

REVUE DE STATISTIQUE APPLIQUÉE

J. J. DAUDIN

F. SIDIBE

Construction et utilisation d'un modèle de simulation de contrôle de la qualité de melons

Revue de statistique appliquée, tome 33, n° 3 (1985), p. 15-27

http://www.numdam.org/item?id=RSA_1985__33_3_15_0

© Société française de statistique, 1985, 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/>

CONSTRUCTION ET UTILISATION D'UN MODÈLE DE SIMULATION DE CONTRÔLE DE LA QUALITÉ DE MELONS

J.J. DAUDIN

*Département de mathématiques et informatique,
Institut National Agronomique Paris Grignon,
16, rue Claude Bernard 75005 Paris*

F. SIDIBE

Société Pomona, 41, rue du Pont Neuf 75004 Paris

RÉSUMÉ

La simulation permet de connaître les caractéristiques de contrôles trop complexes pour être calculés directement, et ainsi de choisir la forme de tri la mieux adaptée, tout en évitant des essais en vraie grandeur, trop coûteux et ne donnant de résultats que pour une seule valeur des paramètres caractérisant la production que l'on cherche à contrôler. Le problème essentiel consiste à trouver un modèle correct permettant de simuler cette production; ce modèle est ici basé sur une analyse des sources de variabilité de la production.

SUMMARY

In some cases, quality control characteristics are complex, and direct mathematical derivation is not possible. In these circumstances, simulation may be worthily substituted to direct experimentation. The main point consists in the construction and the validation of a model of simulation. In this paper, data on variability of melons quality is analysed with a hierarchical random effects anova model whose results are used to build the simulation model.

1. PRÉSENTATION DU PROBLÈME

Le melon est un produit de qualité très variable. Cette caractéristique nuit à son image et à sa vente, notamment à cause de la présence, dans la production courante, de melons de qualité très médiocre, pratiquement impropres à la consommation. Il est donc intéressant pour le développement de ce produit, de fournir au consommateur une catégorie supérieure, lui donnant la garantie d'une qualité dépassant un seuil minimum : tous les melons de cette catégorie doivent présenter un taux de sucre supérieur à une valeur donnée.

Le taux de sucre contenu dans le melon est en effet un paramètre essentiel qui caractérise la qualité : si un melon comportant un taux de sucre acceptable peut présenter d'autres caractéristiques intéressantes telles que l'odeur, le fruité, la jutosité, etc., par contre un melon ne dépassant pas le seuil de sucre minimum ne sera jamais jugé valable par le consommateur. Comme il s'agit ici d'éliminer les mauvais produits, le taux de sucre, mesuré par réfractométrie, est la caractéristique de qualité fondamentale et la seule que nous retenons ici.

Cette mesure se fait depuis peu par un procédé automatique qui utilise un micro prélèvement sur chaque melon, sans le rendre impropre à la consommation, et permet

d'obtenir en quelques secondes la valeur de l'indice réfractométrique de cet échantillon.

Il n'est cependant pas possible d'analyser toute la production apportée par les agriculteurs dans un centre de grossiste, car le volume excède la capacité de la machine de mesure automatique de l'indice réfractométrique. En pratique, seulement un quart de la production de melons peut passer dans le dispositif automatique de tri, constitué de trois appareils de mesures mis en parallèle.

D'autre part, il est nécessaire d'obtenir un tonnage minimum de melons de qualité supérieure, de façon à ne pas pénaliser sa vente par des ruptures de stock et des arrivages irréguliers dans les centres de commercialisation. Cette contrainte est difficile à respecter pour un produit comme le melon, dont la qualité varie beaucoup en fonction du temps; de plus les quantités apportées par les producteurs varient de 35 à 85 tonnes par jour. Le contrôle doit donc agir comme un filtre régulateur donnant à la sortie un tonnage stable d'environ 7 tonnes et une qualité garantie à partir d'un arrivage très irrégulier en qualité et en quantité.

Enfin il se peut qu'une machine automatique de mesure de l'indice réfractométrique tombe en panne, auquel cas le taux de tri passe de 25 % à 16,7 %. Il faut que le dispositif de contrôle assure son service dans cette éventualité, sans diminuer la quantité et la qualité produite.

