

MSP multidimensionnelle

DÉTECTER ET IDENTIFIER “L’INVISIBLE”



Par **Daniel LAFAYE DE MICHEAUX**, directeur R&D de GPC System, et **Didier VIEUX**, directeur des opérations de GPC System, www.gpc-system.com.

La MSP traditionnelle est basée sur la surveillance individuelle de paramètres à contrôler ; cette approche dite “monodimensionnelle” a pour inconvénient de ne pas tenir compte des liens entre les paramètres et tend donc trop souvent à affirmer implicitement « je ne vois rien = il ne se passe rien ».

Notre expérience industrielle montre en fait que les phénomènes graves affectant les lignes de production (arrêts longs, rebuts, retours clients...) présentent des signes précurseurs, qui sont souvent invisibles sur les cartes de contrôle monovariées, mais qui sont nets et clairement identifiables en vision multivariée.

Cette détection précoce permet alors d’augmenter la réactivité du pilote de procédé et, par la suite, d’améliorer la rentabilité de la ligne de production. Quels sont alors ces phénomènes cachés ? Comment les détecter et les identifier ?

Les 2 types de variabilité

La MSP considère 2 types de variabilité : la variabilité ordinaire (ou naturelle, “causes communes” de dispersion du procédé...) et la variabilité exceptionnelle (dysfonctionnements, anomalies, “causes spéciales” de dispersion du procédé...). Nous traiterons ici, dans un premier temps, les phénomènes attachés à la variabilité ordinaire.

Avec les cartes de contrôle monovariées classiques (Shewhart, Cusum ou Ewma), on s’applique principalement à surveiller la variabilité exceptionnelle afin de détecter l’émergence d’une anomalie, qui se traduit par un décalage de moyenne ou une augmentation de la dispersion.

Sur le long terme, on s’efforce de réduire la variabilité ordinaire de certaines variables critiques en mettant en œuvre des actions d’amélioration définies par une analyse technique fine.

Une cause de variabilité qui affecte une seule variable est facile à identifier. Mais, lorsque le nombre de variables est important (plus de 50), il devient difficile de repérer une cause de variabilité qui influence

plusieurs variables à la fois.

On peut s’aider de techniques d’analyse statistique multidimensionnelle pour identifier rapidement les facteurs prépondérants de dispersion communs à plusieurs mesures. Dans ce cas, une action technique permettra d’améliorer simultanément la dispersion de plusieurs variables.

Phénomènes cachés dans la variabilité ordinaire

L’architecture de la ligne de production (cf. figure 1) comprend 3 équipements de même nature produisant en parallèle

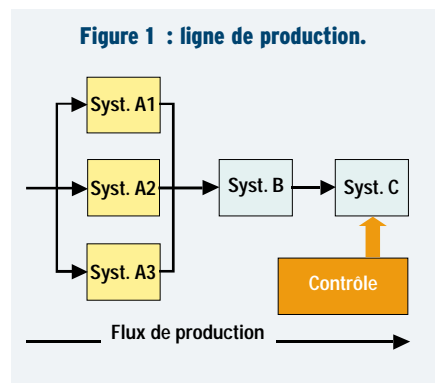


Figure 2 : collecte des données.

A	B	C	D	L	M	N			
Num	Che	Measu1	Measu2	Measu3	Heure	Minute	Fournisr	Machine	P
219	0.6088497	3.72078	-1.2676	3	39				3
220	0.3888424	3.32293	-1.8793	3	39				1
221	0.2618278	2.98843	-1.5489	3	40				2
222	0.7547487	3.72078	-1.4899	3	41				3
223	0.8918112	3.42252	-1.5489	3	42				1
224	-0.1988289	3.12272	-0.2889	3	42				2
225	0.7288782	4.02861	-1.8823	3	43				3
226	0.588849	3.67078	-1.8289	3	43				1
227	-0.0918889	3.36458	-2.2319	3	44				2
228	0.8788712	3.97821	-1.5891	3	41				3
229	0.2878878	3.69987	-1.7914	3	40				1
230	0.1878726	3.18826	-1.8793	3	40				2

(machines A1, A2 et A3) et 2 sous-systèmes produisant en série (B et C). Cette mise en parallèle des moyens de production se retrouve fréquemment dans les usines : stations à postes multiples, moules multiempreintes ou simplement 3 lots ou fournisseurs différents de matière première.

Les mesures (10 dans notre exemple) sont effectuées par un contrôle à 100 % sur le produit en fin de ligne.

