

La place des méthodes chimométriques en chimie analytique *

Max Feinberg

Les relations entre chimie analytique et chimométrie

Chimie et chimie analytique

La chimie analytique, en tant que mode d'investigation d'une science expérimentale qui cherchait à expliquer et à maîtriser le comportement de la matière, a donné naissance à la chimie que nous connaissons aujourd'hui. C'est elle qui a permis à A.L. Lavoisier d'élaborer une nouvelle représentation de la matière en s'appuyant systématiquement sur une première méthode d'analyse qui paraît aujourd'hui désuète : la pesée. Il confirma le phénomène toujours constaté de l'augmentation du poids des métaux par calcination ; une démarche analytique le conduisit à proposer une explication du phénomène de combustion qu'il tenait pour résultant d'une combinaison avec un élément de l'air [1]. De la comparaison analytique de l'état initial et de l'état final était issue une conception économique des transits de matière qu'il résume dans la phrase suivante : "... dans toute opération, il y a une égale quantité de matière avant et après l'opération...". La chimie moderne était née et la découverte des réactions nucléaires ne changera pas grand chose puisqu'il suffira de remplacer le concept de matière par celui d'énergie.

Ainsi, la chimie analytique ne considère pas la description du phénomène comme une fin en soi ; elle en recherche l'explication. Sa méthode est hypothético-déductive, comme celle de toute science expérimentale, l'explication pouvant déboucher à son tour sur un raisonnement inductif, sur une loi. Mais dès Lavoisier, la chimie analytique a d'autres champs d'intérêt que la chimie elle-même. Aussi n'hésite-t-il pas à utiliser la même méthode de pesée pour faire le bilan de l'exploitation de ses propriétés agricoles !

De même, la recherche ne peut avancer qu'un pied dans la théorie et l'autre dans le concret. Si une recherche ne peut être considérée que comme fondamentale tant que le chercheur n'a pas une démarche qui tend vers le plus général, vers la découverte de nouvelles lois, les exemples de A.L. Lavoisier et L. Pasteur montrent qu'une recherche, très appliquée à l'origine, n'exclut nécessairement ni le souci ni le succès d'une recherche fondamentale.

En chimie analytique, nous classerons comme appartenant à la recherche appliquée tout travail qui se borne à trouver une nouvelle méthode d'acquisition de l'information ou à l'améliorer. La

recherche de l'explication d'un phénomène, l'élaboration d'un modèle explicatif n'est pas fondamentale tant que le phénomène abordé n'est pas assez général pour qu'on puisse passer au niveau du principe, de la loi. Dans la majorité des cas, la chimie analytique ne suffira pas seule à découvrir une loi et il faudra disposer d'autres informations, physiques ou biologiques par exemple, pour y parvenir éventuellement. Mais on ne peut plus définir une science expérimentale à partir seulement de l'objet de son étude : la chimie a longtemps été définie comme science de la matière, jusqu'à ce que la physique, avec les méthodes qui lui sont propres, revendique le même objet.

La chimie analytique : une méthodologie

Si la chimie analytique perd de vue son utilité comme méthode d'investigation du monde matériel et des phénomènes naturels, si elle prétend être une fin en soi au lieu d'une méthodologie, c'est-à-dire une façon de mettre en œuvre les méthodes, elle est réduite à ne plus se préoccuper que du développement de ses propres méthodes et à errer dans un labyrinthe d'auto-satisfaction. Puisqu'elle n'est pas la seule détentrice des théories qu'elle utilise, ni la propriétaire exclusive d'une méthode, dès qu'elle est mise au point, des auteurs peuvent alors, à juste titre, se poser la question de son identité et de sa pérennité [2].

Cette recherche d'identité est le résultat d'une crise de croissance. La chimie analytique cesse de se définir par rapport à l'objet étudié et affirme ainsi son existence en tant que discipline, avec son caractère et ses fondements qui lui sont propres. L'information chimique n'a plus à être considérée comme une fin pour connaître la matière, mais comme un moyen dont il faut optimiser le rôle.

La spécificité de la chimie analytique peut se comprendre si on l'aborde comme une science expérimentale. Celle-ci s'organise autour de trois pôles schématisés sur la figure 1 : l'objet étudié (un problème), sa représentation (un modèle) et l'information recherchée. Et ce qui fonde son originalité, c'est que l'information est de l'information chimique, c'est-à-dire une information sur la matière qui exige une connaissance des concepts de base de la chimie et qu'on obtient au moyen de méthodes de mesure chimiques, physiques et physico-chimiques, mises en œuvre lors de l'analyse chimique.

La nature même de cette information conduit la chimie analytique à élaborer une démarche qui lui est propre et qui s'appuie

* Conférence donnée, le 31 mai 1990, à l'Institut National Agronomique Paris-Grignon, dans le cadre de la réunion de la division "Chimie analytique" de la Société Française de Chimie.

Directeur de recherche INRA, Laboratoire de chimie analytique, Institut National Agronomique Paris-Grignon, 16, rue Claude-Bernard, 75231 Paris Cedex 05.

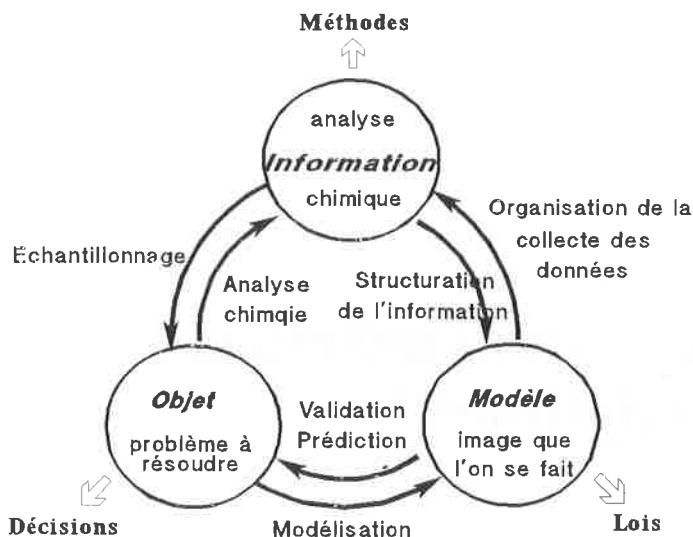


FIGURE 1. - La trinité analytique.

sur les méthodes chimiométriques : méthodes informatiques, statistiques et mathématiques adaptées aux besoins spécifiques des chimistes et dont l'emploi est largement facilité par les ordinateurs.

En un mot, la chimie analytique est donc une méthodologie qui s'appuie sur un ensemble de méthodes chimiques, physiques et physico-chimiques, dites méthodes d'analyse chimique et de méthodes informatiques, statistiques et mathématiques dites méthodes chimiométriques mais qu'on pourrait appeler par analogie mathématico-chimiques. Elle a pour objectif de résoudre des problèmes au moyen d'une information chimique. A cette fin elle doit optimiser l'organisation et améliorer les outils de collecte des données, s'interroger sur le contenu de celles-ci, sur leur validité et leur signification ; elle doit enfin maîtriser leur agencement. C'est une telle définition qui a permis de dire que, "en un sens, la chimie analytique, pour l'analyste, commence là où l'analyse chimique se termine pour le non-analyste" [3].

La figure 1 schématise le cadre de cette méthodologie. Partant d'un problème, la première étape consiste à le traduire sous la forme d'un modèle, en général mathématique. Ce modèle sert à définir la nature de l'information recherchée et organiser sa collecte sur un plan théorique. Ensuite on va concrètement appliquer ce plan de collecte à l'objet de l'étude pour y effectuer un échantillonnage qui servira à obtenir les données aux moyens d'analyses chimiques. Les résultats de ces analyses seront structurés en accord avec le modèle pré-établi afin de vérifier sa validité et ce n'est qu'à ce stade qu'une décision est possible.

En conclusion, la chimie analytique va englober l'analyse chimique et l'ensemble des méthodes d'interprétation des résultats ou de prise de décision : son but est de résoudre un problème posé. Ainsi l'évaluation ou le contrôle de la pollution d'une rivière par le plomb repose sur la mise en œuvre de la chimie analytique, par contre le dosage du plomb dans un échantillon donné relève de l'analyse chimique.

Les méthodes de la chimiométrie

Dès 1970, plusieurs équipes de recherche ont commencé à travailler au développement d'un ensemble de méthodes dont les buts, tels qu'ils apparaissent à travers la définition proposée en 1975 lors de la création de la *Chemometrics Society*, se résument ainsi :

- permettre une meilleure exploitation des données collectées par les analystes et en extraire une information la plus pertinente possible,

- augmenter la qualité des mesures chimiques en améliorant les conditions de leur collecte.

