i, i) = 0$ avec $i \neq j$ ou $I_2(i, j, i, j) = 1$ avec $i \neq j$ et $I_2(i, j, k, i) = 0$ avec $i \neq j \neq k$). Le calcul des moments dépend de l'arrangement des puces sur la plaquette (α_k), de la structure des poids ($w_i(j)$), du nombre de puces défaillantes (N_1) et du nombre total de puces (W) sur la plaquette. Selon Hansen *et al.* (1997), la combinaison linéaire $T_{N_1} + T_{N_0}$ est la statistique de test qui permet d'avoir le test de l'hypothèse d'indépendance spatiale le plus puissant. Toutefois, en pratique, la statistique de test T_{N_1} se comporte aussi de manière satisfaisante. Ainsi, dans un souci de simplification, la statistique T_{N_1} est la statistique qui a été retenue pour tester l'hypothèse d'indépendance spatiale.

La distribution asymptotique de T_{N_1} suit une loi normale (Cliff et Ord, 1981). Asymptotiquement, la loi de la statistique T_{N_1} est donc parfaitement connue. Par définition de la statistique T_{N_1} , l'hypothèse nulle d'indépendance spatiale est rejetée pour une valeur élevée (en valeur absolue) de celle-ci.

Ce test d'hypothèse est à la base de la décomposition des plaquettes en deux sous-plaquettes : l'une dite aléatoire parce qu'elle représente la dispersion

des puces défaillantes sur la plaquette et l'autre dite systématique parce qu'elle représente les regroupements de puces défaillantes (Figure 2). À chaque puce, est associée la proportion de puces défaillantes incluses dans son voisinage. Les puces défaillantes qui ont une proportion élevée de puces voisines défaillantes sont supposées faire partie d'un regroupement de puces défaillantes. Ces puces sont alors associées à la sous-plaquette systématique et les autres puces défaillantes à la sous-plaquette aléatoire. Pour chaque puce, la proportion de puces voisines défaillantes est comparée strictement à un seuil compris entre 0 et 1. Choisir un seuil de 0 revient à considérer que toutes les puces font partie d'un regroupement de puces défaillantes : la plaquette systématique est la plaquette initiale. Choisir un seuil de 1 revient à considérer qu'aucune puce ne fait partie d'un regroupement de puces défaillantes : la plaquette aléatoire est la plaquette initiale. La figure 5 illustre l'influence du choix du seuil sur l'estimation des regroupements de puces défaillantes : plus le seuil est élevé, plus on risque de sous-estimer la taille des regroupements, en revanche, plus le seuil est faible, plus on risque de sur-estimer la taille des regroupements en prenant en compte les puces qui par le hasard du dépôt de particules ont formé un regroupement. Le choix du seuil est donc déterminant pour une bonne estimation des deux sous-plaquettes. Hansen *et al.* (1997) proposent un algorithme pour déterminer objectivement le niveau de seuil. Cet algorithme est basé sur l'application du test statistique d'indépendance spatiale. Au départ, le seuil est fixé à 1. La sous-plaquette aléatoire est la plaquette initiale. Le test statistique est appliqué sur la plaquette aléatoire afin de vérifier l'hypothèse d'indépendance spatiale. Si cette hypothèse n'est pas rejetée, cela signifie qu'effectivement, il n'y a pas de regroupements spatiaux sur la plaquette. L'algorithme ne continue pas. Si l'hypothèse est rejetée, le seuil est diminué. Les puces dont la proportion de puces voisines défaillantes est strictement supérieure au nouveau seuil sont déplacées de la sous-plaquette aléatoire vers la sous-plaquette systématique. Le test est appliqué sur la nouvelle sous-plaquette aléatoire. Si l'hypothèse d'indépendance spatiale n'est pas rejetée, la décomposition de la plaquette est achevée. Si l'hypothèse est rejetée, le seuil est de nouveau diminué et la décomposition de la plaquette est poursuivie jusqu'au non-rejet de l'hypothèse.

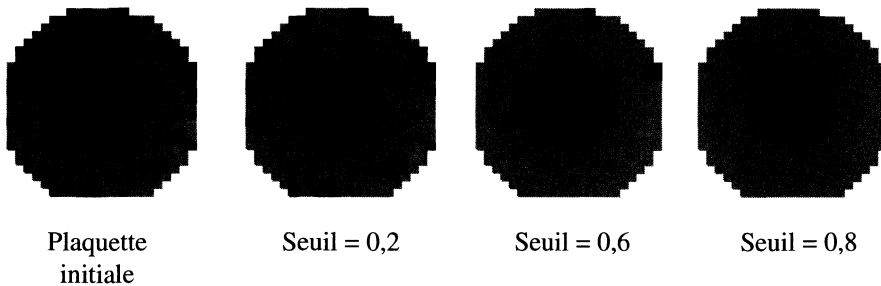


FIGURE 5

Évolution de la sous-plaquette systématique en fonction du seuil

Les rendements aléatoire et systématique sont calculés à partir de la décomposition de la plaquette. Soit N_r , le nombre de puces défaillantes sur la sous-plaquette aléatoire et N_s , le nombre de puces défaillantes sur la sous-plaquette systématique.

Le rendement systématique Y_s s'écrit alors $Y_s = \frac{W-N_s}{W}$ et le rendement aléatoire Y_r correspond à $Y_r = \frac{W-N_s-N_r}{W-N_s}$. On vérifie que le rendement total Y est le produit des deux rendements : $Y = Y_r \times Y_s$.

En pratique, un suivi minutieux des trois rendements Y , Y_r et Y_s va permettre de réagir rapidement à une défaillance du procédé de fabrication. En effet, une baisse du rendement total Y est généralement associée à la baisse de l'un des deux rendements Y_r , Y_s . Les outils classiques appliqués au rendement vont être utilisés sur ces deux composantes, l'avantage étant que les analyses sont alors moins bruitées. Par ailleurs, l'analyse de la sous-plaquette systématique permet d'accélérer le processus d'identification de la défaillance du procédé : une cartographie particulière peut être associée à un mécanisme particulier de défaillance qui a déjà été répertorié par les ingénieurs.

L'application de cette méthode rencontre quelques limites. Lorsque le rendement est proche des extrêmes 0 ou 1, le choix entre hypothèse nulle et alternative est délicat. Si le rendement est très faible, il est difficile de ne pas associer une puce défaillante issue d'un dépôt aléatoire de particules à un regroupement de puces défaillantes issu d'un dysfonctionnement particulier du procédé. La décomposition de la plaquette et du rendement en parties aléatoire et systématique est donc difficile. En revanche, si le rendement est très élevé, la partie aléatoire risque d'être privilégiée aux dépens de la partie systématique. Dans ce contexte, l'application de la méthode peut paraître superflue. Pourtant, malgré un faible nombre de puces défaillantes, des cartographies particulières peuvent être observées. Par exemple, malgré un rendement élevé, une tâche au centre de la plaquette peut être observée. Il est alors judicieux de traiter le problème lorsque les rendements sont élevés afin de prévenir une éventuelle aggravation du mécanisme de défaillance.

Enfin, selon le voisinage choisi, des mécanismes particuliers de défaillance peuvent être assimilés à un phénomène aléatoire alors qu'ils sont issus d'un dysfonctionnement «systématique» du procédé. En particulier, des lignes de puces défaillantes générées par un mécanisme de défaillance répétitif peuvent être attribuées à la sous-plaquette aléatoire. Le voisinage de type 3×3 permet d'identifier les principaux types de regroupement, mais il s'avère peu efficace dans ce type de configuration. Ainsi, il est parfois nécessaire de modifier la structure de voisinage pour être capable d'identifier des mécanismes particuliers de défaillance.