Au regard des différentes contraintes indiquées ci-dessus, il existe plusieurs dispositifs de contrôle respectant tous un taux de tri de 25 % : les melons arrivent par lots de quelques palettes, elles-mêmes constituées de caisses contenant une à deux dizaines de melons. La décision d'affectation à la catégorie supérieure peut être faite à plusieurs niveaux : lot, palette, caisse ou melon. Les deux premiers niveaux impliquent une décision prise sans inspection exhaustive, avec un taux d'échantillonnage de 25 %, à la différence de la décision melon par melon qui permet de n'accepter que les melons effectivement mesurés. Par contre cette dernière méthode présente l'inconvénient de ne passer en revue qu'un quart de l'arrivage, au contraire des deux premières qui filtrent par échantillonnage la totalité de la production.

Le choix du type de contrôle est nécessairement basé sur une analyse de l'importance des sources de variabilité de la qualité du melon qui sont dues aux différences entre lots, entre palettes, entre caisses, à l'intérieur d'une caisse et enfin à l'erreur de mesure, qui est ici non négligeable. Il s'agit de rechercher la plus grande unité (lot, palette, caisse, melon) homogène; on prendra alors la décision d'affectation à ce niveau.

L'étude des sources de variabilité est présentée dans la section 2 de cet article. Cependant elle ne suffit pas pour répondre aux questions suivantes : le dispositif de contrôle est-il suffisamment souple pour répondre aux impératifs de quantité et de qualité exposés plus haut ? Quel est le comportement en cas de panne d'une machine ? Comment en modifier les paramètres pour que le tri s'adapte à la production présentée chaque jour par les agriculteurs ?

La simulation de la production et du fonctionnement du dispositif de contrôle permettent seuls de répondre à ces questions. Elles sont basées sur un modèle statistique fondé sur l'étude des sources de variabilité. Le modèle, ainsi que les principaux résultats de la simulation sont présentés dans la section 3.

La section 4 conclut cet article par la présentation d'éléments de validation du modèle de simulation, basés sur la comparaison de résultats donnés par cette dernière avec ceux d'un essai de contrôle en vraie grandeur, et par quelques réflexions sur l'intérêt d'une démarche statistique globale (de l'analyse de variance à la simulation), dans les problèmes de contrôle de la qualité.

Cet article aborde surtout l'aspect statistique du problème. Une présentation globale de l'étude du contrôle de la qualité des melons est faite par F. SIDIBE et J.J. DAUDIN (1985), que l'on consultera pour plus de détails.

2. ÉTUDE DES SOURCES DE VARIABILITÉ

a) Gradient de l'indice réfractométrique; erreur de mesure

La mesure automatique consiste à prélever un micro échantillon de jus de melon, par l'intermédiaire d'une aiguille, effectuant un mouvement de va et vient d'une profondeur de 3 à 4 cm, puis à en mesurer l'indice réfractométrique. La qualité de la mesure dépend de la représentativité de l'échantillon de jus prélevé. Pour juger de la justesse et de la variabilité de cette mesure, nous avons analysé des prélèvements de pulpe dans huit melons. Chaque melon a été coupé selon son plan équatorial et selon 2 plans verticaux perpendiculaires, passant par l'axe pédoncule-ombilic. Dans chaque plan, 12 fragments de pulpe ont été prélevés de part et d'autre de la cavité interne du fruit, selon 2 axes perpendiculaires et six prélèvements ont été faits sur l'axe pédoncule-ombilic, ainsi que le montre la figure 1.

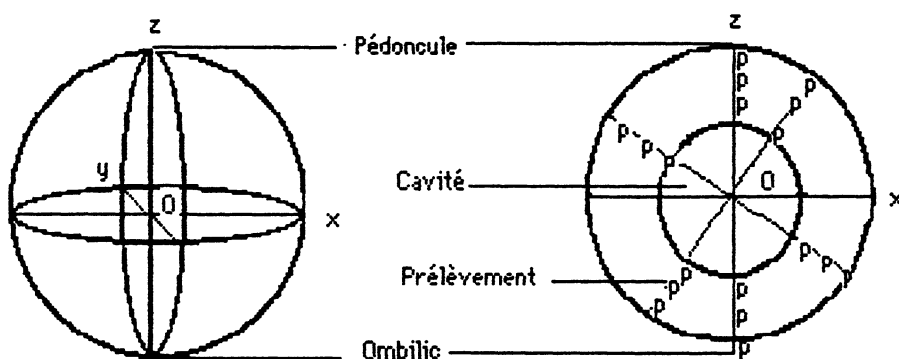


Figure 1. — Représentation des 18 prélèvements dans chaque melon.

