

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE MENTOURI CONSTANTINE
FACULTE DES SCIENCES EXACTES
DEPARTEMENT DE MATHEMATIQUES**

N° d'ordre :

N° série :

THESE

Présentée pour l'obtention du diplôme de :

Magister

Thème

**Etude de quelques plans d'expériences associées
aux surfaces de réponse**

Par :

Kimouche Karima

Directeur de la thèse :

Gheribi Z.

Soutenue le : 2008

Table des matières

Introduction	3
1 Les plans d'expériences	5
1.1 Outils et quelques notions préliminaires	5
1.2 Analyse statistique des plans d'expériences	7
1.2.1 Choix du modèle	7
1.2.2 Estimation	8
1.2.3 Tests d'hypothèses	9
1.3 Critères de qualité d'une matrice d'un plan d'expériences	9
1.3.1 Propriétés associées à la matrice d'un plan d'expériences	10
1.3.2 Propriétés des estimateurs des coefficients du modèle mathématique postulé	11
1.3.3 Qualités prévisionnelles du modèle mathématique	12
2 Les Plans factoriels et fractionnaires	16
2.1 Plans factoriels complets à deux niveaux 2^m	18
2.1.1 Construction des plans factoriels complets	21
2.1.2 Les méthodes pour juger de la signification des effets des facteurs	22
2.1.3 Les avantages et les inconvénients	22
2.2 Les plans factoriels fractionnaires à deux niveaux 2^{m-s}	23
2.2.1 Confusion de ξ_α dans une replicate de 2^{m-s} blocs constitués chacun de 2^s unités	24
2.2.2 Construction de la replicate après choix des ξ_α à confondre	24
2.2.3 Réalisation partielle d'une replicate	26
2.2.4 La résolution d'un plan d'expérience	26
2.2.5 Construction formelle d'un plan fractionnaire	27
2.2.6 Les avantages et les inconvénients	28

	2
3 Plans d'expériences associés aux surfaces de réponse d'ordre 1	29
3.1 Domaine expérimental sphérique	30
3.1.1 Les plans équiradiaux	30
3.1.2 Les plans simplexes	31
3.2 Domaine expérimental cubique	33
3.2.1 Les plans de Plackett et Burman (PB)	33
3.2.2 Les plans de Taguchi	35
4 Plans d'expériences associés aux surfaces de réponse d'ordre 2	40
4.1 Domaine expérimental sphérique	40
4.1.1 Les plans composites	40
4.1.2 Les plans de Box-Behnken (BB)	48
4.1.3 Les plans de Doehlert	51
4.2 Domaine expérimental cubique	54
4.2.1 Les plans composites	54
4.2.2 Les plans de Box-Behnken	54
4.2.3 Les plans de Hoke	54
5 L'analyse canonique	58
5.1 Analyse canonique d'une surface de réponse d'ordre 2	58
5.1.1 Coordonnées du point stationnaire	60
5.1.2 Nature du point stationnaire	62
5.2 Application	65
5.2.1 Les plans de Plackett-Burman	65
5.2.2 Les plans composites centrés	67
Conclusion	75

Introduction

Les principes de planification et d'expérimentation ont été initiés en 1919 par R.A.Fisher à la station expérimentale de Rothamsted et renferment les notions de répétition (replication), de répartition au hasard ou randomisation (randomization), de constitution de blocs (blocking), d'expérience factorielle (factorial experiment), d'effets principaux et interactions et de confusion des effets (confounding) (Fisher, 1925, 1926). Des concepts plus théoriques tels que les notions d'expériences en blocs aléatoires incomplets (incomplete randomized blocks) et d'expérience factorielle fractionnaire (fractional factorial experiment) ont été développés dès 1935 par F.Yates et R.A.Fisher. Les expériences croisées ou avec permutation des traitements (Cross-over ou change-over experiment), ainsi que les expériences factorielles fractionnaires (fractional replication) sont dues aux travaux de W.G.Cochran (1939) et J.Finney (1945 – 1946). Les notions d'expérimentations dans le secteur industriel n'interviennent qu'à partir des années 1935 – 1940. De nouveaux d'autres concepts apparaissent telle que la notion de surface de réponse (response surface) (Box et Wilson, 1951 ; Box, 1952), l'application de cette notion au cas particulier des plans de mélanges (mixture design) (Scheffé, 1958), les plans optimaux (optimal designs) (Keifer, 1959), les plans de Taguchi (1959, 1960, 1987), ... etc. Depuis, les plans d'expériences ont pris un essor considérable avec le développement de l'informatique et la puissance de calcul qui l'accompagne.

L'expérimentateur, quel que soit son domaine d'étude, est toujours confronté au problème difficile de l'organisation optimale de ses essais. Il cherche à obtenir les bonnes informations dans les meilleurs délais et pour le moindre coût pour résoudre son problème. La méthode des plans d'expériences se révèle plus efficace que les démarches habituelles de conduite des essais, c'est-à-dire, qu'elle permet d'arriver aux mêmes résultats avec un nombre d'essais moindre tout en assurant à l'expérimentateur plus de sûreté et de précision sur les résultats attendus.

Dans notre travail, nous nous sommes intéressés à la classe des plans d'expériences associés aux surfaces de réponse. Leur étude nous a semblé d'autant plus intéressante que ces plans s'approprient à une démarche séquentielle, autorisant ainsi une grande souplesse dans l'addition d'autres essais, si les analyses des résultats obtenus s'avèrent non satisfaisantes. Contrairement aux plans d'expériences classiques qui une fois préétablis, ils n'autorisent ni l'addition d'essais supplémentaires ni la suppression d'un essai s'avérant impossible (par exemple une expérience dangereuse!). D'autre part, ces plans s'avèrent très utiles dans une démarche d'optimisation de la réponse Y , si tel est le vœu de l'utilisateur. D'une manière générale, ces plans se caractérisent par deux grandes options qui sont :

- Le choix de l'emplacement des points expérimentaux.
- Le choix de la fonction $f(X_1, X_2, \dots, X_m)$ qui lie la réponse Y aux facteurs quantitatifs X_1, X_2, \dots, X_m contrôlés.

Concernant le premier point, nous avons considéré et étudié ceux dont le domaine expérimental d'intérêt est cubique, la région à explorer est bornée par des contraintes individuelles ou sphériques et ceci lorsque la zone se limite au voisinage d'un point ou d'une cible à atteindre (Feneuille, Mathieu et Phan-Tan-Luu, 1978). Le deuxième point consiste à supposer une certaine régularité de la fonction $f(X_1, X_2, \dots, X_m)$ dans le domaine d'intérêt en l'assimilant à une fonction polynomiale de degrés au plus deux. De ce fait, nous avons structuré notre travail comme suit. Le premier chapitre traite des concepts généraux associés aux plans d'expériences et l'estimation de la surface de réponse. Celui-ci sera suivi par une étude concernant les propriétés statistiques des plans, telles que l'orthogonalité, la presque orthogonalité, l'isovariance ... etc. L'étude théorique de deux types de plans, les plans factoriels complets et les plans fractionnaires, a fait l'objet du chapitre 2. Ces plans se caractérisent par une conception facile et des propriétés intéressantes d'optimalité. Ils servent aussi de plans de base pour plusieurs autres types de plans d'expériences, en particulier ceux qui nous intéressent. Dans le troisième et le quatrième chapitre, nous étudions plus en détail la construction, les propriétés, les avantages et les inconvénients de quelques plans de surface de réponse du premier et deuxième ordre selon que le domaine expérimental d'étude est de forme cubique ou de forme sphérique. Dans le dernier chapitre, nous traitons de la méthode de validation de divers facteurs introduits dans le modèle d'analyse. Quelques outils d'aide à l'interprétation graphique ou analytique, comme l'analyse canonique, seront développés puis illustrés sur des données réelles recueillies auprès du Laboratoire de Génie Microbiologie et Applications (l'Université Mentouri de Constantine) et qui achèveront notre travail.

Chapitre 1

Les plans d'expériences

Avant d'aborder l'étude des plans d'expériences associés aux surfaces de réponse, nous allons présenter quelques notions générales qui régissent leur construction, propriétés et analyse statistique.

1.1 Outils et quelques notions préliminaires

Facteur On appelle facteur, toute série d'éléments de même nature conditionnant le phénomène étudié. Un facteur peut être qualitatif ou quantitatif, continu ou discontinu, contrôlable ou non contrôlable.

- Les facteurs contrôlés sont effectivement étudiés au cours d'une expérience.
- Les facteurs non contrôlés sont les facteurs difficilement maîtrisables et sont la source de variations résiduelles car fréquemment considérés comme aléatoires.
- Les facteurs bruit ont des effets supposés ou réels, et qui ne s'exercent que sur la dispersion ou la qualité des résultats.

Traitement Un traitement est une combinaison spécifique des niveaux de tous les facteurs étudiés.

L'unité expérimentale L'unité expérimentale est l'unité élémentaire qui reçoit un traitement et sur laquelle est faite chaque mesure.

Réponse Une réponse expérimentale ou variable dépendante, est une manifestation mesurable que l'on observe lorsqu'on fait varier les facteurs étudiés.

Bloc Un bloc est le regroupement selon des critères bien définis des unités expérimentales (supposées ou sont) homogènes.

Erreur Variation non expliquée d'un ensemble d'observations.

Interaction Lorsque l'effet d'un facteur dépend de la valeur prise par un autre facteur cela se traduit par une interaction entre les deux facteurs. On appelle interaction d'ordre m une interaction entre m facteurs.

Plan d'expériences Un plan d'expérience est un objet mathématique présenté sous forme de matrice comportant autant de colonnes que de facteurs (m), et autant de lignes que d'expériences (n), de niveaux ou de modalités retenus pour l'expérimentation. L'élément x_{ij} de la matrice correspond au niveau que prend le $j^{\text{ème}}$ facteur à la $i^{\text{ème}}$ expérience.

Domaine expérimental et domaine d'étude Le domaine expérimental est l'espace défini par la variation des facteurs quantitatifs et/ ou par les combinaisons des modalités des facteurs qualitatifs. Un domaine d'étude est défini comme l'ensemble de tous les points de la surface de réponse délimitée par les niveaux bas et haut de chaque facteur ; c'est donc un espace m -dimensionnel, dans lequel chaque point représente un "traitement".

Plan d'expérimentation Il est la traduction d'un plan d'expériences exprimé en fonction des facteurs (ou éventuellement des variables naturelles avant transformation). Pour que le plan permette une interprétation claire en évitant des confusions et assure une précision maximum, l'association des unités expérimentales aux traitements doit obéir aux trois principes de base suivants :

Les trois principes de base

1. La randomisation : l'affectation des niveaux d'un facteur ou des combinaisons des niveaux des facteurs aux unités expérimentales doit être effectuée par un tirage au hasard. Cette opération de randomisation permet d'éviter les biais dans les résultats et garantit l'indépendance des observations d'une unité expérimentale à l'autre.
2. Les répétitions : étant donné le caractère aléatoire de la variable observée Y , il est indispensable d'avoir plusieurs réalisations de Y à chacun de ses niveaux ; ce qui a pour conséquence l'augmentation de la précision des estimateurs des paramètres du modèle.

3. Le contrôle de l'erreur : il peut être obtenu en réduisant le plus possible la part des facteurs non contrôlés par l'expérience. La formation des blocs en unités homogènes rend son estimation plus stable.

1.2 Analyse statistique des plans d'expériences

Dans cette section, nous allons développer des concepts généraux relatifs à l'estimation de la surface de réponse et les propriétés associées.

1.2.1 Choix du modèle

Soit \mathcal{P} un plan d'expérience, \mathcal{U} l'ensemble des n unités expérimentales et $Y_{u,x}$ l'aléatoire observé lorsque le traitement $x = (x_1, \dots, x_m)$ est appliqué à l'unité expérimentale u . Les (x_1, \dots, x_m) sont les réalisations de m **facteurs quantitatifs** (X_1, \dots, X_m) contrôlés.

On suppose que

$$Y_{u,x} = f(x) + \epsilon_{u,x},$$

avec $f(x)$ loi de réponse en x (la réponse en ce point) et ϵ résidu lié aux erreurs expérimentales ou au modèle. f est supposée assez régulière pour admettre un développement limité à l'ordre q dans le domaine expérimental $\mathcal{E} \subset R^m$ contenant l'origine du repère utilisé.

Dans toute la suite, nous nous astreignons aux développements limités polynomiaux d'ordre 1 ou d'ordre 2 pour analyser les plans d'expériences décrits dans le cadre de notre travail ; le nombre de points expérimentaux à réaliser reste raisonnable et la qualité de la prévision est acceptable comme nous le verrons plus loin. Ainsi, le modèle d'ordre $q = 1$ s'écrit pour $u = 1 \dots n, \forall x \in \mathcal{E}$,

$$Y_{u,x} = \beta_0 + \sum_{i=1}^m \beta_i x_i + \epsilon_{u,x},$$

et le modèle d'ordre $q = 2$ s'écrit pour $u = 1 \dots n, \forall x \in \mathcal{E}$,

$$Y_{u,x} = \beta_0 + \sum_{i=1}^m \beta_i x_i + \sum_{i=1}^m \beta_{ii} x_i^2 + \sum_{i < j}^m \beta_{ij} x_i x_j + \epsilon_{u,x}.$$

Désignons par Y le vecteur des n observations, par ϵ le vecteur des n résidus et par β le vecteur des p paramètres tel que $p = \frac{(m+q)!}{m!q!}$. Moyennant une reparamétrisation convenable, les deux modèles peuvent être écrits matriciellement comme suit :

$$Y = X.\beta + \epsilon,$$

où la **matrice des effets** $X_{(n \times p)}$ et le vecteur des coefficients β sont respectivement :

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p-1} \\ 1 & x_{21} & x_{22} & \dots & x_{2p-1} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 1 & x_{n1} & x_{n2} & \dots & x_{np-1} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \cdot \\ \beta_{p-1} \end{bmatrix}.$$

Nous considérons en général les plans complètement randomisés ; nous pouvons alors admettre que pour chaque valeur de Y_x , $E(\varepsilon_x) = 0$, $Var(\varepsilon_x) = \sigma^2$ et les ε_x sont des variables indépendantes.

1.2.2 Estimation

En général, le vecteur des réponses Y est connu puisqu'il contient les réponses expérimentales ; le vecteur des coefficients β est indéterminé et doit être estimé pour fournir une estimation de la surface de réponse $f(x)$, $\forall x \in \mathcal{E} \subset R^m$, de telle manière que l'erreur de modélisation soit minimale.

