Plans d'expériences

par **Jacques GOUPY**
Docteur ès sciences
Ingénieur-conseil
Recherche, conseil et formation

L'expérimentateur, quel que soit son domaine d'étude, est toujours confronté au problème difficile de l'organisation optimale de ses essais. Comment obtenir les bonnes informations dans les meilleurs délais et pour le moindre coût ? Telle est la question à laquelle nous allons nous efforcer d'apporter une réponse dans cet article.

Les scientifiques n'ont abordé ce sujet que depuis peu d'années. Les premiers qui se sont penchés sur ce problème sont des agronomes et des statisticiens. Les techniques et les notions qu'ils ont développées sont si générales qu'elles peuvent être utilisées dans tous les domaines. En particulier, la chimie analytique leur offre un vaste champ d'applications.

Cette science de l'organisation des essais est récente puisqu'on peut la faire démarrer avec les travaux de R.A. Fisher (début du vingtième siècle). Aussi bizarre que cela paraisse, elle ne porte pas encore de nom. Nous avons proposé ***Expérimentique*** *ou* ***Expérimentologie****, mais la communauté scientifique n'a pas encore décidé.*

Le but de cette nouvelle science est l'optimisation du choix des essais et de celui de leur enchaînement au cours de l'expérimentation. Nous verrons que ce but peut être atteint à condition que l'expérimentateur se conforme à une méthode rigoureuse et qu'il accepte d'abandonner certaines habitudes. Lorsqu'il aura apprécié la puissance et le bien-fondé de cette nouvelle technique, il en deviendra un adepte fervent et un chaud défenseur.

Nous avons personnellement constaté que la méthode des plans d'expériences est au moins trois à quatre fois plus efficace que les démarches habituelles de conduite des essais, c'est-à-dire qu'elle permet d'arriver aux mêmes résultats avec trois à quatre fois moins d'essais. Ajoutons que cette méthode apporte à l'expérimentateur un puissant outil de réflexion et d'analyse qui lui permettra de conduire son expérimentation avec sûreté et précision.

Les plans d'expériences ont d'abord été utilisés en agronomie. Puis, peu à peu, ils ont été utilisés dans d'autres domaines techniques. Les chimistes les ont adaptés à leurs problèmes. Récemment, les responsables de la qualité ont découvert ces techniques et ils en font maintenant grand usage. L'universalité de ces méthodes devrait les faire employer dans de nombreux domaines. Nous pouvons déjà signaler des réussites en recherche fondamentale, en recherche appliquée, en développement industriel et même en fabrication. Personnellement, nous avons préconisé leur emploi pour réduire le nombre des passages informatiques lors de simulation sur ordinateur.

Le champ d'applications est extrêmement vaste et l'imagination de chacun est libre de trouver de nouveaux domaines et de nouveaux usages. En chimie analytique, on peut les utiliser pour trouver le réglage optimal d'un appareil, pour découvrir les facteurs influant sur le résultat d'une méthode d'analyse, pour améliorer les essais circulaires, pour détecter des erreurs systématiques, etc.

Pour plus de détails sur les travaux de R.A. Fisher, le lecteur pourra consulter les références [1] [2].

Les domaines d'application des plans d'expériences sont étudiés dans les références [6] [11] [16].

1. Principe de la méthode

1.1 Terminologie

1.1.1 Réponses, facteurs, niveaux

Avant d'aborder l'étude des plans d'expériences, il convient de préciser le vocabulaire que nous allons utiliser. Nous nous inspirerons pour cela de la norme AFNOR X 06-080 et des usages des différentes disciplines auxquelles nous ferons appel. Les termes nouveaux que nous définirons seront en lettres grasses dans la suite du texte.

Un phénomène peut toujours être mis sous la forme mathématique suivante :

$$y = f(x_1, x_2, \ldots, x_n) \quad \textbf{(1)}$$

avec y grandeur à laquelle s'intéresse l'expérimentateur : cette grandeur est appelée **la réponse** ou **la grandeur d'intérêt**,

x_i variables sur lesquelles l'expérimentateur peut agir ; ces variables peuvent être continues ou discontinues ; on les appelle les **facteurs** ; les plans d'expériences permettent d'étudier l'influence d'un grand nombre de facteurs sans multiplier exagérément le nombre des essais,

f fonction mathématique qui explique le mieux les variations de la réponse selon les différentes valeurs données aux x_i. Dans le cas des plans d'expériences, cette fonction mathématique est souvent un polynôme dont nous préciserons la forme dans les paragraphes suivants.

Lorsque l'on étudie un facteur x_i, par exemple une température, on ne le fait pas varier dans de larges proportions. On définit toujours, en fonction des besoins de l'étude, une valeur inférieure et une valeur supérieure. Les variations du facteur sont donc limitées par deux bornes. La borne inférieure retenue par l'expérimentateur est appelée **niveau bas** et la borne supérieure, **niveau haut**. On a l'habitude de désigner par le signe moins (–) le niveau bas d'un facteur et par le signe plus (+) son niveau haut.

1.1.2 Coordonnées centrées réduites

Lorsque l'on attribue la valeur – 1 au niveau bas et la valeur + 1 au niveau haut, on effectue deux modifications importantes.

- **On change l'unité de mesure**

Par exemple, si le niveau bas d'un facteur est 60 °C et le niveau haut 80 °C, il y a 20 °C entre ces deux valeurs, soit 20 fois l'unité courante de température. Entre – 1 et + 1, il y a deux unités nouvelles. La nouvelle unité vaut donc 10 °C, on lui donne le nom de **pas**.

- **On déplace l'origine des mesures**

Dans l'exemple choisi, le milieu de l'intervalle [– 1, + 1] correspond à la température de 70 °C. La nouvelle origine, notée zéro, diffère donc de l'origine exprimée en unités courantes.

Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle **variables centrées réduites** (v.c.r.) (centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité).

Le passage des variables d'origine A aux variables centrées réduites x, et inversement, est donné par la formule suivante (A_0 étant la valeur centrale en unités courantes) :

$$x = \frac{A - A_0}{\text{pas}} \quad \textbf{(2)}$$

L'intérêt des v.c.r. est de pouvoir présenter les plans d'expériences de la même manière quels que soient les domaines expérimentaux retenus et quels que soient les facteurs, ce qui donne une grande généralité de présentation à la théorie des plans d'expériences.

1.1.3 Domaine expérimental et domaine d'étude

Nous allons donner une interprétation géométrique de la relation (1). Cette interprétation nous permettra de mieux comprendre la construction des plans d'expériences et de mieux interpréter les résultats. Pour la simplicité de l'exposé, nous prendrons des exemples à deux facteurs, ce qui conduit à des figures à deux ou trois dimensions. Lorsqu'il y a plus de trois facteurs, ce qui est le cas général, il faut raisonner dans des espaces à plus de trois dimensions. Mais, ayant compris ce qui se passe dans un espace à deux dimensions, le lecteur n'aura aucun mal à extrapoler les raisonnements à des espaces à n dimensions.

Pour fixer les idées, prenons l'exemple de l'étude d'une réaction chimique pour laquelle l'expérimentateur cherche à connaître l'influence de la température et de la pression sur le rendement. Construisons d'abord une représentation géométrique. Le premier axe d'un système d'axes cartésiens est attribué à la température et le second à la pression (figure **1**). Le **domaine expérimental** est le plan température x pression dans lequel toutes les pressions et toutes les températures sont possibles. Mais, en général, un expérimentateur limite les variations des facteurs étudiés. Ici, par exemple, la température varie entre 60 °C et 80 °C, la pression entre 1 et 2 bar. On appellera **domaine d'étude** tous les points de la surface délimitée par les niveaux bas et haut de chaque facteur (surface en tramé bleu de la figure **1**).

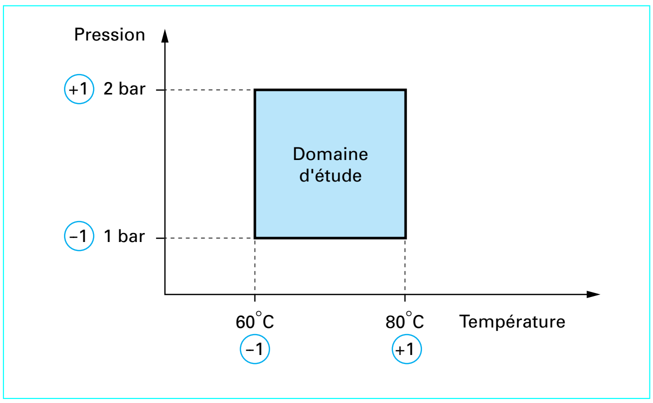


Figure 1 – Définition du domaine d'étude

1.1.4 Facteurs continus, facteurs discrets

Les facteurs peuvent être des grandeurs continues : la température, la pression, les longueurs sont des grandeurs continues. Il est donc possible de leur donner des valeurs comprises entre le niveau bas et le niveau haut et d'en déduire, grâce à un modèle mathématique, toutes les valeurs de la réponse dans le domaine d'étude. Les facteurs continus peuvent prendre plusieurs niveaux d'étude. En v.c.r., les niveaux retenus peuvent être, par exemple, – 2, – 1, 0, + 1, + 2 ou des valeurs non entières comme – 2,6 ou 3,1.

Mais tous les facteurs ne sont pas continus. Si l'on étudie l'influence de différentes personnes sur une réponse, le facteur « personne » n'est pas continu. Il en serait de même d'animaux, d'objets, de couleurs, d'odeurs, etc. Il n'y a aucune valeur intermédiaire entre deux niveaux de ce type de facteur. S'il n'y a que deux niveaux à étudier, on attribuera le signe – à l'un et le signe + à l'autre et cela d'une manière tout à fait arbitraire. Certains facteurs discrets peuvent parfois être ordonnés, par exemple, les trois niveaux « petit », « moyen » et « grand ». Mais ce n'est pas toujours le cas et certains facteurs discrets ne peuvent pas être ordonnés comme, par exemple, les trois personnes Jacques, Pierre et Louis.

Dans certains cas, le même facteur peut se présenter, dans une étude, comme un facteur discret et, dans une autre étude, comme un facteur continu. Par exemple, on peut étudier les ventes de voitures en fonction de leur couleur : bleu, rouge ou vert. Le facteur est ici discret. Mais, si l'on étudie la réfringence de la lumière, on peut classer les couleurs selon leur longueur d'onde. Le facteur couleur est alors continu.

Lorsque l'on effectue une étude, ces différents types de facteur peuvent être simultanément présents. L'espace expérimental est donc constitué d'un mélange d'axes pour facteurs continus et d'axes pour facteurs discrets. Le traitement de ces types de facteur n'étant pas toujours le même, les plans d'expériences doivent être adaptés en conséquence. Dans un premier temps (§ 2, 3, 4), nous étudions les plans factoriels à deux niveaux qui sont applicables à tous les types de facteur sans distinction. Le paragraphe 5 est consacré à certains plans particuliers dont les facteurs ne prennent que deux niveaux. Le paragraphe 6 est consacré aux facteurs continus prenant plus de deux niveaux et dont l'interprétation est basée sur des modèles du second degré. Le paragraphe 7 est consacré aux facteurs discrets ayant plus de deux niveaux. Quant au paragraphe 8 qui traite des mélanges, il ne s'applique qu'aux facteurs continus.

1.2 Méthode

1.2.1 Généralités

Le bon usage des plans d'expériences doit s'inscrire dans une méthode logique et rigoureuse dont la mise en œuvre est de la responsabilité de l'expérimentateur. Si l'on considère le schéma d'acquisition des connaissances (figure **2**), on constate que la première étape consiste à définir le système que l'on se propose d'étudier, puis de poser les questions pour lesquelles on désire des réponses.

Avant de lancer la moindre expérimentation, il est indispensable de vérifier si les solutions que l'on recherche n'existent pas déjà ! Pour cela, il est prudent de faire appel à la théorie, de réaliser une bibliographie, et d'interroger les experts, etc.

Cette première phase étant terminée, l'expérimentateur doit organiser les essais qui lui permettront de répondre aux questions qui restent pendantes. Nous étudierons particulièrement cette étape de la préparation de l'expérimentation, en gardant à l'esprit qu'elle doit faciliter l'interprétation des résultats et permettre une acquisition progressive des connaissances. La méthode des plans d'expériences [6] s'intéresse donc aux trois étapes de réflexion qui sont encadrées dans la figure **2**.

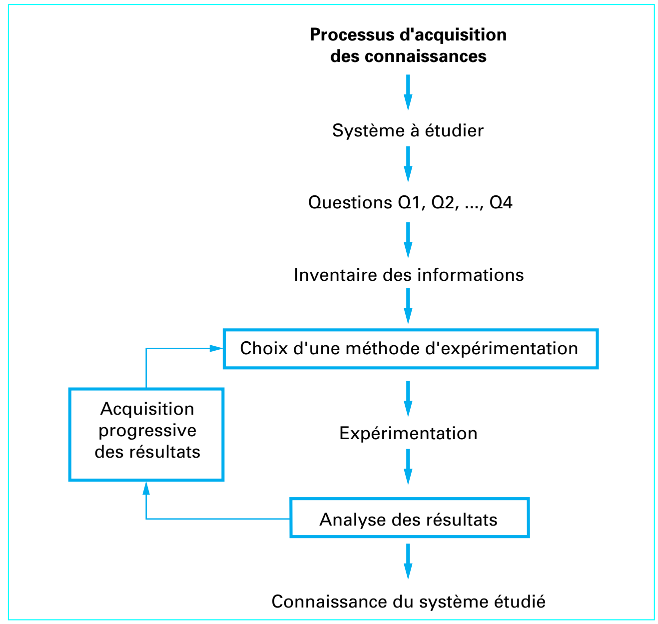


Figure 2 – Schéma d'acquisition des connaissances

1.2.2 Préparation de l'expérimentation

La préparation de l'expérimentation consiste en une réflexion préalable effectuée soit seul, soit en groupe, et permettant :

— de préciser avec soin les réponses qui seront enregistrées à chaque essai ;
— de rechercher tous les facteurs pouvant influer sur le processus étudié ;
— de définir les domaines d'étude de chacun des facteurs ;
— d'envisager les erreurs systématiques possibles pour éventuellement s'en affranchir ;
— de prévoir les contraintes expérimentales possibles ;
— d'organiser des essais supplémentaires pour évaluer l'erreur expérimentale.

Cette réflexion préalable peut durer plusieurs heures au cours desquelles se dégagent avec précision les principaux éléments à prendre en compte pour organiser au mieux les essais. Il est ensuite facile de choisir les plans qui conviennent le mieux au problème posé : nombre de facteurs, ordre des essais (§ 4.2.5), blocking (§ 4.2.2), confusion à éviter, etc.

1.2.3 Analyse des résultats

Lorsque l'on possède les résultats des essais, il faut s'assurer, avant d'entreprendre les calculs :

— qu'il n'y a pas de résultats aberrants ou faux ;
— que les niveaux ont bien été respectés au cours de l'expérimentation ;
— que le modèle mathématique retenu a priori représente bien les résultats d'expériences ;
— que l'on a bien évalué les risques d'ambiguïté.

Étant sûr de la qualité de ses résultats, l'expérimentateur va pouvoir procéder aux calculs et à l'interprétation. En fonction des conclusions, il saura s'il a atteint pleinement son but ou s'il doit envisager une nouvelle série d'essais pour compléter son information.

1.2.4 Acquisition progressive des connaissances

Si l'expérimentateur n'a pas toutes les réponses aux questions posées ou si les premiers résultats soulèvent de nouvelles questions, il va entreprendre des essais supplémentaires. Les conclusions du premier plan lui permettront d'orienter les nouvelles investigations. En particulier, il saura :

— si le domaine d'étude retenu contient les réponses qui l'intéressent ; si oui, il le conservera ; sinon, il saura dans quelle direction il faut aller pour trouver ce qu'il cherche ;

— s'il doit envisager un modèle mathématique différent pour expliquer les résultats des essais ; le modèle du premier degré est parfois insuffisant et des expériences complémentaires devront être entreprises pour établir un modèle du second degré ;

— s'il faut prévoir des essais ou des plans complémentaires pour lever les éventuelles ambiguïtés.