La collecte de données donne un tableau (cf. figure 2). Chacune de ses lignes, représentant une observation (ou un produit), est constituée par le numéro du produit et les mesures contrôlées ; elle est complétée par des informations dites “illustratives” comme le numéro de la machine utilisée, l’heure et la minute de fabrication, le fournisseur, l’équipe (matin, soir, nuit)...

Dans notre exemple, les cartes de contrôle du type Shewhart (cf. figure 3) montrent une variabilité très régulière : aucune mesure n’est en dehors des limites orange, calculées ici à +/- 3 s. Aucun phénomène particulier ne peut être clairement détecté sur ces cartes de contrôle aux données

Figure 3 : carte de contrôle de type Shewhart.

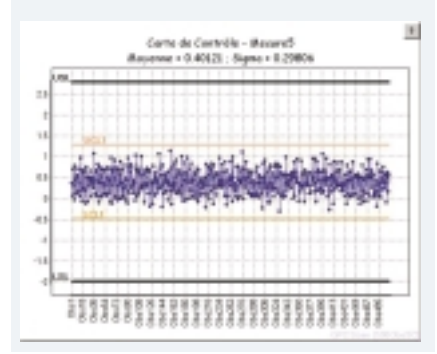
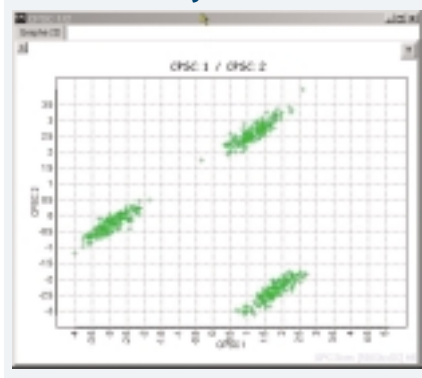


Figure 4a



individuelles, pas plus que sur des cartes aux moyennes d'échantillon (Ewma...).

La vision multivariée par l'analyse en composantes principales (ACP)

Nous allons maintenant nous intéresser à un type d'analyse bien connu des statisticiens : l'analyse en composantes principales ACP. Malgré l'aspect un peu complexe de la théorie, cette analyse et les graphes associés peuvent être très facilement pris en main et interprétés dans les cas les plus courants des besoins du contrôle de procédé.

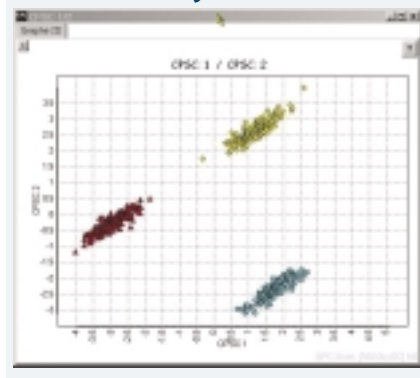
L'ACP permet de mettre en évidence les éléments prépondérants de la variabilité d'un ensemble de mesures ; à partir des mesures initiales, elle calcule des variables "résumées" - appelées composantes principales -, qui sont numérotées par leur ordre d'importance dans la composition de la variabilité.

Ensuite, un simple graphe (x, y) en 2 dimensions, composé des 2 premières composantes principales (ou des suivantes), permet de visualiser graphiquement et très simplement certains phénomènes qui s'avèrent invisibles sur les données initiales.

Si l'on a pris soin de retenir, dans l'historique des mesures sur un procédé de fabrication, uniquement des observations correspondant à un fonctionnement considéré comme correct, nous pouvons alors distinguer par ACP les facteurs prépondérants de la "variabilité ordinaire" de ce procédé.

La figure 4a est une représentation graphique des 2 premières composantes principales calculées à partir des 10 mesures de notre exemple précédent. Sur ce graphe, chaque point représente une observation, c'est-à-dire un produit fabriqué. Les 2 coordonnées de ce point sont données par le calcul d'ACP qui tient compte de toutes les

Figure 4b



mesures initiales associées à ce produit. Il apparaît très clairement sur ce graphique 3 nuages distincts : on peut en conclure qu'un phénomène caché, à 3 états, structure la variabilité ordinaire de notre procédé, en augmentant celle-ci. En effet, en l'absence de ce phénomène, les 3 nuages seraient confondus, ce qui donnerait une variabilité plus faible.

Une telle situation peut avoir pour conséquence un rapprochement dangereux des paramètres du produit vers les limites de contrôle ou de spécification. Cela peut entraîner une diminution de la fiabilité du produit final, ou des difficultés sur les étapes suivantes de production. Il convient donc de rechercher l'origine de cette variabilité ordinaire excessive.