Si $n > p$, la méthode des moindres carrés due à Legendre (1805) permet d'obtenir

$$\widehat{\beta} = \arg \min_{\beta} L(\beta)$$

solution du système linéaire :

$$L(\beta) = \sum_{i=1}^n \epsilon_i^2 = \epsilon' \epsilon = (Y - X\beta)'(Y - X\beta) = Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta,$$

et

$$\frac{\partial L(\beta)}{\partial \beta} = -2X'Y + 2X'X\beta.$$

Puisque $\frac{\partial^2 L(\beta)}{\partial \beta \partial \beta'} = 2X'X$ (définie non négative), la condition de convexité de la fonction $L(\beta)$ étant satisfaite, le minimum est obtenu par annulation des dérivées premières, d'où le système dit des équations normales :

$$X'X\widehat{\beta} = X'Y.$$

Si X est de plein rang, $X'X$ est inversible et la solution est donnée par

$$\widehat{\beta} = (X'X)^{-1}X'Y.$$

Ainsi, la matrice de variance covariance de $\widehat{\beta}$, est telle que :

$$\text{Var}\widehat{\beta} = (X'X)^{-1}\sigma^2,$$

où, l'estimateur de la variance résiduelle $\widehat{\sigma}^2$ est

$$\widehat{\sigma}^2 = \frac{L(\widehat{\beta})}{n - \text{rang}(X)},$$

et la surface de réponse est estimée par $\widehat{Y}_s = X_s\widehat{\beta}$ (voir théorème de Gauss Markov, Coursol (1980)).

Si X est singulière, on peut avoir recours aux inverses généralisées pour trouver une solution aux équations normales (Cf. coursol, 1980).

Propriétés des estimateurs

- L'estimateur de la variance résiduelle $\widehat{\sigma}^2$ et celui des paramètres $\widehat{\beta}$ sont sans biais et non-corrélés.
- Sous l'hypothèse d'indépendance et d'homoscédasticité des résidus, $\widehat{\beta}$ est le meilleur estimateur linéaire sans biais (BLUE) de β .

1.2.3 Tests d'hypothèses

Soit à tester l'hypothèse nulle $H_0 : K\beta = \mu$ contre son alternative, $H_1 : K\beta \neq \mu$, où K' est une matrice (r, p) dont les r lignes sont linéairement indépendantes. Pour réaliser ce test sur les paramètres du modèle, nous allons additionner l'hypothèse forte suivante : $\epsilon \sim N(0, \sigma^2 I_n)$.

Ainsi :

- la distribution de l'estimateur $\widehat{\beta}$ est aussi normale : $\widehat{\beta} \sim N(\beta, (X'X)^{-1}\sigma^2)$;
- la distribution de $K'\widehat{\beta}$. $\left[\text{Var}(K'\widehat{\beta})\right]^{-1} \cdot K'\widehat{\beta} = Q$ est proportionnelle à un Khi-deux ;
- la distribution de $L(\widehat{\beta}) = Y'Y - \widehat{\beta}X'Y \sim \sigma^2\chi^2_{(n-\text{rang}(X))}$;
- la statistique $F(H_0) = \frac{Q/r}{L(\widehat{\beta})/(n-\text{rang}(X))} \sim \mathcal{F}_{(r; (n-\text{rang}(X)))}$. (Rapport de deux variables indépendantes Khi-deux divisée chacune par son nombre de degrés de liberté).

1.3 Critères de qualité d'une matrice d'un plan d'expériences

Le choix de la matrice X du plan d'expérience dépend de la forme du modèle proposé ou plus généralement des objectifs fixés. La diversité des domaines d'application a conduit à mettre en

évidence de nouvelles propriétés (qualitatives ou quantitatives) et à proposer de nouvelles matrices de plans d'expériences. Certaines de ces propriétés ont été quantifiées et ont ainsi constitué des critères de qualité qui sont parfois redondants. On peut, pour chaque critère, comparer des matrices de plans d'expériences et établir ainsi une hiérarchie entre celles-ci. Les critères les plus importants peuvent être rangés dans une des trois classes :

- La première classe regroupe les propriétés des matrices d'expériences elles-mêmes.
- La deuxième classe fait référence à la qualité des estimateurs des coefficients du modèle mathématique postulé.
- La troisième classe porte sur la qualité prévisionnelle du modèle mathématique.

Ces différentes propriétés dépendent de la structure de la matrice X du plan d'expériences choisi et elles peuvent être déterminées avant qu'aucune expérience n'ait été effectuée.

Dans la littérature de nombreux critères mathématiques ont été développés, nous allons en rappeler quelques uns, ceux principalement concernant les plans d'expériences associés aux surfaces de réponses d'ordre 1 et d'ordre 2 et qui font principalement l'objet de notre étude.

1.3.1 Propriétés associées à la matrice d'un plan d'expériences

a) *Nombres d'expériences*

Toute expérimentation est soumise à des contraintes liées au coût de l'expérience et à sa durée. En effet, des modifications des paramètres mesurés peuvent avoir lieu si l'expérience dure dans le temps par exemple.

Il est donc naturel de chercher à réaliser un nombre minimal d'expériences permettant d'obtenir des informations de qualité suffisante pour répondre aux questions posées (Cf. Drobeky, 1997).

b) *Nombre de niveaux de chaque variable*

Les matrices des plans d'expérience peuvent être soit symétriques soit asymétriques. Toujours pour des raisons souvent pratiques, certaines expériences peuvent s'avérer impossibles à réaliser ou même dangereuses et de ce fait, le domaine expérimental n'est plus un polygone régulier. Pour chaque facteur, on choisit le plus petit nombre de niveaux en les espaçant au maximum afin de couvrir une région expérimentale exhaustive (Cf. Drobeky, 1997).

c) *Séquentialité*

Si le plan réalisé s'avère insuffisant pour apporter l'information recherchée, l'addition de nouvelles expériences peut être souhaitable. Si des modifications éventuelles sont prévisibles ou au moins

soupçonnables dès le début de l'expérimentation, il faut choisir des matrices d'expériences qui possèdent la propriété de séquentialité ou complémentarité. Le cas des matrices des plans composites, que nous étudions plus loin, en sont un exemple.

d) *Le "blocage"*

Des difficultés d'ordre pratique, impossibilité de faire toutes les expériences dans les mêmes conditions ou sur la même installation ou dans une durée suffisamment courte pour éviter l'évolution de certains facteurs non contrôlés incitent à partitionner la matrice d'expériences en blocs homogènes de taille suffisamment restreinte et gagner ainsi en précision (Cf. Drobeky, 1997).

1.3.2 Propriétés des estimateurs des coefficients du modèle mathématique postulé

a) *Facteurs d'inflation*

La variance de $\widehat{\beta}_j$ est donnée par $Var(\widehat{\beta}_j) = c_{jj}\widehat{\sigma}^2$ où c_{jj} est l'élément correspondant à la diagonale de la **matrice de dispersion** $(X'X)^{-1}$; celui-ci dépend donc de la taille du domaine expérimental. Le facteur d'inflation VIF_j lui est préférable et il est défini comme l'élément diagonal de la matrice inverse de la matrice de corrélation et est donc une mesure de l'indépendance des coefficients.

Dans le cas où la matrice d'expériences est optimale, les facteurs d'inflations sont tous égaux à l'unité. Une valeur supérieure à 1 indique que la matrice d'expériences perd de son optimalité (Cf. Drobeky, 1997).

b) *Orthogonalité*

La matrice X du plan d'expériences est dite **orthogonale** si sa **matrice d'information** $(X'X)$ est diagonale.

- Cette propriété d'orthogonalité est facile à respecter pour les plans associés aux surfaces de réponse d'ordre 1 et entraîne par conséquent la non corrélation des coefficients du modèle entre eux.

$$(X'X) = \begin{array}{c|cccc} & X_0 & X_1 & \dots & X_m \\ \hline X_0 & n & \sum x_{u1} & \dots & \sum x_{um} \\ X_1 & & \sum x_{u1}^2 & \dots & \sum x_{u1}x_{um} \\ \vdots & & & \dots & \vdots \\ X_m & sym & & & \sum x_{um}^2 \end{array} .$$

C'est le cas des plans factoriels complets et fractionnaires ainsi que les plans de Plackett-Burman et de Taguchi construits à partir des matrices d'Hadamard, qui sont orthogonales et que nous étudierons au chapitre 3.

Pour les plans associés aux surfaces de réponse d'ordre 2, la matrice d'information étant de la forme :

	X_0	X_1 ... X_m	X_1^2 ... X_m^2	X_1X_2 ... $X_{m-1}X_m$
X_0	n	$\sum x_1$... $\sum x_m$	$\sum x_1^2$... $\sum x_m^2$	$\sum x_1x_2$... $\sum x_{m-1}x_m$
X_1		$\sum x_1^2$... $\sum x_1x_m$	$\sum x_1^3$... $\sum x_1x_m^2$	$\sum x_1^2x_2$... $\sum x_1x_{m-1}x_m$
\vdots		\vdots	\vdots ... \vdots	\vdots ... \vdots
X_m		$\sum x_m^2$	$\sum x_1^2x_m$... $\sum x_m^3$	$\sum x_1x_2x_m$... $\sum x_{m-1}x_m^2$
X_1^2			$\sum x_1^4$... $\sum x_1^2x_m^2$	$\sum x_1^3x_2$... $\sum x_1^2x_{m-1}x_m$
\vdots			\vdots	\vdots ... \vdots
X_m^2			$\sum x_m^4$	$\sum x_1x_2x_m^2$... $\sum x_{m-1}x_m^3$
X_1X_2				$\sum x_1^2x_2^2$... $\sum x_1x_2x_{m-1}x_m$
\vdots				\vdots
$X_{m-1}X_m$	<i>sym</i>			$\sum x_{m-1}^2x_m^2$

La propriété d'orthogonalité, qui est un critère d'optimalité, n'est pas toujours vérifiée. Toutefois des transformations sont possibles pour approcher cette condition en introduisant la propriété de "presque orthogonalité".

c) *Presque-orthogonalité*

Une matrice d'expériences X est presque orthogonale si sa matrice d'information $(X'X)$ privée de sa première ligne et de sa première colonne est diagonale, autrement-dit, la propriété de presque-orthogonalité consiste à rapprocher la matrice de dispersion d'une matrice diagonale.

Ainsi, dans ce cas les coefficients estimés sont au plus corrélés avec le terme constant.

1.3.3 Qualités prévisionnelles du modèle mathématique

Si le modèle postulé est validé par un test statistique adéquat (Cf. section2), c'est- à dire, s'il représente "correctement" le phénomène étudié dans le domaine expérimental, il peut donc être utilisé pour prédire en n'importe quel point de ce domaine, la valeur d'une ou plusieurs réponses expérimentales.

Ainsi, l'évaluation de la qualité de la prévision de la surface de réponse en un point $x \in \mathcal{E}$ (sous espace de \mathcal{R}^m), peut être mesurée par $Var(\widehat{Y}_x)$ qui est égale à

$$Var(\widehat{Y}_x) = x' \cdot (X'X)^{-1} x \cdot \widehat{\sigma}^2 = d_x \cdot \widehat{\sigma}^2,$$

qui dépend de la matrice de dispersion $(X'X)^{-1}$ et de la variance résiduelle, d_x est dite **fonction de variance**.

Une matrice de plan d'expériences est considérée de prévision acceptable si la valeur de sa fonction de variance de réponse est inférieure ou voisine de 1 dans tout le domaine expérimental. Ainsi, $Var(\widehat{Y}_x)$ reste inférieure à la variation résiduelle $\widehat{\sigma}^2$.

a) *Isovariance par rotation*

D'une manière générale, un plan d'expériences à m facteurs quantitatifs est dit isovariant par transformations orthogonales P ($P \subset O(\mathcal{R}^m)$, le groupe orthogonal de \mathcal{R}^m) si et seulement si

$$\forall P \subset O(\mathcal{R}^m), \forall x \in \mathcal{E}(\text{sous espace de } \mathcal{R}^m), Var\widehat{Y}(P_x) = Var\widehat{Y}_x.$$

Remarque 1 *Ce résultat implique que si un plan est isovariant alors les prédicteurs \widehat{Y}_x obtenus ont une variance constante sur toute sphère centrée; cette variance ne dépend que de $r^2 = x'x$ (Box et Hunter, 1957).*

Soit un modèle polynomial de degré d . Considérons un élément de la matrice d'information $(X'X)$, que nous pouvons écrire sous la forme

$$\sum x_1^{\delta_1} x_2^{\delta_2} \dots x_p^{\delta_p},$$

avec

$$\delta_1 + \delta_2 + \dots + \delta_p = \delta.$$

Ce terme général divisé par n (le nombre d'observations total) est appelé le moment d'ordre total δ ($\delta = 1, 2, \dots, 2d$) et il est noté :

$$[1^{\delta_1}, 2^{\delta_2}, \dots, p^{\delta_p}] = \frac{1}{n} \sum x_{i1}^{\delta_1} x_{i2}^{\delta_2} \dots x_{ip}^{\delta_p}.$$

Ces moments sont les éléments de la **matrice des moments**

$$M = n^{-1}(X'X),$$

Une condition nécessaire est suffisante pour qu'une matrice d'un plan d'expériences soit isovariante est donnée par le théorème suivant (Myers, 1971; Draper, 1993)

Théorème 1 *La matrice des moments est isovariante si et seulement si tout moment d'ordre*
 $\delta = \sum_{j=1}^p \delta_j \leq 2d$

(i) *est nul si l'un des δ_j est impair.*

(ii) *est égal à*

$$[1^{\delta_1}, 2^{\delta_2}, \dots, p^{\delta_p}] = \lambda_\delta \frac{\prod_{j=1}^p (\delta_j!)}{2^{\delta/2} \prod_{i=1}^p (\delta_i/2)!},$$

sinon, les λ_δ sont certaines constantes.

Nous donnons l'interprétation de ce théorème sur les deux modèles qui nous concernent. Soit :

Le cas d'un modèle d'ordre 1

Dans ce cas, nous devons considérer les moments jusqu'à l'ordre 2. Compte tenu de (i), les moments impaires d'ordre deux sont nuls et les moments purs d'ordre deux sont tous égaux, en d'autres termes :

$$\begin{aligned} [i] &= 0 & (i = 1, 2, \dots, p), \\ [ij] &= 0 & (i \neq j, i, j = 1, 2, \dots, p), \\ [ii] &= \lambda_2 & (i = 1, 2, \dots, p). \end{aligned}$$

Le cas d'un modèle d'ordre 2

Dans ce cas, nous considérons les moments jusqu'à l'ordre 4. Compte tenu de (i), les moments non nuls ne peuvent être que ceux qui correspondent à

- Un $\delta_j = 2$, les autres δ'_j sont nuls. Ils doivent être tous égaux. Soit λ_2 leur valeur commune.
- Deux δ_j égaux à 2, les autres sont nuls. Soit λ_4 leur valeur commune.
- Un $\delta_j = 4$, les autres δ'_j sont nuls. Ils doivent tous être égaux et leur valeur est $3\lambda_4$.

