

# LA MSP

## ► Les situations non standard



► Par **Daniel LAFAYE de MICHEAUX**, directeur R&D, GPC System et **Alain PALSKY**.

*Dans de précédents articles, nous avons exposé différentes méthodes, techniques, témoignages et réflexions sur la maîtrise statistique de processus.*

*Nous avons voulu souligner certains aspects importants qui nous paraissent trop souvent mal identifiés par les utilisateurs de la MSP. Pour ce faire, nous avons tenté d'explicitier les raisonnements basés sur le calcul des probabilités, qui permettent de prendre diverses décisions concernant la conduite du processus et, en particulier, de calculer convenablement les limites à définir sur les cartes de moyennes et les cartes de dispersion.*

Les calculs présentés dans ces articles étaient basés sur des hypothèses qui doivent être vérifiées au préalable, hypothèses sur le fonctionnement considéré comme correct :

- constance dans le temps de la cible et de la variabilité de la grandeur mesurée ;
- indépendance statistique des mesures successives ;
- distribution normale (gaussienne) de la mesure.

Dans la pratique de la MSP on constate parfois que ces hypothèses ne sont pas

vérifiées par les mesures dont on dispose et donc que l'on n'est pas dans la "situation standard" d'utilisation de la MSP.

Doit-on alors renoncer à utiliser les cartes de contrôle ? La norme Afnor (voir Afnor X06 031) ne propose pas de réponse.

Nous allons, dans cet article et dans de prochains articles, présenter des variantes justifiées d'utilisation des cartes de contrôle adaptées à diverses situations "non standard" rencontrées couramment dans la pratique.

### Hypothèses et calculs en situation standard

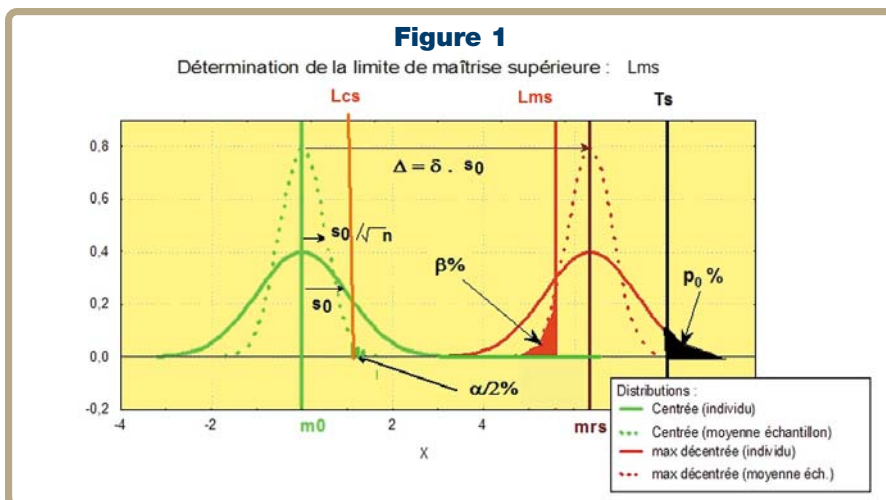
Nous nous sommes placés dans le cadre classique où l'on ne suit qu'une seule grandeur. On reporte alors sur une carte les moyenne et dispersion  $\bar{x}_i$  et  $s_i$  d'un i-ème échantillon de taille  $n$ , prélevé aléatoirement dans une "fenêtre de production" comportant  $N$  éléments fabriqués. La mesure en question peut être celle d'une caractéristique du produit (dimension, dureté, concentration...) ou un paramètre du procédé (température, pression...).

Le processus bien réglé est supposé fonctionner afin que la mesure suive une loi normale centrée sur la valeur cible constante  $m_0$  avec un écart type constant  $s_0$ .