Les résultats de ces mesures peuvent être schématisés dans des représentations graphiques, du type de celle de la figure 2 qui représente 2 plans de prélèvements dans lesquels les points de même indice réfractométrique sont réunis sur des courbes d'iso-indice réfractométrique, et ceci pour un melon parmi les huit analysés. La figure 2 montre une assez grande irrégularité de ces courbes d'iso-indice réfractométrique, ce qui entraîne une variabilité de toute mesure basée sur un prélèvement radial : selon le rayon choisi pour effectuer le prélèvement, on observe une variation du taux de sucre contenu dans ce dernier. De plus le gradient d'indice réfractométrique n'est pas uniforme, mais a tendance à comporter une variation plus rapide vers la cavité et également vers l'écorce.

Ces considérations, ainsi que la répartition empirique de mesures radiales sur le même melon conduisent à modéliser l'erreur de mesure par une loi uniforme de demi-étendue b , où le paramètre b varie entre 0.5 et 1.5 selon le type de melon (voir F. Sidibe, J.J. Daudin (1985)).

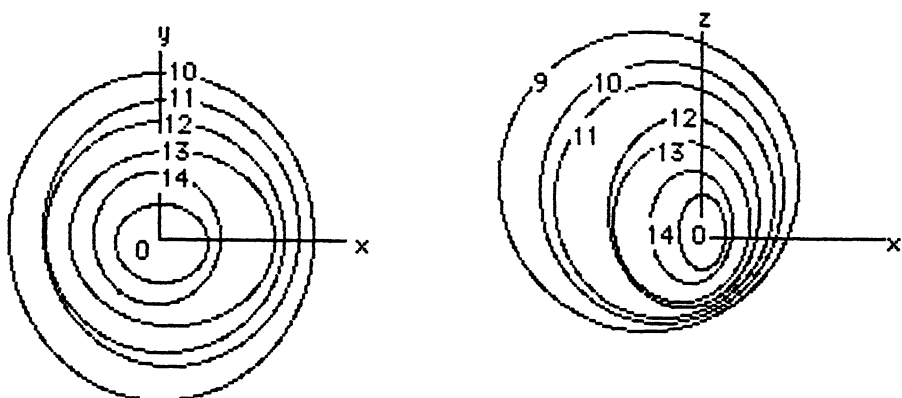


Figure 2. — Courbes d'iso-indice réfractométrique pour 2 plans équatoriaux.

b) Analyse des sources de variabilité intra-lots

Plusieurs indices montrent que les apports des producteurs ne sont pas parfaitement homogènes. Certaines parties de lots, telles que les palettes, les caisses où des groupes de caisses ont une qualité moyenne différente de celle du lot dans son ensemble. L'information recueillie auprès des producteurs et les observations au champ confirment la réalité de ce phénomène. En effet, l'organisation de la cueillette conduit à ce que des melons ayant des caractéristiques communes voisinent à l'intérieur des lots. Pour quantifier ces considérations, plusieurs échantillons, correspondant à différents lots, ont été prélevés. Chacun de ces 3 échantillons comporte les 15 melons de 2 caisses issues d'une même palette et ceci pour 2 palettes différentes. Chacun de ces échantillons contient donc 60 unités, ce qui fait, au total, 180 unités. Un modèle d'analyse de variance hiérarchique à effets aléatoires (SCHEFFÉ, 1967) :

$$Y_{ijkl} = \mu + A_i + B_{ij} + C_{ijk} + E_{ijkl}$$

où Y_{ijkl} est la valeur de l'indice pour le lot i , la palette j , la caisse k et le melon l ,

A_i , B_{ij} , C_{ijk} , E_{ijkl} sont respectivement les effets aléatoires du lot, de la palette, de la caisse et du melon et respectent les hypothèses classiques :

A_i , B_{ij} , C_{ijk} , E_{ijkl} sont des variables aléatoires d'espérance nulle, indépendantes et de variance respective γ_A , γ_B , γ_C , γ_E ,

permet d'obtenir la table d'analyse de variance suivante (tableau 1) :

TABLEAU 1
Analyse de variance

Source	Somme de Carrés	Degrés de Liberté	Carré moyen	E (CM)	Fisher
Lot	198.1	2	99.1	$\gamma_E + 15\gamma_C + 30\gamma_B + 60\gamma_A$	4.1
Palette	72.8	3	24.3	$\gamma_E + 15\gamma_C + 30\gamma_B$	1.5
Caisse	95.8	6	16.0	$\gamma_E + 15\gamma_C$	5.0
Melon	532.6	168	3.2	γ_E	