Ceci peut-être résumé simplement, $\forall i, j = 1 \dots m$ tels que $i \neq j$, par $[ii] = \lambda_2$, $[iijj] = \lambda_4$ et $[iiii] = 3\lambda_4$, avec $[ii] = \frac{1}{n} \sum_{u=1}^n x_{iu}^2$, $[iijj] = \frac{1}{n} \sum_{u=1}^n x_{iu}^2 x_{ju}^2$ et $[iiii] = \frac{1}{n} \sum_{u=1}^n x_{iu}^4$.

Les constantes λ_δ prendront des valeurs particulières pour les plans étudiés ultérieurement dans ce mémoire.

b) *Précision uniforme*

Une matrice d'expériences possède la propriété de précision uniforme si la valeur de la fonction de variance est constante à l'intérieur d'une sphère de rayon égal à la racine carrée du moment d'ordre 2.

Cette propriété assure que la précision de la prédiction est à peu près constante dans tout le domaine expérimental. Elle nécessite que la propriété d'isovariance par rotation soit respectée.

Remarque 2 *Il est important de noter que tous ces critères sont testés a priori, c'est-à-dire indépendamment de tout résultat expérimental. Les plans d'expériences choisis par l'utilisateur peuvent être pris en fonction de certains de ces critères.*

Chapitre 2

Les Plans factoriels et fractionnaires

Définition 1 Soient $\mathcal{U} = \{1, \dots, n\}$ l'ensemble des "unités expérimentales" et \mathcal{F} l'ensemble des traitements munis d'une structure produit canonique $\mathcal{F} = \prod_{l=1}^m \mathcal{F}_l$ et, \mathcal{F}_l désigne le $l^{\text{ième}}$ facteur, l'application $\mathcal{U} \rightarrow \mathcal{F}$ définit un plan **factoriel** \mathcal{P} .

Dans la suite, nous allons nous intéresser au cas où chacun des \mathcal{F}_l est le corps fini à p éléments (p étant une puissance d'un nombre premier), ensuite nous identifions $\mathcal{F} = \prod_{l=1}^m \mathcal{F}_l$ avec l'espace vectoriel \mathcal{F}_p^m et \mathcal{P} est alors noté "plan p^m ", et nous nous intéressons à des combinaisons particulières des éléments de \mathcal{F} , appelées des contrastes plus précisément :

Définition 2 Un contraste $\langle c, t \rangle$ de $t = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_\nu \end{pmatrix}$, est une forme linéaire sur t orthogonale à 1_ν .

Soit $\langle c, 1_\nu \rangle = 0; \nu = p^m$.

Notons que cette identification va nous permettre d'exhiber une décomposition orthogonale de l'espace des contrastes (décomposition reliée aux effets principaux et aux interactions) et choisir éventuellement les sous-espaces à confondre lors de l'étude des plans factoriels fractionnaires.

L'espace projectif

$$\mathcal{P}_{m-1}(\mathcal{F}_p) = \{ \alpha = (\alpha_1, \dots, \alpha_m) \in \mathcal{F}_p^m - \{0\} / \text{le premier } \alpha_i \neq 0 \text{ vaut } 1 \},$$

est l'ensemble des directions de l'espace vectoriel \mathcal{F}_p^m ; $\text{Card} \mathcal{P}_{m-1}(\mathcal{F}_p) = \frac{p^m - 1}{p - 1}$.

A chaque direction α , nous associons une partition de \mathcal{F} en p hyper-plans affines (de p^{m-1} points chacun) orthogonaux à α :

$$\beta_\alpha(l) = \{i \in \mathcal{F}_p^m / \langle i, \alpha \rangle = l\}, l \in \mathcal{F}_p,$$

et un sous espace de contrastes ξ_α de dimension $p - 1$,

$$\xi_\alpha = \left\{ \sum_{l \in \mathcal{F}_p} c_l \left(\sum_{i \in \beta_\alpha(l)} t_i \right) / \sum_{l \in \mathcal{F}_p} c_l = 0 \right\}.$$

Proposition 1 $\{\xi_\alpha\}_{\alpha \in \mathcal{P}_{m-1}(\mathcal{F}_p)}$ constitue une décomposition orthogonale de l'espace des contrastes ξ en $\frac{p^m-1}{p-1}$ sous espace de dimension $p - 1$.

Preuve. Soient α et β deux directions distinctes, ξ_α et ξ_β les deux sous-espaces correspondant sont orthogonaux où

$$\xi_\alpha = \left\{ \sum_{l \in \mathcal{F}_p} c_l \left(\sum_{i \in \beta_\alpha(l)} t_i \right) / \sum_{l \in \mathcal{F}_p} c_l = 0 \right\} \text{ et } \xi_\beta = \left\{ \sum_{l' \in \mathcal{F}_p} d_{l'} \left(\sum_{i \in \beta_\beta(l')} t_i \right) / \sum_{l' \in \mathcal{F}_p} d_{l'} = 0 \right\},$$

respectivement. Si $\langle c, t \rangle \in \xi_\alpha$ et $\langle d, t \rangle \in \xi_\beta$ on a :

$$\langle c, t \rangle = \sum_{l \in \mathcal{F}_p} c_l \left(\sum_{i \in \beta_\alpha(l)} t_i \right) \quad \text{avec} \quad \sum_{l \in \mathcal{F}_p} c_l = 0,$$

et

$$\langle d, t \rangle = \sum_{l' \in \mathcal{F}_p} d_{l'} \left(\sum_{i \in \beta_\beta(l')} t_i \right) \quad \text{avec} \quad \sum_{l' \in \mathcal{F}_p} d_{l'} = 0.$$

Alors,

$$\begin{aligned} \langle c, d \rangle &= \sum_{l \in \mathcal{F}_p} c_l \left(\sum_{i \in \beta_\alpha(l)} t_i \right) \times \sum_{l' \in \mathcal{F}_p} d_{l'} \left(\sum_{i \in \beta_\beta(l')} t_i \right), \\ &= \left(\sum_{i \in \beta_\alpha(l)} t_i \right) \left(\sum_{i \in \beta_\beta(l')} t_i \right) \sum_{l \in \mathcal{F}_p} c_l \sum_{l' \in \mathcal{F}_p} d_{l'}, \\ &= \left(\sum_{i \in \beta_\alpha(l) \cap \beta_\beta(l')} t_i \right) \sum_{l \in \mathcal{F}_p} c_l \sum_{l' \in \mathcal{F}_p} d_{l'}, \\ &= p^{m-2} \sum_{l \in \mathcal{F}_p} c_l \sum_{l' \in \mathcal{F}_p} d_{l'} \quad \text{car } (\text{Card}(\beta_\alpha(l) \cap \beta_\beta(l'))) = p^{m-2} \forall l, l' \text{ et } (\alpha \neq \beta), \\ &= 0. \end{aligned}$$

ξ_α constitue une décomposition orthogonale de l'espace de contraste ξ ; alors $\xi = \bigoplus_{\alpha \in P^{m-1}(\mathcal{F}p)}^\perp \xi_\alpha$. On a

$$\begin{aligned} \dim\left(\bigoplus_{\alpha \in P^{m-1}(\mathcal{F}p)}^\perp \xi_\alpha\right) &= \dim\xi_\alpha \oplus \dim\xi_\alpha \oplus \dots \oplus \dim\xi_\alpha \oplus \dots \text{ tel que } \alpha \in P_{m-1}(\mathcal{F}p), \\ &= \text{Card}(\mathcal{P}_{m-1}(\mathcal{F}p))\dim\xi_\alpha, \\ &= \frac{P^m - 1}{p - 1}(p - 1) \quad \text{car } \text{Card}(\mathcal{P}_{m-1}(\mathcal{F}p)) = \frac{P^m - 1}{p - 1} \text{ et } \dim\xi_\alpha = p - 1, \\ &= p^m - 1, \\ &= \dim\xi. \end{aligned}$$

Interprétation de ξ_α Soit $a_0 = \frac{1}{p^m} \sum_{i \in \mathcal{F}_p^m} t_i$ la moyenne générale des effets traitements.

Le α -effet $X_i^\alpha = A_{i_1}^{\alpha_1} B_{i_2}^{\alpha_2} \dots M_{i_m}^{\alpha_m} = \frac{1}{p^{m-1}} \sum_{j \in B_\alpha(\langle \alpha, i \rangle)} (t_j - a_0)$ est un contraste dans ξ_α pour tout $i \in \mathcal{F}_p^m$. (Il y a seulement pX_i^α distincts car il y a p hyperplans $\beta_\alpha(l)$).

Il s'interprète comme un effet principal ou une composante d'interaction entre les facteurs qui composent l'ensemble des traitements.

Nous allons particulariser l'étude aux plans "plan 2^m ", pour lesquels m facteurs sont étudiés, chacun à deux niveaux.

Ces plans sont les plus simples parmi les plans factoriels complets mais aussi les plus utiles, car ils forment la base de tous les débuts d'étude. En effet, les premiers résultats obtenus grâce à ces plans ou une fraction de ces plans peuvent toujours être complétés par de nouvelles expériences permettant d'atteindre le degré de précision et d'information recherché par l'utilisateur.

2.1 Plans factoriels complets à deux niveaux 2^m

Les sous espaces de contrastes ξ_α au nombre de $2^m - 1$ ($= \text{Card}(\mathcal{P}_{m-1}(\mathcal{F}_2))$) sont de dimension 1 et définissent les effets des m facteurs, toutes les interactions entre les facteurs pris deux à deux, trois à trois, quatre à quatre, ..., m à m . Ainsi, pour un plan 2^2 par exemple, on a

$$\mathcal{P}_1(\mathcal{F}_2) = \{(0, 1); (1, 0); (1, 1)\}.$$

A chacune des 3 directions α , correspond une partition de l'ensemble des traitements $\mathcal{F}_2 = Z_{/2Z} = \{(0, 0); (0, 1); (1, 0); (1, 1)\}$ en 2 hyperplans de 2 points chacun, orthogonaux à α comme suit :

- Pour la direction $\alpha = (1, 0)$,

$$\beta_{(1,0)}(l) = \{i \in \mathcal{F}_2 / i_1 = l \text{ mod}(2)\}.$$

Soient

$$\beta_{(1,0)}(0) = \{(0, 0), (0, 1)\}, \beta_{(1,0)}(1) = \{(1, 0), (1, 1)\},$$

et $\xi_{(1,0)} = \{c_0[(0, 0) + (0, 1)] + c_1[(1, 0) + (1, 1)] / c_0 + c_1 = 0\}$, définit l'effet principal du 1^{er} facteur A .

- Pour la direction $\alpha = (0, 1)$,

$$\beta_{(0,1)}(l) = \{i \in \mathcal{F}_2 / i_2 = l \text{ mod}(2)\},$$

et $\xi_{(0,1)} = \{c_0[(0, 0) + (1, 0)] + c_1[(0, 1) + (1, 1)]\}$ définit l'effet principal du 2^{ème} facteur B .

- Pour la direction $\alpha = (1, 1)$,

$$\beta_{(1,1)}(l) = \{i \in \mathcal{F}_2 / i_1 + i_2 = l \text{ mod}(2)\},$$

et $\xi_{(1,1)} = \{c_0[(0, 0) + (1, 1)] + c_1[(0, 1) + (1, 0)]\}$ définit l'effet interaction $A \circ B$.

Soit, en tenant compte que la moyenne générale $a_0 = \frac{1}{2^2}[(0, 0) + (1, 0) + (0, 1) + (1, 1)]$,

$$\begin{aligned} A &= \frac{1}{2^2} [[(1, 0) + (1, 1)] - [(0, 0) + (0, 1)]], \\ B &= \frac{1}{2^2} [[(0, 1) + (1, 1)] - [(0, 0) + (1, 0)]], \\ A \circ B &= \frac{1}{2^2} [[(0, 1) + (1, 0)] - [(0, 0) + (1, 1)]]. \end{aligned}$$

Géométriquement ces plans sont constitués de 2^m points expérimentaux y compris l'effet moyen a_0 et se situent aux 2^m sommets d'un hypercube à m dimensions.

Remarque 3 *L'analyse statistique de ces plans factoriels s'effectue naturellement par un modèle d'ANOVA à m facteurs à deux niveaux (0 et 1) mais ce modèle peut aussi être considéré comme un modèle de régression linéaire multiple où les variables explicatives prennent les valeurs -1 et*

+1. Conventionnellement, -1 correspond au niveau bas de la variable quantitative et $+1$ au niveau haut. Il est essentiel de remarquer qu'une opération du type

$$x_j = \frac{U_j - U_j^0}{\Delta U_j},$$

où x_j est la valeur de la variable codée j , U_j est la valeur de la variable naturelle j , ΔU_j est le pas de variation, qui transforme n'importe quelle variable quantitative à deux niveaux en variable codée $(-1, +1)$ sans dimension, rendant ainsi comparables les effets des variables naturelles, qui ne sont pas forcément exprimées dans les mêmes unités.

Ainsi, si nous reprenons le système du "plan 2^2 ", Y est la réponse expérimentale, alors

$$Y_{ij} = a_0 + a_i + b_j + (ab)_{ij} \quad i = 1, 2, j = 1, 2$$

avec

- $\sum_{i=1}^2 a_i = 0$, où a_1 et a_2 sont les coefficients de l'effet principal A ;
- $\sum_{j=1}^2 b_j = 0$, où b_1 et b_2 sont les coefficients du deuxième effet principal B ;
- $\sum_i (ab)_{ij} = 0, \forall j = 1, 2$ et $\sum_j (ab)_{ij} = 0, \forall i = 1, 2$, sont les coefficients de l'interaction entre les deux facteurs.

Ces contraintes d'orthogonalité et de régularité du modèle, entraînent l'existence d'un seul coefficient pour l'effet principal A ainsi que pour l'effet principal B et pour l'interaction à savoir :

$$Y_{ij} = a_0 + (-1)^i a + (-1)^j b + (-1)^{i+j} (ab) \text{ pour } i = 1, 2; j = 1, 2, \text{ que l'on peut encore écrire :}$$

$$Y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \text{ avec } x_i = \pm 1 \text{ pour } i = 1, 2.$$

Ainsi, la combinaison du traitement

$$(-1, -1) \text{ se décompose en } a_0 - a_1 - a_2 + a_{12},$$

$$(+1, -1) \text{ se décompose en } a_0 + a_1 - a_2 - a_{12},$$

$$(-1, +1) \text{ se décompose en } a_0 - a_1 + a_2 - a_{12},$$

$$(+1, +1) \text{ se décompose en } a_0 + a_1 + a_2 + a_{12},$$

et nous pouvons résumer ceci par

$$\begin{bmatrix} Y(-1, -1) \\ Y(+1, -1) \\ Y(-1, +1) \\ Y(+1, +1) \end{bmatrix} = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \end{bmatrix}.$$

Et plus généralement pour un "plan 2^m ", le système de 2^m équations à 2^m inconnues,

$$Y = a_0 + \sum_i a_i x_i + \sum_{ij} a_{ij} x_i x_j + \sum_{ijl} a_{ijl} x_i x_j x_l + \cdots + \sum_{ij\dots k} a_{ij\dots k} x_i x_j \dots x_k,$$

peut s'écrire sous une forme plus condensée par

$$Y = X.a,$$

où Y est un vecteur de longueur 2^m ayant pour composantes les réponses de chaque expérience, a est un vecteur de longueur 2^m ayant pour composantes la moyenne, les effets des facteurs et toutes les interactions, qui sont les inconnues que l'on cherche à déterminer et X est une matrice carrée $(2^m, 2^m)$ composée de -1 et $+1$ suivant les valeurs des niveaux x_i , elle est le plan d'expérience du plan factoriel "plan 2^m ".