Nous avons alors distingué clairement :

- les limites de tolérance (limites notées  $T_s$  pour la limite supérieure et  $T_i$  pour la limite inférieure), exigées par le client dans le cas de mesures produit, ou imposées à la conception du procédé pour assurer la sécurité et maintenir la conformité des produits fabriqués ;
- les limites de contrôle de Shewhart (limites notées  $Lcs$  et  $Lci$ ) qui servent à détecter que le processus n'est plus dans son état de fonctionnement normal ;
- les limites de maîtrise (limites notées  $Lms$  et  $Lmi$ ) qui permettent de détecter si l'on commence à fabriquer un taux  $p$  de produits non conformes supérieur au taux maximal  $p_0$  sur lequel un engagement a été pris auprès du client. Pour un taux de fausse alerte  $\alpha$  accepté par le producteur, les limites de contrôle de Shewhart sont alors définies par les formules :

$$\begin{aligned} Lcs &= m_0 + u(1 - \frac{\alpha}{2}) \times \frac{s_0}{\sqrt{n}} \\ Lci &= m_0 - u(1 - \frac{\alpha}{2}) \times \frac{s_0}{\sqrt{n}} \end{aligned} \quad (1)$$



# MSP

avec  $u() = \Phi^{-1}()$  qui désigne l'inverse de la fonction de répartition  $\Phi$  de la loi gaussienne standard disponible dans les tables statistiques.

Pour un taux de non-détection  $\beta$  accepté par le client, les limites de maîtrise (limites "de type Shewhart" selon Afnor X06 031) sont alors définies à partir des moyennes refusables **mrs** et **mri** :

$$\begin{aligned} \text{mrs} &= \text{Ts} - s_0 \cdot u(1 - \beta) \\ \text{mri} &= \text{Ti} + s_0 \cdot u(1 - \beta) \end{aligned} \quad (2)$$

par les formules :

$$\begin{aligned} \text{Lms} &= \text{mrs} - u(1 - \beta) \times \frac{s_0}{\sqrt{n}} \\ \text{Lmi} &= \text{mri} + u(1 - \beta) \times \frac{s_0}{\sqrt{n}} \end{aligned} \quad (3)$$

Lorsque l'aptitude du processus et la taille de l'échantillon sont suffisantes, on se trouve dans une situation conforme à la *figure 1*. On peut alors déclencher une action corrective entre les limites **Lcs** et **Lms** ; si le coût de la correction est faible, on déclenchera l'action dès la valeur **Lcs**, sinon on pourra attendre jusqu'en **Lms** sans danger véritable pour la qualité du produit.

Si le coût de mesure est notable, il peut être intéressant de fixer la taille de l'échantillon au juste nécessaire compatible avec les risques  $\alpha$ ,  $\beta$  et  $p_0$  acceptés, c'est-à-dire choisir **n** tel que **Lcs** et **Lms** ou **Lci** et **Lmi** coïncident. La valeur optimale de **n** vérifie l'équation d'efficacité

$$\delta\sqrt{n} = u(1 - \alpha/2) + u(1 - \beta) \quad (4)$$

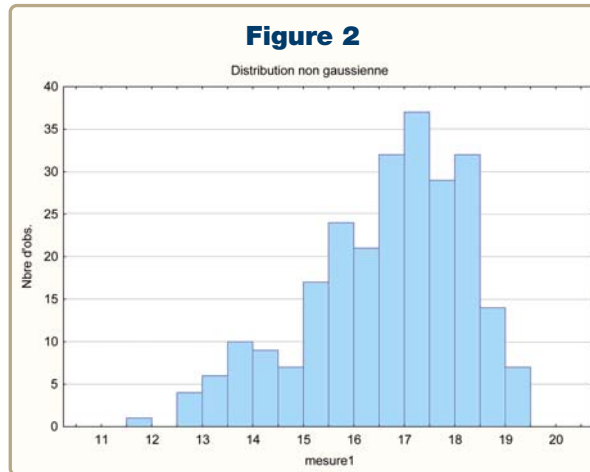
avec

$$\delta = \text{MIN} \left( \frac{\text{mrs} - m_0}{s_0}, \frac{m_0 - \text{mri}}{s_0} \right) \quad (5)$$

On choisira généralement pour valeur de **n** l'entier immédiatement supérieur à la valeur désignée par l'équation d'efficacité, ce qui permettra de revoir légèrement à la baisse l'un des risques  $\alpha$ ,  $\beta$  ou  $p_0$ .