L'organisation des expériences au départ de l'étude est telle que les nouveaux essais viendront s'intégrer harmonieusement aux premiers, évitant ainsi toute perte de temps ou d'argent. Les premiers résultats, s'ils ne répondent pas entièrement aux questions posées, serviront à orienter le choix des nouvelles expériences.

2. Plans factoriels complets à deux niveaux 2^k

Les **plans factoriels complets** à deux niveaux sont les plus simples, ils sont aussi les plus utiles car ils forment la base de tous les débuts d'étude. Les premiers résultats obtenus grâce à ces plans peuvent toujours être complétés par de nouvelles expériences permettant d'atteindre le degré de précision et d'information recherché. Dans le présent paragraphe, nous présentons les plans complets et nous n'abordons les plans fractionnaires que dans le paragraphe 3.

2.1 Plans complets à deux facteurs

Nous allons d'abord décrire le cas simple du plan 2^2. Commençons par expliquer cette notation :

— le 2 en exposant signifie qu'il y a deux facteurs étudiés ;

— l'autre 2 signifie que chaque facteur prend deux niveaux.

Cette notation se généralise immédiatement : pour un plan comportant l'étude de k facteurs prenant chacun deux niveaux, on écrira qu'il s'agit d'un plan 2^k.

Nous avons vu que l'on pouvait donner une représentation géométrique du domaine d'étude. Chaque point de ce domaine représente des conditions opératoires possibles donc une expérience que l'opérateur pourrait réaliser. Le choix des meilleures expériences est le problème fondamental de l'expérimentique. En l'absence de toute information sur la fonction f, on se donne, a priori, une loi d'évolution de la réponse en fonction des variables. Comme on ne désire effectuer, dans un premier temps, que deux essais par facteur, soit deux niveaux par facteur, on adopte une loi du premier degré par rapport à chaque variable. Cette loi est la suivante pour les plans factoriels complets comportant deux facteurs :

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \quad \textbf{(3)}$$

Avec ces hypothèses, on démontre que le meilleur emplacement des points expérimentaux se situe aux sommets du carré représentant le domaine d'étude : points A, B, C et D. La figure **3** illustre les expériences à réaliser et le domaine d'étude. Mais cette représentation géométrique, commode pour comprendre le mécanisme des plans d'expériences, ne peut plus être employée dès que le nombre de facteurs est supérieur à trois. Pour les espaces multidimensionnels, nous adopterons une représentation matricielle. Pour montrer la correspondance entre les deux représentations, géométrique et matricielle, nous allons expliquer la construction de la matrice d'expériences du plan 2^2 associée à la figure **3**.

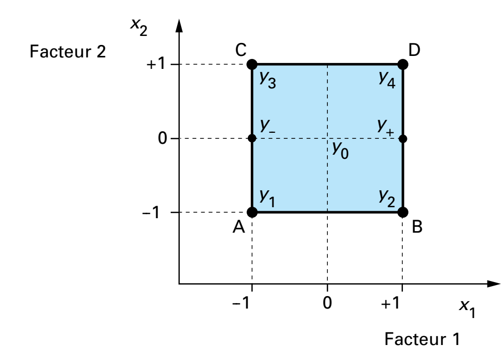


Figure 3 – Meilleur emplacement des points expérimentaux

Tableau 1 – Matrice d'expériences		
Essais à réaliser		
N° essai	**Facteur 1**	**Facteur 2**
1 (A)	– 1	– 1
2 (B)	+ 1	– 1
3 (C)	– 1	+ 1
4 (D)	+ 1	+ 1
Domaine d'étude		
Niveau –	60 °C	1 bar
Niveau +	80 °C	2 bar

La matrice d'expériences est constituée de deux sous-tableaux : le premier définit les essais à réaliser et le second le domaine d'étude (tableau **1**). Le premier sous-tableau comprend trois colonnes ; la première identifie les essais : ici, 1, 2, 3 et 4 ; la seconde et la troisième indiquent les coordonnées (en v.c.r.) des points représentatifs des expériences prévues.

Exemple : l'essai n° 1 est celui pour lequel les deux facteurs étudiés sont aux niveaux bas : – 1 et – 1. Cet essai n° 1 correspond au point A de la figure **3.**

L'essai n° 2 est celui pour lequel le premier facteur est fixé au niveau haut : + 1 et le second facteur est fixé au niveau bas : –1. Cet essai n° 2 correspond au point B.

Le deuxième sous-tableau indique, en unités courantes, les valeurs des niveaux haut et bas de chacun des facteurs. À titre d'exemple, nous avons indiqué des températures et des pressions.

Les deux représentations, géométrique et matricielle, sont équivalentes. Il faut savoir passer de l'une à l'autre pour bien interpréter les résultats des plans d'expériences.

2.2 Effet d'un facteur

L'expérimentateur ayant réalisé les essais est en possession de quatre valeurs de la réponse : y_1, y_2, y_3 et y_4. Il a donc un système de quatre équations à quatre inconnues. Les inconnues étant les coefficients du modèle : a_0, a_1, a_2 et a_{12}. En remplaçant dans la relation (3) les x_i par leur valeur, on obtient :

$$y_1 = a_0 - a_1 - a_2 + a_{12}$$

$$y_2 = a_0 + a_1 - a_2 - a_{12}$$

$$y_3 = a_0 - a_1 + a_2 - a_{12}$$

$$y_4 = a_0 + a_1 + a_2 + a_{12}$$

La résolution de ce système donne :

$$a_0 = \frac{1}{4}\ (+\ y_1 + y_2 + y_3 + y_4) \qquad \textbf{(4)}$$

$$a_1 = \frac{1}{4}\ (-\ y_1 + y_2 - y_3 + y_4) \qquad \textbf{(5)}$$

$$a_2 = \frac{1}{4}\ (-\ y_1 - y_2 + y_3 + y_4) \qquad \textbf{(6)}$$

$$a_{12} = \frac{1}{4}\ (+\ y_1 - y_2 - y_3 + y_4) \qquad \textbf{(7)}$$

■ **Signification de a_0**

Si nous donnons à x_1 et à x_2 la valeur zéro, nous définissons le centre du domaine d'étude. La relation (3) devient alors :

$$y_0 = a_0$$

Le coefficient a_0 est la valeur de la réponse au centre du domaine d'étude. La formule (4) montre également que a_0 peut être considéré comme **la moyenne des quatre réponses**. Nous utiliserons indifféremment ces deux acceptions.

■ **Signification de a_1**

Donnons la valeur zéro à x_2, la relation (3) devient :

$$y = a_0 + a_1\ x_1$$

Puis, donnons maintenant successivement à x_1 les valeurs – 1 et + 1, on obtient les deux réponses y_- et y_+ :

$$y_- = +\ a_0 - a_1$$

$$y_+ = +\ a_0 + a_1$$

d'où :

$$a_1 = \frac{1}{2}(+y_+ - y_-)$$

y_- est la valeur de la réponse pour le point de coordonnées $x_1 = -\ 1$ et $x_2 = 0$, c'est-à-dire celle qui correspond au point milieu du segment AC. Aucune expérience n'a été réalisée en ce point mais, si l'on utilise les relations (4) et (5), on vérifie que y_- est la moyenne des réponses au niveau bas du facteur 1, en effet :

$$y_- = +\ a_0 - a_1$$

$$y_- = \frac{1}{4}\ (+y_1 + y_2 + y_3 + y_4) - \frac{1}{4}\ (-y_1 + y_2 - y_3 + y_4)$$

$$y_- = \frac{1}{2}\ (+y_1 + y_3) \qquad \textbf{(8)}$$

On montrerait de même que y_+ est la moyenne des réponses au niveau haut du facteur 1 :

$$y_+ = \frac{1}{2}\ (+y_2 + y_4) \qquad \textbf{(9)}$$

a_1 est donc la demi-différence entre ces deux moyennes. On peut dire aussi que a_1 représente la moitié de la variation de la réponse quand on passe du niveau bas au niveau haut du facteur 1 (figure **4**). Ce résultat est important car il donne la signification de ce coefficient. C'est la variation de la réponse, variation due au facteur 1 seul, quand on passe du centre du domaine d'étude au niveau haut de ce facteur. a_1 s'appelle l'**effet du facteur 1**.

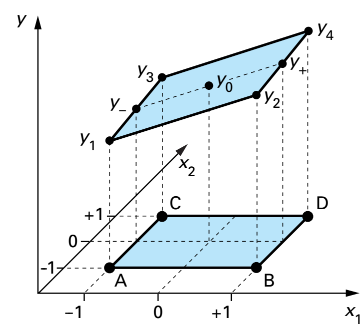


Figure 4 – Représentation géométrique tridimensionnelle d'un plan 2^2 et de la surface de réponse correspondante

On démontrerait de même que a_2 est l'**effet du facteur 2**

Il est commode de représenter l'effet d'un facteur comme l'indique la figure **5** où l'on fait appel au plan de coupe $y\,Ox_1$ passant par $x_2 = 0$ pour le facteur 1.

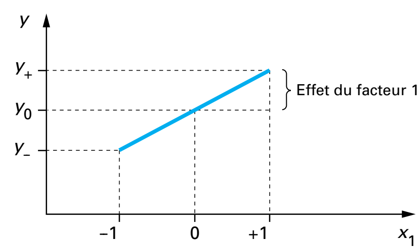


Figure 5 – Représentation de l'effet d'un facteur dans le plan vertical (figure 4) passant par $x_{2\,=\,0}$

2.3 Interaction entre deux facteurs

L'effet d'un facteur a été défini au niveau zéro de l'autre facteur. Mais on peut aussi définir l'effet d'un facteur pour un autre niveau de l'autre facteur. En particulier, on peut introduire l'effet d'un facteur soit au niveau – 1, soit au niveau + 1 de l'autre facteur. L'effet du facteur 1 au niveau – 1 du facteur 2 est la demi-différence entre y_2 et y_1. Et l'effet du facteur 1 au niveau + 1 du facteur 2 est la demi-différence entre y_4 et y_3. Si ces deux effets sont égaux, on dit qu'il n'y a pas d'interaction entre les facteurs. Si ces deux effets sont différents, on dit qu'il y a **interaction** entre les deux facteurs.

Il y a donc interaction lorsque l'effet d'un facteur dépend du niveau de l'autre facteur.

Par définition, la valeur de l'interaction, notée E_{12}, est la demi-différence entre l'effet du facteur 1, e_+, au niveau haut du facteur 2 et l'effet du facteur 1, e_-, au niveau bas du facteur 2.

On a :

$$E_{12} = \frac{1}{2}\,(e_+ - e_-) \qquad \textbf{(10)}$$

$$E_{12} = \frac{1}{2}\left[\frac{1}{2}(+y_4 - y_3) - \frac{1}{2}\,(+y_2 - y_1)\right]$$

En développant :

$$E_{12} = \frac{1}{4}\,(+y_1 - y_2 - y_3 + y_4) \qquad \textbf{(11)}$$

Si l'on compare la valeur de E_{12} à celle de a_{12}, relation (7), on constate qu'elle lui est égale. Si l'on faisait le même calcul pour le facteur 2, en prenant les niveaux haut et bas du facteur 1, on trouverait que l'interaction est la même et qu'elle est égale, elle aussi, à a_{12}. **Ce coefficient a_{12} est donc la mesure de l'interaction entre les deux facteurs.**

2.4 Calcul de l'effet d'un facteur

Reprenons la formule (5) qui donne l'effet du facteur 1 :

$$a_1 = \frac{1}{4}\,(-y_1 + y_2 - y_3 + y_4)$$

On constate :

— que toutes les réponses participent au calcul de l'effet ;

— que chaque réponse est précédée d'un signe et que la suite de ces signes est la même que celle de la colonne du facteur 1 dans la matrice d'expériences, soit – + – + ;

— qu'il y a un coefficient, ici 1/4, dont le dénominateur est égal au nombre d'expériences effectuées.

Le lecteur vérifiera qu'il en est de même pour le facteur 2, la suite des signes étant cette fois – – + +, c'est-à-dire celle de la colonne du facteur 2 dans la matrice d'expériences.

Le calcul pratique d'un effet est le suivant : on multiplie chaque réponse par le signe correspondant de la colonne du facteur ; on additionne les produits et l'on divise la somme par le nombre d'expériences.

2.5 Matrice de calcul des effets

Nous venons de voir que les signes de la matrice d'expériences permettent de calculer les effets. Mais il faudrait pouvoir calculer aussi la moyenne et l'interaction.

Calcul de la moyenne

Le processus de calcul adopté pour les effets peut s'appliquer en utilisant une colonne de signes + puisqu'il n'y a que ce signe dans la formule (4).

Calcul de l'interaction

La suite des signes de la relation (7) est + – – +. Chacun de ces signes provient du produit $x_1\,x_2$ figurant dans la relation (3). On peut retrouver cette suite de signes de la manière suivante : on écrit, en colonne, les signes correspondant à x_1 et à x_2, puis on applique la règle des signes :

x_1	x_2	$x_1\,x_2$
–	–	+
+	–	–
–	+	–
+	+	+

Cette colonne de signes permet de calculer l'interaction par le même mécanisme que celui déjà décrit pour les effets ou la moyenne.

Ayant la matrice d'expériences, il est facile de construire **la matrice de calcul des effets** (tableau **2**) en ajoutant une colonne de signes + pour la moyenne et en calculant celle de l'interaction par la règle des signes.

Tableau 2 – Matrice de calcul des effets

N° essai	Moyenne	Facteur 1	Facteur 2	Interaction 12
1	+ 1	– 1	– 1	+ 1
2	+ 1	+ 1	– 1	– 1
3	+ 1	– 1	+ 1	– 1
4	+ 1	+ 1	+ 1	+ 1

2.6 Notation de Box

La colonne des signes du facteur 1 sera notée par un chiffre gras, soit **1**, celle des signes du facteur 2 par le chiffre gras **2**. La multiplication de ces deux colonnes selon le processus expliqué au paragraphe précédent nous donne la suite des signes de l'interaction entre les facteurs 1 et 2. On peut donc écrire que la colonne des signes de cette interaction est **12**, résultat de la mutiplication de **1** par **2**. On introduit ainsi une algèbre des colonnes de signes, algèbre dont nous venons de définir la multiplication. Si nous multiplions une colonne de signes par elle-même, nous obtenons une colonne qui ne contient que des signes +. Cette colonne sera notée I, d'où :

$$\mathbf{1} \cdot \mathbf{1} = \mathrm{I}$$

On aurait de même : $\mathbf{2} \cdot \mathbf{2} = \mathbf{I}$

Exemple 1 : Un expérimentateur désire étudier le rendement d'une réaction chimique. Après analyse de son problème, il estime qu'il faut retenir deux facteurs : la température et la pression. Il définit également les niveaux haut et bas de chaque facteur (température : 60 °C et 80 °C, pression : 1 et 2 bar). Puis il décide de réaliser les essais selon la démarche des plans d'expériences. Il choisit les quatre essais qui correspondent aux sommets du domaine d'étude (figure **6**).

La matrice d'expériences est analogue à celle que nous avons déjà donnée au paragraphe **2.1**. À chaque expérience, l'expérimentateur enregistre les résultats et il trouve commode d'ajouter à la matrice de calcul des effets une colonne supplémentaire, la colonne « réponses » et une ligne supplémentaire, la ligne des effets. Ainsi, en un seul tableau (tableau **3**), se trouvent résumées toutes les caractéristiques de l'expérimentation.