2.1.1 Construction des plans factoriels complets

Une manière simple de construire les plans factoriels complets est d'étendre celle que nous avons trouvé pour les plans " 2^2 ". La colonne de signes du facteur 1 est

$$- + - + - + - +$$

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.

$$- - + + - - + +$$

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 $+$... etc. Pour chaque facteur, il y a autant de signes $-$ que de signes $+$.

Si les signes $-$ et $+$ de chaque facteur sont disposés comme indiqué, l'ordre des expériences ainsi écrites est couramment appelé ordre standard ou ordre de Yates. La réalisation des expériences ne doit pas suivre cet ordre; il est nécessaire de procéder à une randomisation des essais lors de l'expérimentation.

2.1.2 Les méthodes pour juger de la signification des effets des facteurs

Les techniques d'analyse de la variance permettent de déterminer à partir de quel seuil un effet peut-être considéré comme significatif et avoir une influence sur la réponse Y . Il suffit pour cela de comparer la valeur du rapport F_{obs} (carré moyen du facteur au carré moyen résiduel) au fractile de la loi de Fischer- Snedecor $F_{1-\alpha}$. Il est habituel de résumer les différentes étapes de calculs dans une table dite table d'ANOVA

Tableau 1 - Table d'ANOVA

Source de variation	ddl	Sommes des carrés des écarts	carrés Moyens	F_{obs}	$F_{théo}$	p-valeur
Effet du facteur A	$p - 1$	SCE_a	CM_a	$\frac{CM_a}{CM_r}$	$F_{p-1,r}$	$F_{p-1,r,1-\alpha}$
...
Effet du facteur B	$p - 1$	SCE_b	CM_b	$\frac{CM_b}{CM_r}$	$F_{p-1,r}$	$F_{p-1,r,1-\alpha}$
...
...
Effet interaction $A \circ B$	$(p - 1)^2$	SCE_{ab}	CM_{ab}	$\frac{CM_{ab}}{CM_r}$	$F_{(p-1)^2,r}$	$F_{(p-1)^2, r, 1-\alpha}$
....
Erreur	$r(\text{par différence})$	SCE_r	CM_r	/	/	/
totale	$2^m - 1$	SCE_t	/	/	/	/

2.1.3 Les avantages et les inconvénients

Les avantages des plans factoriels complets sont nombreux et nous en citerons que les principaux :

- Les plans factoriels complets sont faciles à construire.
- Comme chaque facteur ne prend que deux niveaux les essais sont faciles à contrôler et les risques d'erreur sont minimisés.
- Le calcul des effets et des interactions est très simple et ne demande pas d'outils informatiques évolués.

- L'interprétation des résultats est à la portée de tout expérimentateur et ne demande pas de connaissances approfondies en statistiques.
- La modélisation mathématique est immédiate.
- Les résultats obtenus avec un premier plan peuvent être utilisés en partie ou en totalité soit, pour explorer une autre zone du domaine expérimental, soit pour établir un modèle mathématique de degré plus élevé.

Un inconvénient majeur de ces plans factoriels complets est que le nombre d'expériences croît rapidement et cela peut être très coûteux pour l'expérimentateur ou impossible à réaliser. Pour se faire, une fraction de plan judicieusement choisie peut convenir et c'est l'objectif du suivant paragraphe.

2.2 Les plans factoriels fractionnaires à deux niveaux 2^{m-s}

Définition 3 *Un plan fractionnaire 2^{m-s} est un plan factoriel permettant d'étudier m facteurs prenant chacun deux niveaux mais avec un nombre d'essais s fois moins que le plan factoriel complet correspondant.*

Ce plan s'avère donc plus économique du point de vue de nombre d'essais.

Toutefois, un plan fractionnaire de ce type, engendre un système de $n = 2^{m-s}$ équations toujours à p coefficients inconnus ($p > n$) du plan initial ; ce qui introduit une certaine irrégularité du modèle.

Pour résoudre le système d'équations associé, "on regroupe" les coefficients de telle manière qu'il y ait n inconnues. Ces groupes de coefficients sont des contrastes ou des aliasés et les coefficients sont alors dits confondus ou aliasés avec les contrastes.

Mais le problème qui se pose toujours est de savoir si un contraste contient ou non une interaction non négligeable. Pour ce faire nous allons introduire la notion de confusion des effets en utilisant les espaces de contrastes ξ_α et nous exploitons ses résultats pour la construction des plans factoriels fractionnaires.

Dans tous les cas, l'expérimentateur devra mettre en œuvre son bon sens et les connaissances qu'il a du phénomène étudié pour faire le choix du regroupement de sorte que le plan conserve de bonnes propriétés statistiques.

2.2.1 Confusion de ξ_α dans une replicate de 2^{m-s} blocs constitués chacun de 2^s unités

Définition 4 (*confusion*) *Le sous espace des contrastes ξ_α de dimension 1 est dit confondu (dans la replicate), si tous les contrastes dans ξ_α sont totalement confondus avec les blocs.*

Autrement dit, les effets principaux ou interactions sont estimés par la même combinaison linéaire des observations expérimentales que celles des effets blocs. Ainsi, on a les résultats suivants :

R1 ξ_α est confondu si et seulement si, pour tous les traitements i du bloc u , $\langle \alpha, i \rangle$ ne dépend que de u , $u = 1, \dots, 2^{m-s}$; Autrement-dit, $\forall c \in R^2$ avec $\sum_{l=1}^2 c_l = 0$, $\sum_{l=1}^2 c_l (\sum_{i \in B_\alpha(l)} t_i)$ est de la

forme $\sum_{u=1}^{2^{m-s}} a_u (\sum_{j \in \text{Bloc } u} t_j)$ (La partition en blocs de la replicate est plus fine que la partition $\{B_\alpha(l)\}_{l \in \mathcal{F}_2}$).

R2 Si ξ_α et ξ_β sont confondus, $\alpha \neq \beta$ alors $\xi_{\alpha+\lambda\beta}$ est aussi confondu, pour $\forall \lambda \in \mathcal{F}_2$.

Ces propriétés proviennent simplement des règles de calculs dans $\mathcal{F}_2 = Z/2Z$.

2.2.2 Construction de la replicate après choix des ξ_α à confondre

Soient $\alpha^{(l)}, l = 1, \dots, m-s$, $(m-s)$ directions linéairement indépendantes choisies pour le système de confusion d'effets.

Les blocs sont alors définis par

$$\{i \in \mathcal{F}_2^m / \langle \alpha^{(l)}, i \rangle = j_l, l = 1, \dots, m-s\}, (j_1, \dots, j_{m-s}) \in \mathcal{F}_2^{m-s}.$$

Ainsi, supposons par exemple pour un "plan 2^3 " que les conditions expérimentales sont telles ne permettent pas de réaliser des blocs de taille $2^3 = 8$ unités expérimentales, mais il est tout à fait possible de répartir les expériences en 2 blocs de taille 4. Le problème de l'utilisateur qui se pose est de savoir comment répartir les 8 expériences ?

Il est à remarquer que, puisqu'on va rajouter un effet bloc, il y aura donc 9 paramètres à estimer et non plus 8. Un coefficient ne pourra donc être estimé seul et sera confondu nécessairement avec l'effet bloc. Si on décide de confondre l'interaction d'ordre 3, coefficient non interprétable pour l'utilisateur, celle-ci est portée par la direction $\alpha = (1, 1, 1)$ ainsi, le premier bloc est défini par l'équation

$$i_1 + i_2 + i_3 = 0,$$

il est constitué des traitements $\{(1, 1, 0), (1, 0, 1), (0, 1, 1), (0, 0, 0)\}$. Le deuxième bloc est défini par l'équation

$$i_1 + i_2 + i_3 = 1,$$

et il est constitué des traitements $\{(1, 1, 1), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.

De ce fait, le contraste considéré

$$A \circ B \circ C = \frac{1}{2^3} [\{(1, 1, 1) + (1, 0, 0) + (0, 1, 0) + (0, 0, 1)\} - \{(1, 1, 0) + (1, 0, 1) + (0, 1, 1) + (0, 0, 0)\}],$$

est confondu avec la différence des blocs.

Reprenons le "plan 2^3 ".

Si on désire maintenant répartir les 8 traitements dans 4 blocs de taille 2 unités expérimentales. Choisissons 2 directions $\alpha^{(l)}, l = 1, 2$ linéairement indépendantes pour la confusion des effets, par exemple :

$\alpha^{(1)} = (1, 1, 0)$ et $\alpha^{(2)} = (1, 1, 1)$, les blocs étant définis par les systèmes d'équations

$$\begin{cases} i_1 + i_2 = j_1 \text{ mod}(2), \\ i_1 + i_2 + i_3 = j_2 \text{ mod}(2). \end{cases}$$

On en déduit la répartition suivante des traitements dans les blocs :

$$\begin{pmatrix} \text{Bloc1} & \text{Bloc2} & \text{Bloc3} & \text{Bloc4} \\ (0, 0, 0) & (0, 0, 1) & (1, 0, 1) & (1, 0, 0) \\ (1, 1, 0) & (1, 1, 1) & (0, 1, 1) & (0, 1, 0) \end{pmatrix},$$

En pratique, on cherchera le "sous-groupe intrabloc"

$$\mathcal{G} = \{i \in \mathcal{F}_2^m / \langle \alpha^{(l)}, i \rangle = 0, l = 1, \dots, m - s\},$$

qui constituera le premier bloc ; si i_0 appartient à un autre bloc, ce bloc sera $i_0 + \mathcal{G}$...etc. La construction des blocs est simplifiée.

Remarque 4 Dans notre exemple, nous avons choisi de confondre l'interaction $A \circ B$ et $A \circ B \circ C$, pour la construction du plan ; il y a un 3^{ème} effet C (correspondant à la direction $\alpha = (0, 0, 1)$) qui est aussi confondu (d'après R2), c'est donc un effet principal. Ce plan peut être intéressant en pratique, que si nous avons des raisons de penser que cet effet, seul, n'a aucune action sur la réponse mesurée.

2.2.3 Réalisation partielle d'une replicata

Du fait que dans un "plan 2^m " le nombre d'expériences devient vite très grand ($m = 7$ nécessite $2^7 = 128$ expériences!). Si l'expérimentateur décide de ne pas les faire toutes mais à effectuer une réalisation partielle; le problème sera alors de savoir quelles expériences il va réaliser et ce qu'elles lui permettent d'estimer?

Soient $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(s)}$ des directions de \mathcal{F}_2^m linéairement indépendantes; les $\frac{1}{2^s}$ replicates associées sont constituées des 2^{m-s} traitements $i \in \mathcal{F}_2^m$ tels que $\langle \alpha^{(l)}, i \rangle = 0, l = 1, \dots, s$. Pour de tels traitements, $X_i^{\alpha^{(l)}} = \frac{1}{2^{m-1}} \sum_{\{\langle \alpha^{(l)}, i \rangle = 0, l=1, \dots, s\}} (t_i - a_0)$ est donc totalement confondu avec les blocs

$$l = 1, \dots, s, \text{ tous les effets du type } X_i^\alpha = X_i^{\alpha + \sum_{l=1}^s \lambda_l \alpha^{(l)}} = \frac{1}{2^{m-1}} \sum_{\langle \alpha, j \rangle = \langle \alpha, i \rangle} (t_j - a_0), (\lambda_1, \dots, \lambda_s) \in \mathcal{F}_2^s$$

sont mutuellement confondus.

Il est habituel de résumer les équations qui définissent les traitements de la fraction de replicata par $\mathbf{I} \equiv X^{\alpha^{(1)}} \equiv \dots \equiv X^{\alpha^{(s)}}$ et les opérations de calculs dans \mathcal{F}_2 permettent d'exhiber facilement les effets mutuellement confondus entre eux.

Ainsi pour le "plan 2^3 ", parmi les 8 expériences, on décide de ne faire que 4 expériences (la $\frac{1}{2}$ de la replicata), en sacrifiant l'interaction $A \circ B \circ C$ définie par la direction $\alpha = (1, 1, 1)$. L'équation

$$i_1 + i_2 + i_3 = 0$$

définit donc les traitements de la demie replicata; il s'agit de $\{(0, 0, 0), (1, 1, 0), (1, 0, 1), (0, 1, 1)\}$. Les effets mutuellement confondus entre eux, sont alors : A et $B \circ C$, B et $A \circ C$ et enfin C et $A \circ B$ (L'effet associé à une quantité est obtenu en faisant le produit généralisé de cette quantité par celle qui a servi à faire le bloc). Chaque effet principal est associé à l'interaction des 2 autres facteurs. Si les interactions sont négligeables, le dispositif peut être bon.

On peut se demander jusqu'où on peut aller dans le fractionnement du plan? Pour ce faire, la notion de résolution a été introduite et permet de définir pour les plans fractionnaires 2^{m-s} le degré des effets confondus; aidant ainsi au choix du plan d'expériences à réaliser. Les plans fractionnaires de résolution r sont notés 2_r^{m-s} .

2.2.4 La résolution d'un plan d'expérience

Définition 5 *Un plan est dit de résolution r entière si $r = \inf\{\text{nombre de symboles des éléments de la confusion de } \mathbf{I}\}$.*

Le tableau suivant résume quelques cas

Tableau 2 - Résolution - Propriété

Résolution	Propriété
II	Certains effets principaux sont confondus entre eux
III	Certains effets principaux sont confondus avec des interactions d'ordre 2
IV	Certains effets principaux sont confondus avec des interactions d'ordre 3
V	Certains effets principaux sont confondus avec des interactions d'ordre 4

Plus haute est la résolution d'un plan, plus fine est l'information obtenue, mais aussi, plus grand est le nombre d'essais. L'utilisateur doit faire inévitablement un compromis entre l'information désirée et le coût de l'étude.