On les a appelées méthodes chimiométriques. La chimiométrie est donc une discipline encore jeune dont on peut juger de la vigueur dans la création de journaux (*Journal of Chemometrics*, *Journal of Intelligent Laboratory Systems*) ou la tenue régulière de congrès (*Chemometrics in Analytical Chemistry*). La littérature montre que les méthodes chimiométriques s'appuient sur des méthodes mathématiques, statistiques ou informatiques connues mais adaptées à la spécificité de l'information chimique [4]. En ce qui concerne les applications, on peut distinguer quatre champs d'investigation :

1. l'acquisition des signaux à la sortie des appareils de mesure, leur stockage, leur mise en forme et leur restauration,
2. la transformation de ce signal, son expression sous forme d'un résultat par étalonnage et sa validation en termes de précision,
3. l'interprétation des résultats en vue de résoudre le problème posé à l'analyste,
4. enfin, la gestion informatisée des laboratoires d'analyse.

Pour atteindre ses buts, la chimiométrie s'appuie largement sur l'outil informatique. Cependant, ce serait une erreur de penser que ses progrès ne sont liés qu'à ceux de l'informatique ou alors il faudrait aussi dire que ceux de la chromatographie liquide haute performance ne sont liés qu'à l'amélioration des joints d'étanchéité des colonnes. Historiquement, son apparition et son développement font partie intégrante de l'évolution de la chimie. Au moment opportun, la chimiométrie a su proposer des solutions nouvelles, basées sur une réflexion approfondie et une approche mathématique des problèmes qui se posaient aux chimistes et principalement aux analystes.

En fait, les développements de la chimiométrie, dans la mesure où ils affectent la conception même des instruments de mesure, prouvent que c'est une tentative pour définir la spécificité de la chimie analytique qui aura pour conséquence d'influencer l'emploi des appareils, la nature des tâches qui leur seront assignées et le mode de fonctionnement profond des laboratoires. En effet, si l'ordinateur reste un outil de recherche et de développement, un objectif permanent est d'intégrer aux appareils les nouvelles techniques chimiométriques afin de les rendre totalement *transparentes* à l'utilisateur. Par exemple, si aujourd'hui tous les instruments de laboratoire délivrent un signal bien propre, correctement lissé et filtré, c'est qu'avant d'en arriver là tous les travaux théoriques avaient été conduits sur ordinateur.

La technologie informatique a commencé à s'implanter en chimie analytique dès les années 1960, en particulier, en ce qui touche à la conception des instruments de laboratoire. Finalement tous les appareils de mesure disponibles aujourd'hui contiennent des microprocesseurs. Ils viennent essentiellement remplacer les composants électroniques câblés des appareils d'autrefois. Ils augmentent la fiabilité des instruments et y ajoutent des fonctions de filtrage du signal, de calcul statistique ou de représentation graphique. Mais si on parle aujourd'hui d'une instrumentation *intelligente*, cette intelligence proviendra des méthodes informatisées installées dans les appareils et non pas des microprocesseurs eux-mêmes.

La chimiométrie et l'analyse chimique

L'analyse chimique part d'un échantillon : elle a pour but immédiat de donner un résultat, qu'il soit quantitatif comme la concentration d'un constituant, ou qu'il soit qualitatif comme la présence ou l'absence d'un radical sur une molécule.

De nombreux ouvrages décrivent les étapes de l'analyse chimique.

Il est évident que ces méthodes viennent en appui de la démarche décrite avant.

La démarche que nous suggérons peut apparaître, selon le problème posé, comme un investissement important pour un labora-

toire. Cette composante économique conditionne aussi le choix des méthodes. En contrepartie, il faut garder présent à l'esprit qu'une récolte optimisée de l'information n'implique pas toujours le choix de la technique de dosage la plus sophistiquée, et donc la plus onéreuse. Grâce aux méthodes chimiométriques, on montre que la mise en œuvre d'un ensemble de techniques simples permet d'apporter une réponse correcte en exploitant les résultats par des méthodes statistiques. En faisant ça, on intègre à la méthode de dosage une certaine forme d'intelligence dont le coût est comparativement très faible (fig. 2).

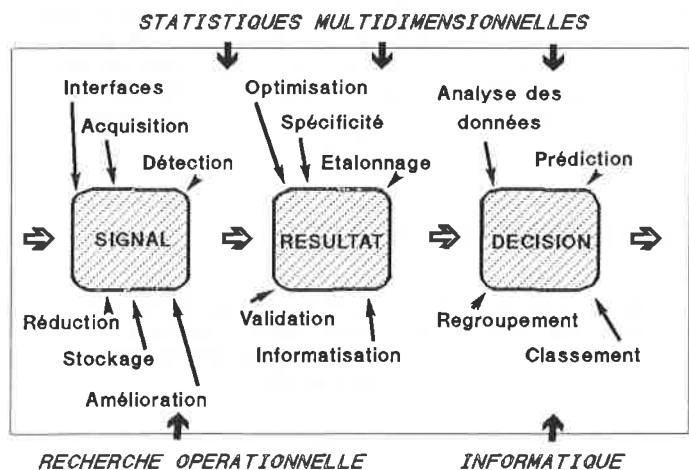


FIGURE 2. - La chimiométrie en chimie analytique.

L'optimisation des méthodes

Les écueils de la méthode "un facteur à la fois"

Une préoccupation fréquente des analystes est d'adapter une méthode existante à un laboratoire ou à un type d'échantillon. Par exemple, il faut ajuster des volumes de réactifs, un temps de réaction ou une température en vue d'atteindre des performances satisfaisantes pour le problème à traiter : il faut optimiser la méthode. En effet, il n'est malheureusement pas de rigueur que les publications fournissent des indications sur l'influence de modifications du protocole opératoire. Même les méthodes normalisées ne sont pas explicites sur ce point. Ainsi, les analystes passent-ils un temps considérable à effectuer des mises au point de méthodes.

Très tôt la chimiométrie s'est posée la question de savoir s'il n'était pas possible, en choisissant une méthodologie appropriée, d'écourter cette étape laborieuse et onéreuse, tout en obtenant des résultats satisfaisants ?

Selon une opinion répandue, la mise au point d'une méthode s'effectue en faisant varier un seul facteur à la fois et en maintenant tous les autres constants. La figure 3 présente une illustration graphique de cette stratégie appliquée à la spectrophotométrie d'absorption atomique sans flamme en tube graphite (SAAsF). L'objectif poursuivi est de trouver les températures de décomposition et d'atomisation qui donnent une absorbance maximum.

Si on ne retient que deux facteurs : (1) la température de décomposition (TD) et (2) la température d'atomisation (TA), la méthode classiquement préconisée consiste, par exemple, à fixer TA à une valeur raisonnable, puis à faire varier régulièrement TD : on étudie un seul facteur à la fois en maintenant l'autre constant. Une première série de points permet de localiser la TD optimale comme l'indique la figure 3. Ensuite, on réalise une deuxième série d'expériences pour lesquelles TD est fixée à son

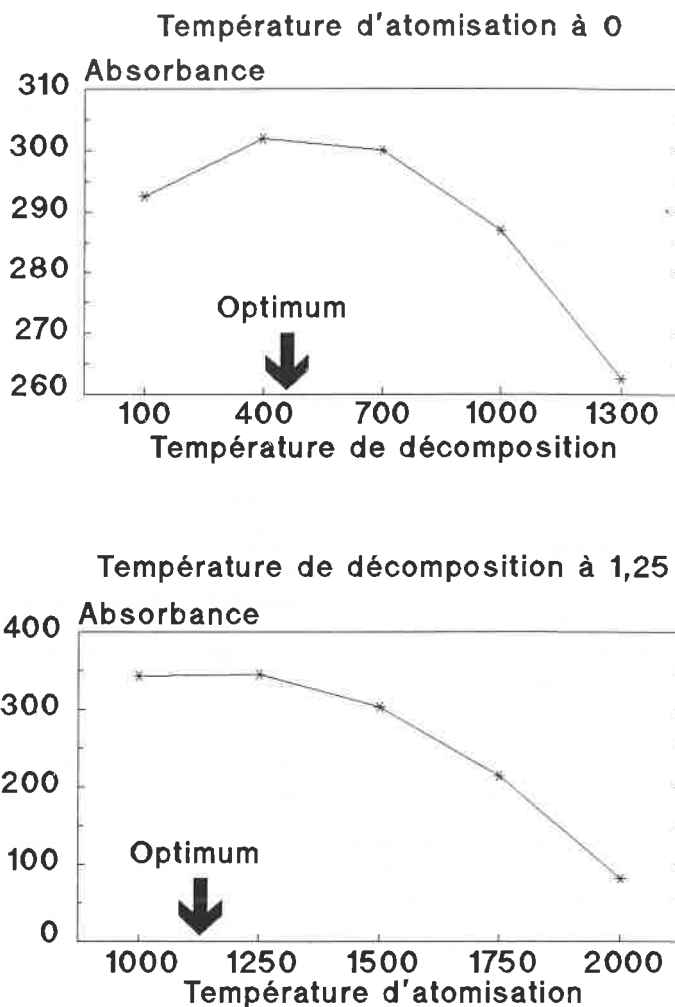


FIGURE 3. - Etude isolée des facteurs.