L'expérimentateur peut maintenant aborder la phase d'interprétation. Il calcule la moyenne, les effets et l'interaction comme nous l'avons indiqué et note ses résultats en bas de chaque colonne. Ces résultats montrent que les deux facteurs étudiés sont influents et que, pour augmenter le rendement, il faut élever la température et la pression. L'absence d'interaction pouvait être prévue en regardant la figure **6** : l'effet de la température pour une pression de 1 bar (niveau –) est le même, cinq grammes de rendement, que pour une pression de 2 bar (niveau +). La modélisation pratique s'obtient en remplaçant les valeurs littérales de la formule **(3)** par les effets calculés :

$$y = 75 + 5\,x_1 + 10\,x_2 \qquad \textbf{(12)}$$

Tableau 3 – Matrice de calcul des effets de l'exemple 1

Numéro de l'essai	Moyenne	Facteur 1 (température)	Facteur 2 (pression)	Interaction 12 (température, pression)	Réponses
1	+ 1	– 1	– 1	+ 1	$y_1 = 60$
2	+ 1	+ 1	– 1	– 1	$y_2 = 70$
3	+ 1	– 1	+ 1	– 1	$y_3 = 80$
4	+ 1	+ 1	+ 1	+ 1	$y_4 = 90$
Effets et interaction	75	5	10	0	
niveau –		60 °C	1 bar		
niveau +		80 °C	2 bar		

Exemple 2 : Supposons qu'un autre expérimentateur soit chargé de la même étude, mais qu'il ait l'idée d'ajouter un catalyseur. Supposons également qu'il choisisse le même domaine d'étude et qu'il réalise, lui aussi, un plan 2^2. L'ensemble de l'expérimentation est résumé dans le tableau **4**. On remarque que les résultats des essais sont différents de l'exemple précédent, fort probablement à cause du catalyseur. Pour calculer la moyenne, les effets et l'interaction, on emploie les mêmes formules et le même processus que précédemment.

Il y a ici une légère interaction entre les facteurs pression et température. La modélisation pratique est :

$$y = 76{,}25 + 6{,}25\,x_1 + 11{,}25\,x_2 + 1{,}25\,x_1\,x_2 \qquad \textbf{(13)}$$

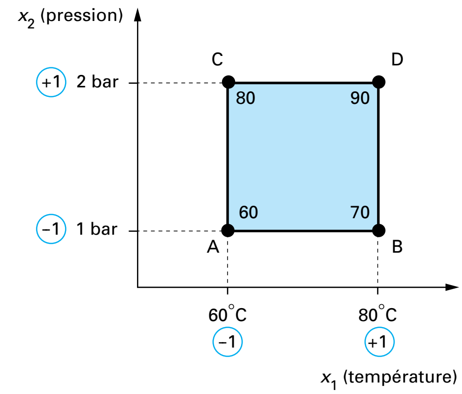


Figure 6 – Représentation géométrique du plan 2^2 correspondant à l'étude d'une réaction chimique

Tableau 4 – Matrice de calcul des effets de l'exemple 2

Numero de l'essai	Moyenne	Facteur 1 (température)	Facteur 2 (pression)	Interaction 12 (température, pression)	Réponses
1	+ 1	– 1	– 1	+ 1	$y_1 = 60$
2	+ 1	+ 1	– 1	– 1	$y_2 = 70$
3	+ 1	– 1	+ 1	– 1	$y_3 = 80$
4	+ 1	+ 1	+ 1	+ 1	$y_4 = 95$
Effets et interaction	76,25	6,25	11,25	1,25	
niveau –		60 °C	1 bar		
niveau +		80 °C	2 bar		

2.7 Plans factoriels 2^k

Il s'agit de plans pour lesquels on étudie k facteurs prenant chacun deux niveaux. Le modèle mathématique adopté a priori est un polynôme prenant en compte la moyenne, les effets de chaque facteur et toutes les interactions entre les facteurs pris deux à deux, trois à trois, quatre à quatre,..., k à k.

$$y = a_0 + \sum a_i x_i + \sum a_{ij}\, x_i x_j + \sum a_{ijl}\, x_i x_j x_l + \ldots + a_{ij\ldots k}\, x_i x_j \ldots x_k \qquad \textbf{(14)}$$

Un plan 2^k comporte 2^k points expérimentaux qui se situent aux 2^k sommets d'un hypercube à k dimensions. Le modèle mathématique contient 2^k coefficients qui sont les inconnues. L'ensemble des résultats d'un plan 2^k conduit donc à un système de 2^k équations à 2^k inconnues, si l'on ne tient pas compte des erreurs expérimentales. Ce système peut se mettre sous forme matricielle :

$$\boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{a} \qquad \textbf{(15)}$$

avec $\boldsymbol{y}$ vecteur ayant pour composantes les réponses de chaque essai, et représenté par une matrice colonne $(2^k, 1)$,

a vecteur ayant pour composantes la moyenne, les effets des facteurs et toutes les interactions, et représenté par une matrice colonne $(2^k, 1)$; ces composantes sont les inconnues que l'on cherche à déterminer,

X matrice carrée $(2^k, 2^k)$ composée de – 1 et + 1 suivant les valeurs des niveaux x_i.

Si nous reprenons le système du plan 2^2 du paragraphe 2.2, la relation (15) prend la forme suivante :

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \end{bmatrix}$$

La matrice ***X*** comporte une colonne de + 1, **I** selon la notation de Box, et trois colonnes ayant chacune autant de signes positifs que de signes négatifs. Si l'on multiplie signe à signe deux quelconques de ces quatre colonnes et que l'on additionne les produits, on trouve zéro. On dit que la matrice est orthogonale. Cette propriété est très importante car, dans ce cas, l'inverse de ***X*** est égale à la transposée de ***X*** divisée par le nombre de lignes n. En effet, d'après Hadamard, on a, pour ce type de matrice, la relation suivante :

$$^t\boldsymbol{X}\,\boldsymbol{X} = n\boldsymbol{I} \tag{16}$$

avec n égal à 2 ou multiple de 4 et ***I*** représentant la matrice unité.

L'opération compliquée de l'inversion d'une matrice se réduit alors à la transposition de ***X***, soit un simple échange de lignes et de colonnes. Le calcul de l'inconnue ***a*** s'effectue à partir de la relation (15), en tenant compte de la relation () :

$$\left.\begin{aligned} {}^t\boldsymbol{X}\,\boldsymbol{y} &= {}^t\boldsymbol{X}\,\boldsymbol{X}\boldsymbol{a} \\ {}^t\boldsymbol{X}\,\boldsymbol{y} &= n\,\boldsymbol{I}\,\boldsymbol{a} \\ a &= \frac{1}{n}\,{}^tX\,y \end{aligned}\right\} \tag{17}$$

Cette relation est valable pour tous les plans factoriels complets. Chaque élément de ***a*** est donc de la forme :

$$a_i = \frac{1}{n}\,[\pm\, y_1 \pm y_2 \pm y_3 \pm \ldots \pm y_n] \tag{18}$$

relation semblable à (4), (5), (6) et (7) qui nous permet de généraliser le processus de calcul que nous avons décrit pour le plan 2^2 à tous les plans 2^k. Nous allons voir, sur un exemple, la construction et l'utilisation de la matrice de calcul des effets d'un plan 2^3.

Exemple 3 : pour l'entretien du réseau routier, les pétroliers sont amenés à préparer des émulsions de bitume. Ces émulsions doivent rester stables depuis leur fabrication jusqu'à leur mise en place. L'étude que nous présentons est la recherche des conditions de stabilité d'une émulsion de bitume en fonction de sa composition. Le responsable a retenu trois facteurs : la teneur en émulsifiant (facteur 1), la teneur en acide chlorhydrique (facteur 2) et la nature du bitume A ou B (facteur 3). Il a fixé les niveaux haut et bas de chacun des facteurs.

Le plan réalisé est un 2^3, trois facteurs et deux niveaux par facteur. Ce plan totalise $2^3 = 8$ essais. La figure **7** donne l'image géométrique du plan d'expériences.
Pour calculer les effets de chacun des facteurs, on construit la matrice de calcul des effets (tableau **5**) selon les règles que nous avons données au paragraphe **2.5**. Une première colonne de signes + pour calculer la moyenne, les trois colonnes de signes de la matrice d'expériences pour calculer les effets principaux des facteurs, trois colonnes pour les interactions d'ordre deux qui sont obtenues par la règle des signes. La huitième colonne, pour l'interaction d'ordre trois, est construite à partir des colonnes **1, 2** et **3**. L'expérimentateur calcule les effets et les interactions.

Il peut rassembler ces résultats dans un tableau (tableau **6**) de présentation dans lequel il pourra ajouter, s'il la possède, la valeur de l'erreur expérimentale (cf. § **4.1.2**, exemple **5**). Ce tableau et les graphiques illustrant les effets (figure **8**) formeront la base de l'interprétation.

À titre d'illustration, nous indiquons une interprétation possible des résultats obtenus. En effet, l'interprétation est de la responsabilité de l'expérimentateur. Cette phase de l'analyse est la plus importante de tout le processus des plans d'expériences. L'expérimentateur engage sa responsabilité, il n'est donc pas question de l'abandonner à une machine ou de vouloir effectuer ce travail automatiquement. C'est l'intelligence et l'expérience du spécialiste qui doivent intervenir ici. C'est à la qualité de l'interprétation que l'on juge l'expérimentateur.
L'interprétation pourrait être la suivante : les interactions sont inférieures à l'erreur expérimentale, elles peuvent donc être considérées comme négligeables. Le facteur 1 a un effet de l'ordre de grandeur de l'erreur, il est considéré comme sans influence. La nature du bitume et la teneur en acide ne peuvent pas être négligées car leur influence sur la stabilité est forte. Cette analyse permet de faire les recommandations suivantes, sachant que l'on cherche une faible valeur de la réponse :

— il faut choisir des bitumes de structure analogue à celle du bitume B ;
— il faut utiliser une forte teneur en acide chlorhydrique ;
— il est conseillé de prendre une faible teneur pour l'émulsifiant.

Cette dernière recommandation, qui tient compte de la non-influence de la teneur en émulsifiant sur la stabilité, a pour but de réduire le coût de fabrication.

Tableau 5 – Matrice de calcul des effets de l'exemple 3

Numéro de l'essai	Moyenne	Facteur 1	Facteur 2	Facteur 3	Interaction 12	Interaction 13	Interaction 23	Interaction 123	Réponses
1	+	–	–	–	+	+	+	–	38
2	+	+	–	–	–	–	+	+	37
3	+	–	+	–	–	+	–	+	26
4	+	+	+	–	+	–	–	–	24
5	+	–	–	+	+	–	–	+	30
6	+	+	–	+	–	+	–	–	28
7	+	–	+	+	–	–	+	–	19
8	+	+	+	+	+	+	+	+	16
Effets et interactions	27,25	– 1	– 6	– 4	– 0,25	– 0,25	0,25	0	
niveau –		faible	faible	A					
niveau +		forte	forte	B					

Tableau 6 – Calcul des effets pour l'étude d'une émulsion de bitume

Moyenne	27,25 ± 0,7
1	– 1 ± 0,7
2	– 6 ± 0,7
3	– 4 ± 0,7
12	– 0,25 ± 0,7
13	– 0,25 ± 0,7
23	0,25 ± 0,7
123	0 ± 0,7

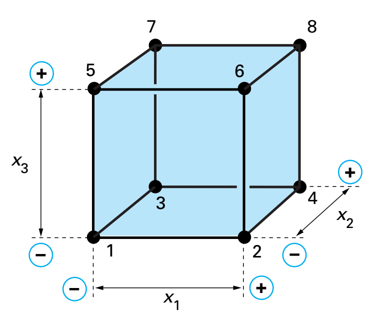
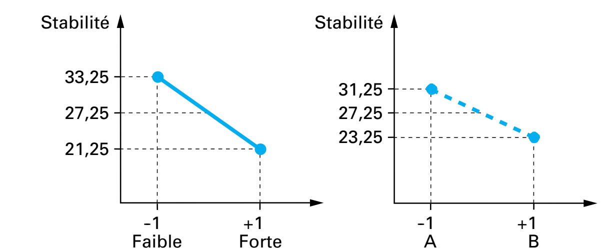


Figure 7 – Représentation géométrique d'un plan 2^3

2.8 Construction des plans factoriels complets

Une manière simple de construire les plans factoriels complets est d'étendre celle que nous avons utilisée pour les plans 2^2 et 2^3.



(a) effet du facteur 2 (concentration en HCl)

(b) effet du facteur 3 (nature du bitume)

Figure 8 – Étude d'une émulsion de bitume

La colonne de signes du facteur 1 est :

$$- + - + - + - + \ldots$$

c'est-à-dire une suite de signes – et + alternés en commençant par un signe –.

La colonne de signes du facteur 2 est constituée de deux signes –, suivis de deux signes +, etc. :

$$- - + + - - + + \ldots$$

Pour le facteur 3, la série de signes est constituée de quatre signes –, suivis de quatre signes +, etc.

Les facteurs suivants, s'il y en a, ont 8, 16, 32 ... signes –, suivis de 8, 16, 32 ... signes +.

Pour chaque facteur, il y a autant de signes – que de signes +.

2.9 Notation des essais des plans factoriels complets

Les signes – et + de chaque facteur étant disposés comme indiqué au paragraphe précédent, les essais sont numérotés dans l'ordre normal des nombres entiers. Cette manière de désigner les essais est appelée la **numérotation classique**.

Il est possible de changer l'ordre des essais : randomisation, dérive ou blocking (§ 4.2). Nous conserverons son numéro à chaque essai quelle que soit sa place dans la présentation. Par exemple, l'essai n° 5 d'un plan 2^3 complet est toujours caractérisé par la suite de niveaux suivante pour les facteurs 1, 2 et 3 :

$$- - +$$

Il existe d'autres manières de construire les plans factoriels et de repérer les essais.

Les plans factoriels complets 2^k deviennent difficiles à utiliser dès que k dépasse la valeur de 3 ou 4. En effet, le nombre d'expériences à réaliser est égal à 2^k, ce qui conduit rapidement à un nombre important d'essais. C'est pour diminuer ce nombre d'essais, sans toucher au nombre de facteurs étudiés, que les plans factoriels fractionnaires ont été mis au point.

3. Plans factoriels fractionnaires à deux niveaux 2^{k-p}

Les **plans factoriels fractionnaires** à deux niveaux sont très utiles car ils permettent de diminuer considérablement le nombre des essais. Mais, pour que les résultats de tels plans soient correctement interprétés, il faut connaître la **théorie des aliases**, théorie que nous allons décrire en nous appuyant sur un plan 2^3.

3.1 Contrastes et relation d'équivalence

Au lieu d'effectuer les huit essais d'un plan 2^3 complet, un expérimentateur ne réalise que quatre expériences. Il choisit quatre expériences de telle manière que les points représentatifs des essais se projettent aux quatre sommets de chaque face du cube définissant le domaine d'étude (figure **9**). Dans un espace à trois dimensions, les quatre points figurent un plan fractionnaire. Si l'un des trois facteurs étudiés est sans influence, on trouvera la même valeur de la réponse pour les niveaux haut et bas de ce facteur et l'on sera ramené à l'étude de deux facteurs. Dans ce cas et avec la disposition proposée, on retrouve toujours un plan complet 2^2 quel que soit le facteur non influent. Si cette représentation géométrique est commode pour comprendre le processus de fractionnement des plans complets, elle devient inutilisable dès qu'il y a quatre facteurs. C'est pourquoi nous allons reprendre la présentation matricielle, valable quel que soit le nombre de dimensions de l'espace expérimental. La traduction matricielle de la disposition géométrique particulière des points d'expériences est l'orthogonalité de la matrice choisie.

Reprenons **l'exemple de l'émulsion de bitume** et écrivons la matrice d'expériences pour quatre essais : 2, 3, 5 et 8. Ces essais sont disposés comme l'indique la figure **9** et leurs coordonnées conduisent à une matrice orthogonale.