2.2.5 Construction formelle d'un plan fractionnaire

On choisit un plan complet et l'on écrit sa matrice de calcul en omettant la colonne de signes plus. On appelle cette nouvelle matrice le plan de base. Dans ce plan de base, on choisit une colonne de signes correspondant à une interaction et on l'attribue à un facteur supplémentaire. Les signes de l'interaction choisie deviennent les niveaux d'étude (haut et bas) de ce facteur supplémentaire. On peut généraliser cette méthode et utiliser toutes les colonnes d'un plan de base. Ainsi par exemple sur le plan de base bâti sur la matrice de calcul d'un "plan 2^3 " on peut étudier sept facteurs. Sur le plan de base d'un plan 2^4 , on peut étudier jusqu'à quinze facteurs. D'une manière générale, on peut étudier autant de facteurs supplémentaires qu'il y a d'interactions dans le plan de base (Box et Hunter, 1978). Le Tableau suivant indique le nombre maximum de facteurs que l'on peut étudier sur différents plans de base.

Tableau 3 - Nombre maximum de facteurs étudiés sur un plan de base

plan de base	Nombre de facteur principaux	Nombre d'interactions	Nombre maximum de facteurs étudiés
2^2	2	1	3
2^3	3	4	7
2^4	4	11	15
2^5	5	26	31
2^6	6	57	63
2^7	7	120	127

Un plan de résolution 3 peut apporter rapidement et à moindre frais une première réponse qui permet de sélectionner des facteurs pertinents pour la suite de l'étude. Lorsqu'on détermine les facteurs importants, on s'oriente vers un plan de résolution 4 ou 5. La résolution 5 est souvent difficile à atteindre à cause du nombre d'expériences à réaliser. La résolution 4 fournit dans de nombreux cas un bon premier compromis pour obtenir des résultats fiables sans être obligé d'écarter a priori des facteurs susceptibles d'avoir un rôle non négligeable. Mais parmi les dispositifs expérimentaux de résolution donnée, il y a souvent de multiples choix possibles. D'autre part, il y a des cas où la connaissance que l'on a des facteurs suggère de prendre en compte certaines interactions et d'en ignorer d'autres.

2.2.6 Les avantages et les inconvénients

Les avantages des plans factoriels fractionnaires sont les mêmes que ceux des plans factoriels complets, en plus du fait qu'ils soient beaucoup plus économiques, en nombre d'essais et plus rapides. Il faut encore ajouter un avantage considérable, celui de pouvoir mener les expériences séquentiellement. En effet, si un "plan 2^{m-s} " s'avère non satisfaisant, on poursuivra les expériences en exécutant un deuxième "plan 2^{m-s} " qui est ajouté au premier, ce qui donnera deux "plan 2^{m-s} " soit $2 \times 2^{m-s}$ ou un "plan 2^{m+1-s} " qui restera plus économique que le "plan 2^m " tant qu'il n'a pas été reproduit 2^s fois. Cette possibilité permet à l'expérimentateur de réaliser le nombre juste suffisant d'essais pour aboutir à la résolution de son problème.

Chapitre 3

Plans d'expériences associés aux surfaces de réponse d'ordre 1

L'analyse statistique de ces plans s'effectue à l'aide des modèles d'ordre 1, pour $u = 1 \dots n$, et $\forall x \in \mathcal{E}$, $Y_{u,x} = \beta_0 + \sum_{i=1}^m \beta_i x_i + \varepsilon_{u,x}$. Les paramètres β_i expriment des effets linéaires par rapport aux m facteurs quantitatifs à étudier. Il est souvent suffisant de ne considérer que deux niveaux par facteur afin que le nombre d'essais reste acceptable. Les plans les plus utilisés sont les plans factoriels complets du type "plan 2^m " ou une fraction de ces plans du type "plan 2^{m-s} ". Ces plans ont été décrits en détail au chapitre précédent, vu leur optimalité et leur utilisation dans la construction des autres plans qui en dérivent. Plusieurs autres plans dans la littérature ont été proposés, nous allons décrire les plus utilisés en pratique et ceci pour trois raisons principales :

- * Leur construction est facile et donc leur mise en œuvre est pratique ;
- * Ils possèdent des propriétés statistiques optimales, notamment l'orthogonalité ;
- * Ils permettent d'identifier rapidement les facteurs les plus influents sur la réponse Y .

Il s'agit :

1. Des plans équiradiaux ;
2. Des plans simplexes ;
3. Des plans de Plackett et Burman ;
4. Des plans de Taguchi.

Cet d'ordre que nous avons adopté, résulte tout simplement de la forme de leur domaine expérimental.

3.1 Domaine expérimental sphérique

3.1.1 Les plans équiradiaux

Définition 6 *Un plan équiradial est un plan qui se compose d'un ensemble de points qui sont répartis uniformément sur un cercle, une sphère ou une hyper-sphère de rayon R (Conventionnellement, le rayon est fixé à l'unité).*

Construction

Pour un nombre de facteurs $m = 2$, seul plan équiradial où le nombre de points peut-être quelconque, il y a une infinité de plans d'expériences constitués des sommets d'un polygone régulier inscrit dans un cercle de rayon R à n_1 points "extérieurs" ($n_1 \geq m + 1 = 3$) et n_0 points au centre. Les coordonnées sont exprimées en fonction des coordonnées polaires :

$$x_{i,1} = R \cos\left(\frac{2\pi i}{n_1}\right), x_{i,2} = R \sin\left(\frac{2\pi i}{n_1}\right), i = 1, 2, \dots, n_1.$$

Le tableau suivant indique les plans équiradiaux pour 2 facteurs :

Tableau 4 - Les plans équiradiaux pour 2 facteurs

	Triangle		carré		pentagone		hexagone	
	X_1	X_2	X_1	X_2	X_1	X_2	X_1	X_2
1	0.000	1.000	-0.707	-0.707	1.000	0.000	1.000	0.000
2	-0.866	-0.500	0.707	-0.707	0.309	0.951	0.500	0.866
3	0.866	-0.500	-0.707	0.707	-0.809	0.588	-0.500	0.866
4			0.707	0.707	-0.809	-0.588	-1.000	0.000
5					0.309	-0.951	-0.500	-0.866
6							0.500	-0.866

Représentation graphique

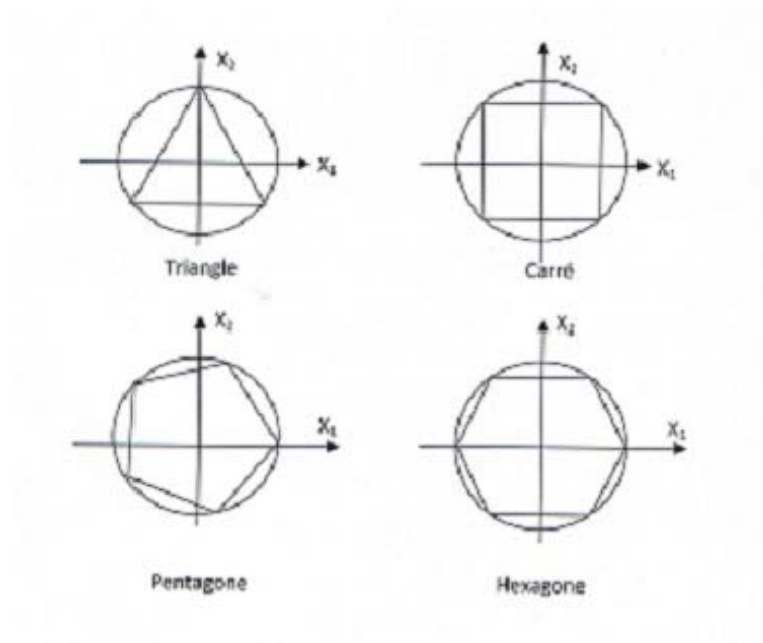


Figure 1 - Les plans équiradiaux pour 2 facteurs

Pour $m = 3$, la théorie limite le nombre de dispositions possibles à cinq (Feneuille, Mathieu et Phan-Tan-Luu, 1978) le tétraèdre, l'octaèdre, le cube, l'icosaèdre et le dodécaèdre. Seuls les trois premiers peuvent être généralisés à $m > 3$, il s'agit du polytope en croix ; (les $2m$ expériences sont disposés symétriquement sur chacun des axes).

3.1.2 Les plans simplexes

Définition 7 *Le plan simplexe (Box, 1952), est un plan orthogonal saturé de premier ordre. Il est constitué de $n = m + 1$ points répartis régulièrement sur une sphère (hyper-sphère). L'angle de deux points quelconques avec l'origine θ est tel que*

$$\cos(\theta) = -\frac{1}{(n-1)} = -\frac{1}{m},$$

Ainsi, par exemple pour $m = 2$, $n = 3$, $\cos(\theta) = -\frac{1}{2}$ et $\theta = 120^\circ$. Les points sont les coordonnées d'un triangle équilatéral inscrit dans un cercle unitaire (Box, Hunter et Hunter, 1978) présenté par

la figure suivante :

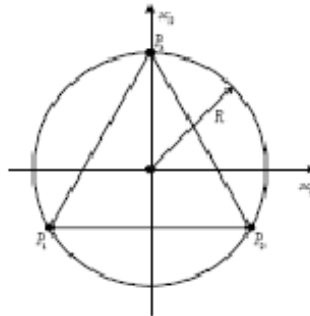


Figure 2 - Plan simplexe pour deux facteurs

Construction

De façon générale, la matrice X pour un plan simplexe à m dimension peut être construite simplement. Si O est une matrice orthogonale d'ordre $n \times n$ dont les éléments de la première colonne sont tous égaux, alors la matrice X telle que

$$X = \sqrt{n}(O),$$

vérifie le critère d'orthogonalité d'un plan d'expériences : $X'X = n(O'O) = nI_n$.

1. Pour une valeur spécifique m , il y a de nombreux plans simplexes qui peuvent être construits. En fait, si on oriente les lignes de la matrice du plan à travers un angle ϕ , le plan d'expériences résultant est encore un simplexe et possède encore la propriété d'orthogonalité du premier ordre. La matrice suivante décrite dans Feneuille, Mathieu et Phan-Tan-Luu (1978) :

N^0 point	x_1	x_2	x_3	x_4	...	x_m
1	$-a_1$	$-a_2$	$-a_3$	$-a_4$...	$-a_m$
2	a_1	$-a_2$	$-a_3$	$-a_4$...	$-a_m$
3	0	$2a_2$	$-a_3$	$-a_4$...	$-a_m$
4	0	0	$3a_3$	$-a_4$...	$-a_m$
5	0	0	0	$4a_4$...	$-a_m$
\vdots	\vdots	\vdots	\vdots	\vdots	...	\vdots
$m+1$	0	0	0	0	...	ma_m

correspondant à une de ces orientations, avec $a_i = \sqrt{\frac{m+1}{m \cdot i(i+1)}}$, $i = 1, \dots, m$, auxquels n_0 points au centre du domaine seront ajoutés pour gagner des degrés de liberté et pouvoir faire les tests statistiques nécessaires à la sélection des facteurs les plus influents.

2. Pour $m = 2$ et $n = 3$, à une rotation près, les plans simplexes et les plans équiradiaux se confondent.

Ces plans simplexes sont utilisés essentiellement dans une démarche séquentielle à la recherche d'un optimum de la réponse mesurée (Cf. Annexe 1).

3.2 Domaine expérimental cubique

Les plans factoriels et fractionnaires sont à la base de tous les plans d'expériences analysables par des modèles polynomiaux de degré 1. Ils ont fait l'objet d'une étude détaillée au cours du chapitre 2 où nous avons en particulier vu leurs définitions et constructions. En particulier,

3.2.1 Les plans de Plackett et Burman (PB)

Ces plans sont utilisés pour sélectionner les facteurs les plus influents sur la réponse Y et où les interactions sont négligeables ou supposées négligeables. Les plans PB sont quelques fois dits plans de criblage ou matrices d'Hadamard. En effet, ces plans se présentent sous forme de matrices carrées orthogonales dont les éléments sont égaux à ± 1 . Plus précisément, on a la définition suivante :

Définition 8 *Un PB est une fraction d'un plan factoriel complet 2^m qui permet d'estimer un modèle d'ordre 1 pour m facteurs quantitatifs ou qualitatifs à deux niveaux ± 1 . Le nombre d'essais n est un multiple de $4 > m + 1$.*

C'est en 1946 que Plackett et Burman ont publié le célèbre article "The designs of optimal multifactorial experiment", dans *Biométrie* dans lequel il est décrit comment construire des plans permettant d'étudier un grand nombre de facteurs avec peu d'essais. Ils indiquent toutes les matrices (n, n) à partir de $n = 8$ et jusqu'à $n = 100$ qu'il est possible de construire sauf la matrice $(96, 96)$, laquelle était plus tard donnée par Baumert, Golomb et Hall en 1962.

Construction d'un plan de Plackett et Burman

Il s'agit d'utiliser la table des signes \pm de Plackett et Burman. Le premier essai correspond à la première ligne de la matrice choisie en fonction du nombre m de facteurs à étudier, les $n - 2$

suiuants sont g n r s par permutation circulaire et on compl te le plan en ajoutant une ligne de signes moins.

Tableau 5 - Table des signes plus et moins pour les plans de Plackett et Burman

m	n	Premier essai du plan d'un plan de Plackett et Burman
7	8	+++ - + - -
11	12	++- + + + - - - + -
15	16	++++ - + - + + - - + - - -
19	20	++- - + + + + - + - + - - - - + + -
23	24	++++ + - + - + + - - + + - - + - + - - - -
35	36	- + - + + + - - - + + + + + - + + + - - + - - - - + - + - + + - - + -

Propri t s

Dans cette partie nous allons donner les diff rents avantages et inconv nients des *PB*.

Les principaux avantages

- Ils sont particuli rement  conomiques pour d tecter les effets principaux. En effet, il suffit de $(k \cdot 2^2)$ exp riences pour tester $((k - 1) \cdot 2^2)$   $(k \cdot 2^2 - 1)$ facteurs   2 niveaux (avec k entier positif).
- Si le nombre d'essais est une puissance de 2, le plan est  quivalent   un plan factoriel fractionnaire et o  le nombre de facteurs m est  gal au nombre d'exp riences moins un (8 essais pour 7 facteurs par exemple).
- Ils sont orthogonaux, permettant ainsi d'estimer les param tres du mod le avec une pr cision optimale.
- Ils sont des plans satur s, car ils comportent autant d'essais que de coefficients (correspondant aux effets principaux)   d terminer dans le mod le math matique et sont de r solution III. Toutefois ils permettent d' tudier les interactions de 2 facteurs si le nombre d'essais n'est pas une puissance de deux sans  tudier les interactions qui sont confondues avec les facteurs simples.

Les inconv nients

- Ces plans n'utilisent que 2 niveaux uniquement, ne peuvent donc  tudier les effets non lin aires.