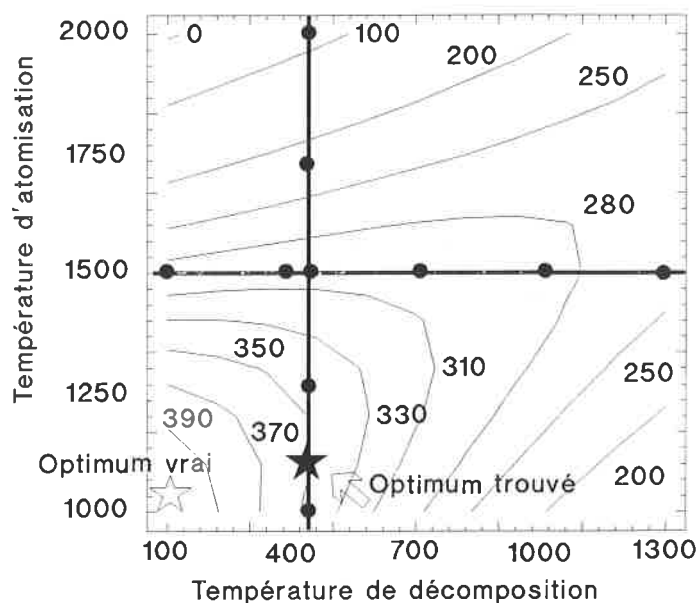


FIGURE 4. - Représentation simultanée des 2 séries de mesure.

optimum et TA varie. Or, si on reporte sur la figure 4, les deux séries de mesure tout en simulant les absorbances théoriques sous la forme de courbes d'isoréponse, il apparaît clairement que

l'optimum trouvé en ne faisant varier qu'un seul facteur à la fois (TD puis TA), ne correspond pas à l'optimum vrai.

Cette erreur tient au fait qu'il y a une interaction entre TA et TD dont la méthode "un seul facteur à la fois" ne tient pas compte. Dès qu'on atteint la ligne de plus faible pente de la surface de réponse on est bloqué. Cette méthode traditionnelle est néfaste car elle conduit à des conclusions d'autant plus erronées qu'il existe des interactions plus fortes entre les facteurs : or, on ne sait pas a priori s'il existe des interactions et cette méthode ne permet pas de les détecter.

Les stratégies d'optimisation

D'autres stratégies ont été proposées qui tiennent compte des interactions entre facteurs. L'une est *déterministe* et part d'équations permettant de décrire les phénomènes mis en cause pour ensuite les contrôler au mieux. Ces équations, construites à partir des lois de la physique ou de la chimie, sont résolues algébriquement de façon à mettre en évidence les solutions les plus favorables. L'autre, *empirique et expérimentale*, consiste à modéliser la méthode étudiée comme une *boîte noire* et à évaluer son comportement à l'aide d'essais planifiés (figure 5). On débouche ainsi sur la méthode des plans d'expériences [5] et la méthode du simplex [6].

Pour la plupart des chimistes, la solution déterministe apparaît comme la plus satisfaisante pour l'esprit et elle est souvent synonyme d'optimisation. Malheureusement, dans bien des cas elle reste inapplicable : il n'y a pas toujours de loi connue qui décrit le phénomène étudié. Il faut regretter que la seconde approche ne soit pas celle que l'analyste choisisse spontanément car elle apporte toujours une solution pratique à la mise au point des méthodes. Cependant, elle exige une certaine planification des essais pour éviter les erreurs d'interprétation.

La méthode à optimiser est vue de l'expérimentateur comme un système fermé (figure 5) dont on va sélectionner une sortie, appelée Y, qui sert de *critère d'optimisation*. On va supposer que les valeurs Y varient lorsqu'un certain nombre de *facteurs contrôlés* par l'expérimentateur prennent des valeurs fixes différentes : ces valeurs fixes sont appelées les *niveaux* des facteurs. La représentation des variations de Y en fonction des niveaux des facteurs forment une *surface de réponse*, spécifique du système étudié. Par convention on notera les p facteurs X_1, \dots, X_p et leurs niveaux pour n essais $x_{11}, x_{21}, \dots, x_{n-1,p}, x_{np}$.

Par exemple, pour une méthode de dosage colorimétrique, la sortie Y sera l'intensité de la réaction, mesurée en hauteur de pic, les facteurs contrôlés seront la température X_1 , la durée de la réaction X_2 et le pH X_3 . Mais on doit aussi supposer qu'il existe des *facteurs connus non contrôlés*, comme les fluctuations de la lampe du colorimètre ou de l'évaporation de la solution dans la cuve de mesure, et des *facteurs inconnus et non contrôlés* qui influencent Y. L'ensemble de ces facteurs non contrôlés for-

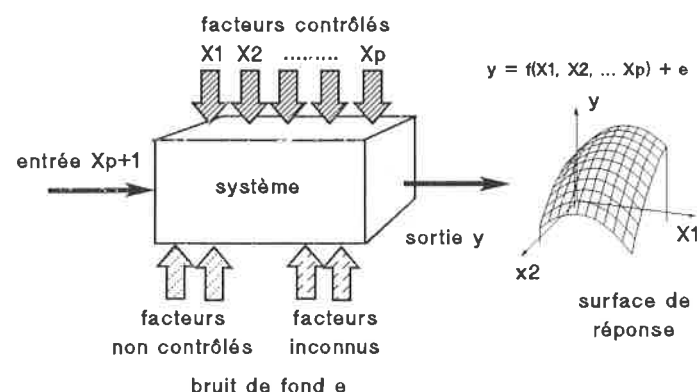


FIGURE 5. - Un système boîte noire.

ment le *bruit de fond* ou l'erreur expérimentale de la méthode, notée e.

A partir de ces deux concepts de facteurs contrôlés et de bruit de fond aléatoire, on peut écrire un modèle mathématique qui les traduise au mieux :

$$Y = f(X_1, X_2, \dots, X_p) + e \quad (1)$$

La fonction f est une fonction quelconque mais, pour se référer aux techniques statistiques existantes, on va supposer que la variable e suit une loi normale. Cette hypothèse semble plausible du fait du théorème de la limite centrale : e est en effet le résultat obtenu par la sommation d'un ensemble de variables aléatoires (les facteurs non contrôlés), de lois quelconques et supposées indépendantes les unes des autres. En se rapportant aux méthodes de régression [7], on peut alors définir Y comme une *variable dépendante* ou expliquée et les facteurs contrôlés comme des *variables indépendantes* ou explicatives.

La régression linéaire multiple

En effet, la méthode de régression linéaire multiple (RLM) permet de proposer une solution algébrique et même une théorie sur l'organisation optimale des essais, appelée méthode des plans d'expériences [8]. Si le modèle (1) est un modèle linéaire, on peut écrire les n essais, sous la forme (2) :

$$\begin{aligned} y_1 &= a_0x_{10} + a_1x_{11} + a_2x_{12} + \dots + a_px_{1p} + e_1 \\ y_2 &= a_0x_{20} + a_1x_{21} + a_2x_{22} + \dots + a_px_{2p} + e_2 \\ &\vdots \\ y_n &= a_0x_{n0} + a_1x_{n1} + a_2x_{n2} + \dots + a_px_{np} + e_n \end{aligned} \quad (2)$$

avec $n > p$

Si on appelle y le vecteur colonne contenant les n résultats y_i , X la matrice des n.p niveaux des facteurs, a le vecteur colonne contenant les coefficients du modèle et e le vecteur colonne contenant les erreurs expérimentales, on peut écrire (2) sous la forme matricielle (3).