La matrice d'expériences du plan fractionnaire est donnée par le tableau **7** où les essais ont été volontairement mis dans un certain ordre.

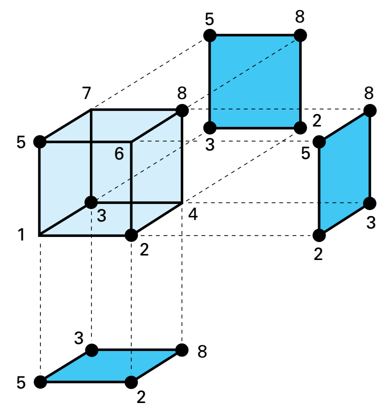


Figure 9 – Projection des points d'expériences d'un plan fractionnaire sur les faces du cube

Tableau 7 – Matrice d'expériences du plan fractionnaire de l'exemple 3

N° essai	Facteur 1	Facteur 2	Facteur 3	Réponses
5	– 1	– 1	+ 1	30
2	+ 1	– 1	– 1	37
3	– 1	+ 1	– 1	26
8	+ 1	+ 1	+ 1	16
Contrastes	– 0,75	– 6,25	– 4,25	
Niveau –	faible	faible	A	
Niveau +	forte	forte	B	

Pour calculer les effets, on applique la même méthode que pour les plans complets : on multiplie les réponses par les signes correspondants des colonnes, on additionne ces produits et on divise leur somme par le nombre d'essais. Dans le cas examiné, nous constatons que les résultats du plan fractionnaire sont tout à fait semblables à ceux du plan complet. Mais, attention, nous n'avons pas calculé les véritables effets de ces facteurs puisque nous n'avons utilisé que la moitié des points expérimentaux. Nous avons calculé ce que l'on appelle un **contraste**. Le contraste ℓ_1 associé au facteur 1 est égal à :

$$\ell_1 = \frac{1}{4}(-y_5 + y_2 - y_3 + y_8) \quad \textbf{(19)}$$

Si nous avions exécuté le plan complet, nous aurions pu calculer l'effet du facteur 1, E_1, et l'interaction 23, E_{23} :

$$E_1 = \frac{1}{8}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$$

$$E_{23} = \frac{1}{8}(+y_1 + y_2 - y_3 - y_4 - y_5 - y_6 + y_7 + y_8)$$

Si l'on additionne ces deux valeurs, on trouve :

$$E_1 + E_{23} = \frac{1}{4}(-y_5 + y_2 - y_3 + y_8) \qquad \textbf{(20)}$$

Le contraste que nous avons calculé est égal à l'effet du facteur 1 augmenté de l'interaction 23. On dit que E_1 et E_{23} sont **aliasés**. Dans le cas de l'émulsion de bitume, l'interaction 23 est très faible et donc :

$$\ell_1 \approx E_1$$

Pour l'interprétation des plans fractionnaires, il est capital de savoir ce que renferme chaque contraste calculé. La notation de Box va nous aider à traiter ce problème. Reprenons la matrice de calcul des effets du plan complet 2^3 et changeons l'ordre des essais (tableau **8**) pour faire apparaître les deux groupes de quatre points de la figure **9**.

Tableau 8 – Matrice de calcul des effets de l'exemple 3

N° essai	I	Facteur 1	Facteur 2	Facteur 3	Interactions 12 13 23 123	Plan
5	+	–	–	+	+ – – +	demi-plan supérieur
2	+	+	–	–	– – + +	
3	+	–	+	–	– + – +	
8	+	+	+	+	+ + + +	
1	+	–	–	–	+ + + –	demi-plan inférieur
6	+	+	–	+	– + – –	
7	+	–	+	+	– – + –	
4	+	+	+	–	+ – – –	

On écrit la matrice de calcul des effets du plan complet et on la divise en deux demi-plans fractionnaires : le **demi-plan supérieur** et le **demi-plan inférieur**.

■ Dans un premier temps, intéressons-nous au **demi-plan supérieur**. Il contient huit colonnes de quatre signes qui sont égales deux à deux. En notation de Box, on peut écrire, par exemple, que la colonne des quatre signes du facteur 1 est égale à celle de l'interaction 23, soit :

$$\mathbf{1} = \mathbf{23} \qquad \textbf{(21)}$$

Si l'on rapproche les relations (20) et (21), on peut énoncer la règle suivante : **dans le demi-plan supérieur, un contraste est la somme des effets et des interactions qui ont la même suite de signes**. On écrira que :

$$\mathbf{1} = \mathbf{23} \text{ est équivalent à } \ell_1 = E_1 + E_{23} \qquad \textbf{(22)}$$

Cette relation d'équivalence est valable dans les deux sens et elle constitue la base de la théorie des aliases. On montrerait que l'on a aussi :

$$\mathbf{2} = \mathbf{13} \text{ est équivalent à } \ell_2 = E_2 + E_{13}$$

$$\mathbf{3} = \mathbf{12} \text{ est équivalent à } \ell_3 = E_3 + E_{12}$$

On peut retrouver ces relations avec la notation de Box : toujours dans le demi-plan supérieur, les deux colonnes de signes + permettent d'écrire :

$$\mathbf{I} = \mathbf{123} \qquad \textbf{(23)}$$

En multipliant par **1** les deux membres de l'équation (23) et en tenant compte des règles de multiplication données au paragraphe 2.5, on retrouve les colonnes identiques :

$$\mathbf{1 \cdot I} = \mathbf{1 \cdot 123}$$

$$\mathbf{1} = \mathbf{23}$$

En multipliant successivement par **2** et par **3** les deux membres de l'équation (23), on obtient les deux relations :

$$\mathbf{2} = \mathbf{13}$$

$$\mathbf{3} = \mathbf{12}$$

La relation (23), **I** = **123**, qui permet de retrouver les égalités de colonnes et les aliases par la relation d'équivalence, s'appelle le **générateur d'aliases.**

■ En considérant le **demi-plan inférieur**, le lecteur vérifiera que les colonnes se correspondent deux à deux, mais avec des signes opposés. On a comme générateur d'aliases :

$$\mathbf{I} = -\,\mathbf{123} \qquad \textbf{(24)}$$

d'où les correspondances de colonnes :

$$\mathbf{1} = -\,\mathbf{23} \quad \mathbf{2} = -\,\mathbf{13} \quad \mathbf{3} = -\,\mathbf{12}$$

et les contrastes calculés avec ce demi-plan indiquent comment les effets et les interactions sont aliasés :

$$\ell'_1 = E_1 - E_{23}$$

$$\ell'_2 = E_2 - E_{13}$$

$$\ell'_3 = E_3 - E_{12}$$

D'une manière analogue à la formule (22), on a la relation d'équivalence suivante pour le demi-plan inférieur :

$$\mathbf{1} = -\,\mathbf{23} \text{ est équivalent à } \ell'_1 = E_1 - E_{23} \qquad \textbf{(25)}$$

3.2 Construction pratique d'un plan fractionnaire

Lorsqu'on examine les quatre premières colonnes du demi-plan supérieur, on constate que l'on retrouve les colonnes de signes d'un plan complet 2^2. On a simplement étudié le facteur supplémentaire en utilisant la colonne des signes de l'interaction 12. La construction pratique des plans fractionnaires est basée sur cette remarque.

■ On choisit un plan complet et l'on écrit sa matrice de calcul des effets. On appelle cette matrice le **plan de base.**

■ Dans ce plan de base, on choisit une colonne de signes correspondant à une interaction pour étudier le facteur supplémentaire. Pour chaque essai, on attribue à ce facteur les niveaux indiqués par les signes de l'interaction choisie.

En notation de Box, on écrira que le facteur 3 est étudié sur les signes de l'interaction 12 :

$$\mathbf{3} = \mathbf{12}$$

et l'on sait que le contraste ainsi calculé est la somme de l'effet du facteur 3 augmenté de l'interaction entre les facteurs 1 et 2.

Si l'on prend la matrice de calcul des effets d'un plan 2^3 comme plan de base, on a quatre interactions disponibles : 12, 13, 23 et 123.

N°	1	2	3	12	13	23	123	I
1	−	−	−	+	+	+	−	+
2	+	−	−	−	−	+	+	+
3	−	+	−	−	+	−	+	+
4	+	+	−	+	−	−	−	+
5	−	−	+	+	−	−	+	+
6	+	−	+	−	+	−	−	+
7	−	+	+	−	−	+	−	+
8	+	+	+	+	+	+	+	+

Les trois premières colonnes permettent d'étudier trois facteurs. Un quatrième facteur peut être étudié sur la colonne de signes de l'interaction 12. On écrira :

$$\mathbf{4 = 12}$$

d'où le générateur d'aliases :

$$\mathbf{I = 124}$$

Ce générateur d'aliases permet de calculer la manière dont les effets sont aliasés avec les interactions. En multipliant ce générateur successivement par **1**, **2**, **3** et **4**, on obtient :

$\mathbf{1 = 24}$ est équivalent à $\ell_1 = E_1 + E_{24}$

$\mathbf{2 = 14}$ est équivalent à $\ell_2 = E_2 + E_{14}$

$\mathbf{3 = 1234}$ est équivalent à $\ell_3 = E_3 + E_{1234}$

$\mathbf{4 = 12}$ est equivalent à $\ell_4 = E_4 + E_{12}$

On aurait pu aliaser le facteur 4 sur une autre interaction, on aurait eu d'autres valeurs des contrastes.

Il est tout à fait possible d'étudier deux facteurs supplémentaires. On choisit deux colonnes de signes. Par exemple, on peut choisir la colonne 12 pour le quatrième facteur et la colonne 13 pour le cinquième facteur. On a :

$$\mathbf{4 = 12}$$

$$\mathbf{5 = 13}$$

d'où les deux générateurs d'aliases indépendants :

$$\mathbf{I = 124}$$

$$\mathbf{I = 135}$$

Si l'on multiplie ces deux générateurs d'aliases indépendants membre à membre, on obtient un troisième générateur :

$$\mathbf{I \cdot I = 124 \cdot 135}$$

$$\mathbf{I = 2345}$$

Deux facteurs supplémentaires introduisent donc un **groupe de générateurs d'aliases** ou **GGA** comportant quatre termes :

$$\mathbf{I = 124 = 135 = 2345}$$

On utilise ce GGA pour savoir comment les facteurs et les interactions sont aliasés dans les contrastes que l'on calcule avec ce plan fractionnaire. Par exemple, le contraste ℓ_1 sera déterminé en multipliant tous les termes du GGA par la colonne **1** qui a servi à le calculer.

$$\mathbf{1 \cdot I = 1 \cdot 124 = 1 \cdot 135 = 1 \cdot 2345}$$

En simplifiant :

$$\mathbf{1 = 24 = 35 = 12345}$$

La relation d'équivalence donne les quatre termes du contraste :

$$\ell_1 = E_1 + E_{24} + E_{35} + E_{12345}$$

On peut généraliser cette méthode et utiliser toutes les colonnes d'un plan de base. Par exemple, sur le plan de base bâti sur la matrice de calcul des effets d'un 2^3, on peut étudier sept facteurs (voir exemple 4) et, s'il s'agit de la matrice de calcul des effets d'un plan 2^4, on peut étudier quinze facteurs.

Exemple 4 : étude du réglage d'un spectrofluorimètre.

Un technicien se propose d'améliorer le réglage de la sensibilité d'un spectrofluorimètre. Une discussion approfondie avec des spécialistes lui a montré qu'il fallait envisager l'examen de sept facteurs pouvant être influents sur la sensibilité.

Facteur 1 : largeur de la fente d'excitation.
Facteur 2 : largeur de la fente d'émission.
Facteur 3 : température de l'échantillon.
Facteur 4 : vitesse de balayage.
Facteur 5 : gain de l'appareil.
Facteur 6 : tension du photomultiplicateur.
Facteur 7 : amortissement de la plume d'enregistrement.

Il désire passer le moins de temps possible à la mise au point de ce réglage, il décide donc d'employer un plan d'expériences fractionnaire construit à partir d'un plan de base 2^3. Les quatre interactions y sont utilisées pour étudier les facteurs supplémentaires. Les facteurs supplémentaires sont aliasés ainsi :

$$\mathbf{4 = 123}$$

$$\mathbf{5 = 12}$$

$$\mathbf{6 = 23}$$

$$\mathbf{7 = 13}$$

■ **Calcul des contrastes**

Les quatre générateurs d'aliases indépendants sont :

$$\mathbf{I = 1234 = 125 = 236 = 137}$$

Les générateurs dépendants se calculent à partir des générateurs indépendants en les multipliant 2 à 2, 3 à 3 et 4 à 4. Le nombre de générateurs dépendants obtenus pour la multiplication 2 à 2 est donné par le nombre de combinaisons de 4 objets 2 à 2, soit C_4^2.

Multiplication 2 à 2 ($C_4^2 = 6$) :

$$\mathbf{1234 \cdot 125 = 345}$$

$$\mathbf{1234 \cdot 236 = 146}$$

$$\mathbf{1234 \cdot 137 = 247}$$

$$\mathbf{125 \cdot 236 = 1356}$$

$$\mathbf{125 \cdot 137 = 2357}$$

$$\mathbf{236 \cdot 137 = 1267}$$

Multiplication 3 à 3 ($C_4^3 = 4$) :

$$\mathbf{1234 \cdot 125 \cdot 236 = 2456}$$

$$\mathbf{1234 \cdot 125 \cdot 137 = 1457}$$

$$\mathbf{125 \cdot 236 \cdot 137 = 567}$$

$$\mathbf{1234 \cdot 236 \cdot 137 = 3467}$$

Multiplication 4 à 4 ($C_4^4 = 1$) :

$$\mathbf{1234 \cdot 125 \cdot 236 \cdot 137 = 1234567}$$

Le plan complet 2^7 contient 128 effets et interactions, le plan $2^{7-4} = 2^3$ ne permettra de calculer que huit contrastes. Il y aura donc seize termes dans chaque contraste, termes que l'on va retrouver grâce au groupe des générateurs d'aliases (seize termes) :

$$\mathbf{I = 1234 = 125 = 236 = 137 = 345 = 146 = 247 = 1356}$$

$$\mathbf{= 2357 = 1267 = 2456 = 1457 = 567 = 3467 = 1234567}$$

Le calcul complet est présenté ici pour le facteur 1. On multiplie chaque terme du GGA par **1** et l'on additionne les seize termes trouvés pour avoir le contraste ℓ_1 :

$$\ell_1 = \mathbf{1} + \mathbf{234} + \mathbf{25} + \mathbf{1236} + \mathbf{37} + \mathbf{1345} + \mathbf{46} + \mathbf{1247} + \mathbf{356} + \mathbf{12357}$$
$$+ \mathbf{267} + \mathbf{12456} + \mathbf{457} + \mathbf{1567} + \mathbf{13467} + \mathbf{234567}$$

Si l'on néglige les interactions d'ordre supérieur à deux, on peut écrire :

$$\ell_1 = \mathbf{1} + \mathbf{25} + \mathbf{37} + \mathbf{46} + \ldots$$

Le calcul est le même pour les autres contrastes :

$$\ell_2 = \mathbf{2} + \mathbf{15} + \mathbf{36} + \mathbf{47} + \ldots$$

$$\ell_3 = \mathbf{3} + \mathbf{17} + \mathbf{26} + \mathbf{45} + \ldots$$

$$\ell_4 = \mathbf{4} + \mathbf{16} + \mathbf{27} + \mathbf{35} + \ldots$$

$$\ell_5 = \mathbf{5} + \mathbf{12} + \mathbf{34} + \mathbf{67} + \ldots$$

$$\ell_6 = \mathbf{6} + \mathbf{14} + \mathbf{23} + \mathbf{57} + \ldots$$

$$\ell_7 = \mathbf{7} + \mathbf{13} + \mathbf{24} + \mathbf{56} + \ldots$$

Après avoir calculé les contrastes, le technicien établit le plan d'expériences, réalise les essais et rassemble les résultats dans une matrice de calcul des effets qui, dans ce cas, est identique à la matrice d'expériences puisque toutes les interactions ont été utilisées (tableaux **9** et **10**).