2.4. Modélisation par arbre de classification

Lorsque le rendement d'une plaquette est relativement faible, l'ingénieur responsable des rendements visualise sa cartographie pour détecter la présence d'une signature spatiale particulière. Le volume élevé de la production rend ce travail fastidieux, notamment lorsqu'un dysfonctionnement du procédé de fabrication impacte des dizaines de plaquettes. Par ailleurs, lorsque le rendement d'une plaquette est relativement élevé, l'ingénieur ne va pas systématiquement visualiser la cartographie. Pourtant, une plaquette avec un rendement élevé peut présenter un début de signature spatiale qui peut s'aggraver avec le volume de production. En d'autres termes, il n'est pas superflu de vouloir automatiser la classification des plaquettes selon leur cartographie. Duvivier (1999) propose une méthode simple de classification. Elle est

basée sur l'application d'un arbre binaire de type CART (Breiman *et al.*, 1984). Toutefois, avant de procéder à cette classification, il préconise le calcul d'un indice de présence d'une dépendance spatiale sur la plaquette. Selon, la valeur de cet indice, la présence d'une dépendance spatiale est plus ou moins probable. Le calcul est basé sur la décomposition du rendement total Y en un rendement systématique Y_s et un rendement aléatoire Y_r , tel que $Y = Y_r \times Y_s$. Duvivier (1999) propose d'utiliser la méthode basée sur le modèle de Poisson pour procéder à la décomposition. De cette décomposition sont déduits plusieurs indices intermédiaires :

1. Pourcentage de puces défailtantes issues **exclusivement** d'un processus générant des regroupements de puces défailtantes : $(1 - Y_s)Y_r$.
2. Pourcentage de puces défailtantes issues **exclusivement** d'un processus générant des puces défailtantes de manière dispersée : $Y_s(1 - Y_r)$.
3. Pourcentage de puces défailtantes issues **indistinctement** des deux processus précédents : $(1 - Y_s)(1 - Y_r)$.

Dans le calcul du dernier pourcentage (3), les contributions des deux mécanismes de rejets sont indissociables et supposées équiprobables. Il est donc nécessaire de corriger les deux précédents pourcentages (1 et 2) :

- 1'. Pourcentage corrigé de puces défailtantes issues **exclusivement** d'un processus générant des regroupements de puces défailtantes : $K_s = (1 - Y_s)Y_r + \frac{(1 - Y_s)(1 - Y_r)}{2}$.
- 2'. Pourcentage corrigé de puces défailtantes issues **exclusivement** d'un processus générant des puces défailtantes de manière dispersée : $K_r = Y_s(1 - Y_r) + \frac{(1 - Y_s)(1 - Y_r)}{2}$.

Les puces d'une plaquette sont forcément fonctionnelles ou défailtantes. Ainsi, $1 = Y + K_r + K_s$. Le ratio RR défini par $\frac{K_r}{K_r + K_s}$ quantifie la contribution du mécanisme de dispersion des puces défailtantes sur l'ensemble des puces défailtantes de la plaquette. En remplaçant K_r et K_s par leur formule respective, le ratio RR s'écrit alors :

$$RR = \frac{(1 + Y_s)(Y_s - Y)}{2Y_s(1 - Y)}.$$

Selon l'importance de chaque mécanisme de défaillance, le ratio varie de 0 à 1. Si toutes les puces défailtantes sont issues d'un mécanisme dispersé, le rendement systématique vaut 1 ($Y_s = 1$) et le ratio RR prend la valeur 1. En revanche, si toutes les puces défailtantes sont regroupées sur la plaquette, le rendement systématique Y_s est équivalent au rendement total Y , et le ratio RR prend la valeur 0. Plus le ratio RR est proche de 0, plus la présence d'un regroupement significatif de puces défailtantes est donc fort probable. Dans ce cas, Duvivier (1999) propose alors d'identifier la signature du regroupement par un arbre de classification. L'état de la puce (fonctionnelle ou défailtante) est la variable à expliquer. C'est une variable binaire qui prend 1 si la puce est défailtante et 0 sinon. Les variables explicatives représentent différents paramètres permettant de caractériser la signature spatiale d'une plaquette :

- Coordonnées cartésiennes des puces sur la plaquette.
Il s'agit de deux variables quantitatives : (diex, diey). Dans le cadre des illustrations de la section 3, $diex \in [1, 22]$ et $diey \in [1, 24]$.
- Coordonnées polaires (r, θ)
- Distance r des puces par rapport au centre de la plaquette.
Il s'agit d'une variable quantitative. Pour faciliter l'interprétation de l'arbre, les puces sont regroupées par anneaux autour du centre de la plaquette. Le nombre d'anneau dépend du nombre de puces sur la plaquette. Dans le cadre des illustrations de la section 3, l'expérience a montré que le découpage en 6 anneaux était suffisant.
- Position angulaire θ des puces sur la plaquette.
Il s'agit d'une variable quantitative : $\theta \in [0, 2\pi]$.
- Position sur le réticule. Un réticule permet de reproduire le motif d'une puce sur la plaquette. Pour accélérer le temps de cycle, un réticule comprend plusieurs représentations du même motif. Ainsi, plusieurs motifs sont reproduits simultanément sur la plaquette. Le réticule est très fragilisé par les nombreuses manipulations. Certaines positions peuvent donc devenir défectueuses au cours du temps et générer des puces défailtantes régulièrement ordonnées sur la plaquette. Ce type de défaillance répétitif est mal interprété par les calculs

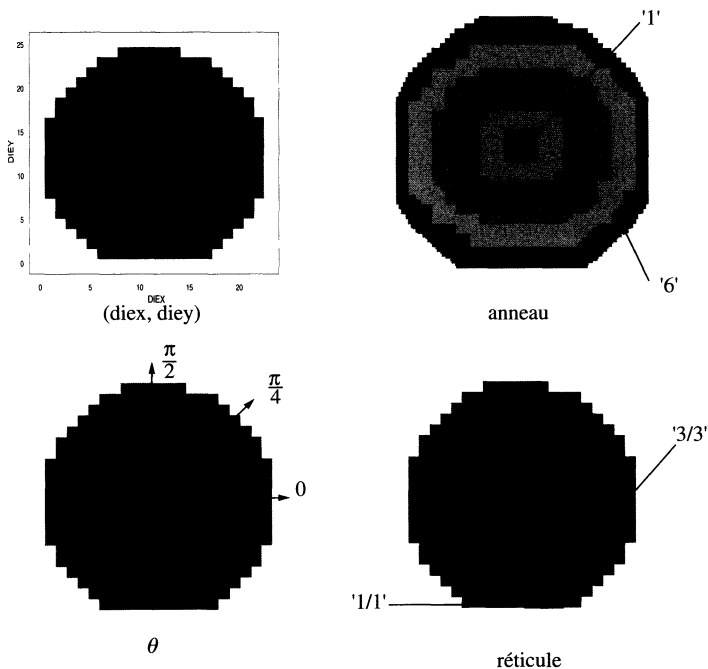


FIGURE 6

Variables explicatives de l'arbre de classification

3/1	3/2	3/3
2/1	2/2	2/3
1/1	1/2	1/3

FIGURE 7
Réticule

de décomposition du rendement. Le calcul du ratio RR ne le reflète donc pas toujours. Le nombre de positions sur le réticule dépend du motif de la puce. Par la suite, les illustrations se feront à partir d'un réticule 3×3 . La Figure 7 en est une représentation. Il s'agit donc d'une variable qualitative dont les modalités sont $\{1/1, 1/2, 1/3, 2/1, 2/2, 2/3, 3/1, 3/2, 3/3\}$.