$$y = X a' + e \quad (3)$$

Dans la mesure où le nombre d'essais n est supérieur au nombre de variables p, il faut utiliser la méthode des moindres carrés pour calculer des estimations des coefficients a_i [9]. Si on appelle \hat{y}_i les valeurs prédites par le modèle, une fois que ses coefficients a_i seront connus, le principe de cette méthode consiste à rechercher les estimations des a_i sous la contrainte (4) :

$$\Sigma (y - \hat{y})^2 \text{ doit être minimum} \quad (4)$$

La solution de ce problème classique conduit aux équations (5) :

$$\text{Estimateurs des coefficients} \quad (5)$$

$$\hat{a} = [X' X]^{-1} X' y$$

$$\text{Estimateur de la variance résiduelle} \quad (5)$$

$$s^2 = \Sigma (y - \hat{y})^2 / (n - p)$$

Matrice des variances-covariances des coefficients

$$\text{Var}(\hat{a}) = [X' X]^{-1} s^2$$

Cette solution implique que l'on suppose les erreurs e distribuées selon des lois normales indépendantes, de moyennes 0 et de variances égales à σ^2 , estimées par la variance résiduelle s^2 . La matrice de variance-covariance des coefficients donne des indications sur la précision des estimations. A partir de là on peut définir la notion de plan d'expériences optimales. Un plan d'expériences optimal donnera la meilleure précision possible pour les estimateurs des coefficients, c'est-à-dire des variances et covariances les plus petites possibles. On observe une telle circonstance lorsque la matrice X est orthogonale [10].

La surface de réponse et les plans d'expériences

Si on définit l'optimisation comme la recherche des valeurs des X_i qui donnent la valeur Y la plus élevée (ou la plus basse), le

modèle boîte noire permet une définition mathématique de l'optimum : c'est le ou les points où la dérivée première de la fonction f est nulle ; c'est-à-dire les maximums, les minimums et les points stationnaires. Cet optimum théorique n'est pas toujours accessible. Par exemple, un modèle peut conduire à un rendement de réaction dont l'optimum est supérieur à 100 % [11]. Il va falloir définir un optimum pratique : ce sont les valeurs des facteurs X_i donnant la meilleure valeur de Y dans le cadre d'un *domaine expérimental* pré-défini.

Selon cette approche, le problème de l'optimisation est déplacé vers celui de l'exploration de la surface de réponse (SR) : c'est le but de la méthodologie de la surface de réponse (MSR) [10]. Trois stratégies existent,

1. Un maillage régulier. Cette technique laborieuse consiste à faire des expériences réparties selon une maille régulière puis à repérer l'optimum par interpolation. Cette approche pose le problème du choix de la taille de la maille.
2. L'ajustement d'un modèle choisi *a priori*. La méthode des plans d'expériences correspond à ce cas et sera traitée plus complètement ci-après.
3. Une exploration itérative. La méthode du simplexe permet de localiser selon un algorithme itératif un optimum. On progresse pas-à-pas, comme un alpiniste, perdu dans le brouillard qui n'aurait qu'un altimètre et une boussole pour atteindre le sommet.

Ces deux dernières stratégies ont connu des applications très fécondes en chimie analytique. Cependant, quelle que soit la méthode choisie, un acquit fondamental est de *démontrer de façon formelle qu'une optimisation doit être conduite en faisant varier tous les facteurs à la fois* et non pas, comme il est trop souvent dit, en faisant varier un seul facteur et en maintenant tous les autres constants.

L'utilisation d'un plan d'expériences pour explorer une SR et trouver un optimum se fait en 6 étapes dont il faut scrupuleusement respecter la séquence.

1. Choix d'un critère d'optimisation.
2. Sélection des facteurs contrôlés et délimitation exacte du domaine expérimental.
3. Choix d'un modèle, de préférence linéaire.
4. Choix d'un plan d'expériences.
5. Réalisation des essais, calcul des coefficients et localisation de l'optimum pratique.
6. Exécution d'un essai complémentaire à l'optimum pour valider le modèle.

L'exemple suivant montre une mise en œuvre pratique de la méthode des plans d'expériences.

L'optimisation d'un programme d'atomisation

Sélection du critère et des facteurs

Dans l'exemple, l'objectif est de trouver les conditions optimales d'atomisation du plomb en SAAsF. Le critère d'optimisation choisi est la hauteur du pic d'absorbance à 283,3 nm, exprimée en milli-absorbances (mA) ; il correspond à la sensibilité, puisque la concentration de l'échantillon est constante.

Il est apparu que 3 facteurs pouvaient jouer un rôle sur la hauteur du pic :

X_1 : La température de décomposition exprimée en °C (TD),

X_2 : La rampe de décomposition (RD) qui représente le laps de temps nécessaire pour passer de la température de séchage de 100 °C à la température de décomposition ; cette durée s'exprime en secondes,

X_3 : La température d'atomisation exprimée en °C (TA) ; la rampe d'atomisation étant constante et égale à 2 secondes.

Les limites de variation de ces facteurs, reportées au *tableau I*, définissent le domaine expérimental. Elles ont été choisies d'après une connaissance bibliographique du sujet. Ainsi, les

TABLEAU I. - *Limites du domaine expérimental.*

Facteur	Unité		Mini	Maxi	Centre
Décomposition	TD	X_1 °C	130	1 270	700
Rampe	RD	X_2 sec.	5	35	20
Atomisation	TA	X_3 °C	1 020	2 700	1 860

valeurs au centre du domaine sont celles conseillées par le constructeur pour le dosage du plomb à cette longueur d'onde.

Choix du modèle

Pour guider le choix d'un modèle, il est utile de posséder quelques renseignements sur l'allure de la surface de réponse. On sait, par exemple, que l'augmentation d'un facteur amène la réponse à un plateau, suivi éventuellement d'une diminution (*figure 1*). Il est alors intéressant de tenir compte de cette information et de l'introduire dans le modèle. Les modèles dits du second degré qui se présentent sous la forme générale (5) le permettent.

$$Y = a_0 + \sum a_i X_i + \sum a_{ij} X_i X_j + \sum a_{ii} X_i^2 + e \quad (5)$$

avec $i < j$

Le problème que soulève ce type de modèle du second degré est qu'il y a obligatoirement une dépendance linéaire entre les variables X_i et leurs carrés X_i^2 . Les plans d'expériences qui sont proposés tiennent compte de cet inconvénient et essaient d'y pallier par différents artifices.

Construction du plan d'expériences

Aussi n'existe-t-il qu'un nombre limité de plans d'expériences adaptés aux modèles du second degré [12]. Le plan central composite (PCC) apparaît comme un bon choix car le coût de chaque expérience est faible et peut être réalisée rapidement. En se référant aux valeurs conseillées dans la littérature [13], on débouche sur un PCC de 20 essais, indiqués au *tableau II* en valeurs codées. La *figure 6* est la représentation graphique de ce plan : 8 essais dits factoriels forment un cube, 6 essais en étoile sont situés sur la sphère du rayon 1,682 circonscrite à ce cube et 6 répétitions du point central sont au centre de la sphère.

Chaque facteur prend donc 5 niveaux différents indiqués au

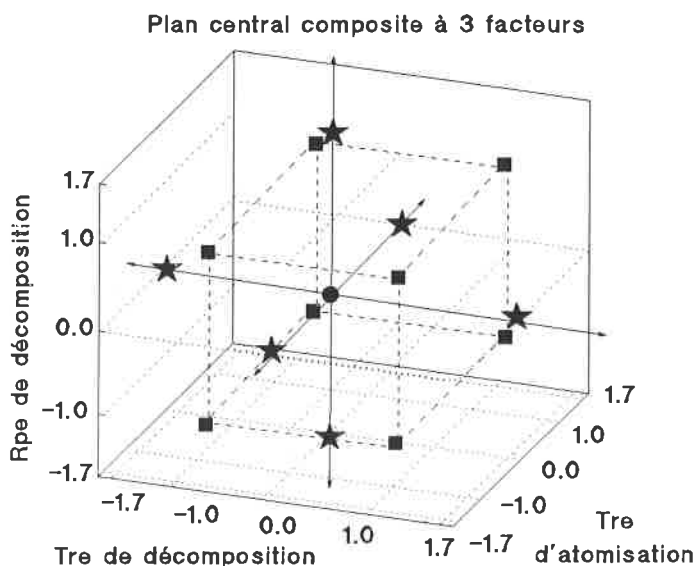


FIGURE 6. - *Plan central composite à 3 facteurs.*

TABLEAU II. - Plan central composite à 3 facteurs.