L'identification du phénomène caché par une variable illustrative

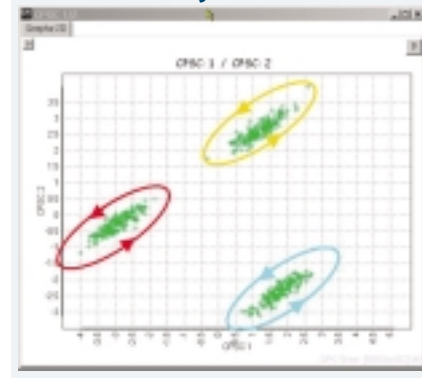
Plusieurs démarches peuvent être utilisées pour identifier l'origine de ce phénomène. La plus simple consiste à tenter de mettre en correspondance la structuration des observations (3 groupes) constatée sur le graphe avec des informations qualitatives ou discrètes dont on dispose sur chaque produit observé (appelées aussi variables illustratives). Cette mise en correspondance peut faire appel à des outils statistiques quelque peu complexes, mais peut aussi être simplement graphique.

Pour une variable illustrative donnée, chaque modalité aura une couleur et une forme dédiées ; pour la variable Numéro Machine, on obtient par exemple :

- machine 1 = ●
- machine 2 = ▲
- machine 3 = ◆

La figure 4b propose le même graphe que la figure 4a mais en plus, chaque observation a été colorée selon son appartenance à la modalité de la variable Numéro Machine.

Figure 4c



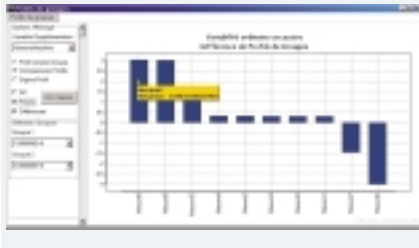
On voit que les modalités 1, 2 et 3 de cette variable permettent de discriminer les 3 nuages ; nous avons donc identifié la source de la variabilité excessive : elle est induite par des décalages existant entre les points de fonctionnement des 3 machines disposées en parallèle. Est-ce une question de niveaux d'usure différents, de mauvais réglages... ? Le problème est maintenant bien circonscrit, le champ de l'investigation technique s'est considérablement rétréci.

L'identification du phénomène caché par la création d'une variable illustrative

Lorsqu'on ne dispose pas de variable illustrative caractéristique de la dispersion constatée, on va la créer, puis essayer de l'interpréter. Il s'agit ici d'associer à chacun des groupes de la figure 4a un numéro (1, 2 ou 3), puis d'associer à chaque observation le numéro du groupe dans lequel il appartient sur la figure 4a : nous avons ainsi défini une nouvelle variable discrète à trois valeurs possibles.

Pour ce faire, on peut utiliser des techniques de classification de l'analyse statistique multivariée (Classification ascendante hiérarchique, nuées dynamiques, Isodata...). Plus simplement, il est possible de créer l'information graphiquement : on sélectionne les observations d'un nuage du graphe d'ACP par un outil graphique de type lasso (cf. figure 4c) puis on associe une valeur de la variable créée à l'ensemble des observations ainsi sélectionnées. On effectue cette manœuvre pour chaque groupe identifié dans le plan d'ACP. Le phénomène de dispersion peut être alors être précisé par la mise en correspondance de la variable de groupe ainsi créée avec d'autres informations sur les produits (date de production...) ; on peut aussi étudier la liste des variables mesurées qui sont fortement sensibles à ce phénomène, c'est l'objet du paragraphe suivant. 8

Figure 5



Évaluation de l'intensité du phénomène sur les mesures

Il s'agit maintenant d'évaluer sur quelles grandeurs mesurées ce phénomène a une incidence et si ce phénomène caché a un réel impact sur la qualité de la production. Plusieurs statistiques et plusieurs graphiques permettent d'appréhender cette question. On peut simplement calculer la moyenne de chaque mesure pour chaque nuage et comparer ces moyennes entre elles.

Ainsi, la figure 5 présente la différence des moyennes des groupes 1 et 3 pour chaque variable. La disposition des valeurs par ordre d'importance (Pareto), en tenant compte du signe, montre que l'impact le plus fort porte sur la variable 8, avec une différence positive de 3.046 en valeur réelle mesurée. Cette représentation des décalages permet de rapidement identifier :

- les variables les plus touchées par le phénomène ;
- la valeur réelle de l'impact sur ces variables.