La représentation graphique de ces variables explicatives est donnée Figure 6. Ces variables permettent d'identifier la plupart des regroupements de puces défailtantes :

- Les coordonnées cartésiennes permettent d'identifier des formes géométriques particulières à n'importe quel endroit de la plaquette. Par exemple, il peut s'agir d'une ligne en bas de plaquette, ou encore d'un rectangle en haut de plaquette, ...
- La distance par rapport au centre permet d'identifier des tâches au centre de la plaquette ou encore des couronnes en bord de plaquette.
- La position angulaire permet d'identifier des cadrans de la plaquette.
- La position sur le réticule permet d'identifier une position défectueuse du réticule.

La cartographie des puces défailtantes d'une plaquette peut être le résultat de l'interaction de plusieurs dysfonctionnements du procédé. Par exemple, la plaquette peut être simultanément victime d'un mauvais étalement de résine qui va générer une couronne de puces défailtantes en bord de plaquette et d'une position défectueuse du réticule. Par construction, l'arbre de classification est capable d'identifier les deux phénomènes qui interagissent : le premier niveau de l'arbre peut mettre en valeur le problème de couronne et le deuxième niveau, le problème de la position défectueuse du réticule. En revanche, l'arbre peut avoir tendance à révéler plus d'interactions qu'il n'y en a en réalité. L'élagage de l'arbre apparaît donc nécessaire pour ne retenir que les effets les plus significatifs. Duvivier (1999) préconise de ne retenir que les deux ou trois premiers niveaux de l'arbre. L'expérience montre qu'effectivement, les niveaux inférieurs de l'arbre sont rarement informatifs. Toutefois, compte tenu de la complexité croissante des procédés de fabrication, des cartographies particulières de défaillance peuvent nécessiter plusieurs niveaux de décomposition pour être totalement identifiées. Selon l'utilisation finale des résultats obtenus, l'élagage par validation croisée peut donc apparaître moins restrictif.

3. Application

Un échantillon de plaquettes avec différentes cartographies a été sélectionné. Les plaquettes ont toutes la même architecture : mêmes coordonnées cartésiennes, mêmes coordonnées polaires, même réticule et même nombre de puces (449). L'objectif est de comparer les résultats de chaque approche en terme de performance et de limite dans différentes configurations.

3.1. Échantillon

L'échantillon de plaquettes est décrit Figure 8 et Table 1. Les plaquettes 1, 2 et 3 représentant une dispersion aléatoire des puces défaillantes sur la plaquette sont issues de simulations. Ainsi, la disposition aléatoire des puces défaillantes est assurée. Les plaquettes 15 et 16 sont aussi des plaquettes simulées : à partir d'une plaquette ayant des puces défaillantes aléatoirement dispersées, un mécanisme réel de défaillance a été reproduit pour simuler des lignes horizontales de défaillance. Ces plaquettes vont permettre d'identifier les limites de plusieurs méthodes face à cette configuration. L'ensemble des plaquettes restantes est issu de la production. Les plaquettes issues de la production ont été choisies pour leurs caractéristiques spatiales (puces défaillantes en bord de plaquette avec une couronne plus ou moins grande, puces défaillantes regroupées au centre de la plaquette, ...).

TABLE 1
Table des cartographies étudiées

	Rend. élevé $Y \geq 80\%$	Rend. moyen $50\% \leq Y < 80\%$	Rend. faible $Y < 50\%$
Puces défaillantes majoritairement (1) dispersées sur la plaq.	P1	P2	P3
Puces défaillantes majoritairement (2) regroupées sur la plaq.	P4 et P7	P5 et P8	P6
Puces défaillantes pour position (3) réticule défectueuse	P9	P10	P11
Puces défaillantes à cause de (4) (2)+(3)			P12, P13 et P14
Puces défaillantes regroupées (5) par ligne		P15	P16

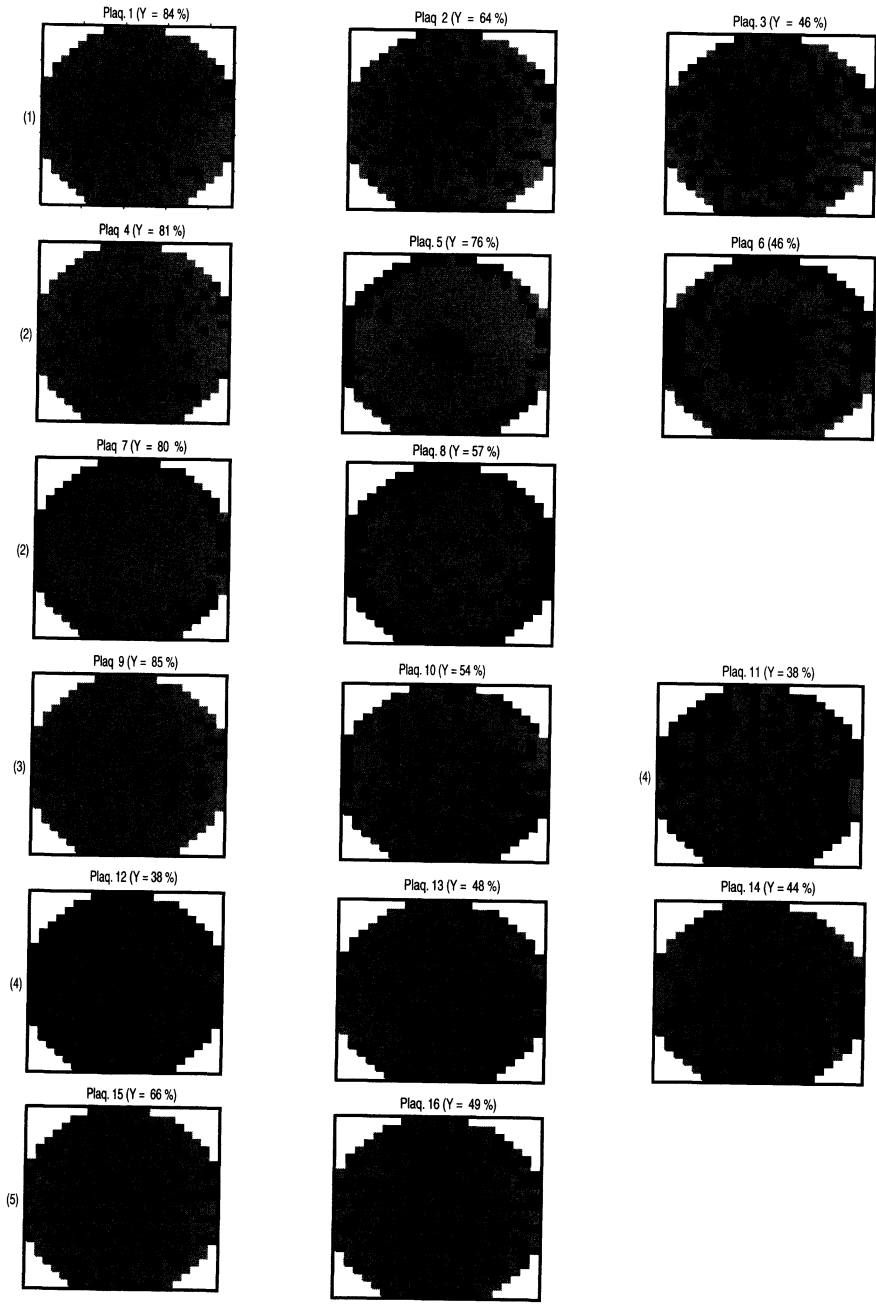


FIGURE 8
Cartographies des plaquettes analysées

3.2. Résultats

3.2.1. Modélisation markovienne

La Table 2 donne l'estimation des 10 paramètres canoniques issus de la modélisation markovienne de chaque plaquette étudiée. La comparaison des 10 paramètres est censée permettre d'identifier leur signature spatiale.