Deux facteurs seulement semblent influents sur la sensibilité (figure **10**) :

— la fente d'émission (2) ;
— la tension du photomultiplicateur (6).

Dans le domaine d'étude et pour l'appareil considéré, on peut faire les recommandations suivantes :

— comme on pouvait s'y attendre, il faut élargir la fente d'émission pour augmenter la sensibilité de l'appareil ;
— par contre, et cela est beaucoup plus inattendu, dans les conditions de l'expérimentation, il faut diminuer la tension du photomultiplicateur pour augmenter la sensibilité de l'appareil ;
— les cinq autres facteurs sont sans influence et peuvent être réglés à des valeurs intermédiaires situées entre le niveau haut et le niveau bas de chacun des facteurs considérés.

Tableau 9 – Matrice d'expériences/matrice de calcul des effets pour le réglage d'un spectrofluorimètre

N° de l'expérience	Fente d'excitation (1)	Fente d'émission (2)	Température de l'échantillon (3)	Vitesse de balayage (4) 4 = 123	Gain (5) 5 = 12	Tension du photomultiplicateur (6) 6 = 23	Amortissement (7) 7 = 13	Moyenne	Réponses (sensibilité)
1	–	–	–	–	+	+	+	+	1,22
2	+	–	–	+	–	+	–	+	0,90
3	–	+	–	+	–	–	+	+	5,33
4	+	+	–	–	+	–	–	+	5,64
5	–	–	+	+	+	–	–	+	3,89
6	+	–	+	–	–	–	+	+	3,88
7	–	+	+	–	–	+	–	+	2,82
8	+	+	+	+	+	+	+	+	2,33
Niveau –	2,5 nm	2,5 nm	20 °C	20 nm/min	1	310 V	2		
Niveau +	7,5 nm	7,5 nm	40 °C	100 nm/min	10	460 V	4		
Sensibilité	– 0,06	0,78	– 0,02	– 0,14	0,02	– 1,43	– 0,06	3,25	

Tableau 10 – Valeur des effets pour le réglage de la sensibilité d'un spectrofluorimètre

Moyenne	3,25
1	– 0,06
2	0,78
3	– 0,02
4	– 0,14
5	0,02
6	– 1,43
7	– 0,06

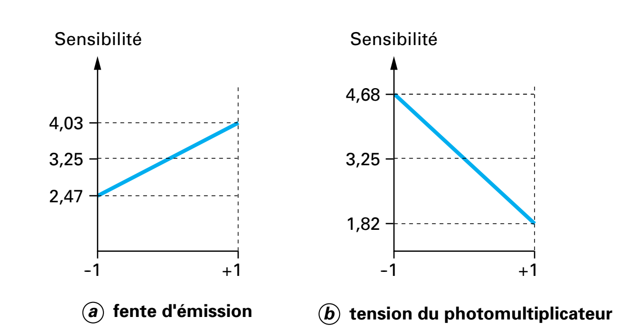


Figure 10 – Effets des facteurs fente d'émission et tension du photomultiplicateur sur la sensibilité du spectrofluorimètre

3.3 Notation des plans factoriels fractionnaires

Pour trois facteurs prenant deux niveaux, le plan complet est noté 2^3. Il comporte huit essais. Le plan fractionnaire, moitié du plan complet, n'a que quatre essais soit 1/2 2^3 ou 2^{3-1} essais. Chaque chiffre de cette notation a une signification :

— le 3 signifie qu'il y a trois facteurs étudiés ;

— le 2 signifie que chaque facteur prend deux niveaux ;

— le 1 signifie qu'il y a un facteur supplémentaire par rapport au plan complet sur lequel est construit le plan de base.

Le plan de base 2^3 peut avoir un facteur supplémentaire, on le notera 2^{4-1} : quatre facteurs étudiés, deux niveaux par facteur et un facteur supplémentaire.

Le plan de base 2^3 peut avoir deux facteurs supplémentaires, on le notera 2^{5-2} : cinq facteurs étudiés, deux niveaux par facteur et deux facteurs supplémentaires.

Un plan fractionnaire à deux niveaux avec lequel on étudie k facteurs dont p supplémentaires se note 2^{k-p}.

Les plans fractionnaires permettent d'étudier un grand nombre de facteurs avec un nombre très restreint d'essais. Certains techniciens pourront s'inquiéter de ne pas avoir tous les résultats. La démarche que nous avons préconisée est « **l'acquisition progressive des connaissances** ». Cela signifie qu'après une première série d'essais obtenue avec un plan fractionnaire l'expérimentateur sait analyser ses résultats et détecter les ambiguïtés (aliases entre les effets et les interactions). Il est alors en mesure de choisir les nouveaux essais pour lever ces ambiguïtés. Deux outils sont à la disposition de l'expérimentateur pour analyser les premiers résultats : la théorie des aliases que nous venons de présenter et les hypothèses d'interprétation que nous allons indiquer.

3.4 Hypothèses d'interprétation

Tous les plans fractionnaires posent le même problème d'interprétation des résultats. Les hypothèses de travail le plus souvent retenues sont les suivantes :

— les interactions du troisième ordre ou d'ordre plus élevé sont considérées comme négligeables ;

— si un contraste est nul, cela peut signifier :
a) que les effets et les interactions aliasés sont tous nuls,
b) que les effets et les interactions aliasés se compensent.

L'hypothèse *a*) est la plus probable et c'est elle qui est généralement retenue ;

— si deux effets sont faibles, on supposera que leur interaction l'est aussi ;

— si deux effets sont forts, on se méfiera de leur interaction qui peut également être forte.

Pour interpréter les résultats d'un plan fractionnaire, il est essentiel de savoir comment sont aliasés les effets et les interactions. Les hypothèses présentées ici sont valables dans la plupart des cas mais il est toujours possible, et même parfois recommandé, d'en adopter d'autres en fonction du problème et des risques encourus.

4. Erreurs expérimentales

4.1 Définition et estimation des erreurs expérimentales

4.1.1 Erreurs aléatoires et erreurs systématiques

Si l'on effectue plusieurs fois la même mesure, on ne trouve pas toujours le même résultat. Il y a une dispersion des mesures. On caractérise le plus souvent la série de mesures par deux chiffres : la moyenne et l'écart-type qui est un indice de la dispersion des mesures autour de la moyenne. Les erreurs ainsi constatées sont appelées les erreurs aléatoires.

À côté de l'erreur aléatoire, il se peut qu'il y ait des variations d'ensemble des mesures. Tous les résultats peuvent être plus forts ou plus faibles d'une valeur donnée. Il y a un décalage constant de toutes les mesures. Cette erreur n'est plus aléatoire, elle est systématique.

L'erreur totale est la somme de ces deux types d'erreur :

Erreur totale	=	Erreur aléatoire	+	Erreur systématique

Si nous reprenons la formule (1), nous voyons que la réponse dépend de nombreux facteurs. Lorsque l'on réalise une mesure, on fixe certains facteurs à des niveaux bien précis. On dit que ces facteurs sont contrôlés. Mais l'expérimentateur ne peut pas contrôler tous les facteurs. Il reste toujours des **facteurs** non contrôlés. Ce sont eux qui sont à l'origine des erreurs :

— les variations aléatoires des facteurs non contrôlés entraînent des erreurs aléatoires ;

— les variations systématiques des facteurs non contrôlés entraînent des erreurs systématiques.

L'erreur expérimentale est égale à l'erreur totale, mais comme il est difficile de détecter les erreurs systématiques, il arrive bien souvent que l'on ne retienne que l'erreur aléatoire comme valeur de l'erreur expérimentale.

4.1.2 Calcul des erreurs sur les effets

Il s'agit ici uniquement des erreurs aléatoires. Le problème est de savoir comment les erreurs qui affectent chaque réponse d'un plan se répercutent sur la précision de l'effet calculé. Reprenons la formule (18) en notant l'effet (ou l'interaction) par E au lieu de a_i.

$$E = \frac{1}{n}[\pm y_1 \pm y_2 \pm \ldots \pm y_n]$$

Les théories statistiques nous enseignent que la variance sur E, $V(E)$, est égale à la somme des variances sur les réponses y_i, $V(y_i)$, divisée par n^2, soit :

$$V(E) = \frac{1}{n^2}[V(y_1) + V(y_2) + \ldots + V(y_n)] \quad \textbf{(26)}$$

Si l'on suppose que la variance est la même pour toutes les réponses, la formule précédente se simplifie :

$$V(E) = \frac{1}{n^2}[nV(y)] \quad \textbf{(27)}$$

$$V(E) = \frac{1}{n}V(y) \quad \textbf{(28)}$$

en prenant les racines carrées des deux membres de (29), on obtient l'écart-type :

$$\sigma(E) = \frac{1}{\sqrt{n}}\sigma(y) \quad \textbf{(29)}$$

Exemple 5 : Dans l'exemple de l'émulsion de bitume, chaque effet (ou interaction) est calculé avec huit réponses. Selon la formule (28), on a :

$$\sigma(\text{effet}) = \frac{1}{\sqrt{8}}\,\sigma(\text{réponse})$$

L'écart-type sur la réponse est connu de l'expérimentateur et vaut deux points de réponse :

$$\sigma(\text{réponse}) = \pm 2$$

$$\sigma(\text{effet}) = \frac{\pm 2}{\sqrt{8}} = \pm 0{,}7$$

4.1.3 Comparaison erreur-effet

Nous allons maintenant aborder un problème qui mérite une attention particulière : quand peut-on dire d'un effet qu'il est significatif ou non ? Sur quels critères peut-on baser son raisonnement ? La méthode consiste à comparer l'erreur $\sigma(E)$ commise à l'effet E lui-même. Trois cas sont possibles.

- L'effet est bien plus grand que l'erreur :

$$E >> \sigma(E)$$

Dans ce cas, la conclusion est aisée, l'**effet est influent.**

- L'effet est plus petit que l'erreur :

$$E << \sigma(E)$$

Dans ce cas, la conclusion est, dans la plupart des situations, que l'**effet est sans influence**.

- L'effet et l'erreur sont du même ordre de grandeur :

$$E \approx \sigma(E)$$

Dans ce cas, la conclusion n'est pas toujours facile : l'effet peut être **sans influence ou légèrement influent**. Il est nécessaire de faire appel à son bon sens, à ses connaissances du phénomène et aux tests statistiques pour donner un avis pertinent. Si l'effet ne joue pas un grand rôle dans l'étude et si une mauvaise décision a peu ou pas de conséquences, on pourra ne pas aller plus loin. Mais s'il y a de gros enjeux financiers ou des problèmes de sécurité importants liés à cette conclusion, des études statistiques devront être effectuées et des essais complémentaires devront être envisagés en vue d'évaluer les risques.

4.1.4 Estimation de l'erreur expérimentale

Pour estimer l'erreur expérimentale, il faut effectuer plusieurs mesures en un même point tout en contrôlant les mêmes facteurs que ceux du plan. La meilleure solution est de choisir le point central du domaine d'étude à chaque fois que cela est possible. Dans ce cas, si l'on effectue n mesures, l'écart-type est donné par la formule :

$$\sigma(y) = \sqrt{\frac{1}{n-1}\sum_i \left(y_i - \bar{y}\right)^2} \qquad \textbf{(30)}$$

avec $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ moyenne de toutes les mesures effectuées en un même point expérimental.

Il existe d'autres moyens d'estimer l'erreur expérimentale, leur description dépasse le cadre de cet article et nous renvoyons le lecteur intéressé aux ouvrages de statistiques.

On notera que nous avons calculé l'erreur expérimentale en ne prenant en compte que l'erreur aléatoire.

4.2 Lutte contre les erreurs systématiques

4.2.1 Origine des erreurs

Les plans d'expériences permettent de s'affranchir de certaines erreurs systématiques et de traiter les autres comme des erreurs aléatoires. L'art de l'expérimentateur est d'organiser au mieux l'ensemble de ses essais pour profiter de ces avantages.

Une réponse y dépend d'une multitude de facteurs x_i :

$$y = f(x_1, x_2, ..., x_i, ..., x_n ...)$$

Tous ces facteurs ne jouent pas le même rôle, et nous allons les classer en cinq catégories :

— les facteurs que l'on étudie et auxquels l'on attribue deux niveaux : ils sont parfaitement contrôlés et leurs niveaux sont définis avec précision ; ces facteurs sont supposés ne pas introduire d'erreur ;

— les facteurs non contrôlés qui restent à un niveau fixe pendant toute l'expérimentation : ces facteurs introduisent un décalage constant dans la mesure de y ; ils sont à l'origine d'erreurs systématiques mais n'introduisent pas d'erreur aléatoire.

Si l'expérimentation est réalisée en deux campagnes, ces facteurs non contrôlés peuvent avoir des niveaux différents au cours des deux campagnes. Pour éviter que le décalage se répercute dans les résultats, il faut prendre une précaution : le **blocking** ;

— les facteurs non contrôlés et dont le niveau évolue régulièrement au cours de l'expérimentation : ces facteurs sont à l'origine d'une dérive de la réponse. Pour lutter contre ces erreurs systématiques, on utilise des **plans antidérives** ;

— les facteurs non contrôlés dont les niveaux se fixent à une valeur quelconque mais constante au cours d'un essai et différente d'un essai à l'autre : ils introduisent des erreurs systématiques contre lesquelles il est difficile de se prémunir. On les assimile à des erreurs aléatoires grâce à la **randomisation** ;

— les facteurs non contrôlés et dont les niveaux se modifient d'une manière quelconque au cours de l'expérimentation : ils sont à l'origine des erreurs aléatoires.

4.2.2 Blocking

On suppose qu'un facteur non contrôlé introduit un écart constant entre deux séries de mesures. Par exemple, deux laboratoires se partagent les essais pour étudier l'influence de certains facteurs sur le résultat d'une méthode d'analyse. Comment s'affranchir des erreurs introduites par les appareils et les opérateurs des deux laboratoires ?

Les deux laboratoires opèrent dans le même domaine d'étude. Les résultats du premier sont entachés d'une erreur systématique ε_1 et ceux du second d'une erreur systématique ε_2, c'est-à-dire que les réponses s'écrivent :

$$y_{1,i} = y_i + \varepsilon_1$$

$$y_{2,i} = y_i + \varepsilon_2$$

avec y_i réponse sans erreur du i^e essai,
$y_{1,i}$ réponse du i^e essai obtenue par le laboratoire 1,
$y_{2,i}$ réponse du i^e essai du second laboratoire.

S'il n'y a pas d'erreur systématique, l'effet d'un facteur est donné par la formule :

$$E = \frac{1}{n}(\pm y_1 \pm y_2 \pm ... \pm y_n)$$

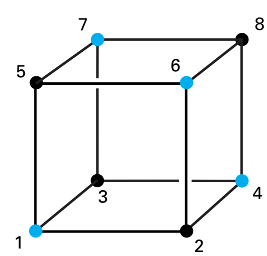
Mais si l'on tient compte des erreurs systématiques entre les laboratoires, il faut remplacer les y_i par les $y_{1,i}$ et les $y_{2,i}$. Si l'on s'y prend mal, les erreurs ne se compensent pas et la valeur de l'effet sera fausse. Si l'on s'y prend bien, on peut éliminer cette erreur.

Prenons le cas d'un plan 2^3, l'effet du facteur 1 est :

$$E_1 = \frac{1}{8}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$$

et choisissons l'emplacement des points expérimentaux de telle manière que les erreurs ε_1 et e_2 apparaissent autant de fois avec le signe – qu'avec le signe +. On peut prendre (figure **11**) :

— laboratoire 1 : essais 1, 4, 6 et 7 ;
— laboratoire 2 : essais 2, 3, 5 et 8.