Retour au monovarié

Après avoir identifié les causes du phénomène caché, et repéré les variables les plus touchées, il est intéressant de regarder de près ces variables. La figure 6a illustre la carte de la mesure la plus touchée. Comme pour les graphes d'ACP, nous pouvons colorer les mesures initiales par les modalités de la variable illustrative Numéro Machine ou de la variable de groupe créée. Il apparaît alors sur la figure 6a 3 niveaux distincts qui se répètent régulièrement. Un zoom sur ce graphe (cf. figure 6b) permet de visualiser une alternance qui correspond effectivement à l'organisation du flux de production sur la ligne. Ce type particulier de variabilité de la variable 8 aurait pu être effectivement détecté par un examen attentif de la carte de contrôle par un œil averti. La démarche préconisée permet de mettre en évidence très rapidement cette variabilité ordinaire excessive ainsi que son origine au milieu des variations de plus de 100 variables ; elle peut être aussi efficace

Figure 6a

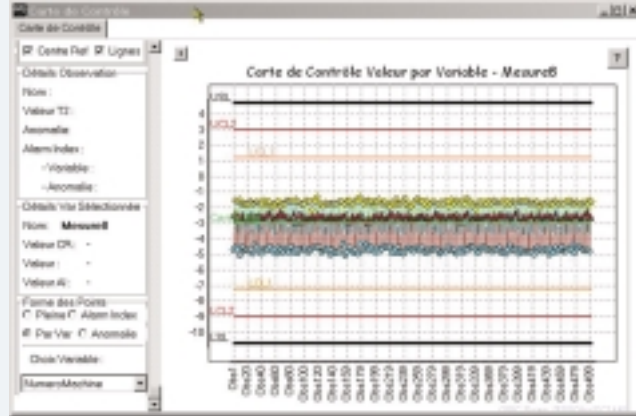
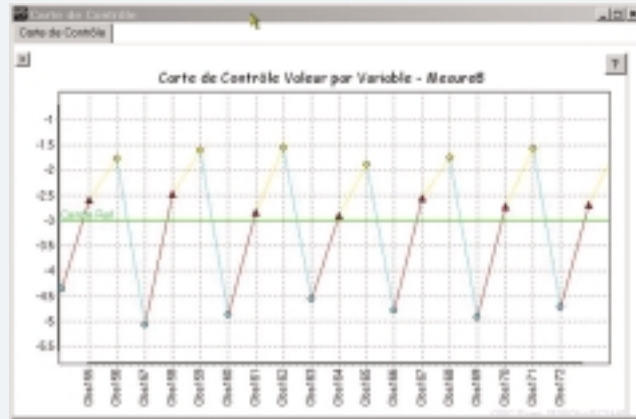


Figure 6b



alors que les effets sur les variables mesurées sont moins perceptibles.

Réduction de la variabilité ordinaire

Nous sommes donc face à un procédé au départ peu dispersé, mais qui voit sa variabilité augmenter de façon importante à cause du dérèglement entre sous-systèmes parallèles de la ligne.

Deux possibilités de corrections sont envisageables pour réduire la variabilité :

- correction sur la ligne ;
- correction des variables.

Ayant identifié que les 3 machines ont des points de fonctionnement différents, la correction la plus satisfaisante consiste à régler plus finement les 3 machines afin que leurs points de fonctionnement coïncident. Cette correction opérée sur la ligne n'est parfois pas possible : on a affaire à 3 machines comportant des niveaux d'usure différents ; seul leur changement par 3 machines neuves permettrait techniquement de résoudre le problème. On peut alors avoir recours à la correction

des variables : on centre les variables du produit sur les moyennes de la machine qui a fabriqué le produit. Cela ne change pas la dispersion ordinaire des caractéristiques du produit, mais cela permet au moins de réduire la variabilité des mesures et ainsi de détecter plus tôt les anomalies éventuelles.

Conclusion

Les données de l'exemple pédagogique ci-dessus ont été obtenues par simulation.

Il rend compte d'une situation où la cause principale de variabilité ordinaire, de nature discrète, peu ou pas visible sur les cartes monovariées, donne lieu à la formation de groupes dans des plans de représentation de l'ACP.

Dans le prochain article, nous illustrerons la situation précédente par des exemples industriels : construction mécanique et moulage de pièces. Nous évoquerons, par la suite, le cas de causes de variabilité ordinaire de type continu et non plus discret. Nous traiterons enfin le cas des anomalies, non détectables en contrôle monovarié mais bien détectables et identifiables en multivarié ■