Si les indices i, j, k et l varient respectivement de 1 à r, s, t et u , les espérances des carrés moyens sont données par les formules suivantes :

$$\begin{aligned} E(\Sigma (Y_{ijkl} - Y_{ijk})^2 / rst (u - 1)) &= \gamma_E \\ E(\Sigma (Y_{ijk} - Y_{ij.})^2 / rs (t - 1)) &= \gamma_E + u\gamma_C \\ E(\Sigma (Y_{ij.} - Y_{i..})^2 / r (s - 1)) &= \gamma_E + u\gamma_C + ut\gamma_B \\ E(\Sigma (Y_{i..} - Y_{...})^2 / (r - 1)) &= \gamma_E + u\gamma_C + ut\gamma_B + stu\gamma_A \end{aligned}$$

On a alors les estimateurs suivants pour les paramètres qui nous intéressent, avec $r = 3, s = 2, t = 2$ et $u = 15$:

$$\gamma_A = 1.2, \gamma_B = 0.3, \gamma_C = 0.9, \text{ et } \gamma_E = 3.2$$

La variance de l'effet lot est ici artificielle dans la mesure où les trois lots considérés n'ont pas été tirés au sort mais choisis pour représenter des cas assez différents; il est donc normal que cet effet soit important. Par contre, l'effet palette possède une faible variance. Nous avons procédé à d'autres analyses du même type dans des conditions différentes : melons de différentes variétés, cultivés en serre ou en plein champ... etc. Des analyses de variances semblables à celle présentée ci-dessus, nous ont permis d'obtenir d'autres estimations pour $\gamma_A, \gamma_B, \gamma_C$ et γ_E . A partir de cet ensemble de résultats, nous avons adopté de façon empirique les estimations moyennes suivantes :

Variance de l'erreur de mesure : 0.35
 Variance intra caisse : 2.5
 Variance intra palette : 3.5
 Variance intra lot : 3.8

Il faut noter que la variance intra-caisse varie beaucoup selon la qualité des lots considérés : elle peut aller de 1.5, pour des lots uniformément bons ou uniformément mauvais, à 3.5 pour des lots très hétérogènes. Il en est de même pour les variances des ensembles de niveaux supérieurs, la palette et le lot.

Ces résultats conduisent aux conclusions suivantes :

Les variances des ensembles de niveau supérieur à l'unité sont assez élevées; les décisions prises à des niveaux supérieurs risquent de conduire à une perte de précision du tri.

Si on veut choisir, pour des raisons de souplesse de tri, un niveau de décision supérieur à l'unité, la caisse semble être le meilleur choix, la variance intra-palette étant nettement supérieure à celle intra-caisse.

La variance intra-lot n'est pas de beaucoup supérieure à celle de la palette; la décision au niveau du lot ne peut pourtant pas être envisagée, car il est psychologiquement difficile de faire passer auprès des producteurs le rejet de l'ensemble des melons d'un lot.

A l'issue de l'étude de la décomposition de la variabilité, on est conduit à limiter le choix entre 2 méthodes de contrôle, celle du tri melon par melon et celle du tri caisse par caisse, basée sur un échantillon de la caisse. Cependant on ne possède que peu d'éléments pour choisir. Le critère de la précision favorise le premier type de tri mais il est difficile de chiffrer précisément l'écart. Le critère de la souplesse, face aux contraintes de stabilité de la quantité de melons mis en catégorie supérieure, favorise le second type de tri, mais, là encore, on ne peut pas avoir de résultats chiffrés. La simulation sur ordinateur de la production et des 2 types de contrôles va nous permettre de préciser ces 2 points fondamentaux.

3. SIMULATIONS; MODÈLE ET RÉSULTATS

a) Modélisation de la population de melons

Elle est basée sur la modélisation de lots de 3 000 melons. Chaque lot possède une moyenne μ ; un melon d'un lot de moyenne μ , a un indice réfractométrique réel X , donné par la formule suivante :

$$X = \mu + \gamma_1^{1/2} T_1 + (\gamma_2 + U_1)^{1/2} T_2$$

où γ_1 est la variance inter-caisse égale à $\gamma_C + \gamma_B$

γ_2 est la variance intra-caisse moyenne, égale à γ_E

T_1 est une variable aléatoire normale centrée réduite, de même que T_2 .

U_1 est une variable aléatoire distribuée selon une loi uniforme dans l'intervalle $[-a, +a]$.

Le modèle précédent donnant X , décrit la composition d'un lot formé de caisses de moyennes $\mu + \sqrt{\gamma_1} T_1$, elles-mêmes composées de melons distribués selon une loi normale, centrée autour de cette moyenne et de variance $\gamma_2 + U_1$. Cette dernière est constituée d'une partie fixe, γ_2 , et d'une partie aléatoire nécessitée par la variabilité de ce paramètre, telle qu'elle a été observée au cours des études précédentes.