Les points noirs forment un premier bloc et les points bleus en forment un second

Figure 11 – Choix de l'emplacement des points expérimentaux pour le blocking

Calculons l'effet E'_1 avec les résultats des deux laboratoires :

$$E'_1 = \frac{1}{8}\,[-y_1 - \varepsilon_1 + y_2 + \varepsilon_2 - y_3 - \varepsilon_2 + y_4 + \varepsilon_1 - y_5 - \varepsilon_2 + y_6 + \varepsilon_1 - y_7 - \varepsilon_1 + y_8 + \varepsilon_2]$$

Simplifions :

$$E'_1 = \frac{1}{8}\,[-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8] = E_1$$

Malgré les erreurs systématiques ε_1 et ε_2, on obtient une valeur correcte pour l'effet du facteur 1.

La disposition des points expérimentaux est favorable au calcul de l'effet du premier facteur mais qu'en est-il des autres effets et des interactions ? Le lecteur vérifiera que les effets des facteurs 2 et 3 et les interactions 12, 13 et 23 sont également obtenus sans que les erreurs systématiques faussent les résultats.

Qu'en est-il de l'interaction 123 ?

$$E'_{123} = \frac{1}{8}\,[-y_1 - \varepsilon_1 + y_2 + \varepsilon_2 + y_3 + \varepsilon_2 - y_4 - \varepsilon_1 + y_5 + \varepsilon_2 - y_6 - \varepsilon_1 - y_7 - \varepsilon_1 + y_8 + \varepsilon_2]$$

$$E'_{123} = \frac{1}{8}\,[y_2 + y_3 + y_5 + y_8 + 4\varepsilon_2] - \frac{1}{8}\,[y_1 + y_4 + y_6 + y_7 + 4\varepsilon_1]$$

$$E'_{123} = E_{123} + \frac{1}{2}\,[\varepsilon_2 - \varepsilon_1]$$

E'_{123} est donc égal à l'interaction E_{123} augmentée de la demi-différence $(\varepsilon_2 - \varepsilon_1)$. Que signifie cette demi-différence ? ε_1 et ε_2 mesurent les variations de la réponse dues aux deux laboratoires. Cette demi-différence mesure donc l'effet « laboratoire ». Tout se passe comme si l'on avait introduit un facteur supplémentaire, le facteur « laboratoire » (figure **12**). Ce quatrième facteur a été étudié avec les niveaux + et les niveaux – de l'interaction 123 : les niveaux + pour les essais 2, 3, 5 et 8 et les niveaux – pour les essais 1, 4, 6 et 7.

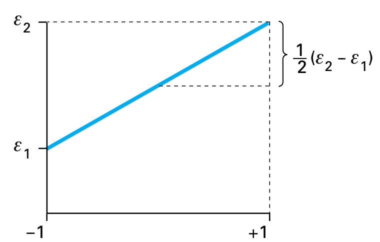


Figure 12 – Introduction d'un facteur supplémentaire : erreur systématique liée à la différence entre les laboratoires

Si l'interaction 123 est nulle, E'_{123} mesure l'effet des laboratoires.

Nous pouvons utiliser les notions introduites lors de l'étude des plans fractionnaires. Il y avait trois facteurs 1, 2 et 3 et un quatrième a été introduit, nous l'appellerons 4. Le blocking a eu pour conséquence d'aliaser ce facteur 4 avec l'interaction 123.

En notation de Box, le blocking s'exprime par :

$$\mathbf{4 = 123}$$

c'est-à-dire qu'il introduit le générateur d'aliases :

$$\mathbf{I = 1234}$$

On a réalisé un plan 2^{4-1} dans lequel le facteur 4 est aliasé à l'interaction 123.

La technique du blocking est utilisée chaque fois que l'on redoute des variations systématiques entre groupes d'essais.

Nota

1) la réalisation du blocking est très simple : dans un plan de base, on choisit une interaction d'ordre élevé, car elle a bien des chances d'être quasi nulle, et l'on utilise ses signes + pour former un premier bloc d'essais et ses signes – pour former le second bloc. Le plan de base est ainsi divisé en deux demi-plans ;

2) il est possible de bloquer sur plusieurs interactions.

4.2.3 Variations systématiques : cas de la dérive

Examinons maintenant le cas de la dérive. Il y a dérive lorsque la réponse de chaque expérience est augmentée (ou diminuée) d'une quantité croissante par rapport à la réponse sans dérive. Pour fixer les idées, nous présenterons la dérive linéaire.

La réponse du premier essai est y_1 quand il n'y a pas de dérive. Quand il y a dérive, elle est y'_1 telle que :

$$y'_1 = y_1 + h$$

où h est appelé l'incrément de la dérive.

Pour le second essai, la réponse sans dérive est y_2, la réponse avec dérive est y'_2, telle que :

$$y'_2 = y_2 + 2h$$

et ainsi de suite.

Peut-on obtenir des effets et des interactions qui soient les mêmes que s'il n'y avait pas de dérive ? La matrice de calcul des effets d'un plan 2^3 va nous permettre de calculer l'incidence de la dérive sur chacun des effets et sur chacune des interactions. Le tableau **11** illustre ce calcul.

Tableau 11 – Influence de la dérive sur les effets d'un plan 2^3

N° essai	1	2	3	12	13	23	123	I	Augmentation de la réponse
1 (1)	–	–	–	+	+	+	–	+	1 h
2 (2)	+	–	–	–	–	+	+	+	2 h
3 (3)	–	+	–	–	+	–	+	+	3 h
4 (4)	+	+	–	+	–	–	–	+	4 h
5 (5)	–	–	+	+	–	–	+	+	5 h
6 (6)	+	–	+	–	+	–	–	+	6 h
7 (7)	–	+	+	–	–	+	–	+	7 h
8 (8)	+	+	+	+	+	+	+	+	8 h
Influence dérive	0,5 h	1 h	2 h	0	0	0	0	4,5 h	

Lorsque l'on réalise un blocking ou un plan antidérive ou que l'on randomise, l'ordre d'exécution des expériences n'est plus l'ordre habituel de numérotation des essais. Il faut distinguer le nom de l'essai et l'ordre dans lequel les expériences ont été réalisées. On a pris comme convention d'indiquer l'essai par son numéro et de figurer l'ordre d'exécution par un chiffre entre parenthèses situé à côté du numéro de l'essai.

L'examen du tableau **11** permet de conclure que les trois effets principaux et la moyenne sont faux mais que les quatre interactions ne sont pas influencées par la dérive. Ne peut-on pas s'arranger pour que ce soit l'inverse ?

Il suffit de remarquer que c'est l'ordre des signes + et – dans chaque colonne qui annule l'influence de la dérive. On change l'ordre des essais pour que les colonnes de signes qui annulent l'erreur de dérive soient celles des effets principaux. L'exemple d'un tel ordre est donné dans le tableau **12**. Il y a beaucoup d'autres ordres qui donnent le même résultat. Pour le plan 2^3, il y a 144 façons d'ordonner les essais pour que les effets principaux ne soient pas faussés par la dérive [7].

Tableau 12 – Effets principaux non influencés par une dérive linéaire

N° essai	1′	2′	3′	1′2′	1′3′	2′3′	1′2′ 3′	I	Augmentation de la réponse
7 (1)	–	+	+	–	–	+	–	+	1 *h*
6 (2)	+	–	+	–	+	–	–	+	2 *h*
2 (3)	+	–	–	–	–	+	+	+	3 *h*
3 (4)	–	+	–	–	+	–	+	+	4 *h*
4 (5)	+	+	–	+	–	–	–	+	5 *h*
1 (6)	–	–	–	+	+	+	–	+	6 *h*
5 (7)	–	–	+	+	–	–	+	+	7 *h*
8 (8)	+	+	+	+	+	+	+	+	8 *h*
Influence dérive	0	0	0	+ 2 *h*	+ 0,5 *h*	0	+ 1 *h*	+ 4,5 *h*	

Si l'on a pris la précaution d'ordonner les essais suivant l'un de ces ordres, il est possible de détecter une dérive en examinant les interactions qui, dans cette éventualité, sont dans les proportions 1, 2 et 4, à condition, toutefois, que ces interactions soient faibles devant l'erreur de dérive.

4.2.4 Randomisation

L'expérimentateur sait que de petites variations dues à des facteurs non contrôlés peuvent légèrement modifier les réponses qu'il va mesurer, mais il ne les connaît pas. Ces petites variations vont introduire des erreurs à chaque mesure ; si les variations sont purement aléatoires, l'ordre des essais peut être quelconque mais, si elles ne le sont pas, elles introduiront des erreurs systématiques non contrôlées. Les tests statistiques ne sont alors plus valables car ils prennent pour hypothèse de base la distribution au hasard des erreurs. Il est quand même possible d'utiliser ces tests si l'on peut répartir au hasard les erreurs systématiques inconnues : il faut exécuter les expériences dans un ordre choisi au hasard. On dit que l'on **randomise les essais**.

La technique de randomisation est très simple, on inscrit les numéros d'essais sur des morceaux de papier différents que l'on mélange bien, puis on les tire au sort, le premier numéro tiré étant la première expérience à réaliser et ainsi de suite.

Il est recommandé d'utiliser la randomisation chaque fois que l'on craint des erreurs systématiques que l'on ne peut pas déceler ou mesurer. C'est, par exemple, souvent le cas en agriculture où la fertilité du sol varie de place en place, mais où il n'est pas possible de contrôler les facteurs agissant sur celle-ci. Ces facteurs non contrôlés entraînent des erreurs qui dépendent des variations de leurs niveaux.

4.2.5 Ordre des essais

L'ordre des essais doit être choisi en deux temps. On commence par lutter contre les erreurs systématiques : blocking ou plans antidérives. Puis, dans un deuxième temps, on répartit au hasard les erreurs systématiques contre lesquelles on ne peut pas lutter : on randomise.

Cette démarche a l'avantage de minimiser l'erreur aléatoire en éliminant les erreurs systématiques que l'on peut maîtriser.

5. Autres plans à deux niveaux

5.1 Objectifs des autres plans à deux niveaux

Les plans factoriels complets et les plans factoriels fractionnaires sont basés sur des modèles mathématiques du premier degré. Ils couvrent la plupart des besoins des expérimentateurs lors d'une étude de dégrossissage. Ce sont eux qui seront employés dans la majorité des cas.

D'autres plans à deux niveaux et basés également sur un modèle mathématique du premier degré ont été mis au point pour répondre à des situations particulières. Il s'agit, soit de rechercher et d'accepter un modèle mathématique simple, soit de sélectionner rapidement les quelques facteurs influents parmi un grand nombre de facteurs.

Nous examinons ici d'abord les plans de Rechtschaffner qui répondent au premier objectif. Puis nous présentons les plans de Plackett et Burman ainsi que les plans supersaturés qui répondent au deuxième objectif.

5.2 Plans de Rechtschaffner

Les plans de Rechtschaffner sont des plans factoriels fractionnaires simplifiés qui permettent de déterminer les effets des facteurs et les interactions d'ordre deux. Toutes les autres interactions sont supposées nulles avant même l'expérimentation. Le modèle mathématique adopté au départ de l'étude est donc :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j$$

Il suffit de choisir un plan fractionnaire de résolution III pour obtenir un plan de Rechtschaffner. Mais l'idée de ne déterminer que les effets principaux et les interactions a été étendue par Rechtschaffner aux plans du second degré et aux facteurs prenant trois niveaux. Ces plans spéciaux sont indiqués dans des tables auxquelles il faut se référer pour les utiliser [8].

5.3 Plans de Plackett et Burman

Les plans de Plackett et Burman [9] sont des plans factoriels fractionnaires encore plus simplifiés, qui ne permettent de déterminer que les effets principaux des facteurs. Ils ne permettent pas d'évaluer les interactions. Toutes les interactions sont donc supposées nulles avant même l'expérimentation. Le modèle mathématique adopté au départ de l'étude est :

$$y = a_0 + \sum a_i x_i$$

Les plans de Plackett et Burman sont construits sur des matrices orthogonales d'Hadamard. Ces matrices n'existent que lorsque le nombre d'essais est multiple de 4. Ils coïncident donc avec les plans fractionnaires pour un nombre d'essais égal à 4, 8, 16, 32, etc. Mais ils peuvent traiter des situations où le nombre d'essais est 12, 20, 24, 28, etc. Nous indiquons la matrice à douze essais (tableau **13**) qui permet de connaître les effets principaux de onze facteurs. S'il existe une interaction que l'expérimentateur ne soupçonne pas, elle sera aliasée à un ou plusieurs des facteurs étudiés. L'interprétation de ces plans est donc très délicate.

Tableau 13 – Plan de Plackett et Burman pour étudier 11 facteurs 12 en essais

N° essai	1	2	3	4	5	6	7	8	9	10	11
1	+	+	+	+	+	+	+	+	+	+	+
2	–	+	–	+	+	+	–	–	–	+	–
3	–	–	+	–	+	+	+	–	–	–	+
4	+	–	–	+	–	+	+	+	–	–	–
5	–	+	–	–	+	–	+	+	+	–	–
6	–	–	+	–	–	+	–	+	+	+	–
7	–	–	–	+	–	–	+	–	+	+	+
8	+	–	–	–	+	–	–	+	–	+	+
9	+	+	–	–	–	+	–	–	+	–	+
10	+	+	+	–	–	–	+	–	–	+	–
11	–	+	+	+	–	–	–	+	–	–	+
12	+	–	+	+	+	–	–	–	+	–	–

5.4 Plans sursaturés

Un plan saturé est un plan qui comporte autant d'essais que de coefficients à déterminer dans le modèle mathématique. Les plans fractionnaires, les plans de Rechtschaffner, les plans de Plackett et Burman sont souvent des plans saturés. Un plan sursaturé est un plan qui comporte moins d'essais que de coefficients à déterminer dans le modèle mathématique [10]. Ces plans sont utiles lorsqu'il y a beaucoup de facteurs à examiner et lorsqu'on est sûr que peu d'entre eux sont influents sur la réponse.

Exemple : il est possible d'étudier 66 facteurs en 12 essais ou 272 facteurs en 24 essais.

6. Plans du second degré

Nous avons considéré jusqu'ici des facteurs continus ou discrets mais ne prenant que deux niveaux. Dans ce paragraphe, nous allons étudier les facteurs continus prenant plus de deux niveaux. Le modèle mathématique adopté dans les paragraphes précédents était du premier degré par rapport à chacune des variables. Si ce modèle du premier degré est valide, il est rarement nécessaire de poursuivre l'expérimentation. Par contre, s'il n'est pas valide, il faut passer à un modèle de degré plus élevé. Nous nous limitons ici au modèle du second degré qui se révèle suffisant dans la plupart des cas.

6.1 Validation du modèle du premier degré

Le modèle du premier degré permet de calculer la valeur de la réponse au centre du domaine d'étude. Il est facile de comparer cette valeur calculée à la mesure réelle effectuée en ce point. Si ces deux valeurs diffèrent peu, on peut considérer le modèle du premier degré comme valide ; si elles diffèrent trop, il faut adopter un modèle du second degré. On réalise alors des expériences supplémentaires pour déterminer tous les coefficients de ce nouveau modèle.

Au lieu d'effectuer une seule mesure au centre du domaine, on peut en faire plusieurs. Cela permet d'obtenir une estimation de l'erreur expérimentale (§ 4.1.2).

On remarquera que des mesures au point central introduisent un niveau supplémentaire d'étude des facteurs. Il y a maintenant trois niveaux par facteur : – 1, 0 et + 1.