- Les propriétés des plans du Plackett–Burman sont basées sur celles des matrices d’Hadamard (Cf. Annexe 2), cela contraint toutefois l’expérimentateur à un nombre d’essais n égale à un multiple de 4.

Remarque 5 *Pour $m = n - 1$, les matrices d’Hadamard engendrent des simplexes, c’est à dire des plans équiradiaux saturés.*

3.2.2 Les plans de Taguchi

La méthode de Taguchi vise à combiner un ensemble de techniques pour obtenir des améliorations rapides de la qualité et des coûts de production. Dans un premier temps, elle a connu un succès dans les secteurs industriels et en particulier dans le domaine agroalimentaire, puis elle a suscité l’intérêt de la communauté scientifique pour un développement et une étude plus larges. De nombreux statisticiens, Hunter (1985), Kackar (1985), Box (1988), ont donc essayer de retrouver dans sa présentation les idées générales de la statistique sous-jacentes à sa démarche dont les justifications étaient un peu confuses au départ. Finalement, la "méthode Taguchi" vient enrichir les méthodes de plans d’expériences en apportant une amélioration considérable aux plans factoriels complets et fractionnaires. Elle vise essentiellement à simplifier la mise en œuvre des plans d’expériences. Quel plan d’expériences choisir ? Comment prendre en compte telle ou telle interaction ? ... etc. Elle propose un recueil de tables et des outils d’aide au choix de la table la plus adaptée.

Dénomination des tables

Chaque table peut être identifiée par la forme $L_g(p^f)$ avec

- g : nombre de lignes
- p : nombre de modalités (niveau des facteurs) pour l’ensemble des facteurs
- f : nombre de colonnes (nombre de facteurs + interactions)

Tables de Taguchi

Elles précisent le contenu d’un plan d’expériences qui se présentent sous la forme d’une table orthogonale, avec éventuellement un ou des graphes linéaires et un triangle des interactions. Ces tables orthogonales sont au nombre de 18 et permettent de répondre à un bon nombre de problèmes industriels dans le contrôle de la qualité. Elles peuvent être classées en trois groupes :

- Interactions impossibles $L_{12}(2^{11})$ et $L_{36}(2^{11}x3^{12})$,

- Interactions limitées $L_{18}(2^1x3^7)$, $L_{32}(2^1x4^9)$, $L_{50}(2^1x5^{11})$,
- Interactions possibles $L_4(2^3).L_8(2^7)$, $L_{16}(2^{15}).L_{32}(2^{31})$, $L_{64}(2^{63})$, $L_9(3^4).L_{27}(3^{13})$, $L_{81}(3^{40})$, $L_{36}(2^3x3^{13})$, $L_{54}(2^1x3^{25})$, $L_{16}(4^5).L_{64}(4^{21})$, $L_{25}(5^6)$.

Exemple

Tableau 6 - La table de Taguchi $L_{27}(3^{13})$

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	2	2	2	2	2	2	2	2	2
3	1	1	1	1	3	3	3	3	3	3	3	3	3
4	1	2	2	2	1	1	1	2	2	2	2	2	2
5	1	2	2	2	2	2	2	3	3	3	3	3	3
6	1	2	2	2	3	3	3	1	1	1	2	2	2
7	1	3	3	3	1	1	1	3	3	3	3	3	3
8	1	3	3	3	2	2	2	1	1	1	3	3	3
9	1	3	3	3	3	3	3	2	2	2	1	1	1
10	2	1	2	3	1	2	3	1	2	3	1	2	3
11	2	1	2	3	2	3	1	2	3	1	2	3	1
12	2	1	2	3	3	1	2	3	1	2	3	1	2
13	2	2	3	1	1	2	3	2	3	1	3	1	2
14	2	2	3	1	2	3	1	3	1	2	3	1	2
15	2	2	3	1	3	1	2	3	1	3	2	3	1
16	2	3	1	2	1	3	3	1	2	3	1	2	3
17	2	3	1	2	2	3	1	2	3	1	3	1	2
18	2	3	1	2	3	1	2	3	1	2	3	1	2
19	3	1	3	2	1	3	2	1	3	2	1	3	2
20	3	1	3	2	2	3	1	3	2	1	3	2	1
21	3	1	3	2	3	1	2	3	1	3	2	1	3
22	3	2	1	3	1	3	2	2	1	3	3	2	1
23	3	2	1	3	2	3	1	3	2	1	3	2	1
24	3	2	1	3	3	1	2	3	1	3	2	1	3
25	3	3	2	1	1	3	2	3	1	2	3	1	3
26	3	3	2	1	2	3	1	3	2	1	3	1	2
27	3	3	2	1	3	1	2	3	1	3	2	1	3
	a	b	c	d	e	f	g	h	i	a	b	c	d
		b	c	d	e	f	g	h	i	a	b ²	c ²	d ²
			c	d	e	f	g	h	i	a	b	c ²	d ²
				d	e	f	g	h	i	a	b	c	d ²
					e	f	g	h	i	a	b	c	d
						f	g	h	i	a	b	c	d
							g	h	i	a	b	c	d
								h	i	a	b	c	d
									i	a	b	c	d
										a	b	c	d
											a	b	c
												a	b
													a

Il est possible de séparer le choix des tables en plusieurs étapes, notamment on peut :

- Choisir le nombre de colonnes minimum $f = \text{nombre facteurs} + \text{nombre d'interactions sélectionnées}$;
- Choisir la modalité des facteurs (p) ;
- Choisir parmi les tables comportant le nombre de modalités sélectionnées, celles permettant de respecter le nombre de facteurs, le nombre d'interactions choisies, et la disposition relative des interactions par rapport aux facteurs ;
- Affecter les colonnes aux facteurs et interactions (lecture du graphe linéaire choisi).

Remarque 6

- *il n'est pas nécessaire d'utiliser tous les facteurs et les interactions présents dans la définition de la table.*
- *il existe et on peut modifier des tables existantes de façon à obtenir des tables avec différents nombres de modalités pour les facteurs (exemples : $L_{36}(2^3x3^{13})$, $L_{50}(2^1x5^{11})$)*

Table d'interaction

Elle permet de savoir dans quelles colonnes se trouvent les interactions que l'on a négligées.

Graphes linéaires associés à un modèle

La méthode Taguchi est basée sur une représentation graphique du modèle qui distingue 4 groupes de facteurs en fonction de leur difficulté à être modifiés :

1. Rond vide : facteurs très difficiles à modifier donc changement de modalité doit être rare.
2. Cercle concentrique : facteurs moyennement difficiles à modifier donc changement de modalité doit être peu fréquent.
3. Cercle entourant un rond plein : facteurs assez faciles à modifier donc changement de modalité peut être assez fréquent.
4. Rond plein : facteurs très faciles à modifier donc changement de modalité peuvent être très fréquent.

Les sommets des graphes indiquent les facteurs et les interactions qui sont représentés par un trait entre 2 nœuds (facteurs). Les numéros sur le graphe linéaire sont ceux des colonnes associées aux facteurs et interactions.



Exemple de graphe

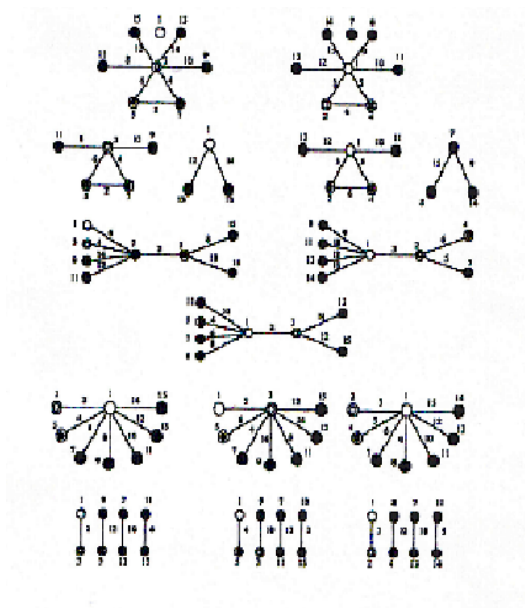


Figure 3 - Graphes linéaires de Taguchi

Les avantages des plans de Taguchi

Les avantages majeurs de l'approche Taguchi sont :

- Les plans associés sont préférés aux méthodes standards des plans d'expériences, d'utilisation simple et ne demandant pas des connaissances particulières en techniques statistiques.
- Ils offrent une stratégie unique des systèmes avec les caractéristiques dynamiques (variables) de la réponse.
- Ils favorisent l'analyse graphique et évitent l'analyse complexe des résultats.

Les inconvénients

La méthode Taguchi est jugée peu efficace par la plus part des statisticiens intéressés et de plus :

- Elle ne tient compte d'aucun progrès réalisé dans l'emploi des plans d'expériences depuis 1950 (transformation des réponses, emploi des modèles non-linéaires, ... etc.).
- Les tables ne contiennent qu'un nombre limité de plan d'expériences, ce qui oblige certains utilisateurs non avertis à modifier leur problème en fonction d'un plan cité.

- Les plans proposés ne sont pas toujours "optimaux" (résolution faible).
- Les méthodes d'analyses sont superficielles et peu "académiques".
- Quelques améliorations sont proposées dans la littérature récente, mais pratiquement tous les plans proposés sont de résolution III.

Les plans que nous avons vu précédemment n'avaient que deux niveaux d'étude par facteur et les modèles mathématiques utilisés étaient du premier ordre (avec ou sans interactions) par rapport à chaque facteur. Ces plans sont les plus employés car ils permettent le criblage des facteurs et conduisent parfois à des modélisations simples mais suffisantes. Pourtant, il existe de nombreux cas où il est nécessaire d'avoir une bonne modélisation des phénomènes étudiés et où il faut passer à des modèles mathématiques du deuxième ordre.

Chapitre 4

Plans d'expériences associés aux surfaces de réponse d'ordre 2

4.1 Domaine expérimental sphérique

4.1.1 Les plans composites

Box et Wilson (1951) ont proposé de nouveaux plans spécialement adaptés à l'ajustement des modèles du 2^{ème} ordre. Leurs premiers plans, appelés les plans composites, sont de différentes classes :

- Si les points en étoile sont sur les faces du cube (ou hypercube), ($\alpha = \pm 1$), le plan est dit plan composite à faces centrées.
- Si les points en étoile sont à l'extérieur du domaine cubique ($\alpha > 1$), le plan est composite centré extérieur (*CCE*).
- Si les points en étoile sont à l'intérieur du domaine cubique ($\alpha < 1$), le plan est composite centré intérieur (*CCI*).

Parmi les nombreux types de plans permettant de construire des surfaces de réponses du deuxième ordre, les plans composites centrés (*CCD*) sont très utilisés et ont une grande popularité auprès des praticiens.

Définition 9 *Le plan composite est une extension du plan factoriel à deux niveaux ou du plan factoriel fractionnaire.*

Construction du plan composite

Le plan composite se prête bien au déroulement séquentiel d'une étude, la première séquence étant la réalisation d'un plan factoriel 2^m (ou fractionnaire 2^{m-s}). Si la validité du modèle du 1^{er} ordre n'est pas retenue, c'est-à-dire, la réponse mesurée statistiquement est différente de celle calculée au même point et s'il semble que le centre de l'expérience soit proche de l'optimum, on ajoutera d'autres essais pour établir un modèle du 2^{ème} ordre. Les nouvelles combinaisons des traitements sont représentées par les points d'expériences à une distance α du centre du domaine expérimental et appartiennent aux axes du plan factoriel, dits points en étoile et par n_0 répétitions au centre du domaine conduisant à n essais égale à la somme :

- des essais du plan factoriel (ou fractionnaire), notés n_f ;
- des essais du plan en étoile, notés n_α ;
- des essais au centre, notés n_0 .

Pour $m = 2, 3, 4$, par exemple, cela donne 9, 15 et 24 essais au lieu de 9, 27 et 81 essais par les plans 3^m .

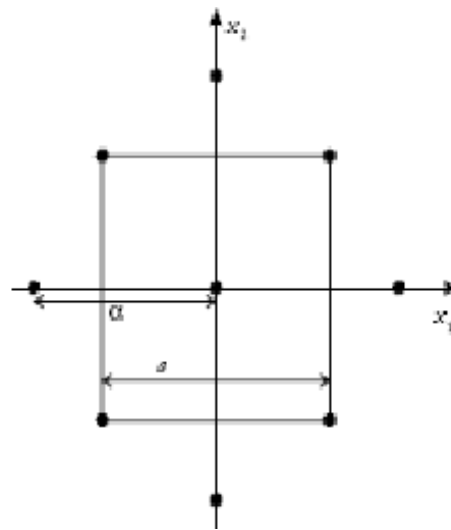


Figure 4 - Plan composite pour 2 facteurs

Remarque 7 Si le plan d'expériences factoriel est un plan fractionnaire, il est préférable qu'il soit de résolution V .

Propriétés

Généralement les critères d'optimalités d'un plan d'expériences sont définis à partir de sa matrice d'information $X'X$, ou sa matrice de dispersion ; ce qui revient à chercher des valeurs particulières de α et n_0 afin d'obtenir les propriétés intéressantes recherchées.