La valeur mesurée de l'indice réfractométrique, XM , est la valeur réelle plus l'erreur de mesure, ce qui donne :

$$XM = X + U_2$$

où U_2 est une variable aléatoire distribuée selon une loi uniforme dans l'intervalle $[-b, +b]$.

Les valeurs types des paramètres sont celles données par l'étude de variabilité. D'autres valeurs peuvent leur être données pour étudier la robustesse des résultats du contrôle face à la variation de ces paramètres.

Enfin la population totale d'une journée est constituée de lots dont les moyennes sont réparties de façon variable selon les jours. Pour simplifier, on utilise 3 jours types :

- Les « gros jours », appelés ainsi car il s'agit de jours où les producteurs apportent un volume important de melons, la plupart de mauvaise qualité.
- Les « jours moyens », où tant la qualité que la quantité sont d'un niveau moyen.
- Les « bons jours », où les apports des producteurs sont bons en qualité et faibles en volume.

Cette typologie a été établie à l'aide des statistiques faites sur les melons de la campagne 1982. Il s'agit de jours types extrêmes, pour lesquels il est important de connaître le comportement de la procédure de contrôle. Les « gros jours » sont particulièrement sensibles : les conditions y sont défavorables à l'obtention de la quantité voulue de melons de catégorie supérieure. Ils constituent un bon test de la fiabilité du contrôle. Le tableau 2 donne le nombre de lots pour les différentes classes de moyenne et pour les 3 types de jours. Il faut enfin noter que la quantité minimum exigée en catégorie supérieure est, quels que soit les jours de 7 tonnes, soit environ 10 000 melons.

TABLEAU 2
Répartition des lots par tranches de moyenne

Moyenne de l'indice	≤ 9	9-10	10-11	11-12	12-13	≥ 13	total	total melons
Jours moyens	3	3	3	4	5	4	22	66000
Gros jour	17	3	8	5	3	1	37	111000
bon jour	1	2	2	3	7	4	19	57000

b) Modélisation des contrôles

Contrôle melon par melon : il consiste à simuler une proportion donnée de chaque lot (25 % ou 16,7 %), et à garder en qualité supérieure tous les melons dont l'indice réfractométrique mesuré, XM , dépasse un seuil donné S_1 .

Contrôle caisse par caisse : il consiste à simuler 6 ou 4 melons par caisse de 24 (chiffres correspondant aux taux de sondage de 25 % et 16,7 %) et à utiliser la règle suivante, basée sur la moyenne m et l'écart-type s des 6 (ou 4) valeurs mesurées (XM), de l'échantillon :

- acceptation en qualité supérieure de la caisse si $m - 2s \geq S_2$;
- refus de la caisse dans le cas contraire, mais acceptation individuelle des melons qui, parmi les 6 (ou 4) mesurés sont tels que $XM \geq S_3$.

On observe qu'il y a 2 seuils pour ce type de contrôle : un seuil pour le tri de la caisse entière et un autre pour le tri, individuel des melons de l'échantillon, dans le cas où la caisse n'est pas acceptée en qualité supérieure.

Le choix des seuils S_1 , pour le tri melon par melon, S_2 et S_3 pour le tri caisse par caisse, est déterminé en fonction de la population du jour, de façon à satisfaire les contraintes de quantité (environ 10 000 melons) et de qualité de la catégorie supérieure. Pour la simulation, plusieurs valeurs ont été utilisées, de façon à mieux connaître l'influence de ce choix sur les résultats des 2 types de contrôle.

c) Résultats des simulations

Le contrôle crée une population spécifique, dite de qualité supérieure, appelée population triée. On peut en calculer trois caractéristiques pour chaque type de tri : son tonnage ou sa taille, la proportion de bons melons, c'est-à-dire tels que leur indice dépasse 12, et celle de mauvais, dont l'indice est inférieur à 10.

Les 2 types de contrôles sont effectués sur la même population simulée, afin que la comparaison ne soit pas entachée d'une variabilité supplémentaire due au changement de population.

Les résultats d'une simulation, pour des valeurs fixées des paramètres caractérisant la population de melons et des seuils caractérisant les contrôles, sont présentés sous la forme des 3 histogrammes contenus dans les figures 3, 4 et 5. Elles représentent respectivement les distributions de la population totale, de celle sélectionnée par le tri melon par melon et caisse par caisse. Il s'agit d'un « gros jour »; les valeurs des paramètres et des seuils sont précisées dans la légende. Les seuils ont été déterminés par tâtonnements, de façon à obtenir un peu plus de 10 000 melons pour la population triée. On observe que la population triée par le contrôle caisse par caisse est de meilleure qualité que celle triée par le contrôle melon par melon. Ceci est confirmé par les 2 critères de qualité de la table 3, tant pour le pourcentage de mauvais melons que pour celui des bons, pour lequel la différence est très sensible.