6.2 Modèle du second degré

Le modèle mathématique est analogue à la relation (14) à laquelle on ajoute un terme carré :

$$\begin{aligned} y = {} & a_0 + \Sigma a_i x_i + \Sigma a_{ij} x_i x_j \\ & + \Sigma a_{ijl} x_i x_j x_l + \ldots + a_{i..k} x_i \ldots x_k + \Sigma a_{ii} x_i^2 \end{aligned} \quad \textbf{(31)}$$

Dans le cas d'un plan à deux facteurs, la formule est la suivante :

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 \quad \textbf{(32)}$$

il y a six inconnues à déterminer.

Le plan d'expériences fournit un certain nombre de valeurs de la réponse y. L'interprétation du plan consiste donc à trouver les coefficients et, par suite, à résoudre un système de n équations (s'il y a n réponses) et p inconnues (s'il y a p coefficients). Il est commode d'écrire ce système sous forme matricielle en tenant compte des erreurs expérimentales :

$$\underset{(n,1)}{\boldsymbol{y}} = \underset{(n,p)}{\boldsymbol{X}} \; \underset{(p,1)}{\boldsymbol{a}} + \varepsilon \quad \textbf{(33)}$$

La résolution de ce système est généralement conduite selon la méthode des moindres carrés (cf. [22]). et la solution est notée $\hat{\boldsymbol{a}}$. Cette solution est donnée par la formule suivante [11] :

$$\hat{\boldsymbol{a}} = ({}^t\boldsymbol{X}\boldsymbol{X})^{-1}\, {}^t\boldsymbol{X}\boldsymbol{y} \quad \textbf{(34)}$$

Le calcul de l'inconnue $\hat{\boldsymbol{a}}$ est beaucoup moins facile que pour les plans 2^k. L'emploi d'un logiciel se révèle absolument nécessaire (cf. § 9).

Il existe plusieurs types de plans du second degré. Nous allons décrire les principaux et indiquer leurs avantages et leurs inconvénients.

6.3 Plans composites

Un plan composite est constitué de trois parties :

1) un plan factoriel à deux niveaux par facteur analogue à ceux que nous avons précédemment décrits ;

2) au moins un point expérimental situé au centre du domaine expérimental ;

3) les points axiaux. Ces points expérimentaux sont situés sur les axes de chacun des facteurs.

La figure **13** représente un plan composite pour deux facteurs. Les points A, B, C et D sont les points d'un plan 2^2. Le point E est le point central. Ce point peut avoir été répliqué une ou plusieurs fois. Les points F, G, H et I sont les points axiaux. Ces quatre derniers points forment ce que l'on appelle le plan en étoile. Dans cet exemple, l'expérimentateur réalise 9 essais et doit déterminer 6 coefficients. Il faut donc résoudre un système de 9 équations à 6 inconnues. Le calcul est effectué à l'aide d'un logiciel approprié (cf. § 9).

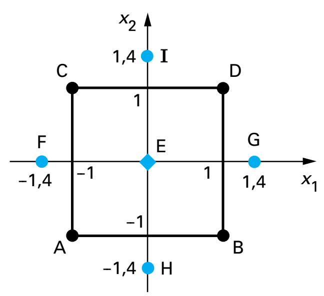


Figure 13 – Représentation d'un plan composite pour l'établissement d'un modèle du second degré

L'intérêt des plans composites réside dans le fait qu'ils prennent facilement la suite d'un premier plan factoriel dont les résultats sont inexplicables par un modèle du premier degré. Il suffit d'effectuer les expériences correspondant aux points en étoile et de faire les calculs sur l'ensemble de toutes les expériences. Les plans composites sont parfaitement adaptés à l'acquisition progressive des résultats. Le nombre de niveaux d'un plan composite est de cinq par facteur : le point central, les deux niveaux du plan factoriel et les deux niveaux des points en étoile.

Les points en étoile sont sur les axes des facteurs. Mais quelles coordonnées faut-il leur donner ? Mettons-nous dans le cas idéal où tous les emplacements sont possibles et où les contraintes expérimentales ne gênent pas. La disposition des points expérimentaux dépend alors du critère d'optimalité que l'on choisit. En général, on s'arrange pour que les erreurs sur les coefficients du modèle soient les plus petites possible ou qu'elles soient les mieux réparties possible. Les solutions de ce problème sont multiples et nous n'indiquons ici que les principales.

6.4 Critères d'optimalité

Les statisticiens ont établi la formule qui donne l'erreur sur les coefficients du modèle lorsque l'on connaît l'erreur sur les réponses mesurées expérimentalement. Cette formule, sous sa forme la plus simple, est la suivante :

$$V(\hat{\boldsymbol{a}}) = \sigma_y^2({}^{\mathrm{t}}\boldsymbol{XX})^{-1} \tag{35}$$

c'est-à-dire que les variances des coefficients sont égales à la variance des réponses multipliée par une matrice dépendant de la position des points expérimentaux et du modèle mathématique choisi par l'expérimentateur. Les variances des coefficients sont les termes diagonaux de la matrice $V(\hat{\boldsymbol{a}})$. On obtient ces variances par identification avec les termes diagonaux de la matrice $({}^{\mathrm{t}}\boldsymbol{XX})^{-1}$. Les principaux critères d'optimalité sont les suivants.

■ **Critère de O-optimalité**

La matrice $\boldsymbol{X}$ est une matrice orthogonale d'Hadamard. Il en résulte que la matrice $({}^{\mathrm{t}}\boldsymbol{XX})^{-1}$ est une matrice diagonale. Seuls les termes diagonaux de cette matrice sont différents de zéro et l'on démontre que ces termes diagonaux ne peuvent pas être plus petits. La variance des coefficients est donc, à coup sûr, la plus faible possible.

Lorsque l'on ne peut pas utiliser une matrice orthogonale, il faut choisir l'un des critères suivants.

■ **Critère de D-optimalité**

Ce critère garantit une variance des coefficients du modèle telle que l'erreur de prévision de la réponse soit minimale. On obtient ce résultat en maximisant le déterminant de la matrice ${}^{\mathrm{t}}\boldsymbol{XX}$. Le critère correspondant s'appelle le critère de D-optimalité.

■ **Critère de A-optimalité**

La somme des variances des coefficients peut être minimisée. Dans ce cas, on parle de critère de A-optimalité. Un plan est A-optimal si la position des points expérimentaux minimise la trace de la matrice $({}^{\mathrm{t}}\boldsymbol{XX})^{-1}$.

■ **Critère de G-optimalité**

Parmi les variances des coefficients, il y en a une qui est plus grande que toutes les autres. On peut vouloir que cette forte variance soit la plus faible possible. Le critère correspondant s'appelle le critère de G-optimalité.

■ **Critère de rotabilité**

On désire que les réponses calculées avec le modèle issu du plan d'expériences aient une erreur identique pour des points situés à la même distance du centre du domaine expérimental. Dans ce cas, on parle de plans rotatables [12]. La distance α des points expérimentaux au centre du domaine est donnée, pour un plan sans réplication, par la formule :

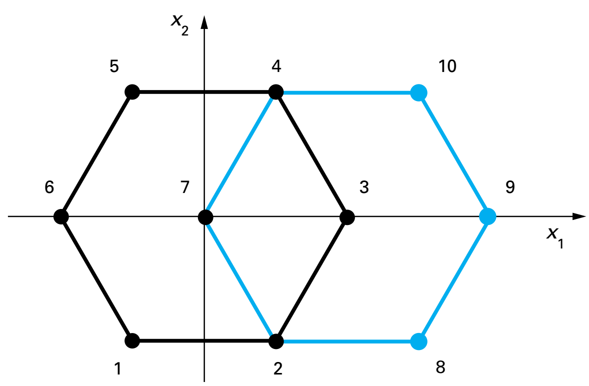
$$\alpha = (n_c)^{\frac{1}{4}}$$

formule dans laquelle n_c est le nombre des sommets du domaine d'étude défini par le cube ou l'hypercube des points expérimentaux.

6.5 Plans de Doehlert

La caractéristique principale des plans de Doehlert [13] est d'avoir une répartition uniforme des points expérimentaux dans l'espace expérimental. La figure **14** donne la disposition de ces points pour un plan à deux facteurs. Tous les points sont à la même distance du centre du domaine expérimental et sont situés sur le cercle trigonométrique. Ils forment un hexagone régulier. La matrice d'expériences se construit en prenant les coordonnées de chaque point expérimental (tableau **14**)

Tableau 14 – Plan de Doehlert pour établir un modèle du second degré avec deux facteurs		
N° essai	Facteur 1	Facteur 2
1	0	0
2	−1	0
3	+1	0
4	−1/2	$-\sqrt{3}/2$
5	+1/2	$-\sqrt{3}/2$
6	−1/2	$+\sqrt{3}/2$
7	+1/2	$+\sqrt{3}/2$



Les cercles noirs (de 1 à 7) illustrent un premier plan de Doehlert. Les trois cercles bleus illustrent les expériences supplémentaires qui permettent d'explorer l'espace expérimental avec un deuxième plan de Doehlert (points 2, 8, 9, 10, 4, 7, et 3).

Figure 14 – Plan de Doehlert

Si l'expérimentateur désire explorer le domaine expérimental, il peut facilement ajouter des points d'expériences supplémentaires et retrouver une disposition identique à celle de départ. La figure **14** montre qu'avec trois points d'expériences on peut obtenir un nouveau plan de Doehlert.

Ce type de plan est utilisable pour un nombre quelconque de facteurs. Nous indiquons la matrice correspondant à un plan de Doehlert pour trois facteurs (tableau **15**). On constate que le facteur 3 est au niveau zéro pour les sept premiers essais, c'est-à-dire que l'étude du troisième facteur ne nécessite que l'exécution de six essais si l'on a déjà réalisé l'étude des deux premiers facteurs. Cette facilité est très utile si l'on désire obtenir rapidement des renseignements sur les deux premiers facteurs. Il n'est pas nécessaire de réaliser l'ensemble du plan (13 essais) pour connaître l'influence de ces deux facteurs. Cette propriété est toujours vraie quel que soit le nombre de facteurs. Elle facilite l'acquisition progressive des résultats.

Tableau 15 – Plan de Doehlert pour établir un modèle du second degré avec trois facteurs			
N° essai	Facteur 1	Facteur 2	Facteur 3
1	0	0	0
2	− 1	0	0
3	+ 1	0	0
4	− 1/2	$-\sqrt{3}/2$	0
5	+ 1/2	$-\sqrt{3}/2$	0
6	− 1/2	$+\sqrt{3}/2$	0
7	+ 1/2	$+\sqrt{3}/2$	0
8	− 1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
9	0	$+ 1/\sqrt{3}$	$-\sqrt{2/3}$
10	+ 1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
11	− 1/2	$+1/2\sqrt{3}$	$\sqrt{2/3}$
12	0	$-1/\sqrt{3}$	$\sqrt{2/3}$
13	+ 1/2	$+1/2\sqrt{3}$	$\sqrt{2/3}$

6.6 Plans de Box-Behnken

Les plans de Box-Behnken [14] répondent à un critère d'optimisation particulier : l'erreur de prévision des réponses est la même pour tous les points d'une sphère (ou d'une hypersphère) centrée à l'origine du domaine expérimental. C'est le critère de rotabilité. Le plus connu des plans de Box-Behnken est celui qui permet d'étudier trois facteurs. Les points expérimentaux sont au milieu des arêtes de chacun des côtés du cube (figure **15**). Ce plan comporte douze essais auxquels on peut ajouter un (ou plusieurs) point(s) central(aux). La matrice du tableau **16** indique ces douze essais accompagnés d'un seul point central.

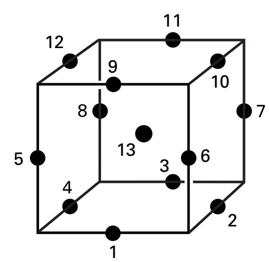


Figure 15 – Plan de Box-Behnken pour trois facteurs

Tableau 16 – Plan de Box-Behnken pour établir un modèle du second degré avec trois facteurs			
N° essai	Facteur 1	Facteur 2	Facteur 3
1	0	– 1	– 1
2	+ 1	0	– 1
3	0	+ 1	– 1
4	– 1	0	– 1
5	– 1	– 1	0
6	+ 1	– 1	0
7	– 1	+ 1	0
8	+ 1	+ 1	0
9	0	– 1	+ 1
10	+ 1	0	+ 1
11	0	+ 1	+ 1
12	– 1	0	+ 1
13	0	0	0

6.7 Plans hybrides

Les plans hybrides ont été mis au point par Roquemore [15]. Leur objectif est d'essayer d'approcher les deux critères d'optimalité, orthogonalité et rotabilité, sans en atteindre aucun. L'orthogonalité garantit la meilleure précision possible sur les coefficients du modèle et la rotabilité conduit à des erreurs de prévisions identiques à une même distance du centre du domaine. Si l'expérimentateur recherche ces deux propriétés, il doit penser à utiliser un plan hybride.

Les plans hybrides se désignent de la manière suivante : on indique le nombre de facteurs, puis le nombre de points expérimentaux dont un seul point central, enfin une lettre pour distinguer deux plans ayant le même nombre de facteurs et le même nombre de points expérimentaux. Par exemple, le plan 311B est l'un des deux plans hybrides (le 311A existe) qui permet d'étudier trois facteurs en onze essais. Nous en donnons la matrice dans le tableau **17**.

Tableau 17 – Plan hybride 311B			
N° essai	Facteur 1	Facteur 2	Facteur 3
1	0	0	$+\sqrt{6}$
2	0	0	$-\sqrt{6}$
3	–0,7507	2,1063	+ 1
4	2,1063	0,7507	+ 1
5	0,7507	– 2,1063	+ 1
6	– 2,1063	– 0,7507	+ 1
7	0,7507	2,1063	– 1
8	2,1063	– 0,7507	– 1
9	– 0,7507	– 2,1063	– 1
10	– 2,1063	0,7507	– 1
11	0	0	0

Les autres plans hybrides sont, par exemple, les plans 310, 416A, 416B, 416C ou 628A.

6.8 Plans quadratiques gigognes

Les plans quadratiques gigognes ont été imaginés par Mozzo [16] pour allier les avantages séquentiels des plans de Doehlert sans en avoir les inconvénients. En effet, les niveaux des plans de Doehlert sont nombreux et souvent difficiles à atteindre en milieu industriel où la souplesse de réglage n'est pas toujours facile à obtenir. Les plans quadratiques gigognes n'ont que trois niveaux quel que soit le nombre de facteurs. On peut commencer par étudier trois facteurs en douze essais. Puis, si l'on désire étudier un quatrième facteur, il suffit de réaliser douze essais supplémentaires. Les 24 essais sont pris en compte dans le calcul du modèle à 4 facteurs. On peut arrêter l'étude à ce moment ou l'on peut vouloir continuer et étudier un cinquième facteur. Dans ce dernier cas, il faut faire 16 essais de plus pour établir le modèle à 5 facteurs. Le tableau **18** donne les deux premiers plans gigognes pour 3 et 4 facteurs.