Les trois premières plaquettes représentent une dispersion aléatoire des puces défaillantes. Dans ce cas, les probabilités qu'une puce soit défaillante alors que ses puces voisines ne le sont pas, devraient donc être les plus élevées de tout l'échantillon avec une estimation positive du paramètre α . Seule l'estimation de la plaquette 2 est positive et elle est relativement faible par rapport aux estimations des plaquettes 15 et 16. Pour ces trois plaquettes, le paramètre estimé le plus élevé témoigne de la présence d'une interaction spatiale : interaction double (θ) pour les plaquettes 1 et 2 et effet vertical (μ) pour la plaquette 3. Dans le cas de la plaquette 4, l'estimation la plus influente correspond à l'interaction triple (κ). Cette interaction est révélatrice de regroupements de puces défaillantes sur la plaquette. Les estimations ne permettent pas de préciser le nombre de regroupements et leur étendue. Les estimations de la plaquette 5 révèlent la présence simultanée d'interactions doubles (η et ξ) et d'effets verticaux (μ) et horizontaux (λ). Ces résultats peuvent laisser pressentir la présence d'une couronne (lignes horizontales et verticales en bord de plaquette) et d'une tâche (combinaisons d'interactions doubles) qui caractérisent la plaquette 5. Toutefois, ces résultats peuvent aussi s'interpréter autrement. Par exemple, un réticule défectueux peut aussi générer des effets verticaux et/ou horizontaux. En d'autres termes, ces résultats ne sont pas suffisants pour identifier une signature spatiale unique. La plaquette 6 a la même signature spatiale que la plaquette 5 accentuée, toutefois, par des puces défaillantes plus nombreuses. Les résultats devraient donc être similaires. En fait, les estimations des paramètres les plus influents sont relativement moins élevées et ne pointent pas exactement les mêmes interactions. Il est même difficile de conclure sur la nature de la signature spatiale. Les résultats de la plaquette 7 témoignent de la présence simultanée d'effets verticaux et horizontaux (λ et μ). Ces résultats peuvent être synonymes de la présence d'une couronne comme sur la plaquette 7 (lignes verticales et horizontales en bord de plaquette) mais ils peuvent aussi témoigner d'un réticule défectueux. La signature spatiale de la plaquette 8 est identique à celle de la plaquette 7 accentuée par une proportion de puces défaillantes plus importante. Les résultats ne sont pourtant pas similaires et les effets pointés le sont dans de plus faibles proportions. Les plaquettes 9, 10 et 11 sont victimes d'un réticule plus ou moins défectueux. Cela se traduit par des lignes verticales de puces défaillantes plus ou moins discontinues selon l'état du réticule. La valeur élevée de l'estimation $\hat{\mu}$ des plaquettes 9 et 11 témoignent bien de la présence d'un effet vertical. En revanche, pour la plaquette 10, les estimations ne permettent pas de conclure à la présence d'un effet vertical. Les plaquettes 12, 13 et 14 ont été victimes à la fois, d'un problème de réticule défectueux et d'un dysfonctionnement du procédé de fabrication qui génère des regroupements particuliers de puces défaillantes. Pour la plaquette 12, seul l'effet vertical est identifié pour caractériser sa signature spatiale. Les plaquettes 13 et 14 ont des signatures spatiales très proches, pourtant leurs résultats sont différents. En effet, la présence d'une tâche sur la plaquette 13 est identifiée par les valeurs élevées des trois interactions doubles ξ , θ et ι . En revanche, l'interprétation des résultats de la plaquette

14 est beaucoup plus difficile, même si l'effet estimé le plus important correspond à l'effet vertical. Les plaquettes 15 et 16 ont été créées pour évaluer la capacité des méthodes à détecter des lignes de puces défaillantes sur la plaquette. On constate que sur ces deux exemples, cette méthode a quelques difficultés. En effet, les estimations du paramètre α sont d'une part les plus élevées de tout l'échantillon et d'autre part représentent les effets estimés les plus importants de chaque plaquette. On devrait donc en déduire l'absence d'une dépendance spatiale particulière sur les plaquettes. Pourtant, certaines interactions doubles ont aussi une valeur estimée élevée. De plus, dans les deux cas, l'estimation du paramètre λ ne permet pas de conclure à la présence de lignes horizontales de puces défaillantes.

TABLE 2

Valeurs estimées des 10 paramètres canoniques. Les paramètres en gras indiquent les effets estimés les plus importants numériquement selon les valeurs estimées de chaque plaquette

	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\mu}$	$\hat{\nu}$	$\hat{\pi}$	$\hat{\eta}$	$\hat{\xi}$	$\hat{\theta}$	$\hat{\iota}$	$\hat{\kappa}$
P1	-1.39	-0.71	-0.25	-0.27	-0.21	0.16	0.46	2.34	-37	-31
P2	0.39	-0.72	-0.63	-0.57	-0.26	0.28	0.03	1.19	0.6	-1.28
P3	-0.25	0.27	0.68	0.04	-0.07	-0.03	-0.38	0.17	-0.25	-0.19
P4	-1.23	-0.38	0.30	-1.68	0.60	-0.09	-13	2.65	0.47	13
P5	-3.66	1.24	1.85	0.18	-1.35	1.98	1.28	0.23	-0.44	-1.92
P6	-1.70	0.11	0.52	-0.26	0.25	0.03	0.63	0.87	0.32	-0.74
P7	-3.69	1.45	2.15	-0.11	0.16	0.23	-0.6	0.86	-0.64	-2.38
P8	-1.65	0.59	0.28	-0.15	0.49	-0.51	-0.36	0.33	-0.33	0.97
P9	-2.07	0.06	1.27	-1.12	-0.03	-1.11	0.77	0.59	0.34	0.27
P10	-0.67	-0.19	-0.04	0.19	-0.34	0.81	0.22	0.07	0.52	-0.75
P11	-0.81	-0.06	0.83	0.07	0.41	0.42	-1.29	-0.06	0.16	0.44
P12	-0.25	-0.03	0.96	-0.56	0.58	-0.8	-0.6	0.53	-0.14	0.59
P13	-1.11	0.08	0.21	-0.45	0.07	0.21	0.66	0.80	0.72	-1.21
P14	-1.35	0.71	0.98	-0.05	0.66	-0.42	-0.7	-0.55	0.1	0.74
P15	1.68	-0.31	-1.86	-1.51	-1.25	-0.45	1.44	0.66	1.17	-2.28
P16	1.77	-0.04	-0.94	-1.01	-0.35	-0.42	0.48	0.18	1.19	-1

En conclusion, on peut donc retenir que cette méthode présente plusieurs limites. Elle est certes capable d'identifier certaines signatures spatiales mais lorsque celles-ci sont aggravées par une proportion de puces défaillantes plus importante, les résultats ne convergent pas toujours. Elle ne permet pas d'identifier une plaquette où les puces défaillantes sont dispersées aléatoirement sur la plaquette. Elle présente aussi certaines difficultés pour détecter des lignes de puces défaillantes répétées sur

la plaquette. Enfin, elle ne permet pas toujours de conclure à une seule signature spatiale et l'interprétation des résultats est souvent difficile. En fait, l'estimation des paramètres canoniques ne semble pas adaptée à une interprétation individuelle des résultats mais plutôt à une classification globale des plaquettes, par exemple à l'aide d'un réseau de neurones comme proposent Baron *et al.* (2001). En effet, il n'y a pas de critère statistique qui permette de juger de l'importance d'un effet par rapport à un autre. Le choix des paramètres en gras dans la table 2 reflète donc une interprétation, certes objective, mais personnelle.