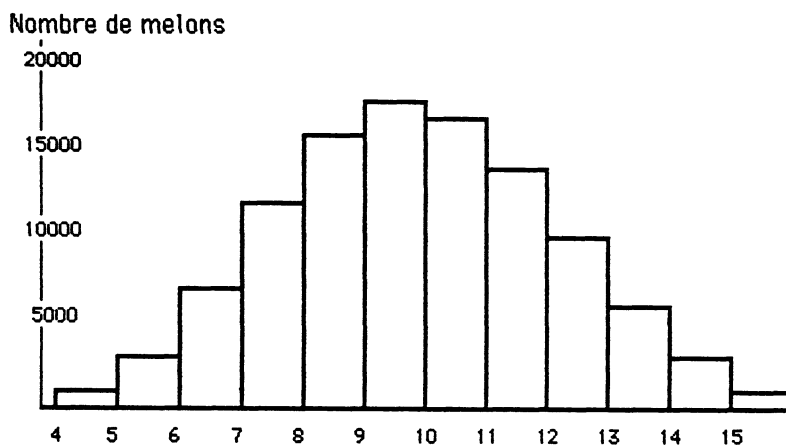


Figure 3. — Population totale d'un « gros jour » ($\gamma_1 = 1$ $\gamma_2 = 2,5$ $a = 1$ $b = 0,8$).

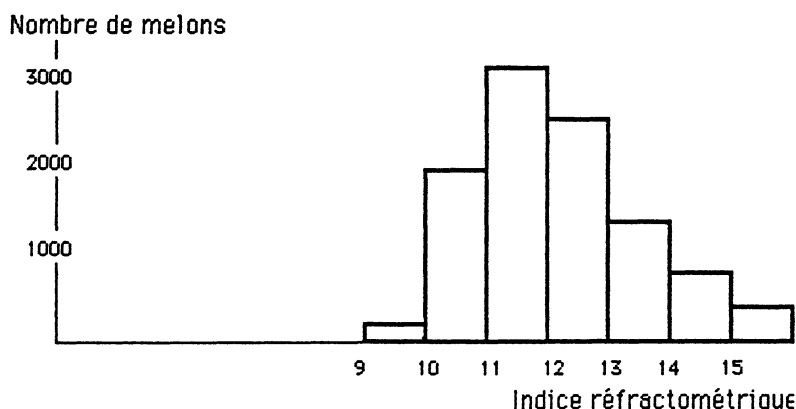


Figure 4. — Population triée melon par melon (taux de sondage de 25 % et seuil à 10.5).

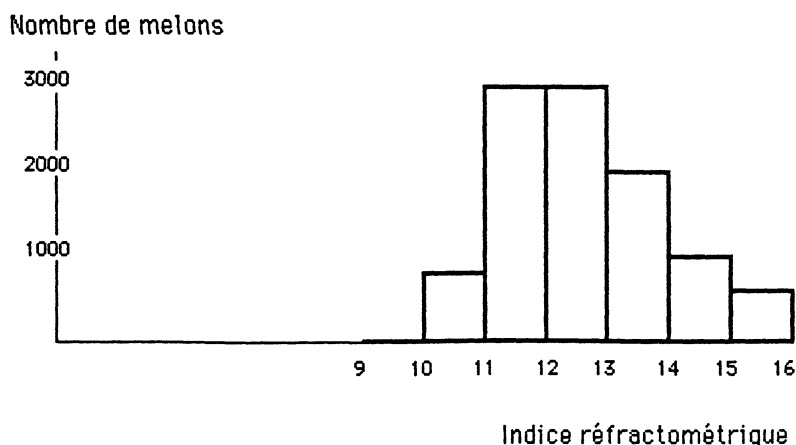


Figure 5. — Population triée caisse par caisse (taux de sondage 25 % et seuils 12.5 et 11).

Si une des machines de tri automatique tombe en panne, le tri melon par melon, dont on est conduit à abaisser le seuil pour respecter la norme de quantité, donne des résultats très mauvais, alors que le tri caisse par caisse reste acceptable comme le montre le tableau 3.