Tableau 18 – Plan gigogne pour trois facteurs, puis quatre facteurs				
N° essai	Facteur 1	Facteur 2	Facteur 3	Facteur 4
1	– 1	– 1	0	0
2	+ 1	– 1	0	0
3	– 1	+ 1	0	0
4	+ 1	+ 1	0	0
5	0	– 1	– 1	0
6	– 1	0	– 1	0
7	+ 1	0	– 1	0
8	0	+ 1	– 1	0
9	0	– 1	+ 1	0
10	– 1	0	+ 1	0
11	+ 1	0	+ 1	0
12	0	+ 1	+ 1	0
13	0	0	– 1	– 1
14	0	– 1	0	– 1
15	– 1	0	0	– 1
16	+ 1	0	0	– 1
17	0	+ 1	0	– 1
18	0	0	+ 1	– 1
19	0	0	– 1	+ 1
20	0	– 1	0	+ 1
21	– 1	0	0	+ 1
22	+ 1	0	0	+ 1
23	0	+ 1	0	+ 1
24	0	0	+ 1	+ 1

6.9 Plans non conventionnels

Tous les plans précédemment décrits cherchent à répondre à l'un des critères d'optimalité du paragraphe 6.4. Mais les contraintes expérimentales ne permettent pas toujours d'être dans ces conditions idéales. Par exemple, les réglages de l'appareil ne permettent pas d'atteindre les niveaux préconisés par la théorie ou des combinaisons de niveaux peuvent se révéler dangereuses : réaction explosive pour les chimistes, concentration toxique pour les médecins, etc. Faut-il alors se priver de la puissance des plans d'expériences ? Eh bien, non, Goupy a montré [17] que des plans qui ne s'éloignent pas trop des plans optimaux sont encore de bons plans et que les erreurs sur les coefficients du modèle et sur les réponses prédites sont, dans la plupart des cas, tout à fait acceptables et peu différentes de la plus basse valeur. Ces plans peuvent avoir autant de points expérimentaux que l'on veut et les facteurs peuvent prendre n'importe quelle valeur. La figure **16** illustre un plan non conventionnel.

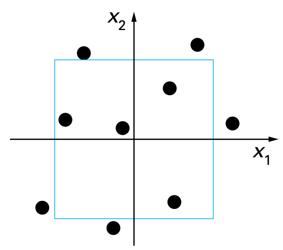


Figure 16 – Exemple de plan non conventionnel

L'interprétation de ces plans se fait en déterminant les coefficients du modèle mathématique à l'aide de la formule :

$$\hat{\boldsymbol{a}} = ({}^{t}\boldsymbol{XX})^{-1}\,{}^{t}\boldsymbol{Xy}$$

Mais, dans ce cas particulier, il est essentiel de calculer les erreurs de chaque coefficient et de s'assurer que les corrélations entre coefficients ne sont pas trop grandes. La formule donnant les variances et les covariances est la suivante :

$$V(\hat{\boldsymbol{a}}) = \sigma_y^2\,({}^{t}\boldsymbol{XX})^{-1}$$

Si les erreurs sur les coefficients sont faibles et si les corrélations entre coefficients le sont aussi, on pourra utiliser le modèle prédictif. Les réponses prédites sont données par :

$$\hat{\boldsymbol{y}} = \hat{\boldsymbol{a}}$$

Ces plans peuvent être utilisés lorsque les contraintes expérimentales ne permettent pas d'employer un plan optimal. Dans ce cas, l'expérimentateur essayera toujours de se rapprocher le plus près possible d'un plan optimal.

Ces plans peuvent également servir à interpréter des résultats qui n'avaient pas été obtenus avec un plan préalablement établi. Dans ce cas, la matrice $\boldsymbol{X}$ risque d'être mal conditionnée et le déterminant de ${}^{t}\boldsymbol{XX}$ d'être petit, d'où de fortes erreurs sur les coefficients du modèle. Il est alors prudent de n'utiliser les résultats de l'interprétation que comme une première approche donnant une orientation pour les futurs essais. On peut se servir de ces premiers essais comme base de départ d'un nouveau plan mieux organisé. Il faudra y ajouter de nouveaux essais tels qu'ils permettent d'obtenir une matrice bien conditionnée. On parle alors de **réparation** d'un mauvais plan d'expérience.

7. Analyse de la variance

Abordons maintenant le cas des facteurs discrets prenant plus de deux niveaux. Pour leur étude, il faut adopter une technique bien connue des statisticiens : l'**analyse de la variance** ou ANOVA (cf. [23]). Il n'est donc pas question de décrire en détail cet outil, que l'on trouvera dans les bons ouvrages de statistiques, mais simplement de montrer les liens avec les plans d'expériences.

7.1 Définition de l'effet d'un facteur

Prenons l'exemple d'un facteur « personne ». S'il y a trois niveaux de personne, cela signifie qu'il y a trois individus : Jacques, Louis et Pierre. S'ils mettent au point la même méthode d'analyse et si toutes les conditions sont les mêmes pour une détermination faite par chacun d'eux, ils obtiendront trois résultats légèrement différents : y_J pour Jacques, y_L pour Louis et y_P pour Pierre. Si l'on note la valeur moyenne des trois réponses par y_0, on peut modéliser le résultat de Jacques par :

$$y_J = y_0 + a_J \qquad \textbf{(36)}$$

où a_J est la variation de la réponse par rapport à la moyenne, variation due à Jacques. C'est donc, par analogie avec les plans d'expériences, l'effet de Jacques

On définit de même les effets individuels des trois personnes : a_J, a_L et a_P. Mais ces variations peuvent aussi être regardées comme les écarts des statisticiens. En termes de plans d'expériences, on parle d'effets individuels et en termes de statistiques, on parle d'écarts à la moyenne. Les formules établies en statistiques peuvent être reprises dans la théorie des plans d'expériences. Étant définis par rapport à la moyenne, les trois effets individuels ne sont pas indépendants (figure **17**). Il existe la relation :

$$a_J + a_L + a_P = 0 \qquad \textbf{(37)}$$

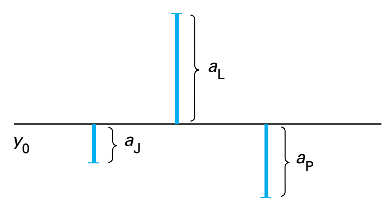


Figure 17 – Évaluation des effets individuels par rapport à la moyenne des réponses

Par définition, le carré de l'effet a des trois personnes prises globalement est identique à la variance de l'échantillon :

$$a^2 = \frac{1}{3}(a_J^2 + a_L^2 + a_P^2) \qquad \textbf{(38)}$$

7.2 Somme des carrés des réponses

Les trois effets individuels a_J, a_L et a_P associés à la formule (36) permettent d'écrire le système suivant :

$$y_J = y_0 + a_J$$
$$y_L = y_0 + a_L$$
$$y_P = y_0 + a_P$$

On remplace a_P par sa valeur en fonction de a_J et a_L, en tenant compte de la relation (37) :

$$y_J = y_0 + a_J$$

$$y_L = y_0 + a_L$$

$$y_P = y_0 - a_J - a_L$$

et l'on met ce nouveau système sous forme matricielle :

$$\begin{bmatrix} y_J \\ y_L \\ y_P \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} y_0 \\ a_J \\ a_L \end{bmatrix}$$

ou sous forme condensée :

$$\boldsymbol{Y} = \boldsymbol{XA}$$

La somme des carrés des réponses est égale à :

$$^{t}\boldsymbol{Y}\,\boldsymbol{Y} = {}^{t}\boldsymbol{A}\,{}^{t}\boldsymbol{X}\,\boldsymbol{X}\,\boldsymbol{A}$$

La matrice $^{t}\boldsymbol{X}\,\boldsymbol{X}$ peut être décomposée en sous-matrices :

$$^{t}\boldsymbol{XX} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \quad (39)$$

$$^{t}\boldsymbol{YY} = 3y_0^2 + [a_J \; a_L] \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} a_J \\ a_L \end{bmatrix}$$

$$^{t}\boldsymbol{YY} = 3\,y_0^2 + 3\,a^2 \quad (40)$$

La somme des carrés des réponses est décomposable en deux carrés : trois fois celui de la moyenne y_0, et trois fois celui de l'effet a des personnes.

Si l'on étudie plusieurs facteurs, par exemple les trois facteurs suivants : les appareils, les laboratoires et les opérateurs, on définit les effets individuels et les effets globaux comme précédemment. Un calcul identique à celui que nous avons indiqué permet de décomposer la somme des carrés des réponses en quatre carrés : celui de la moyenne, celui des appareils, celui des laboratoires et celui des opérateurs.

Si l'on réalise plus d'expériences qu'il n'y a de coefficients dans le modèle mathématique adopté, on peut introduire un facteur supplémentaire correspondant à la dispersion des mesures. L'analyse mathématique est tout à fait analogue à celle que nous avons décrite mais elle fait apparaître un carré correspondant à la dispersion des mesures. Il est alors possible de construire le tableau classique de l'analyse de la variance.

L'analyse de la variance décrite ici pour les facteurs discrets est également valable pour les facteurs continus. Mais cette technique a pour principal inconvénient de faire disparaître le signe de l'effet global puisque l'on utilise le carré de cet effet. Le signe de l'effet est une information capitale pour l'expérimentateur qui désire connaître comment la réponse évolue en fonction du sens de variation de la variable continue. Il est donc tout à fait déconseillé d'employer l'analyse de la variance pour les facteurs continus. Dans le même ordre d'idée et pour faciliter l'interprétation des résultats, il est recommandé de calculer les effets individuels des facteurs discrets, car ces effets individuels conservent leur signe et sont ainsi une source d'informations utiles pour l'expérimentateur.

8. Plans de mélange

Les plans de mélange sont des plans d'expériences que l'on utilise lorsque l'on étudie des produits composés de plusieurs constituants [18]. L'objectif est de trouver la loi qui régit une ou plusieurs réponses en fonction de la composition du mélange.

8.1 Modèle mathématique

Les facteurs sont les concentrations x_i de chaque constituant i du mélange. Les réponses y sont exprimées en fonction de ces concentrations. Il faut tenir compte de la contrainte suivante : la somme des concentrations des constituants d'un mélange est égale à cent pour cent. Il existe donc la relation suivante entre les x_i :

$$\sum x_i = 1 \quad (41)$$

C'est la relation fondamentale des mélanges.

Le modèle mathématique exprimant la valeur d'une réponse en fonction des concentrations x_i ne peut pas s'écrire selon une formule analogue aux plans factoriels ou aux plans composites, car il faut tenir compte de la relation (41). On montre facilement que cette relation impose dans le modèle mathématique :

— la disparition du terme constant ;
— la disparition des termes carrés.

Il ne reste donc, dans le modèle mathématique, que les termes du premier degré et les termes rectangles.

$$y = \Sigma\, a_i x_i + \Sigma\, a_{ij}\, x_i x_j + \Sigma\, a_{i...l}\, x_i ... \,x_l \quad (42)$$

8.2 Emplacement des points expérimentaux

La représentation graphique doit également tenir compte de la contrainte imposée par la relation (41). Celle-ci fait disparaître une dimension dans l'espace des x_i. Pour un mélange composé de deux produits, on utilise une droite (figure **18*a***) et, pour un mélange de trois produits, on adopte la représentation du triangle équilatéral bien connue des chimistes (figure **18*b***).

Scheffé [19] a montré que les meilleurs points expérimentaux se situent aux extrémités du domaine (les produits purs), au milieu des arêtes (les mélanges de deux produits à égales concentrations) et au centre de gravité (mélange de tous les produits à égales concentrations).

8.3 Difficultés soulevées par les mélanges

La disposition des points expérimentaux est souvent loin d'être aussi simple. En effet, deux difficultés apparaissent fréquemment :

a) les propriétés du mélange ne sont pas des fonctions continues de la composition : il y a des précipitations, des changements de phase, des démixtions, etc. ;

b) toutes les concentrations ne sont pas toujours réalisables.

Afin de pouvoir utiliser un modèle mathématique simple, l'expérimentateur est amené à subdiviser le domaine expérimental en plusieurs domaines d'étude dont la forme peut être compliquée. Un exemple est donné (figure **19**) pour trois constituants. Déjà, dans ce cas, le positionnement des meilleurs points expérimentaux est

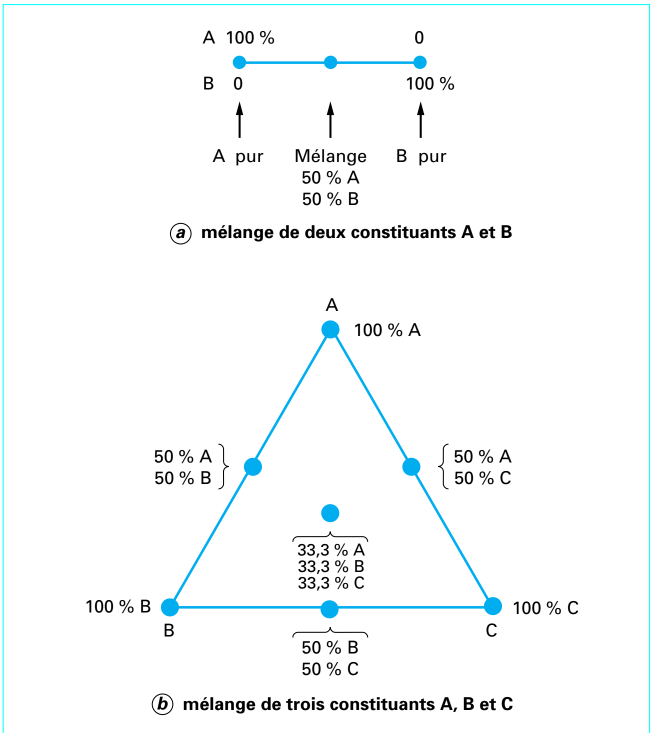


Figure 18 – Représentation de mélanges à 2 et 3 constituants

compliqué et nécessite le recours à des algorithmes d'optimisation comme celui de Mitchell [20] ou celui de Fedorov [21]. On imagine la difficulté du problème lorsqu'il y a dix constituants et qu'il faut travailler dans un espace à neuf dimensions. L'emploi d'ordinateurs puissants est alors indispensable.

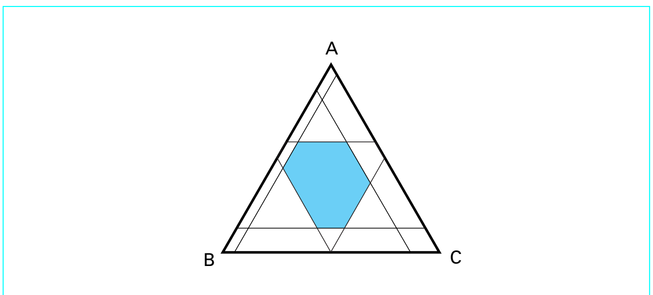


Figure 19 – Représentation géométrique d'un domaine d'étude (partie bleue) lorsqu'il y a des contraintes sur les concentrations des constituants d'un mélange

9. Logiciels

La plupart des logiciels qui traitent des plans d'expériences sont inclus dans des logiciels de statistiques. Cette introduction s'est faite petit à petit sous la pression de la demande. Les premiers logiciels de plans d'expériences étaient très pauvres et mal adaptés aux besoins des expérimentateurs. Depuis peu, un effort considérable a été entrepris par les informaticiens et les statisticiens pour que ces logiciels répondent mieux à l'esprit et aux besoins des expérimentateurs. Ces logiciels sont encore très marqués par leur origine statistique et peuvent rebuter certains utilisateurs. Mais l'on constate un progrès constant vers l'amélioration, c'est-à-dire vers une meilleure prise en compte des exigences des expérimentateurs. Néanmoins, ces logiciels nécessitent tous une bonne connaissance de la méthode des plans d'expériences et ne peuvent pas être utilisés sans une formation solide aux plans d'expériences.

Ces logiciels comportent, en général, les chapitres suivants :

— construction des plans d'expériences :

— plans factoriels complets, plans factoriels fractionnaires, plans à plus de deux niveaux (surface de réponse), plans de mélanges, plans D-optimaux ;

— interprétation :

— calcul des effets, des interactions, des coefficients du modèle mathématique, modélisation, calcul des réponses prédites, analyse des résidus ;

— représentations graphiques :

— diagramme des effets, diagramme des interactions, diagramme des résidus, diagramme de Daniel, courbes isoréponses en 2-D ou 3-D ;

— aide :

— tutorial, aide en ligne, possibilité d'importer et d'exporter des fichiers de données dans différents formats, parfois hot-line.

On trouvera en [Doc. P 230 (tableaux A et B)], les principales caractéristiques de quelques logiciels analysés uniquement du point de vue des plans d'expériences, ainsi qu'une étude comparative des principaux d'entre eux.