3.2.2. Modélisation du rendement selon un modèle de Poisson

TABLE 3
Résultats de la modélisation du rendement selon un modèle de Poisson

	Y	Y_r	Y_s
P1	84 %	86 %	98 %
P2	64 %	64 %	> 100 %
P3	46 %	56 %	82 %
P4	81 %	87 %	93 %
P5	76 %	> 100 %	35 %
P6	46 %	84 %	54 %
P7	80 %	99 %	81 %
P8	57 %	75 %	76 %
P9	85 %	87 %	98 %
P10	54 %	92 %	58 %
P11	38 %	76 %	50 %
P12	37 %	100 %	37 %
P13	48 %	52 %	91 %
P14	44 %	81 %	54 %
P15	66 %	71 %	93 %
P16	49 %	49 %	> 100 %

La Table 3 donne la décomposition du rendement total en rendements aléatoire et systématique de chaque plaquette. Les plaquettes 1, 2 et 3 représentent une dispersion aléatoire des puces défailtantes. En conséquence, le rendement aléatoire doit être très proche du rendement total et le rendement systématique proche de 100 %. C'est effectivement le cas pour la plaquette 1. Pour la plaquette 2, l'estimation du rendement systématique engendre un rendement supérieur à 100 % ce qui conduit à fixer arbitrairement le rendement aléatoire au rendement total alors que l'estimation lui

donne la valeur 53 %. Le rendement systématique de la plaquette 3 est le plus faible des trois plaquettes, sans doute en raison du nombre important de puces défailantes et de la difficulté à distinguer les regroupements dus au hasard, des regroupements provoqués par un dysfonctionnement. Toutefois, le rendement aléatoire de cette plaquette reste largement plus faible que le rendement systématique et révèle donc bien la présence d'un mode de défaillance majoritairement aléatoire. La décomposition relative à la plaquette 4 montre un rendement aléatoire plus faible que le rendement systématique relativement élevé (93 %). Cette décomposition ne permet donc pas d'alerter sur la présence d'un regroupement particulier de puces défailantes au centre de la plaquette. Le rendement systématique de la plaquette 5 est si faible (35 %) que l'estimation du rendement aléatoire est supérieur à 100 %. La plaquette est certes essentiellement composée de regroupements de puces défailantes (tâche au centre et couronne au bord), mais cela ne justifie pas un rendement systématique plus faible que le rendement total. La plaquette 6 a la même signature spatiale que la plaquette 5 accentuée par un nombre de puces défailantes plus important. Dans ce cas, le rendement systématique est estimé à 54 %. Ce résultat est plus cohérent que celui de la plaquette 5. La décomposition du rendement de la plaquette 7 permet de conclure à la présence de regroupements de puces défailantes sur la plaquette avec un rendement aléatoire estimé à 99 % et un rendement systématique estimé à seulement 81 %. La plaquette 8 a une signature spatiale identique à celle de la plaquette 7, mais elle est accentuée par un plus grand nombre de puces défailantes. Les rendements aléatoires et systématiques estimés sont très proches. Cela suggère la présence simultanée sur la plaquette de puces défailantes regroupées et de puces défailantes aléatoirement dispersées. Les plaquettes 9, 10 et 11 représentent une signature spatiale issue d'un réticule plus ou moins défectueux. Pour la plaquette 9 qui a le rendement total le plus élevé des trois, l'aspect aléatoire est privilégié sur l'aspect systématique. En revanche, pour les deux autres plaquettes, l'effet systématique est prépondérant avec des rendements systématiques inférieurs à 60 %. Les résultats de la plaquette 9 ne sont pas surprenants, puisque bien qu'ordonnées régulièrement les puces défailantes sont dispersées sur la plaquette. Pour les deux autres plaquettes, la proportion plus importante de puces défailantes semble influencer les résultats et le faible rendement systématique relève plus de la présence massive de puces défailantes que d'une influence du réticule défectueux. La décomposition du rendement de la plaquette 12 est conforme aux attentes avec un rendement systématique largement plus faible (37 %) que le rendement aléatoire (100 %). Toutefois, il semble abusif de considérer qu'aucune puce ne soit disposée sur la plaquette de manière aléatoire. Les plaquettes 13 et 14 ont des signatures spatiales très similaires, pourtant leur décomposition sont complètement opposées : l'une met l'accent sur l'aspect aléatoire avec un rendement aléatoire de 52 % et l'autre met l'accent sur l'aspect systématique avec un rendement systématique de 54 %. Les résultats des plaquettes 15 et 16 sont similaires : dans les deux cas, le regroupement des puces défailantes selon des lignes horizontales est interprété comme provenant d'un mode de défaillance aléatoire.

En conclusion, cette méthode donne globalement des résultats satisfaisants. Toutefois, elle apparaît assez instable dans plusieurs cas de figures, notamment pour les plaquettes 13 et 14. Par ailleurs, elle a aussi tendance à accentuer l'aspect systématique au détriment de l'aspect aléatoire dès que le rendement total est faible. Enfin, elle considère que des regroupements en ligne de puces défailantes sont issus d'un mécanisme de défaillance aléatoire.

3.2.3. Modélisation par un processus spatial

La Table 4 donne la décomposition du rendement total en rendements aléatoire et systématique de chaque plaquette. La colonne Indép. (pour indépendance) est le résultat du test statistique de l'hypothèse d'indépendance spatiale appliqué sur la plaquette initiale («non» si l'hypothèse est rejetée et «oui» sinon). Le niveau de test est fixé arbitrairement à 0.05. La Figure 9 représente les décompositions des plaquettes.

TABLE 4
Résultats de la modélisation par un processus spatial

	Y	Nbre Défauts Total	Indép.	Y _r	Y _s	Nbre Défauts Groupés	Seuil = 0.6		
							Y _r	Y _s	Nbre Défauts Groupés
P1	84 %	73	Oui	84 %	100 %	0	.	.	.
P2	64 %	163	Oui	64 %	100 %	0	.	.	.
P3	46 %	242	Oui	46 %	100 %	0	.	.	.
P4	81 %	86	Non	82 %	99 %	6	84 %	96 %	16
P5	76 %	107	Non	93 %	82 %	81	82 %	93 %	33
P6	46 %	242	Non	64 %	72 %	126	71 %	65 %	157
P7	80 %	89	Non	94 %	85 %	67	84 %	95 %	20
P8	57 %	195	Non	71 %	80 %	89	76 %	75 %	113
P9	85 %	68	Non	86 %	99 %	4	85 %	100 %	0
P10	54 %	208	Non	54 %	98 %	7	64 %	84 %	70
P11	38 %	280	Non	40 %	93 %	32	61 %	62 %	172
P12	37 %	281	Non	43 %	87 %	56	61 %	61 %	176
P13	48 %	235	Non	54 %	89 %	51	66 %	72 %	126
P14	44 %	252	Non	48 %	90 %	43	65 %	67 %	147
P15	66 %	154	Oui	66 %	100 %	0	.	.	.
P16	49 %	229	Oui	49 %	100 %	0	.	.	.