## Situations non standard

Que faire lorsque les hypothèses qui justifient ces calculs ne sont pas vérifiées ? Il faut commencer par identifier cor-



rectement le type de situation dans lequel on se trouve pour obtenir une solution plus efficace.

On peut avoir un fonctionnement centré de façon stable mais qui présente une distribution dissymétrique, comme le montre l'histogramme de la *figure 2*.

La *figure 3* montre un graphe de mesures qui serpentent, l'indépendance entre mesures successives n'est alors pas vérifiée, la normalité non plus. La *figure 4* est caractéristique d'un phénomène d'usure ou d'encrassement : la dérive lente est inévitable, le calcul habituel des limites de Shewhart ne peut pas convenir.

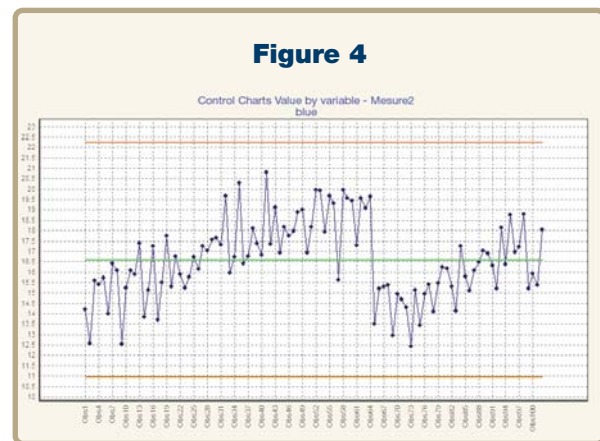
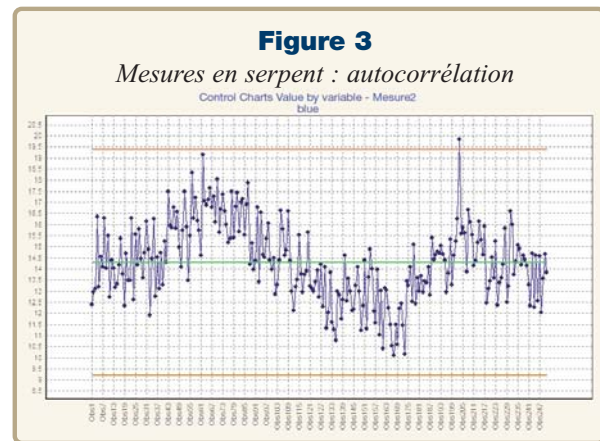
Au préalable, il faut rappeler que les calculs de limites doivent être effectués sur l'historique constitué à partir d'une période de référence la plus représentative de ce que l'on

convient de considérer comme le "fonctionnement correct" : toute situation d'anomalie que l'on veut détecter par la suite doit absolument être éliminée de l'historique.

## Situation 1 : cartes de maîtrise pour mesures individuelles

Dans de nombreux cas, il est possible d'avoir un contrôle à 100 % grâce à des

dispositifs de mesure automatique. Il est alors utile de construire des cartes aux mesures individuelles lorsque l'on craint des anomalies éphémères, c'est-à-dire qui entachent une mesure mais pas la précédente ni la suivante (phénomène de contamination par exemple). Supposons toujours que les hypothèses 1, 2 et 3 soient vérifiées. En posant la taille de l'échantillon  $n = 1$ , la formule (1) est encore valable, mais le raisonnement développé pour établir les autres formules doit



être simplifié : il n'y a plus d'incertitude d'échantillonnage sur la valeur de la moyenne instantanée, l'incertitude sur la vraie valeur est uniquement l'incertitude de mesure  $s_{mes}$ .

