

# Optimiser un procédé : avec ou sans données ?

L'industrie chimique et pharmaceutique doit constamment accroître tant la qualité de ses produits que sa productivité. Une optimisation régulière des procédés est donc nécessaire. Les méthodes de la statistique industrielle constituent pour cela des moyens éprouvés, car elles permettent d'obtenir une très bonne compréhension des procédés à partir de données qui les décrivent. Mais quelle méthode est-elle donc la plus appropriée selon la quantité de données disponible ?

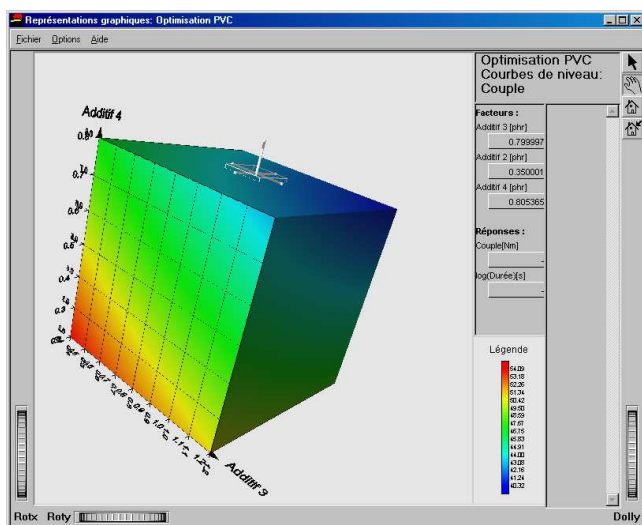
Lorsque la structure d'un procédé a été définie, il « ne reste plus qu'à » régler ses paramètres de manière idéale pour l'optimiser. De fréquentes variables de réponse sont alors le rendement et des caractéristiques de qualité, telle la pureté du produit fini. Les paramètres considérés – les « facteurs » - sont la plupart du temps des températures, des valeurs de pression, des durées opératoires, les quantités de matière employées... Occasionnellement, des facteurs catégoriels apparaissent également, comme la centrifugeuse utilisée ou le fournisseur d'une matière première.

La recherche de la combinaison idéale pour le réglage des paramètres se déroule en deux phases. Tout d'abord, on identifie les facteurs qui, pris isolément ou en considérant leurs interactions, ont une influence significative, par exemple sur le rendement. C'est seulement après cette phase de criblage que l'optimisation commence réellement. Elle vise à déterminer la meilleure valeur de chacun de ces facteurs importants.

## Pas de données à disposition

Si on ne dispose pas encore de données, par exemple au début du développement, l'idéal est d'avoir recours aux plans d'expériences statistiques, car ils garantissent une collecte de données efficace pour l'optimisation de procédés. Concrètement, on génère à partir des facteurs et de leurs domaines de variation une liste minimale de points expérimentaux optimalement répartis, pour rendre possible un calcul fiable du réglage idéal du procédé sur la base des mesures expérimentales.

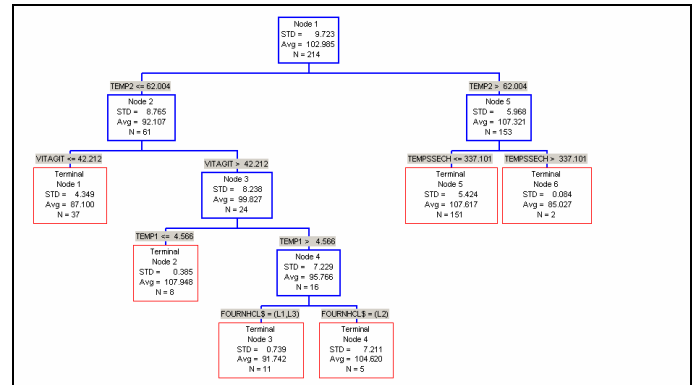
Dans une application parmi beaucoup d'autres, cette méthode a été utilisée pour déterminer la composition en additifs permettant d'obtenir des valeurs optimales de couple lors d'essais de malaxage d'un mélange de PVC. Si nécessaire, des spécificités de la problématique peuvent être prises en considération, entre autres au moyen de restrictions telles que « la quantité totale de deux additifs déterminés n'excède pas une limite donnée à l'avance ».



1: À chaque additif correspond un axe. La réponse est représentée par un dégradé de couleurs.

Dans la pratique, la mise en oeuvre des plans d'expériences est grandement facilitée par l'utilisation de puissantes solutions

logicielles. Par exemple, STAVEX guide confortablement l'utilisateur de la définition du problème à l'analyse des résultats, en passant par la définition du plan d'expériences. Même si des expériences préliminaires ont été réalisées sans avoir recours à un plan, il est possible de les exploiter en complétant cette liste d'expériences pour obtenir un plan approprié. En outre, de nombreuses représentations graphiques en deux, trois ou quatre dimensions autorisent une interprétation aisée des résultats : pour le cas décrit précédemment, la figure 1 montre comment le couple se comporte selon la quantité de chacun des trois additifs.



2: Extrait de l'arbre de la connaissance fourni par CART.

## Même en cas de déluge de données

À l'opposé de la situation discutée ci-dessus, d'immenses quantités de données sont générées dans les environnements de production et ne sont que trop rarement analysées. Même s'ils ne sont pas parfaits en raison d'erreurs, de tels cimetières de données recèlent d'informations capitales pour l'optimisation de procédés, par exemple lorsque le rendement varie trop d'un lot à l'autre. Les méthodes modernes de Data Mining, tels les arbres de classification et de régression, permettent d'extraire facilement du flux de données les dépendances essentielles entre variables de réponse et facteurs.

Le fonctionnement de cette technique est expliqué en utilisant l'exemple de l'optimisation du rendement d'un procédé, lequel comprend deux phases de réaction, une centrifugation et un séchage. On détermine d'abord le facteur dont les variations expliquent le mieux les fluctuations du rendement : dans la figure 2, il s'agit de la température de la deuxième phase de réaction. Une valeur limite de cette température – ici 62°C – est alors fixée : c'est elle qui sépare le plus clairement les lots de rendement « élevé » de ceux de « bas » rendement. Le jeu de données est ainsi séparé en deux parties auxquelles on applique le même principe, etc. Dans l'arbre qui en résulte, les connexions montrent que les meilleures valeurs de rendement sont obtenues lorsque la température de la deuxième phase de réaction est inférieure à 62°C, la vitesse de mélange supérieure à 42 min<sup>-1</sup> et la température de la première phase de réaction inférieure à 4,5°C.

Des outils logiciels conviviaux sont indispensables pour la mise en oeuvre de cette technique. Le produit CART combine efficacité de calcul, clarté des représentations (figure 2) et flexibilité. Ainsi, même de très grandes quantités de données sur les procédés peuvent être analysées sans peine.

## Conclusion

Indépendamment de la quantité de données disponibles, les méthodes statistiques peuvent apporter une contribution significative à l'optimisation de procédés. Qui croit malgré tout que la statistique n'est applicable qu'avec peine ou que par des spécialistes, réduit involontairement sa compétitivité.

Auteur : Dr. Philippe Solot, AICOS Technologies SA