D'après leurs expérimentations, Hansen *et al.* (1997b) jugent que fixer arbitrairement à 0.6, le seuil à partir duquel une puce est censée faire partie d'un regroupement, permet, d'une manière générale, une bonne estimation des sous-plaquettes aléatoires et systématiques. Compte tenu de la structure de voisinage choisie (Figure 3), cela revient à supposer qu'une puce fait partie d'un regroupement de puces défaillantes si au moins cinq de ses puces voisines sont défaillantes. La Table 4 permet de comparer les résultats obtenus en fixant arbitrairement le seuil à 0.6 et les résultats obtenus

en laissant l'algorithme fixer le seuil à l'aide du test d'indépendance spatiale. L'hypothèse d'indépendance n'est pas rejetée pour les plaquettes 1, 2 et 3. Ce résultat est cohérent avec la signature spatiale de ces plaquettes qui ont été créées artificiellement pour représenter une dispersion aléatoire des puces défailtantes. Dans ce cas, la sous-plaquette aléatoire est confondue avec la plaquette initiale et le rendement systématique est évalué à 100 %. L'hypothèse d'indépendance spatiale n'est pas rejetée non plus pour les plaquettes 15 et 16. Ce non-rejet s'explique par la structure de voisinage qui ne permet pas de considérer des lignes de puces défailtantes comme des regroupements systématiques. En choisissant une structure de voisinage basée seulement sur les deux puces voisines (gauche et droite pour les lignes horizontales, en haut et en bas pour les lignes verticales), ce type de signature spatiale est mieux identifié.

Pour les plaquettes 4 à 14, l'hypothèse d'indépendance spatiale est rejetée. Ce résultat est conforme à leur signature spatiale qui témoigne d'une dépendance spatiale entre plusieurs puces défailtantes. Le choix du seuil à 0.6 a tendance à accroître l'estimation du nombre de puces défailtantes faisant partie d'un regroupement. Le rendement systématique en est donc affecté avec des valeurs estimées plus faibles. Comparées à celles de la méthode précédente, les estimations des rendements sont relativement stables pour les plaquettes qui ont des signatures similaires. C'est le cas notamment pour les plaquettes 13 et 14. Par ailleurs, les estimations des rendements sont globalement conformes à leur cartographie. Toutefois, les plaquettes 4, 5, 7 et 9 qui se distinguent des autres par des rendements relativement élevés, présentent certaines particularités. Ainsi, on observe que pour les plaquettes 5 et 7, le nombre estimé de puces défailtantes faisant partie d'un regroupement est plus faible lorsque le seuil est fixé à 0.6. Cela signifie que l'algorithme conduit au choix d'un seuil plus bas. Pour la plaquette 4, bien que la présence d'une dépendance spatiale soit acceptée, peu de puces défailtantes sont identifiées comme faisant partie d'un regroupement. Ces 3 plaquettes (4, 5 et 7) traduisent les difficultés de la méthode à traiter le cas de plaquettes avec des rendements élevés. Dans ce contexte, on observe que plus la proportion de puces défailtantes aléatoirement dispersées sur la plaquette est faible, plus le seuil estimé par l'algorithme est faible et plus les regroupements sont estimés dans leur quasi-globalité (plaquettes 5 et 7). En revanche, plus cette proportion est élevée, plus les regroupements de puces défailtantes sont sous-estimés (plaquette 4). La plaquette 9, avec un rendement systématique relativement élevé (> 90%), illustre la difficulté de la méthode à identifier des puces défailtantes régulièrement positionnées sur la plaquette même si l'hypothèse d'indépendance spatiale est rejetée.

En conclusion, cette méthode donne donc globalement de bons résultats, en outre ils sont stables pour les plaquettes qui ont des cartographies semblables. Toutefois, le choix du seuil apparaît déterminant puisque les résultats obtenus sont très différents pour certaines plaquettes. En choisissant un seuil de 0.6, on peut s'inquiéter d'une sur-estimation des regroupements de puces défailtantes. La figure 9 permet de comparer à partir de la plaquette initiale, les sous-plaquettes systématiques obtenues avec le seuil de 0.6 et le seuil fixé par l'algorithme. Excepté pour les plaquettes 5, 7 et 9, le seuil de 0.6 conduit à des regroupements plus importants de puces défailtantes. Toutefois, ces regroupements ne semblent pas témoigner d'une sur-estimation. Concernant la plaquette 9, la différence entre les deux seuils est seulement de 4 puces. En revanche, la différence est plus remarquable pour les plaquettes 5 et 7 où la couronne de bord n'est estimée qu'en partie. Choisir un seuil de 0.6 semble donc

présenter le risque de sous-estimer les regroupements de puces défailtantes dans le cas de plaquettes avec un rendement élevé et un faible nombre de puces défailtantes dispersées. Néanmoins, la plus grande limite de cette méthode semble plutôt être son incapacité à identifier un mécanisme de défaillance générant des puces défailtantes de manière régulière sur la plaquette (plaquette 9 ou de manière plus accentuée, plaquettes 15 et 16). Dans ce cas, la modification du voisinage permet tout de même d'améliorer l'estimation.

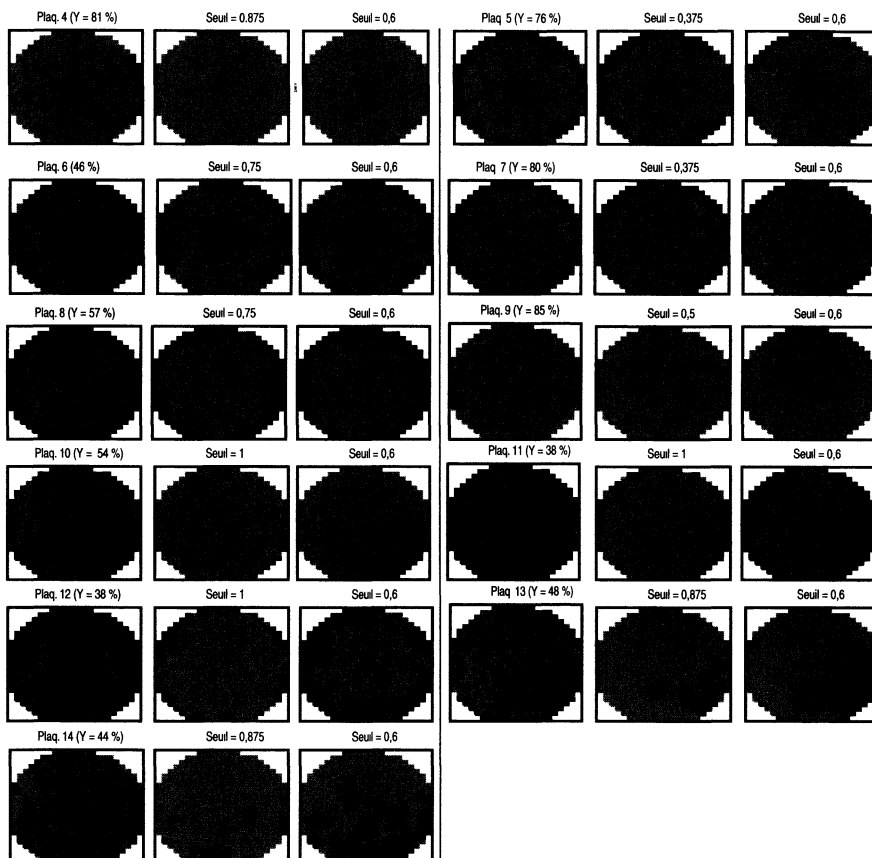


FIGURE 9

Décomposition des plaquettes 4 à 14 selon le seuil fixé par l'algorithme (deuxième plaquette) et un seuil de 0.6 (troisième plaquette)

3.2.4. Modélisation par arbre de classification

Avant de procéder à la modélisation des plaquettes par un arbre de classification, Duvivier (1999) préconise de calculer l'indice RR de présence d'une dépendance spatiale sur la plaquette. La Table 5 donne pour chaque plaquette la valeur de l'indice

calculé à partir des deux méthodes de décomposition du rendement. La colonne Indice 1 est obtenue à partir de la décomposition du rendement selon un modèle de Poisson. La colonne Indice 2 est obtenue à partir de la décomposition issue de la modélisation par un processus spatial. Le seuil est déterminé par l'algorithme de choix basé sur le test d'indépendance spatiale. La colonne Indice 3 est aussi issue de la décomposition selon la modélisation par un processus spatial mais avec un seuil fixé arbitrairement à 0.6.