N°	X ₁	X ₂	X ₃
Essais factoriels			
1	- 1	- 1	- 1
2	+ 1	- 1	- 1
3	- 1	+ 1	- 1
4	+ 1	+ 1	- 1
5	- 1	- 1	+ 1
6	+ 1	- 1	+ 1
7	- 1	+ 1	+ 1
8	+ 1	+ 1	+ 1
Essais en étoile			
9	- 1,682	0	0
10	+ 1,682	0	0
11	0	- 1,682	0
12	0	+ 1,682	0
13	0	0	- 1,682
14	0	0	+ 1,682
Répétitions du point central			
15	0	0	0
16	0	0	0
17	0	0	0
18	0	0	0
19	0	0	0
20	0	0	0

TABLEAU III. - Niveaux des facteurs en valeurs naturelles arrondies.

Facteur	Valeurs codées				
	Z _{min} -1,682	Z ₋₁ -1	Z ₀ 0	Z ₊₁ +1	Z _{max} +1,682
X ₁	130	360	700	1 040	1 270
X ₂	5	11	20	29	35
X ₃	1 020	1 360	1 860	2 360	2 700

tableau III. Pour passer des valeurs codées aux valeurs naturelles permettant l'exécution pratique des essais on applique une règle de trois. Les symboles x_{min} , x_{max} et x_{cent} représentent respectivement les valeurs naturelles des facteurs aux niveaux minimum, maximum et au centre et Z_{min} et Z_{max} les limites codées du domaine expérimental (implicitement on choisit 0 comme valeur de Z au centre). Le tableau IV a été obtenu en remplaçant les valeurs codées du tableau II par les niveaux du tableau III : c'est la liste des essais à réaliser effectivement. Dans la dernière colonne du tableau IV apparaissent les résultats obtenus, exprimés en mA.

Résultats et calculs

Pour calculer sur ordinateur les coefficients on utilise un programme classique de régression linéaire multiple [14]. On obtient ainsi les résultats du tableau V. L'écart type résiduel est d'environ 22 mA, ce qui représente l'erreur expérimentale moyenne sur chaque essai. Le coefficient de détermination est

TABLEAU IV. - Plan central composite à 3 facteurs et résultats.

N°	X ₁ °C	X ₂ sec.	X ₃ °C	y _j mA
1	360	11	1 360	38,0
2	1 040	11	1 360	60,0
3	360	29	1 360	30,0
4	1 040	29	1 360	22,0
5	360	11	2 360	96,0
6	1 040	11	2 360	150,0
7	360	29	2 360	94,0
8	1 040	29	2 360	116,0
9	130	20	1 860	92,0
10	1 270	20	1 860	12,0
11	700	5	1 860	10,0
12	700	35	1 860	114,0
13	700	20	1 020	8,0
14	700	20	2 700	162,0
15	700	20	1 860	114,0
16	700	20	1 860	116,0
17	700	20	1 860	114,0
18	700	20	1 860	112,0
19	700	20	1 860	110,0
20	700	20	1 860	110,0

TABLEAU V. - Coefficients du modèle.

	Coefficient	Ecart type	t	Prob.
a ₀	112,78	9,134	12,346	< 0,001
a ₁	- 3,26	6,060	- 0,538	0,60
a ₂	- 4,27	6,060	- 0,706	0,49
a ₃	41,36	6,060	6,826	< 0,001
a ₁₂	- 7,75	7,918	- 0,978	0,35
a ₁₃	7,75	7,918	0,978	0,35
a ₂₃	1,25	7,918	0,157	0,87
a ₁₁	- 22,22	5,899	- 3,767	< 0,001
a ₂₂	- 2,78	5,899	- 0,471	0,64
a ₃₃	- 10,55	5,899	- 1,789	0,10

égal à 0,867 ; le modèle explique 86,7 % de la variance totale. Ces résultats sont satisfaisants, étant donné les hypothèses très limitatives faites au départ.

A partir de l'écart type de chaque coefficient on peut vérifier si celui-ci est statistiquement différent de 0. Pour réaliser ce test, on utilise le critère t de Student. Dans cette étude on a décidé qu'un coefficient était significatif si la probabilité associée était inférieure ou égale à 10 %. On retient donc les coefficients a₀, a₃, a₁₁ et a₃₃ et on conclut que la rampe de décomposition n'a aucune influence sur le phénomène étudié. Au contraire, les températures d'atomisation et de décomposition ont un rôle au premier et/ou au second degré ; enfin, on n'observe aucune interaction significative entre les facteurs.

Définition de l'optimum et prise de décision

Le modèle finalement conservé est le suivant (6).

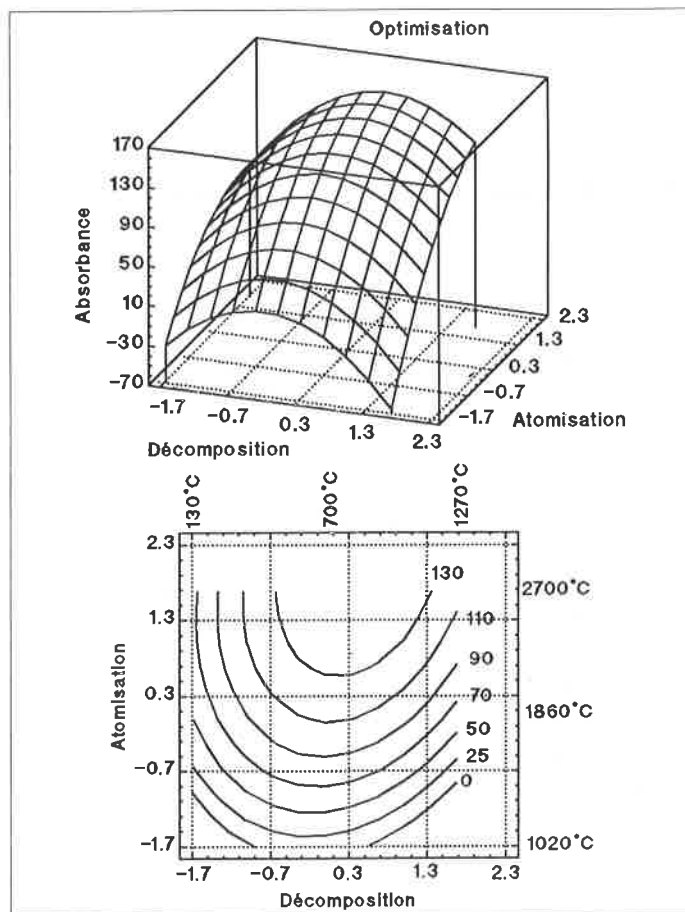


FIGURE 7. - Surface de réponse.

$$\hat{Y} = a_0 + a_2x_2 + a_{11}x_1^2 + a_{33}x_3^2 \quad (6)$$

$$\hat{Y} = 112,78 + 41,36.TA - 22,22.TD^2 - 10,55.TA^2$$

L'optimum mathématique a pour coordonnées TA = 2 943 °C (+ 1,96 en valeur codée) et TD = 700 °C (0). L'absorbance prédite est alors de 155,4 mA. Techniquement, ces réglages opératoires sont impossibles à atteindre, la température TA maximum étant limitée à 2 700 °C par le constructeur et de toutes façons il ne faut, en aucun cas, extrapoler en dehors du domaine expérimental. La représentation graphique du modèle (6) confirme que l'optimum mathématique se situe bien en dehors du domaine expérimental (figure 7). On va donc conserver comme optimum pratique TA = 2 700 °C et TD = 700 °C.

A travers cet exemple, on a vu comment l'optimisation au moyen des plans d'expérience a permis de sélectionner les facteurs dont le rôle est prépondérant pour focaliser l'attention sur eux. Ceci sera très important dans le cas d'un transfert ou d'une standardisation de la méthode entre plusieurs laboratoires ; on sait, par avance, quel facteur risque d'avoir un rôle important sur la reproductibilité.

Les modèles multidimensionnels de contrôle de la qualité

L'approche monodimensionnelle classique

Un autre exemple, destiné à montrer l'apport des méthodes chimiométriques en chimie analytique, est emprunté à l'industrie agro-alimentaire. Il présente la mise en œuvre d'une stratégie de

contrôle de la qualité, basée sur l'analyse chimique des produits et sous-produits de sucrerie [15]. En effet, le sucre blanc de betterave est un produit industriel de haute pureté qui doit légalement contenir au moins 99,76 % de saccharose. Une concentration de 20 mg/kg d'impuretés correspond à 1 point de cendres et baisse d'autant la valeur marchande. Ces impuretés sont présentes dans ce qu'on appelle le "non-sucre" et consistent en diverses espèces minérales et organiques qui, lorsqu'elles sont partiellement éliminées, gênent la cristallisation du saccharose.