Isovariance par rotation

En fonction de ce critère, et à partir de la matrice d'information $X'X$:

	X_0	X_1	...	X_m	X_1^2	...	X_m^2	X_1X_2	...	$X_{m-1}X_m$
X_0	n	0	...	0	$n_f + 2\alpha^2$...	$n_f + 2\alpha^2$	0	...	0
X_1		$n_f + 2\alpha^2$...	0	0	...	0	0	...	0
\vdots				\vdots	\vdots	...	\vdots	\vdots	...	\vdots
X_m				$n_f + 2\alpha^2$	0	...	0	0	...	0
X_1^2					$n_f + 2\alpha^4$...	n_f	0	...	0
\vdots							\vdots	\vdots	...	\vdots
X_m^2							$n_f + 2\alpha^4$	0	...	0
X_1X_2								n_f	...	0
\vdots										\vdots
$X_{m-1}X_m$	<i>sym</i>									n_f

tous les moments qui doivent être nuls le sont effectivement par construction, et les moment impaires imposent

$$[iiii] = 3[iijj],$$

alors

$$n_f + 2\alpha^4 = 3n_f,$$

donc la valeur de α est

$$\alpha = \sqrt[4]{n_f}.$$

Presque orthogonalité

On pose dans la matrice d'information $X'X$

$$a = n_f + 2\alpha^2, \quad b = n_f, \quad c = n_f + 2\alpha^4,$$

donc la matrice de dispersion $(X'X)^{-1}$ (à σ^2 près) est :

	X_0	X_1 ... X_m	X_1^2 ... X_m^2	X_1X_2 ... $X_{m-1}X_m$
X_0	D	0 ... 0	C ... C	0 ... 0
X_1		E ... 0	0 ... 0	0 ... 0
\vdots			\vdots ... \vdots	\vdots ... \vdots
X_m		E	0 ... 0	0 ... 0
X_1^2			A ... B	0 ... 0
\vdots			\vdots	... \vdots
X_m^2			A	0 ... 0
X_1X_2				F ... 0
\vdots				\vdots
$X_{m-1}X_m$	<i>sym</i>			F

où le déterminant Δ de la matrice d'information est

$$\Delta = (c - b)^{m-1} \cdot [n(c + (m - 1)b) - ma^2],$$

et les éléments de la matrice de dispersion $(X'X)^{-1}$ sont

$$\begin{aligned} A &= \frac{n(c + (m - 2)b) - (m - 1)a^2}{(c - b) \cdot [n(c + (m - 1)b) - ma^2]}, \\ B &= \frac{a^2 - n \cdot b}{(c - b) \cdot [n(c + (m - 1)b) - ma^2]}, \\ C &= \frac{-a}{n(c + (m - 1)b) - ma^2}, \\ D &= \frac{c + (m - 1)b}{n(c + (m - 1)b) - ma^2}, \\ E &= \frac{1}{a}, \\ F &= \frac{1}{b}. \end{aligned}$$

La condition d'orthogonalité ne peut être assurée car le terme C ne peut jamais être nul vu que

$$a = n_f + 2\alpha^2 > 0.$$

Comme la propriété de presque-orthogonalité consiste à rapprocher la matrice de dispersion d'une matrice diagonale, en éliminant la première ligne et la première colonne de $(X'X)^{-1}$ et en posant

$$B = \frac{a^2 - n.b}{(c - b).[n + (c + (m - 1)b) - ma^2]} = 0,$$

ce qui implique

$$\begin{aligned} a^2 &= n.b, \\ (n_f + 2\alpha^2)^2 &= n.n_f, \\ n_f + 2\alpha^2 &= \sqrt{n}.\sqrt{n_f}, \\ 2\alpha^2 &= \sqrt{n}.\sqrt{n_f} - n_f, \\ \alpha &= \sqrt{\frac{\sqrt{n}.\sqrt{n_f} - n_f}{2}}. \end{aligned}$$

Pour respecter la propriété de presque orthogonalité, si les valeurs de n_f et n_0 sont fixées, on peut imposer une relation entre la valeur α et n_0 comme dans le tableau suivant :

Tableau 7 - Valeurs α suivant le nombre de points au centre pour satisfaire la propriété de presque orthogonalité

m	2	3	4	5	5	6	6
plan	2^2	2^3	2^4	2^{5-1}	2^5	2^{6-1}	2^6
n_f	4	8	16	16	32	32	64
n_α	4	6	8	10	10	12	12
$n_0 = 1$	1.000	1.215	1.414	1.547	1.596	1.724	1.761
$n_0 = 2$	1.078	1.287	1.483	1.607	1.662	1.784	1.824
$n_0 = 3$	1.147	1.353	1.547	1.667	1.724	1.841	1.885
$n_0 = 4$	1.210	1.141	1.607	1.719	1.784	1.896	1.943

Les blocs orthogonaux

On peut écrire le modèle d'ordre 2 pour m facteurs et b blocs sous la forme :

$$Y_u = \beta_0^* + \sum_{i=1}^m \beta_i x_{ui} + \sum_{i=1}^m \beta_{ii} x_{ui}^2 + \sum_{i < j}^m \beta_{ij} x_{ui} x_{uj} + \sum_{k=1}^b \delta_k (z_{uk} - \bar{z}_k), u = 1, \dots, n,$$

où $z_{uk} = 1$ si la $u^{\text{ème}}$ observation est dans la $k^{\text{ème}}$ bloc, on dit que les effets bloc sont orthogonaux si :

$$\begin{aligned} \sum_{u=1}^n x_{ui}(z_{uk} - \bar{z}_k) &= 0, i = 1, 2, \dots, m; k = 1, 2, \dots, b, \\ \sum_{u=1}^n x_{ui}x_{uj}(z_{uk} - \bar{z}_k) &= 0, i \neq j; k = 1, 2, \dots, b, \\ \sum_{u=1}^n x_{ui}^2(z_{uk} - \bar{z}_k) &= 0, i = 1, 2, \dots, m; k = 1, 2, \dots, b. \end{aligned}$$

Box et Hunter (1957), ont développé les conditions générales de l'orthogonalité des blocs :

1. Pour un bloc spécifique k , les moments d'ordre 1 et les moments impaires d'ordre 2 sont nuls, en d'autres termes

$$\forall i, j = 1 \dots m \text{ tels que } i \neq j, [i] = 0, [ij] = 0, \text{ pour le bloc } k.$$

2. Pour chaque bloc et chaque facteur i , la somme des carrés des niveaux codés du $i^{\text{ème}}$ facteur dans le bloc doit être proportionnelle au nombre d'observations dans le bloc

$$\sum_{u=1}^n x_{ui}^2 z_{uk} = \bar{z}_k \sum_{u=1}^n x_{ui}^2, i = 1, 2, \dots, m,$$

où \bar{z}_k est la fraction de l'ensemble des points qui sont dans le $k^{\text{ème}}$ bloc.

Cas de 2 blocs Si $n_{0\alpha}$ points sont au centre du bloc axial et n_{0f} points au centre du bloc factoriel, la 1^{ère} condition est satisfaite lorsque nous considérons les points factoriels et les points axiaux comme les deux blocs. La 2^{ème} condition est satisfaite si

$$\frac{\sum_{ax.bloc} x_{ui}^2}{\sum_{fac.bloc} x_{ui}^2} = \frac{n_{\alpha} + n_{0\alpha}}{n_f + n_{0f}}, i = 1, 2, \dots, m,$$

d'où

$$\frac{2\alpha^2}{n_f} = \frac{n_{\alpha} + n_{0\alpha}}{n_f + n_{0f}},$$

donc la valeur de α pour que le plan *CCD* respecte l'orthogonalité du bloc est :

$$\alpha = \sqrt{\frac{n_f(n_{\alpha} + n_{0\alpha})}{2(n_f + n_{0f})}}.$$

Pour un plan CCD qui satisfait à la fois les propriétés de l'isovariance par rotation et de l'orthogonalité des blocs, la valeur de n_{0f} est donnée par :

$$n_{0f} = (\sqrt{n_f}/2)(n_\alpha + n_{\alpha 0}) - n_f.$$

Pour l'orthogonalité des blocs, il est possible pour un plan composite centré de décomposer les deux blocs : le bloc axial ne peut pas être subdivisé, mais le bloc factoriel peut être divisé en 2^m blocs factoriels, donc le nombre total possible des blocs est : 2, 3, 5, 9, 17, ...etc. Si le nombre de points factoriels centrés n_{0f} est divisé par 2^m , donc les points factoriels centrés peuvent être divisés également parmi les 2^m blocs factoriels.

Tableau 8 - Récapitulatif des paramètres des plans bloqués orthogonalement

m	2	3	4	5	$5(\frac{1}{2})rep$	6	$6(\frac{1}{2})rep$	7	$7(\frac{1}{2})rep$
bloc factoriel									
n_f	4	8	16	32	16	64	32	128	64
nombre de blocs	1	2	2	4	1	8	2	16	8
le nombre de points dans chaque bloc	4	4	8	8	16	8	16	8	8
n_{0f}	3	2	2	2	6	1	4	1	1
le nombre total de points dans chaque bloc	7	6	10	10	22	9	20	9	9
bloc axial									
n_a	4	6	8	10	10	12	12	14	14
n_{0a}	3	2	2	4	1	6	2	11	4
le nombre total de points dans le bloc	7	8	10	14	11	18	14	25	18
le nombre total de points pour le plan	14	20	30	54	33	90	54	169	90
valeur de α pour l'orthogonalité du bloc	1.4142	1.6330	2.0000	2.3664	2.0000	2.8284	2.3664	3.3636	2.8284
valeur de α pour l'isovariance par rotation	1.4142	1.6818	2.0000	2.3784	2.0000	2.8284	2.3784	3.3333	2.8284

Les avantages et les inconvénients

L'intérêt des plans composites réside dans le fait qu'ils sont faciles à construire en ajoutant des points à un plan factoriel complet (ou fractionnaire). Il suffit alors de réaliser seulement quelques expériences supplémentaires pour pouvoir estimer la surface de réponse du critère étudié. Ils sont cependant moins économiques en nombre d'essais par rapports à d'autres types de plans, mais ce

nombre reste raisonnable lorsque le nombre de facteurs étudiés reste faible (entre 2 et 4 facteurs). En effet, ce type de plans nécessite 5 niveaux différents pour chacun des facteurs $(-\alpha, -1, 0, +1, \alpha)$, au lieu de trois dans le cas des expériences factorielles $(-1, 0, +1)$; cela peut parfois constituer un inconvénient majeur.

4.1.2 Les plans de Box-Behnken (BB)

Box et Behnken (1960) ont introduit un type différent de plans d'expériences pour les modèles du deuxième ordre qui permettent l'estimation de certaines interactions.

Définition 10 *Un plan de Box et Behnken est une fraction d'un plan factoriel complet 3^m qui permet d'estimer un modèle du second ordre.*

Construction

Un plan de Box-Behnken pour m facteurs est composé d'un plan en blocs incomplets équilibrés BIBD suivant un arrangement particulier de m traitements dans b blocs de taille k , et un plan factoriel à deux niveaux, complété par des points au centre. Dans chaque bloc, un certain nombre de facteurs est utilisé dont les combinaisons constituent le plan factoriel, pendant que les autres facteurs sont maintenus à leurs valeurs centrales. Dans ce plan les points expérimentaux ne se trouvent pas aux sommets du cube ou hypercube mais ils sont aux milieux des arêtes ou au centre des faces ou au centre des cubes, donc sur une sphère ou hypersphère de rayon constant R , le tableau suivant donne certaines valeurs de R

m	3	4	5	6	7
R	$m\sqrt{2}$		$m\sqrt{3}$		

Remarque 8 *Les plans de Box-Behnken pour 3 facteurs impliquent trois blocs, dans chacun, un plan factoriel 2^2 qui est représenté par chaque paire de traitement pendant que le troisième facteur reste fixé à 0. Donc le plan de Box-Behnken de 3 facteurs possède 15 essais (12 arêtes et 3 points au centre).*

En conséquence, les plans de Box-Behnken pour $m = 3$ sont donnés par

X_1	X_2	X_3
$\pm b$	$\pm b$	0
$\pm b$	0	$\pm b$
0	$\pm b$	$\pm b$
0	0	0

dont la représentation graphique est

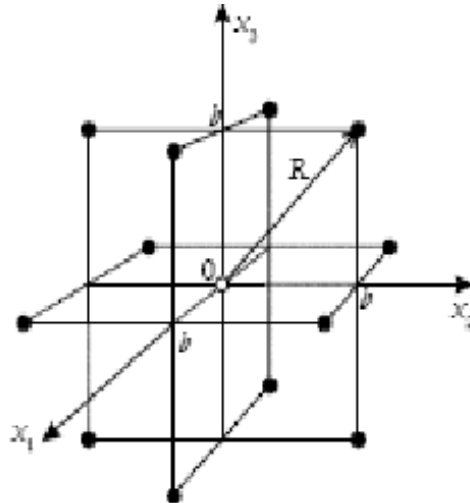


Figure 5 - Plan de Box-Behnken pour 3 facteurs

Il en est de même pour un nombre de facteurs $m = 4$ et $m = 5$. Dans le cas où $m = 6$ et $m = 7$, les BB sont construits à partir d'un plan d'expérience 2^3 avec les $(m - 3)$ autres facteurs fixés à la valeur 0 et répartis suivant un plan en blocs incomplets équilibrés pour $m = 7$ et un plan en blocs incomplets partiellement équilibrés pour $m = 6$, ce qui donne dans ce cas-là, un plan d'expériences qui n'est pas symétrique pour tous les facteurs.

Remarque 9

- *Il n'existe pas de plan de Box-Behnken à deux facteurs. si on peut l'écrire, il ne serait pas un plan très utile.*
- *Les plans de Box-Behnken sont disponibles pour les modèles avec 3 – 12 et 16 facteurs.*

Le tableau suivant indique le nombre d'expériences nécessaires pour la réalisation d'un plan BB.

Tableau 9 - Nombre d'expériences nécessaires pour la réalisation d'un plan BB

m	nombre de facteurs dans chaque bloc	nombre de blocs	n_f dans le bloc	n avec un point au centre	n avec les points au centre
3	2	3	4	13	15
4	2	6	4	25	27
5	2	10	4	41	46
6	3	6	8	49	54
7	3	7	8	57	62
8	4	14	8	113	120
9	3	15	8	121	130
10	4	10	16	161	170
11	5	11	16	177	188
12	4	12	16	193	204
16	4	24	16	385	396

Propriétés des plans de Box-Behnken

Les plans de Box-Behnken possèdent les propriétés suivantes :

- Ils nécessitent 3 niveaux pour chacun des facteurs ; $-b$, 0 et $+b$ pour assurer la condition d'orthogonalité.
- Ils possèdent la propriété de séquentialité.
- Les plans de Box-Behnken pour 4 et 7 facteurs sont isovariants et les autres sont presque isovariants.
- Ils peuvent respecter le critère de presque orthogonalité, si on met 4 points au centre pour le plan à 3 facteurs et 12 points au centre pour le plan à 4 facteurs.
- Les erreurs de prédiction sur les réponses calculées sont plus faibles que les erreurs expérimentales.
- Les plans d'expériences de Box-Behnken sauf pour $m = 3$ et $m = 11$ peuvent être partitionnés en blocs orthogonaux.

Remarque 10 Comme la construction des plans de Box-Behnken n'étant pas figée, pour une valeur de m à une autre, les matrices de dispersion sont différentes donc chaque plan possède des propriétés différentes de l'autre (Cf. Box et Behnken, 1960 ou Goupy, 1999).

Les plans de Box-Behnken sont également des plans conçus pour l'étude des surfaces de réponse d'ordre 2. Ils exigent relativement peu d'essais. Par exemple, un plan de 12 facteurs à 192 arêtes avec 12 points au centre, soit un total de 204 essais ce qui les rend appropriés. Leur principal avantage réside à éviter que les points d'expériences ne soient aux sommets d'un cube ou un hypercube, mais placés à la même distance du centre du domaine.

4.1.3 Les plans de Doehlert

Doehlert (1970) propose des plans d'expériences où les points expérimentaux se répartissent de manière uniforme dans l'espace expérimental.