TABLE 5

Calcul de l'indice de présence d'une dépendance spatiale. L'indice 1 correspond à la modélisation du rendement selon un modèle de Poisson. L'indice 2 correspond à la modélisation selon un processus spatial avec le seuil fixé par l'algorithme. L'indice 3 correspond à la modélisation selon un processus spatial avec le seuil fixé à 0.6. Plus l'indice est proche de 1, plus l'indépendance spatiale est probable

	Y	Indice 1	Indice 2	Indice 3
P1	84 %	0.88	1	1
P2	64 %	1	1	1
P3	46 %	0.74	1	1
P4	81 %	0.65	0.95	0.81
P5	76 %	.	0.28	0.73
P6	46 %	0.21	0.57	0.45
P7	80 %	0.056	0.27	0.77
P8	57 %	0.51	0.60	0.49
P9	85 %	0.88	0.94	1
P10	54 %	0.12	0.97	0.71
P11	38 %	0.29	0.92	0.51
P12	37 %	0	0.85	0.50
P13	48 %	0.87	0.84	0.55
P14	44 %	0.25	0.87	0.51
P15	66 %	0.82	1	1
P16	49 %	1	1	1

Sur l'ensemble de l'échantillon, les valeurs de l'indice sont différentes selon la méthode de calcul. En outre, les difficultés ou différences observées entre les méthodes se reportent sur ce calcul. Ainsi, compte tenu des difficultés à estimer correctement les rendements aléatoires et systématiques de la méthode basée sur le modèle de Poisson, l'indice 1 n'est pas toujours calculable (pour la plaquette 5, le résultat est négatif). Par ailleurs, il est relativement instable pour des plaquettes qui

ont des signatures semblables (plaquettes 13 et 14). Les résultats de l'autre méthode de calcul sont beaucoup plus stables. En revanche, on note des écarts importants selon le choix du seuil. En particulier pour les plaquettes 5, 7, 8, 11, 12, 13 et 14, l'écart est si important qu'il peut conduire à des interprétations différentes. Le choix de l'estimation des rendements aléatoires et systématiques apparaît donc déterminant pour une bonne estimation de l'indice de présence d'une dépendance spatiale sur la plaquette. Par ailleurs, si on analyse les résultats à partir de l'indice 3, on constate qu'il n'est pas toujours facile de conclure à la présence d'une dépendance spatiale. Par exemple, pour les plaquettes 4 et 10, peut-on conclure à une dépendance spatiale avec un indice au delà de 0.7, sachant qu'un indice proche de 1 est censé traduire une indépendance spatiale? L'application du test statistique d'indépendance spatiale semble donc beaucoup plus fiable.

TABLE 6

Résultats de la modélisation par arbre de classification. La première colonne Nbre Défauts représente le nombre total de défauts sur la plaquette. Les autres colonnes Nbre Défauts représentent le nombre de défauts classifiés par les divisions de l'arbre selon les variables décrites par les colonnes Var

	Nbre Défauts	Première Division		Deuxième Division		Conclusion sur la Signature spatiale
		Var	Nbre Défauts	Var	Nbre Défauts	
P4	86	anneau = 1	15	reticule	44	tâche centrale + reticule defectueux
P5	107	anneau = 1 et 6	82	anneau = 3, 4 et 5	16	tâche centrale + grande couronne
P6	242	anneau = 1,2 et 6	131	anneau = 3 et 5	93	grande tâche centrale + grande couronne
P7	89	anneau = 6	69	anneau = 3, 4 et 5	20	couronne
P8	195	anneau = 6	98	diex < 20.5	87	couronne
P9	68	reticule	48	reticule	19	reticule defectueux
P10	208	angle > 264°	81	reticule	86	cadran Sud-Est + reticule defectueux
P11	280	reticule	61	anneau = 5 et 6	77	reticule defectueux + grande couronne
P12	281	anneau = 5 et 6	178	reticule	71	grande couronne + reticule defectueux
P13	235	angle > 246°	113	reticule	83	cadran Sud-Est + reticule defectueux
P14	252	angle > 196°	157	reticule	67	cadran Sud + reticule defectueux
P15	154	reticule	102	reticule	52	reticule defectueux
P16	229	reticule	102	reticule	127	reticule defectueux

L'application d'un arbre de classification dans ce contexte a pour objet de déterminer la ou les variables ainsi que leurs valeurs qui permettent d'identifier le positionnement des puces défailtantes sur la plaquette. Dans la plupart des cas, il s'agit ainsi d'identifier la structure spatiale des regroupements de puces défailtantes. Il peut s'agir aussi d'identifier un réticule défectueux. Dans ce cas, les puces défailtantes ne sont pas nécessairement regroupées. Elles sont plutôt disposées de manière régulière sur la plaquette et selon la qualité du réticule, elles peuvent être regroupées sous forme de lignes plus ou moins discontinues. Ce type de défaillance est difficilement identifié par les méthodes précédentes. Toutefois, plus les puces défailtantes sont dispersées sur la plaquette, plus l'arbre a aussi des difficultés à identifier un mode de défaillance particulier. Dans ce cas, il peut même avoir tendance à pointer à tort certains mécanismes de défaillance, comme par exemple, un réticule défectueux. La table 6 résume les résultats obtenus à l'issue des arbres de classification construits pour les plaquettes 4 à 16. Seuls les deux premiers niveaux de l'arbre sont représentés. La Figure 10 est un exemple d'arbre de classification.

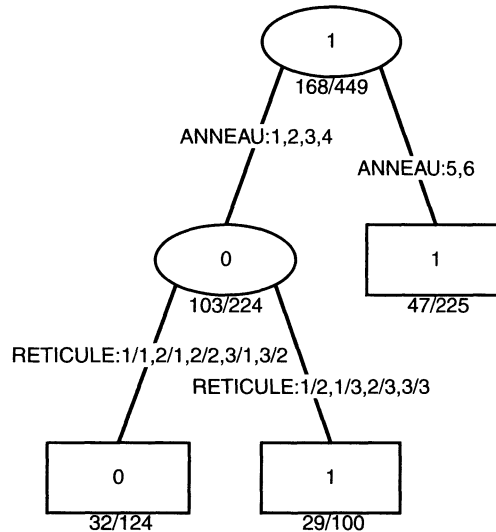


FIGURE 10

Arbre de classification élagué de la plaquette 12. La plaquette comprend 449 puces. Les puces fonctionnelles sont codées 0 et les puces défailtantes sont codées 1. Chaque noeud de branche (ellipse) ou chaque feuille (rectangle) indique le type de puces majoritaires (0 ou 1). En dessous est renseigné le taux de mal classés par rapport au type majoritaire. Ainsi, 178 puces défailtantes ont été identifiées sur les anneaux 5 et 6 et 71 autres aux positions 1/2, 1/3, 2/3 et 3/3 du réticule