Traditionnellement, la qualité des produits intermédiaires et finals de sucrerie s'évalue en déterminant leurs teneurs en matière sèche par réfractométrie, exprimée en °Brix, et en saccharose par polarimétrie [16]. Comme il est délicat de manipuler deux mesures séparément, on calcule leur rapport (en %) qui s'appelle la pureté. Or, la seule mesure de la pureté, ne suffit plus aujourd'hui pour mettre en place un programme efficace d'assurance qualité. C'est pourquoi, une étude analytique plus approfondie a été nécessaire pour développer un modèle permettant de mettre en pratique ce programme d'assurance qualité. Cette étude s'est faite en deux étapes :

1. déterminer les teneurs "moyennes" de plusieurs espèces chimiques indésirables dans les produits finals et intermédiaires de sucrerie, et établir un "atlas" qui serve de référentiel de composition,
2. à partir de cet atlas, bâtir des modèles descriptifs et prédictifs qui permettent de structurer un plan de contrôle de la qualité et améliorer le suivi du procédé de fabrication.

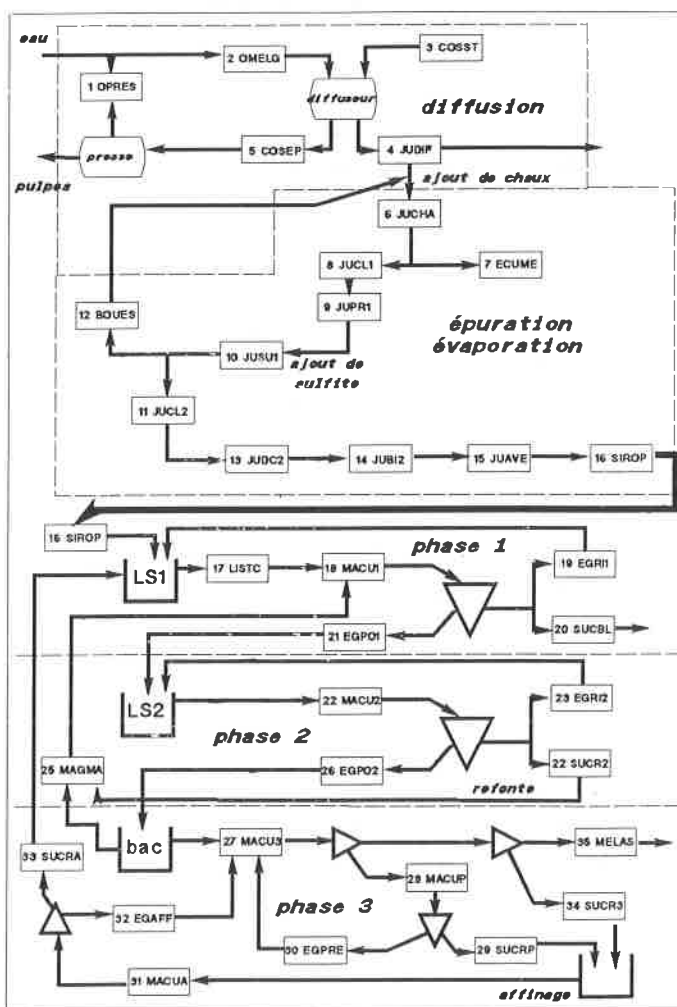


FIGURE 8. - Schéma simplifié de la fabrication du sucre de betterave.

La collecte des données

La figure 8 présente un schéma succinct du procédé de fabrication du sucre, décomposé en 6 étapes, et le tableau VI liste les produits et leurs codes selon une numérotation qui suit leur ordre d'apparition [17].

En partant de différents résultats obtenus dans la littérature, on

TABLEAU VI. - Liste et codes des produits prélevés.

N°	Code	Nom du produit
		diffusion
1	OPRES	eaux de presse
2	OMELG	eaux mélangées
3	COSST	cossettes
4	JUDIF	jus de diffusion
5	COSEP	cossettes épuisées
		épuration
6	JUCHA	jus chaulé
7	ECUME	écumes
8	JUCL1	jus clair de première carbonatation
9	JUPR1	jus clair 1 pré-éaporé
10	JUSU1	jus clair 1 sulfité
11	JUCL2	jus clair deuxième carbonatation
12	BOUES	boues
		évaporation
13	JUDC2	jus clair 2 décalcifié
14	JUBI2	jus clair 2 bisulfité
15	JUAVE	jus avant évaporation
16	SIROP	sirop
		crystallisation 1
17	LISTC	liqueur standard concentrée
18	MACU1	masse cuite de premier jet
19	EGRI1	égout riche de premier jet
20	SUCBL	sucre blanc
21	EGPO1	égout pauvre de premier jet
		crystallisation 2
20	MACU2	masse cuite de deuxième jet
23	EGRI2	égout riche de deuxième jet
24	SUCR2	sucre de deuxième jet
25	MAGMA	magma
26	EGPO2	égout pauvre de deuxième jet
		crystallisation 3
27	MACU3	masse cuite de troisième jet
28	MACUP	masse cuite de pré-turbinage
29	SUCRP	sucre de pré-turbinage
30	EGPRE	égout de pré-turbinage
31	MACUA	masse cuite affinée
32	EGAFF	égout d'affinage
33	SUCRA	sucre affiné
34	SUCR3	sucre de troisième jet
35	MELAS	mélasse

TABLEAU VII. - Liste et codes des constituants analysés.

Code	Nom	Méthode
cations minéraux		
POT	Potassium	photométrie de flamme
SOD	Sodium	
CAL	Calcium	
anions minéraux		
PHO	Phosphate	chromatographie par exclusion d'ions
CHL	Chlorure	
NIT	Nitrate	
SUL	Sulfate	
anions organiques		
CIT	Citrate	chromatographie par échange d'ions
OXA	Oxalate	
MAL	Malate	
LAC	Lactate	
FOR	Formiate	
ACE	Acétate	
FUM	Fumarate	
PCA	Pyrrolidone-carboxylate	

a retenu 15 espèces chimiques qui pouvaient servir de marqueurs de la qualité et qui ont été regroupées en trois sous-ensembles d'après les méthodes analytiques qu'il était nécessaire de développer pour les déterminer, à savoir les cations, les anions minéraux et les anions organiques (tableau VII). Un important travail préalable de mise au point et validation des méthodes analytiques a été entrepris [18]. En effet, si le dosage des cations est effectué depuis longtemps en sucrerie par photométrie de flamme, celui des anions par chromatographie est relativement récent et le protocole non normalisé.

Finalement, la détermination photométrique du potassium (POT) et du sodium (SOD), s'effectue avec une flamme air-butane et celle du calcium (CAL) avec une flamme air-acétylène. Le phosphate (PHO), le nitrate (NIT), le sulfate (SUL) et le chlorure (CHL) sont déterminés par chromatographie ionique sur une silice greffée NH_4^+ , par échange d'anions, suivie d'une détection en spectrophotométrie par diminution d'absorbance [19, 20]. Une autre technique de chromatographie ionique, par exclusion d'ions, est employée pour les anions minéraux accompagnée d'une détection par mesure d'absorbance.

Sur 2 ans, 718 échantillons ont été prélevés sur lesquels ont été réalisées près de 10 000 déterminations. Ces données brutes ont été organisées en une banque de données qui a servi à construire l'atlas de composition des produits de sucrerie.

Devant les problèmes soulevés par l'interprétation de telles volumes de données, une première préoccupation est de mettre en évidence les mesures corrélées et proportionnelles afin d'éliminer la redondance et évaluer plus précisément les évolutions concomitantes. L'analyse en composantes principales (ACP) permet parfaitement d'atteindre ce but [21]. Une des caractéristiques essentielles de cette méthode est de déboucher sur une description approfondie des données et ceci en dehors de toute hypothèse sur la nature de leur dispersion.