Ceci conduit aux formules :

$$\begin{aligned} \text{Lcs} &= m_0 + u(1 - \frac{\alpha}{2}) \times s_0 \\ \text{Lci} &= m_0 - u(1 - \frac{\alpha}{2}) \times s_0 \end{aligned} \quad (1')$$

$$\begin{aligned} mrs &= Ts - s_{mes} \cdot u(1 - p_0) \\ mri &= Ti + s_{mes} \cdot u(1 - p_0) \end{aligned} \quad (2^1)$$

$$\begin{aligned} Lms &= mrs \\ Lmi &= mri \end{aligned} \quad (3^1)$$

Les formules (4) et (5) de calcul de  $n$  optimal sont alors sans objet.

L'écart type  $s_0$  à prendre en compte est l'écart type des mesures, calculé sur la totalité de l'historique de référence correspondant à un fonctionnement correct.

Voici quelques mises en garde concernant l'utilisation des cartes aux mesures individuelles :

1. Les formules (1<sup>1</sup>), (2<sup>1</sup>) et (3<sup>1</sup>) ne peuvent être appliquées que si elles conduisent à des limites vérifiant  $Lmi \leq Lci$  et  $Lcs \leq Lms$ , ce qui nécessite une aptitude importante. Sinon, il convient d'abandonner le suivi de mesures individuelles et de construire les cartes aux moyennes décrites au point suivant.
2. Rappelons que, pour augmenter la capacité de détection de faibles dérives, il convient d'associer à une carte aux valeurs individuelles une carte aux moyennes : moyennes mobiles ("glissante") ou moyennes de groupements de  $n$  mesures successives (dans ce dernier cas  $n = N$ ).
3. Il faut également surveiller la stabilité de la dispersion par une carte aux étendues ou aux écarts types glissants.
4. L'opérateur doit se garder de modifier le réglage de l'équipement lorsqu'une anomalie éphémère se manifeste. Il convient alors d'en rechercher la cause pour éviter qu'elle ne se reproduise par la suite. Ainsi toute anomalie détectée sur une carte aux mesures individuelles doit entraîner une ou plusieurs autres mesures pour confirmer la pertinence d'un réglage.

### Situation 2 : hypothèses 1 et 2 vraies, hypothèse 3 fausse

Pour identifier cette situation, il faut avoir effectué au moins un test de normalité sur les mesures individuelles, non sur les moyennes d'échantillon. De tels tests statistiques sont proposés par les logiciels spécialisés de statistique et maintenant par la plupart des logiciels de MSP. Les tests les plus courants sont la droite de Henry (test "à vue"), les tests du Khi2 et de Kolmogorov.

Si les tests ne peuvent accepter l'hypothèse de normalité, il est bon de s'interroger sur la raison d'une telle situation.

Cela peut venir de la nature de la grandeur elle-même, dans certains cas il est même possible de connaître sa loi de distribution théorique : loi lognormale pour des défauts de forme, de Weibull pour des mesures de fiabilité, loi de Rayleigh...

Cela peut également venir d'un mélange de plusieurs lois gaussiennes différentes : par exemple les produits sont fabriqués par des équipements différents n'ayant pas tous un réglage parfait, ou par des équipes différentes n'appliquant pas de la même façon les procédures, ou encore ils sont fabriqués à partir de lots différents de matières premières. Par des analyses appropriées de la variance, il convient alors de faire la chasse à de telles sources d'hétérogénéité avant de reconstituer un historique de référence.

On peut rencontrer des distributions non gaussiennes de lois inconnues lorsque la valeur cible est proche d'une "butée" : pour un produit chimique, ce pourra être une pureté ciblée sur 99 % ou une siccité ciblée sur 0,1 %.