Afin d'identifier les divisions de l'arbre les plus pertinentes, il paraît judicieux de s'appuyer sur les résultats de la modélisation selon un processus spatial. Par exemple, si l'hypothèse d'indépendance spatiale n'est pas rejetée, la construction de l'arbre

n'est pas souhaitée. C'est le cas des plaquettes 1, 2 et 3. Toutefois, il faut aussi s'assurer que les puces défaillantes ne sont pas regroupées sous la forme de ligne en prenant le voisinage adéquat pour tester la présence d'une telle dépendance spatiale. C'est le cas par exemple, des plaquettes 15 et 16. Par ailleurs, l'estimation selon la modélisation par un processus spatial, du nombre de puces défaillantes faisant partie d'un regroupement permet de juger de la pertinence des divisions successives de l'arbre. Par exemple, pour la plaquette 4, la première division de l'arbre classe 15 puces défaillantes au centre de la plaquette. Or, le nombre de puces défaillantes faisant partie d'un regroupement a été précédemment estimé à 16 pour un seuil de 0.6. Il en est de même avec la plaquette 5 où 82 puces sont identifiées comme faisant partie soit d'une tâche au centre de la plaquette soit d'une couronne en bord, alors que le nombre de puces défaillantes faisant partie d'un regroupement a été estimé au plus à 81. Pour ces deux plaquettes, les sous-divisions de l'arbre apparaissent donc superflues. Par ailleurs, lorsque les divisions successives d'un arbre sont construites à partir de la même variable, la pertinence des divisions inférieures peut être aussi remise en cause. Par exemple, on peut se demander pourquoi les anneaux proposés dans les deuxièmes divisions des plaquettes 6 et 7 n'ont pas été pris en compte dans la première division si elles témoignent d'une réelle défaillance. La plaquette 8 a aussi une deuxième division peu pertinente puisqu'elle identifie les puces dont l'abscisse est inférieure à 20.5 alors que l'abscisse maximale est de 22 et que le dernier anneau de la plaquette a été identifié par la première division. Les plaquettes 9 à 11 ont été choisies pour représenter des exemples de cartographie dans le cas d'un réseau déficient. Cet effet réseau est clairement identifié pour la plaquette 9 avec une redondance de la variable réseau aux deux premières divisions de l'arbre. Pour la plaquette 11, l'effet réseau est identifié par la première variable de division avec 61 puces défaillantes classifiées selon cette défaillance. La deuxième variable de division témoigne d'un effet couronne en bord de plaquette regroupant 77 puces défaillantes. Sachant que la modélisation par un processus spatial n'est pas vraiment adaptée pour identifier les puces défaillantes issues d'un réseau déficient, on peut considérer que ces 77 puces défaillantes font partie des 172 (seuil = 0.6) puces défaillantes identifiées comme faisant partie d'un regroupement. Cette division apparaît donc pertinente. Pour la plaquette 10, l'effet réseau n'est démontré que par la deuxième division de l'arbre. Auparavant, l'accent est mis sur le cadran Sud-Est de la plaquette qui représente une zone de regroupement massif de puces défaillantes avec 81 puces. Ce nombre est proche de l'estimation du nombre de puces faisant partie d'un regroupement (70 puces pour un seuil de 0.6). Les plaquettes 12, 13 et 14 représentent l'interaction d'un dysfonctionnement du procédé générant un ou des regroupements massifs de puces défaillantes avec un réseau déficient qui a plutôt tendance à disperser les puces défaillantes. Pour les trois plaquettes, cette interaction est clairement mise en valeur par les deux premières divisions de l'arbre : l'accent est mis en priorité sur les regroupements massifs de puces défaillantes, l'effet réseau n'étant pointé qu'en deuxième position. Enfin, les regroupements de puces défaillantes selon des lignes plus ou moins discontinues (plaquettes 15 et 16) sont bien identifiés à l'aide de la variable réseau. Plus précisément, les 102 puces défaillantes générées manuellement en choisissant deux positions déficientes du réseau sont toutes identifiées par la première division de l'arbre.

En conclusion, cette approche par arbre de classification semble donc pertinente pour l'identification des signatures spatiales des plaquettes. Dans l'ensemble, les deux premières divisions apparaissent largement suffisantes pour classer la plupart des puces défailtantes, la deuxième division apparaissant même superflue dans certains cas.

4. Discussion

L'analyse des différentes techniques statistiques proposées dans cet article est motivée par une application concrète dans une unité de fabrication de circuits intégrés. Il est donc nécessaire de proposer un outil à la fois robuste et facile d'application. Au vu des résultats obtenus par les différentes techniques, la modélisation du rendement par un processus spatial semble être la méthode la plus robuste pour décomposer le rendement en des rendements systématique et aléatoire. Cette décomposition permet de hiérarchiser les actions correctives à mettre en place : un faible rendement systématique traduit un dysfonctionnement du procédé alors qu'un faible rendement aléatoire traduit plutôt une concentration en particules trop importante. Un rendement systématique faible se traduit souvent sur la plaquette par une signature spatiale particulière. Caractériser cette signature permet d'accélérer le processus d'amélioration du rendement : selon la position des regroupements de puces défailtantes sur la plaquette, l'expertise des ingénieurs permet de cibler les dysfonctionnements du procédé qui en sont à l'origine. Baron *et al.* (2001) proposent d'utiliser un réseau de neurones multi-couches pour classer les plaquettes à partir des paramètres canoniques de la modélisation markovienne. Malgré l'efficacité démontrée des réseaux de neurones dans plusieurs applications, leur complexité ainsi que les difficultés rencontrées pour interpréter les paramètres canoniques ne sont pas favorables à l'utilisation d'une telle méthode. En revanche, l'utilisation conjointe d'un arbre de classification et des résultats de la modélisation par un processus spatial (test d'indépendance spatiale et estimation du nombre de puces défailtantes affectées à un regroupement) se révèle particulièrement efficace pour identifier une grande majorité de signatures spatiales issues de différents mécanismes de défaillance.

Références

- BARON M., LAKSHMINARAYAN C. K. et CHEN Z. (2001), Markov Random Fields in Pattern Recognition for Semiconductor Manufacturing. *Technometrics*, vol. 43, n° 1, pp. 66-72.
- BESAG J. (1974), Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Korean Statistical Society*, vol. 2, pp. 192-236.
- BREIMAN J., FRIEDMAN J., OLSHEN R. et STONE C. (1984), *Classification and Regression Trees*. Wadsworth & Brooks.
- CLIFF A.D., ORD J.K. (1981), *Spatial Processes, Models and Applications*, London: Pion Limited.

- COMETS F. (1992), On Consistency of a Class of Estimators for Exponential Families of Markov Random Fields on the Lattice. *The Annals of Statistics*, vol. 20, pp. 455-468.
- CRESSIE N. A. C. (1991), *Statistics for Spatial Data*. New York, Wiley.
- DUVIVIER F. (1999), Automatic Detection of Spatial Signature on Wafermaps in a High Volume Production. *Internal Symposium on Defect and Fault Tolerance in VLSI Systems*, pp. 61-66.
- HANSEN M. H., NAIR V. N. et FRIEDMAN D. J. (1997), Monitoring Wafer Map Data From Integrated Circuit Fabrication Processes for Spatially Clustered Defects. *Technometrics*, vol. 39, n° 3, pp. 241-253.
- HANSEN M. H., NAIR V. N., JAMES D. A. et FRIEDMAN D. J. (1997b), Model-Free Estimation of Defect Clustering in Integrated Circuit Fabrication. *IEEE Transactions on Semiconductor Manufacturing*, vol. 10, n° 3, pp. 344-359.
- LONGTIN M. D., WEIN L. M., WELSCH R. E. (1996), Sequential Screening in Semiconductor Manufacturing, I: Exploiting Spatial Dependence. *Operations Research*, vol. 44, pp. 173-195.
- MORAN P. A. P. (1948), The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society, Series B*, vol. 10, pp. 243-251.
- PAZ O., LAWSON T. R., Jr (1977), Modification of Poisson statistics : Modeling defects induced by diffusion. *IEEE Journal of Solid-State Circuits*, vol. SC-12, pp. 540-546.
- STAPPER C. H., ARMSTRONG F. M., et SAJI K. (1983), Integrated Circuit Yield Statistics. *Proceeding of the IEEE*, vol. 71, n° 4, pp. 453-470.