L'analyse en composantes principales

L'ACP peut avant tout être considérée comme une méthode descriptive. C'est avant tout l'interprétation géométrique de cette

méthode qui permet de comprendre son principe. L'ACP revient à projeter les individus dans un nouvel espace, soit à un changement de coordonnées, que l'on fait précéder par une transformation des données qui supprime les effets d'échelle dus aux diverses amplitudes des variables. Cette opération est comparable au codage des niveaux des facteurs des plans d'expériences. Si on appelle x_{ij} les valeurs brutes initiales, leurs transformées z_{ij} , dites *centrées réduites*, s'obtiennent selon la formule (7) où m_j et s_j sont respectivement la moyenne et l'écart type de la variable j .

$$z_{ij} = (x_{ij} - m_j) / s_j \quad (7)$$

L'ACP cherche un nouveau système d'axes qui assure la dispersion la plus étendue possible des points. C'est-à-dire que les variances des points, calculées avec leurs nouvelles coordonnées, seront maximum. Ces nouvelles coordonnées f_{ij} s'obtiennent par des combinaisons linéaires des anciennes coordonnées z_{ij} . En notation matricielle, le vecteur f_i représente la *composante principale* i et sa variance, qui doit être maximum, est donnée par (8).

$$\begin{aligned} f_i \cdot f_i &= c_i' Z' Z c_i \\ V &= Z' Z \\ f_i \cdot f_i &= c_i' V c_i \end{aligned} \quad (8)$$

Dans cette formule, c_i est le vecteur propre associé à la valeur propre λ_i , Z la matrice des données centrées réduites et V la matrice des variances-covariances de X . Pour obtenir un seul ensemble de solutions, on pose que $c_i' c_i = 1$, ceci afin que les vecteurs propres soient de norme égale à 1 [22].

Les vecteurs propres représentent les coefficients des combinaisons linéaires des variables initiales et s'interprètent comme leur *poids* ou *inertie* dans la composante principale : on va ainsi évaluer l'importance de chaque variance. Par ailleurs, la valeur propre associée à chaque composante va être interprétée comme une fraction de la variance totale : elle permet de connaître l'importance relative de la composante principale.

Définition d'un plan d'échantillonnage minimum de contrôle

A partir de l'atlas de composition, on a construit une base de données des teneurs moyennes contenant 525 valeurs (soit 35 produits \times 15 champs analytes). Le calcul des composantes principales montre que près de 99 % de la variance totale est expliquée par les 3 premières composantes seulement. La figure 9,

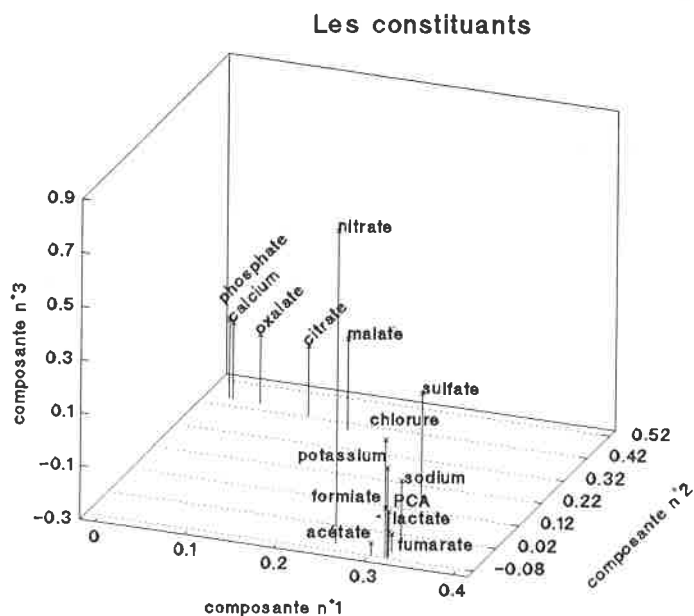


FIGURE 9. - Projections des variables sur les 3 premières composantes.

qui présente les projections des variables (constituants) dans l'espace des 3 premières composantes, révèle qu'il existe trois groupes bien individualisés.

1. Le groupe 1, forme un angle faible avec la composante 1 et rassemble le potassium (POT), le sodium (SOD), le lactate (LAC), le formiate (FOR), l'acétate (FOR), le fumarate (FUM) et le pyrrolidone-carboxylate (PCA).
2. Le groupe 2 associe le phosphate (PHO), le calcium (CAL), le citrate (CIT), l'oxalate (OXA) et le malate (MAL). Ce qui est intéressant, c'est que les directions indiquées de ces groupes, par rapport au centre du repère, sont perpendiculaires ; cette disposition s'interprète comme une absence de corrélation entre les 2 groupes de variables.
3. Quant au sulfate, nitrate et, à moindre titre, chlorure leurs positions à mi-distance des 2 groupes montrent leur rôle particulier, représentatif de la composante 3.

On peut interpréter la première composante (donc le premier groupe de constituants) comme représentative de la pureté. Pour la deuxième composante, on remarque que les ions présents sont ceux qui sont éliminés à l'étape d'épuration. En effet, le poids du phosphate, évaluée par ses coordonnées, est très fort tandis que celui du sulfate est moyen et celui du potassium quasiment nul. Le troisième groupe traduit le rôle particulier du sulfate et du nitrate. En fait, cet unique traitement statistique permet déjà de retrouver la plupart des tendances dégagées par des études individuelles de bilan massique des constituants [18].

La figure 10 représente les projections des produits, et non plus des constituants, sur les axes 1, 2 et 3. Il est tout à fait remarquable qu'à part les écumes (ECUME), la mélasse (MELAS) et l'égot pauvre de deuxième jet (EGPO2), tous les produits soient parfaitement alignés sur une droite située dans le plan 1,3.

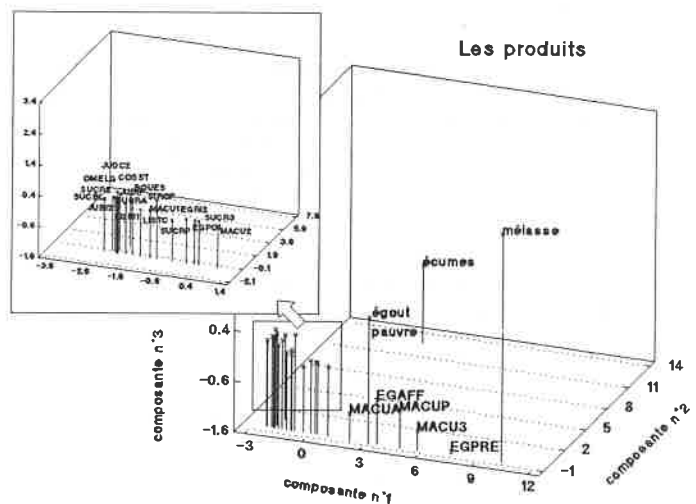


FIGURE 10. - Projections des individus sur les 3 premières composantes.

Comme conséquence de cet alignement exceptionnel, on peut d'abord proposer d'utiliser ce modèle en composantes principales pour établir un classement relatif des produits. En effet, on peut calculer la droite qui passe par ces points et les ranger sur en fonction de leurs distances à un produit de référence, comme le sucre blanc par exemple (figure 11).

On peut ainsi clairement voir que les concentrations de EGPRE sont dans un rapport de 1 à 10 avec celles de SUCBL. L'intérêt de ces résultats est de permettre l'établissement d'un plan de pré-

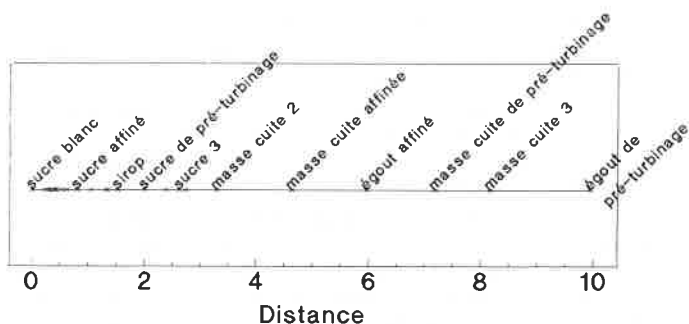


FIGURE 11. - Classement des produits relativement au sucre blanc.

lèvement raisonné. Ainsi on peut, dès à présent, dire qu'il est nécessaire de prélever les 3 produits qui ont des compositions atypiques, à savoir les écumes (ECUME), la mélasse (MELAS) et l'égout pauvre de deuxième jet (EGPO2). Puis doivent apparaître le sucre blanc (SUCBL) et les produits les plus dispersés sur la droite. Pour le reste du programme de contrôle, tout dépend du coût envisagé, mais on a à sa disposition des éléments objectifs pour établir un choix.

Réduction du programme analytique

Une autre possibilité qui est offerte par les méthodes statistiques à variables multiples est d'établir des modèles qui permettent de prévoir les valeurs de certaines variables en fonction des autres et ainsi réaliser des économies sur le plan analytique en réduisant le nombre de mesures. On pourrait employer une technique de régression linéaire multiple mais, dans ce contexte, la co-linéarité des variables biaise le modèle. C'est pourquoi on lui a préféré une méthode de régression sur composantes principales (RCP) [23].