Il est possible de contourner cette difficulté par une transformation des données. On ne suit plus les mesures  $x_i$  mais les mesures transformées  $y_i = f(x_i)$ , la fonction  $f$  étant choisie de telle sorte que la distribution des  $y_i$  soit approximativement gaussienne. La fonction logarithme convient souvent et conduit à construire des cartes manuelles sur papier logarithmique.

#### ► Cas où $n$ est "assez grand"

Lorsque la non-normalité doit être acceptée, il convient de distinguer le cas où la taille  $n$  de l'échantillon est "suffisamment grande" pour pouvoir admettre que la moyenne d'échantillon suive une loi approximativement gaussienne d'après le théorème de la limite centrale. À partir de  $n = 5$  l'utilisation de cette approximation semble raisonnable.

Dans ce cas les formules (1), (3), (4) et (5) sont inchangées, seule la formule (2) de calcul des moyennes refusables doit être adaptée, puisqu'elle fait référence à la distribution des mesures individuelles. Il faut alors la remplacer par

$$\begin{aligned} mrs &= Ts - (F^{-1}(1 - p_0) - m_F) \\ mri &= Ti + (m_F - F^{-1}(p_0)) \end{aligned} \quad (2^{2a})$$

où  $F$  désigne la fonction de répartition de la distribution des mesures individuelles et  $m_F$  sa moyenne (espérance mathématique). Si cette distribution est connue (lognormale, Weibull...) tout va bien, sinon on peut essayer de trouver une loi classique paramétrée qui conviendrait pour la représenter : loi du khi2, loi beta...

Cette recherche est possible grâce aux "tests d'ajustement" disponibles dans les logiciels statistiques et appliqués sur les mesures de l'historique de référence. Les résultats de ces tests doivent être interprétés avec discernement car ils sont sensibles à la taille de l'historique : avec 50 mesures plusieurs distributions différentes pourront satisfaire le test d'ajustement aux données de l'historique (laquelle choisir ?), avec 5000 aucune ne conviendra.

En fait, l'utilisation de la formule ne nécessite pas la connaissance de la distribution dans sa globalité mais uniquement celle des quantiles  $F^{-1}(1 - p_0)$  et  $F^{-1}(p_0)$ . Une estimation de ces "quantiles extrêmes" est possible en utilisant les approximations de Johnson à partir des mesures expérimentales. Ces approximations sont disponibles sur plusieurs logiciels généralistes de statistique et dans certains logiciels MSP.

#### ► Cas des mesures individuelles

Si  $n = 1$ , il convient, dans les équations, de faire intervenir la fonction  $F^{-1}$  au lieu de  $u$  dans les formules (1<sup>1</sup>), (2<sup>1</sup>) et (3<sup>1</sup>), ce qui donne les formules suivantes :

$$\begin{aligned} Lcs &= F^{-1}\left(1 - \frac{\alpha}{2}\right) \\ Lci &= F^{-1}\left(1 - \frac{\alpha}{2}\right) \end{aligned} \quad (1^{2b})$$

$$\begin{aligned} mrs &= Ts - (F^{-1}(1 - p_0) - m_F) \\ mri &= Ti + (m_F - F^{-1}(p_0)) \end{aligned} \quad (2^{2a})$$

$$\begin{aligned} Lms &= mrs \\ Lmi &= mir \end{aligned} \quad (3^1)$$

#### ► Cas où $1 < n < 5$

Il est difficile d'être rigoureux dans ce cas car il faudrait estimer la distribution de la moyenne des échantillons de taille  $n$  pour pouvoir calculer les limites. On s'efforcera plutôt de trouver une transformation ramenant la distribution à une loi normale à défaut de pouvoir augmenter  $n$ .

### Qualité références

► Nous examinerons dans le prochain article d'autres cas où les hypothèses des cartes classiques de Shewhart ne sont pas vérifiées.