TABLEAU 3

Caractéristiques des populations triées,
pour une population d'un « gros jour » de paramètres suivants :
 $\gamma_1 = 1$; $\gamma_2 = 2,5$; $a = 1$; $b = 0,8$.

Type de tri	Seuil	Taux de sondage	Quantité	% de mauvais	% de bons
melon/melon	10.5	25%	10900	1,2	49,5
caisse / caisse	12.5 et 11	25%	11300	0,04	66
melon/melon	9.5	16,7%	10150	16	33,5
caisse / caisse	12 et 11	16,7%	10250	1,4	66

Les résultats précédents sont obtenus pour une population donnée. L'analyse des simulations obtenues en faisant varier les différents paramètres, γ_1 , γ_2 , a et b, ainsi que le type de jour, permet d'établir les conclusions suivantes :

1) Le tri caisse par caisse est toujours supérieur au tri melon par melon. Cette supériorité s'affirme surtout dans les cas difficiles : « gros jour », arrivage de mauvaise qualité, ou taux de sondage tombant à 16,7 %, suite à une panne de l'un des trois appareils de mesure automatique de l'indice réfractométrique. Les « bons jours », l'écart entre les deux méthodes est faible; cependant le tri caisse par caisse permet alors de passer à un taux de sondage de 16,7 %, économisant ainsi un appareil de mesure, tout en respectant les normes de quantité et de qualité.

2) Les variations du paramètre γ_1 influent peu sur les résultats; la diminution de γ_2 et l'augmentation de a, augmente l'efficacité du tri caisse par caisse, mais de façon générale, ces paramètres ont peu d'influence sur les résultats, tant qu'ils restent dans des intervalles raisonnables. L'augmentation de l'erreur de mesure, liée au paramètre b, diminue l'efficacité du tri melon par melon, alors qu'il modifie très peu ceux du tri caisse par caisse.

3) La répétition d'une simulation pour des valeurs inchangées des paramètres, permet d'obtenir des éléments sur la variabilité des résultats. On observe que cette dernière est faible, ce qui signifie que le système de tri est fiable, et présente peu de risques de donner des résultats très éloignés des valeurs attendues. Le tableau 4 donne une idée des ordres de grandeur des écart-types des 3 caractéristiques des systèmes de contrôle : la quantité, les pourcentages de bon et de mauvais dans la catégorie supérieure.

4) Les simulations permettent non seulement de choisir le système de tri le mieux adapté, mais aussi de connaître son comportement en face des différentes

TABLEAU 4

Ecart-types des caractéristiques des populations triées,
(le coefficient de variation est entre parenthèses)

Type de tri	Seuil	Taux de sondage	Quantité	% de mauvais	% de bons
melon/melon	10.5	25%	128 (1,2%)	0.0006	0.006
caisse / caisse	12.5 et 11	25%	141 (1,2%)	0.0003	0.005
melon/melon	9.5	16,7%	203 (2%)	0.004	0.02
caisse / caisse	12 et 11	16,7%	247 (2,4%)	0.002	0.01

situations dans lesquelles il sera placé. Par conséquent, les résultats des simulations donnent une information très utile pour choisir les seuils de tri, et le taux de sondage adaptés à l'arrivage de la journée. Le pilotage du système pourra donc se baser sur les résultats des simulations.

d) Aspects informatiques des simulations

Le programme de simulation d'une journée de contrôle comporte environ 200 instructions FORTRAN. Les lois uniformes sont générées par des méthodes linéaires de congruence (Phillips, 1976) dont l'initialisation est basée sur l'horloge interne de l'ordinateur, pour éviter la répétition des mêmes séries de nombres pseudo-aléatoires. Les lois normales centrées réduites, sont générées comme sommes de 12 variables aléatoires indépendantes et uniformes dans l'intervalle $[-1/2, 1/2]$.

Le programme contient un module de saisie conversationnelle des paramètres de la simulation, et donne en sortie, l'histogramme de la population simulée, ainsi que ceux des populations triées par les 2 types de contrôle, et les 3 caractéristiques de ces derniers (nombre de melons en catégorie supérieure, pourcentage de bons et de mauvais fruits).

Le temps d'exécution du programme, est d'environ 15 minutes sur un Mini 6. Pour un « gros jour », 111 000 melons sont simulés, ce qui nécessite la génération de 1,5 millions de nombres pseudo-aléatoires.

4. VALIDATION ET CONCLUSIONS

a) Validation

Les seules données disponibles, sur un ensemble assez important pour permettre une comparaison intéressante avec la simulation, se présentent de la façon suivante :

Deux lots, comportant respectivement 1 867 et 3 263 melons, ont été analysés de façon complète. Les caisses contiennent 8 melons, et l'échantillon est de taille 2. La règle de tri de la caisse consiste à l'accepter si aucun des 2 melons n'a un indice inférieur à 12.5. Dans le cas contraire, la caisse est rejetée, sauf les melons analysés qui ont un indice supérieur à 12.