Le choix s'est porté sur la prédiction des cations, à savoir SUL, PHO, CHL et NIT, car ils demandent un changement de colonne chromatographique qui représente un handicap au cas où un autocontrôle, au niveau de l'usine, serait mis en place. En outre, ce choix est prudent car la plupart de ces analytes sont représentatifs du troisième groupe de variables qui n'expliquent qu'une faible part de la variance totale.

Dans ce but, on a refait une ACP en ne conservant que 11 constituants. Cette réduction de l'information ne diminue pas significativement la qualité de la représentation des produits ou des constituants. En partant des coefficients c_{ij} des variables pour les 11 nouvelles composantes principales obtenues par ACP, il est possible de construire des modèles de RCP. Par exemple, pour le sulfate on obtient la fonction de prédiction (9) dont le coefficient de corrélation est supérieur à 0,95.

$$\text{SUL} = -3,930 + 0,126 \text{ POT} + 0,774 \text{ SOD} + 0,546 \text{ ÇAL} + 0,157 \text{ CIT} + 0,179 \text{ OXA} + 0,072 \text{ MAL} - 0,149 \text{ LAC} - 0,336 \text{ FOR} + 0,328 \text{ ACE} - 10,800 \text{ FUM} + 0,075 \text{ PCA} \quad (9)$$

La qualité de prédiction de ce modèle est tout à fait remarquable et pour la vérifier on calcule les intervalles de confiance des valeurs prédites pour chaque produit (figure 12). Ainsi on s'aperçoit qu'il est déjà possible de mettre en place des cartes de contrôle très efficaces à l'aide de ces résultats. En outre, si l'on compare les valeurs calculées aux valeurs observées, par exemple, dans le cas de la masse cuite de premier jet, l'atlas de composition donne une teneur moyenne en sulfate de 1 160 mg/L avec une variation allant de 936 à 1 391, alors que le calcul donne une moyenne de 1 145 mg/L et intervalle de confiance, au niveau de 95 %, allant de 1 130 à 1 160.

Finalement, partant d'une étude lourde comportant plus de 10 000 données, on a mis en évidence une redondance de l'information grâce à l'ACP. On a pu ainsi définir un plan d'échantillonnage réduit à moins de dix produits, au lieu de 35, sur lesquels on détermine 11 constituants au lieu de 15, ce qui évite la mise

intervalles de confiance à 95%

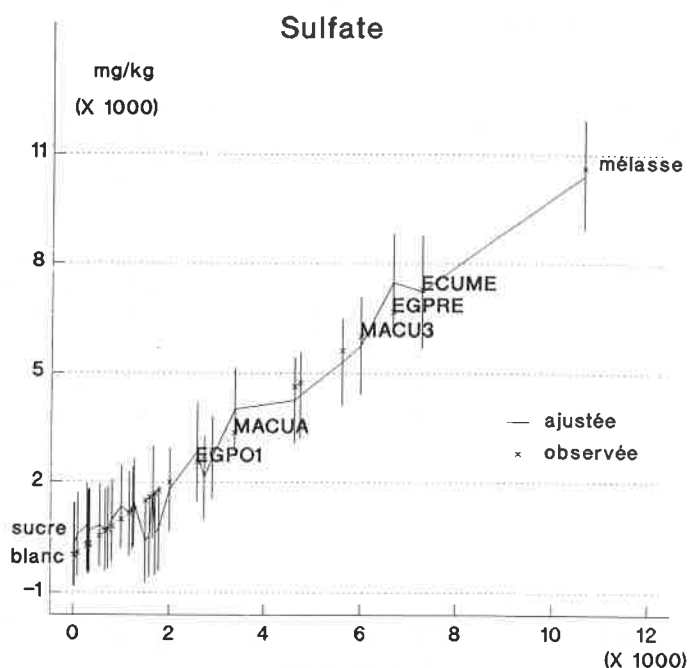


FIGURE 12. - Intervalles de confiance des valeurs prédites du sulfate.

en œuvre d'une méthode. Ce contrôle de qualité mis en place, il suffira de faire chaque année quelques mesures, en début de campagne, pour actualiser les paramètres du modèle mathématique qui permet de prédire le reste des données.

Par rapport à des méthodes monodimensionnelles classiques, comme un bilan massique, l'ACP apporte une vision globale qui facilite les prises de décision. L'autre intérêt de l'ACP est de permettre la construction de modèles prédictifs, soutenant un suivi du procédé et une stratégie de contrôle de la qualité.

Conclusion

Que le but soit d'optimiser une méthode ou de définir un plan de contrôle, on a pu montrer qu'il existe des méthodes plus efficaces, plus rapides et par conséquent moins onéreuses, pour répondre à quelques questions quotidiennes des analystes. Ainsi, pour l'optimisation on a vu qu'étudier un facteur à la fois en maintenant tous les autres constants conduit à des erreurs d'interprétation, alors que les plans d'expériences, en faisant varier tous les facteurs à la fois, évitent cet écueil. En ce qui concerne l'interprétation des données, l'intérêt des méthodes multidimensionnelles est clair : ce sont les seules techniques disponibles pour élaborer des modèles mathématiques qui permettent d'appréhender des mécanismes complexes. En outre, elles démontrent un principe souvent affirmé en chimométrie : il est souvent plus rentable de combiner, à l'aide d'un modèle des mesures simples, plutôt que de mettre en œuvre une technique analytique onéreuse et apparemment plus spécifique.

Nous sommes dans une société d'information, la chimie analytique est riche en informations, il est logique qu'elle puise une part active de son développement dans les nouveaux concepts issus des méthodes de gestion de l'information. Ainsi, on peut penser que les développements à court terme de l'analyse chimique reposent sur un emploi plus général des méthodes de calcul numérique informatisées. A l'heure actuelle il n'existe pas

encore de logiciel polyvalent réellement adapté au besoin du laboratoire, parce que ce besoin ne s'est pas clairement exprimé. Cette demande ne trouvera une réponse que si les analystes savent préciser ce qu'ils veulent et se former aux méthodes automatisées nécessaires à leur pratique : ceci est l'objectif de la chimiométrie.

Références

- [1] Lavoisier A.L., *C.R. Acad. Sciences*, 1775.
- [2] Findeis A.F., Wilson M.K., Meinke W.W., *Anal. Chem.*, 1970, 42, 26A-38A.
- [3] Kateman G., Pijpers F.W., *Quality Control In Analytical Chemistry*, Wiley, New York, 1981.
- [4] Massart D.L., et al., *Chemometrics : a textbook*, Elsevier, Amsterdam, 1988.
- [5] Hunter J.S., *Experimental Designs : Fractionals*, in *Chemometrics* (Kowalski B., Ed.), Reidel Pub. Co., Dordrecht, 1984.
- [6] Deming S.N., *Experimental Designs : Response Surfaces*, in *Chemometrics* (Kowalski B., Ed.), Reidel Pub. Co, Dordrecht, 1984.
- [7] Draper N.R., Smith H., *Applied Regression Analysis*, Wiley, New York, 1966.
- [8] Deming S.N., Morgan S.L., *Experimental Design : A Chemometrical Approach*, Elsevier, Amsterdam, 1987.
- [9] Lest-squares and observations.
- [10] Box G.E.P., Hunter W.G., Hunter J.S., *Statistics for experimenters*, Wiley, New York, 1978.
- [11] Cheyner V., Feinberg M., Ducauze C., *Appl. Env. Microbiol.*, 1983, 45, 634-639.
- [12] Afnor Norme X 060-080. *Plans d'expériences : vocabulaire et indications générales*, Afnor, Paris, 1989.
- [13] Draper R.N., *Technometrics*, 1982, 24, 127.
- [14] SAS, *User's Guide : Statistics, version 5*, SAS Institute Inc., Cary N.C., 1985.
- [15] Foissac S., *Thèse de doctorat*, Institut National Agronomique Paris-Grignon, 1990.
- [16] IRIS, *Méthodes d'Analyses : tome 1*, Syndicat National des Fabricants de Sucre de France, Paris, 1984.
- [17] Mc Ginnis R., *Beet Sugar Technology*, Beet Sugar Development Corporation, USA, 1982.
- [18] Foissac S., Feinberg M., *Science des Aliments* (soumis pour publication).
- [19] Diop A., Jardy A., Caude M., Rosset R., *Analisis*, 1986, 14, 67.
- [20] Jardy A., Rosset R., *Analisis*, 1979, 7, 259.
- [21] Feinberg M., *Anal. Chim. Acta.*, 1986, 191, 75.
- [22] Lefebvre J., *Introduction aux analyses statistiques multidimensionnelles*, Masson et Cie, Paris, 1980.
- [23] Martens H., Naes T., *Multivariate Calibration*, Wiley, New York, 1989.