Le programme de simulation a été modifié, pour tenir compte des modalités de tri caisse par caisse réellement utilisées dans cette expérience. Les résultats des simulations sont donnés dans le tableau 5 pour la population totale, et dans le tableau 6 pour les populations triées.

Le tableau 5 révèle un léger écart entre le lot A et ses simulations; il provient du fait que la distribution est bimodale; en fait elle est formée de deux sous populations : d'une part des melons cueillis trop tôt, et d'autre part des melons bien mûrs.

TABLEAU 5

Pourcentages de mauvais, moyens et bons melons
dans les lots A et B réels et simulés
(l'écart-type des valeurs simulées est entre parenthèses)

	IR $\leq$ 10	10 < IR $\leq$ 12	IR > 13
observé	23%	27,5%	49,5%
Lot A simulé	23,3% (1,5%)	30,5% (2%)	46,2% (2,6%)
observé	25,5%	31,5%	43%
Lot B simulé	25,9% (1,9%)	31,9% (0,9%)	42,2% (1,8%)

TABLEAU 6

Caractéristiques des populations triées réelles
et simulées pour les lots A et B

	Tri melon par melon			Tri caisse par caisse		
	Nombre	% mauvais	% bons	Nombre	% mauvais	% bons
observé	200	0	87	489	4,5	76,5
Lot A simulé	213 (13)	0 (0)	82 (3)	497 (53)	4,3 (1,8)	77,1 (2,2)
observé	330	---	---	479	5,5	74,5
Lot B simulé	335 (15)	0 (0)	84 (1,7)	518 (47)	4,6 (2,7)	77 (4,1)

Le tableau 6 montre une bonne concordance entre les chiffres expérimentaux et ceux prédits par le modèle. Il faut remarquer que la simulation d'un lot donne des résultats plus variables que celle d'une journée, ce qui est naturel, car cette dernière est basée sur plusieurs dizaines de lots. Par suite d'une défaillance de l'appareil de mesure, les résultats réels du tri melon par melon du lot B sont complètement faux, et n'ont pas été reportés dans le tableau 6.

Il semble y avoir une contradiction entre ces résultats et ceux de la section 3, qui montraient une supériorité du tri caisse par caisse; cette dernière n'est pas vérifiée sur les données expérimentales, au contraire, puisque les pourcentages de mauvais melons dans la catégorie supérieure sont non négligeables. Ce fait provient du seuil trop bas du tri caisse par caisse, qui accepte trop de melons. Pour améliorer les résultats de ce dernier, il faut en relever le seuil d'acceptation. Il est facile de refaire une simulation avec un seuil S_2 de 14, en gardant les melons d'indice supérieur à 12.5 (S_3). On obtient alors pour le lot B par exemple, 386 melons en catégorie supérieure, avec 0,1 % de mauvais et 88 % de bons.

b) Conclusions

La simulation est un instrument très souple, qui s'adapte à des problèmes complexes de contrôle de la qualité. Elle permet de tenir compte de tous les facteurs aléatoires importants, ainsi que des contraintes imposées au contrôle. Elle permet d'éviter des expérimentations coûteuses et de répondre à beaucoup de questions du style « Que se passe-t-il si... ? », qui naissent toujours de l'analyse primaire.

La partie la plus complexe et la plus longue, est généralement la modélisation, qui doit être menée avec soin; elle nécessite de faire des expériences, de recueillir des données pour préciser tel ou tel aspect du modèle. En contrôle de la qualité, il est particulièrement important d'analyser en détail les sources de variabilité et d'utiliser les méthodes statistiques appropriées. Cependant, ces dernières ne sont généralement pas suffisantes pour décider des modalités du contrôle. C'est pourquoi, la simulation est un complément logique aux études de variabilité, classiques en contrôle de la qualité.

RÉFÉRENCES

- D.T. PHILLIPS, A. RAVINDRAN (1976). — Operation research principles. Wiley.
H. SCHEFFE (1967). — The analysis of variance. Wiley.
F. SIDIBE, J.J. DAUDIN (1985). — Variabilité de l'indice réfractométrique et modélisation du tri industriel des melons en fonction de leur qualité gustative. Ind. Aliment. Agric. IAA 102, 1-2, p. 27-34 et IAA 102, 3, p. 219-27.