Définition 11 *Un plan de Doehlert est un plan formé de n_0 essais au centre du domaine et de $m^2 + 2m$ essais répartis " le plus uniformément possible " sur une sphère de rayon 1.*

Construction

La construction des plans de Doehlert s'effectue en utilisant une matrice de coordonnées de points issus de plans simplexes. Ces plans sont classés en trois groupes : $D - 1$, $D - 2$ et $D - 3$, selon leur orientation dans l'espace.

1. Pour les plans d'expériences de Doehlert de type 1, le simplexe de départ est orienté de telle façon qu'un de ses cotés soit parallèles à l'axe du facteur qui semble le plus important. Le simplexe de départ est

$N^\circ \text{point}$	x_1	x_2	x_3	...	x_{m-1}	x_m
1	0	0	0	...	0	0
2	1	0	0	...	0	0
3	0.5	0.866	0	...	0	0
4	0.5	0.289	0.816	...	0	0
\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
m	0.5	0.289	0.204	...	$\sqrt{\frac{m}{2(m-1)}}$	0
$m + 1$	0.5	0.289	0.204	...	$\frac{1}{\sqrt{2(m-1)m}}$	$\sqrt{\frac{m+1}{2m}}$

2. Pour les plans d'expériences de Doehlert de type 2, il suffit de faire subir des rotations aux plans d'expériences $D - 1$ au moyen de matrices de rotation générées à partir des matrices d'Hadamard (Doehlert et Klee, 1972).

3. Pour les plans d'expériences de Doehlert de type 3, un sommet est situé à l'origine des axes donc on utilise comme simplexe de départ

N° point	x_1	x_2	x_3	...	x_{m-1}	x_m
1	0	0	0	...	0	0
2	p	q	q	...	q	q
3	q	p	q	...	q	q
\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
m	q	q	q	...	p	q
$m + 1$	q	q	q	...	q	p

où

$$p = \frac{i-1 + \sqrt{i+1}}{i\sqrt{2}}, q = \frac{\sqrt{i+1} - 1}{i\sqrt{2}}, i = 2, \dots, m.$$

Les coordonnées des points du plan sont alors calculées par soustraction de toutes les lignes du simplexe entre elles, 2 à 2, pour toutes les combinaisons possibles.

Exemple Le plan $D - 1$ pour $m = 2$, les coordonnées du simplexe initial sont

N	X_1	X_2
1	0.000	0.000
2	1.000	0.000
3	0.500	0.866

donc le plan d'expériences est

N	X_1	X_2	obtenu par
1	0.000	0.000	
2	1.000	0.000	
3	0.500	0.866	
4	-1.000	0.000	(1) - (2)
5	-0.500	-0.866	(1) - (3)
6	0.500	-0.866	(2) - (3)
7	-0.500	0.866	(3) - (2)

Tous les points sont à la même distance du centre du domaine d'étude et sont situés sur le cercle trigonométrique. Ils forment un hexagone régulier.

Représentation graphique

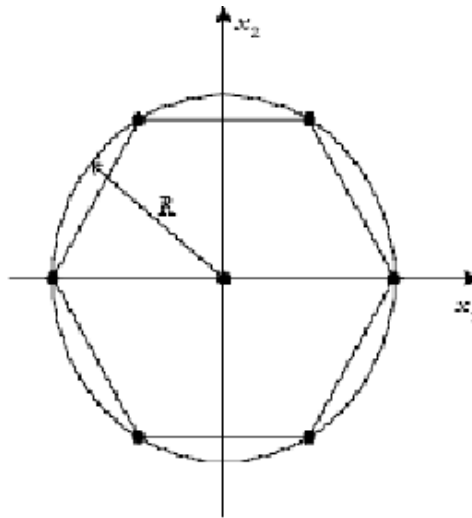


Figure 6 - Plan de Doehlert pour 2 facteurs

Propriétés

Les caractéristiques principales des plans de Doehlert sont :

- Les points expérimentaux sont répartis uniformément dans l'espace des facteurs normés.
- Ils sont extensibles dans n'importe quelle direction.
- Il y a possibilité d'ajouter de nouvelles variables en cours d'expérimentation, sans que le plan d'expériences ne perde de sa qualité.
- Le nombre de points distincts n d'un plan de Doehlert est égal à $n = m^2 + m + 1$.
- Les plans d'expériences asymétriques (nombre de niveau différent) sont :
 - 1- Les plans $D - 1$: le 1^{er} facteur à 5 niveaux, le $m^{\text{ème}}$ facteur à 3 niveaux, les autres facteurs à 7 niveaux.
 - 2- Les plans $D - 2$: le nombre de niveaux est minimal.
 - 3- Les plans $D - 3$: tous les facteurs sont à 7 niveaux.

Remarque 11 *Les plans de Doehlert ne possèdent aucune des propriétés classiques des plans d'expériences de surface de réponses (isovariance par rotation, orthogonalité, précision uniforme etc.). En effet :*

- * Certains moments impairs de la matrice $X'X$ ne sont pas nuls.
- * Il existe des moments pairs qui ne sont pas constants.
- * La deuxième condition d'isovariance par rotation n'est pas vérifiée.

Avantages et inconvénients :

- Les plans de Doehlert sont de qualité suffisante car ils sont économiques en nombre d'essais, ils existent pour un nombre quelconque de facteurs et possèdent la facilité de la recherche d'un optimum ou d'une valeur cible.
- Leur seul inconvénient est que toutes les expériences sont regroupées en un seul bloc.

4.2 Domaine expérimental cubique

4.2.1 Les plans composites

Si on prend $\alpha = 1$ les points axiaux sont situés au centre de la face du cube ou hypercube et ils peuvent alors dans ce cas être considérés comme des plans à domaine expérimental cubique.

4.2.2 Les plans de Box-Behnken

Les plans de Box-Behnken pour les facteurs $m = 3, \dots, 7$ ont a un domaine expérimental cubique, si $b = 1$.

4.2.3 Les plans de Hoke

Hoke (1974) propose des plans pour un modèle du deuxième ordre qui sont des simplifications des plans factoriels à 3 niveaux et qui comportent un nombre plus faible d'expériences.

Construction

Les plans sont construits par regroupement de blocs $S_{r(j)}$, ensemble des points d'un plan factoriel à 3 niveaux, où toutes les expériences qui ont r coordonnées non nulles (± 1) sont partitionnées en blocs de j coordonnées égales à $+1$, ($0 \leq j \leq r$) et les $(r - j)$ autres coordonnées égales à -1 et les $(m - r)$ restantes sont égales à 0. Ces blocs sont regroupés en 4 classes, en fonction du nombre

d'expériences qu'ils comprennent

Classes	Blocs	Nombre d'expériences des blocs
C_1	$-1, 0, 1$	1
C_2	$S_{1(0)}, S_{1(1)}, S_{m-1(m-1)}, S_{m(1)}, S_{m(m-1)}$	m
C_3	$S_{2(0)}, S_{2(2)}, S_{m-2(m-2)}, S_{m(2)}, S_{m(m-2)}$	$m(m-1)/2$
C_4	$S_{3(0)}, S_{3(3)}, S_{m-3(m-3)}, S_{m(3)}, S_{m(m-3)}$	$m(m-1)(m-2)/6$

Remarque 12

- Les trois blocs sont formés par un seul point, noté par (Cf. Hoke, 1974)

$$-1 = S_{m(0)} = (-1, -1, \dots, -1),$$

$$0 = S_{0(0)} = (0, 0, \dots, 0),$$

$$1 = S_{m(m)} = (1, 1, \dots, 1).$$

- Le nombre d'expériences $n_{r(j)}$ du bloc $S_{r(j)}$ est donné par la relation

$$n_{r(j)} = \frac{m!}{(j!(m-r)!(r-j)!)}.$$

Les plans proposés par Hoke (1974) sont regroupés en blocs pour réaliser un nombre total d'expériences minimum pour un modèle du 2^{ème} ordre ($\frac{(m+1)(m+2)}{2}$), ils sont consignés dans le tableau suivant :

	Blocs et classes concernés		
Plan	C_1	C_2	C_3
D_1	-1	$S_{m(m-1)}, S_{1(1)}$	$S_{m(2)}$
D_2	-1	$S_{m(m-1)}, S_{1(0)}$	$S_{m(2)}$
D_3	0	$S_{m(m-1)}, S_{m-1(0)}$	$S_{m(2)}$
D_4	-1	$S_{m(m-1)}, S_{1(1)}, S_{1(0)}$	$S_{m(2)}$
D_5	-1	$S_{m(m-1)}, S_{1(1)}, S_{m-1(0)}$	$S_{m(2)}$
D_6	-1	$S_{m(m-1)}, S_{1(0)}, S_{m-1(-1)}$	$S_{m(2)}$
D_7	0	$S_{m(m-1)}, S_{m-1(0)}, S_{1(1)}$	$S_{m(2)}$

Remarque 13 Dans certains cas particuliers pour $m = 3$, $S_{m(m-1)} = S_{m(2)}$, et pour éliminer la singularité de la matrice d'information $X'X$, on peut remplacer le bloc $S_{3(2)}$ par le bloc $S_{3(1)}$.

Exemple Le plan d'expérience $D4$ de Hoke pour $m = 3$

$$\begin{array}{c|c}
 & \begin{array}{c} \hline -1 \quad -1 \quad -1 \\ +1 \quad +1 \quad -1 \\ +1 \quad -1 \quad +1 \\ -1 \quad +1 \quad +1 \\ \hline +1 \quad 0 \quad 0 \\ 0 \quad +1 \quad 0 \\ 0 \quad 0 \quad +1 \\ \hline +1 \quad -1 \quad -1 \\ -1 \quad +1 \quad -1 \\ -1 \quad -1 \quad +1 \\ \hline -1 \quad 0 \quad 0 \\ 0 \quad -1 \quad 0 \\ 0 \quad 0 \quad -1 \end{array} \\
 D4 & \begin{array}{l} -1 \\ \\ S_{3(2)} = \prod(+1, +1, -1) \\ \\ S_{1(1)} = \prod(+1, 0, 0) \\ \\ S_{3(1)} = \prod(+1, -1, -1) \\ \\ S_{1(0)} = \prod(-1, 0, 0) \end{array} .
 \end{array}$$

Représentation graphique

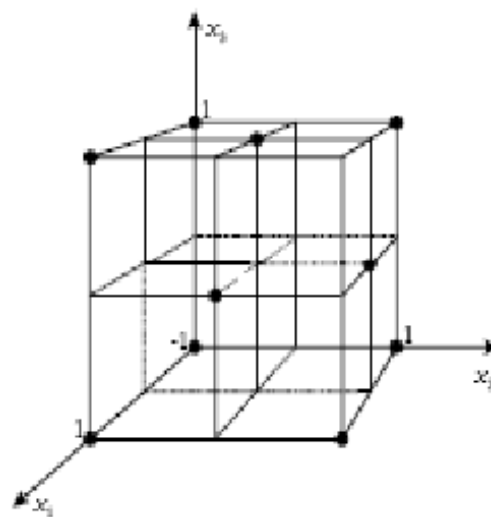


Figure 7 - Plan de Hoke pour 3 facteurs

Propriétés

- Les plans de Hoke sont des plans non orthogonaux, basés sur des plans saturés de fraction irrégulière de plans factoriels 3^m qui est partiellement équilibrée et un sous-ensemble de points axiaux.
- Ils sont économiques et ils existent pour tout facteur $m \geq 3$.
- Il y a 7 types différents de plan de Hoke, $D1, \dots, D7$.
- les plans $D1, D2$ et $D3$ sont des plans saturés.
- les plans $D4, D5, D6$ et $D7$ sont proposés pour augmenter l'efficacité de ces plans et pour obtenir quelques degrés de liberté pour l'estimation de l'erreur expérimentale.

Chapitre 5

L'analyse canonique

5.1 Analyse canonique d'une surface de réponse d'ordre 2

La recherche des points stationnaires est une étape importante dans l'analyse des surfaces d'ordre 2. En effet, une fois le modèle validé par les techniques de l'analyse de la variance, nous pouvons d'une part l'utiliser pour calculer en tout point x du domaine expérimental la réponse prédite \hat{Y} ; il n'est toutefois pas souhaitable d'extrapoler en dehors du domaine puisque la variance de la prévision au point x_0 étant égale à $Var(\hat{Y}_{x_0}) = x'_0(X'X)^{-1}x_0 = d_{x_0}.\sigma^2$ et dépend donc des coordonnées de ce point x_0 . D'autre part, l'interprétation graphique de l'équation de la surface de réponse est très utile dans la recherche d'un optimum. Il s'agit alors de déterminer les valeurs (x_1, x_2, \dots, x_m) des m facteurs quantitatifs qui optimisent la réponse \hat{Y} . Plusieurs cas peuvent se présenter :

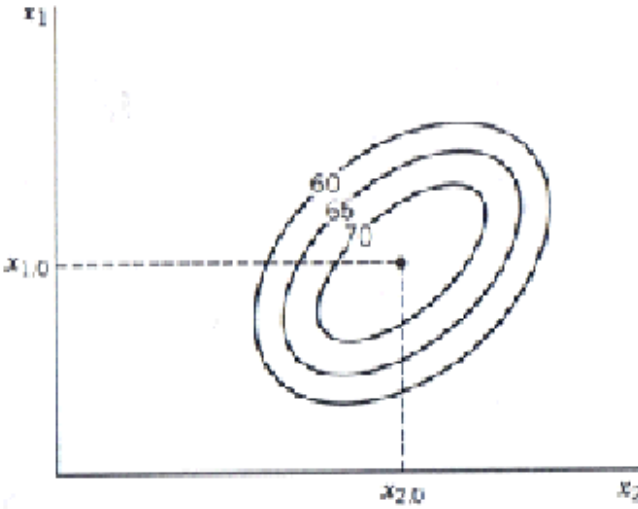


Figure 8 - Le cas d'un point maximum,

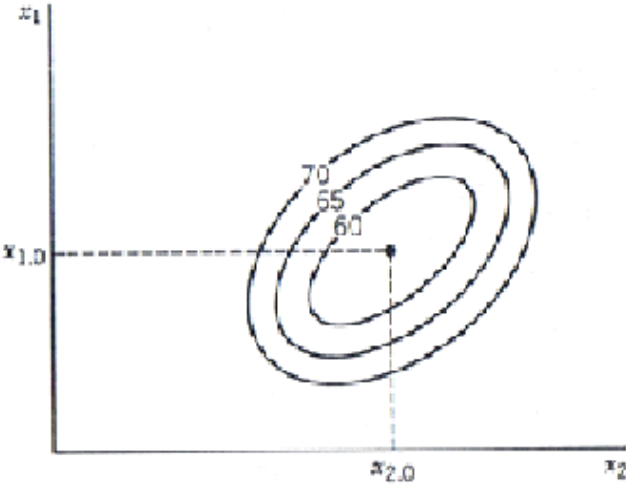


Figure 9 - Le cas d'un point minimum,

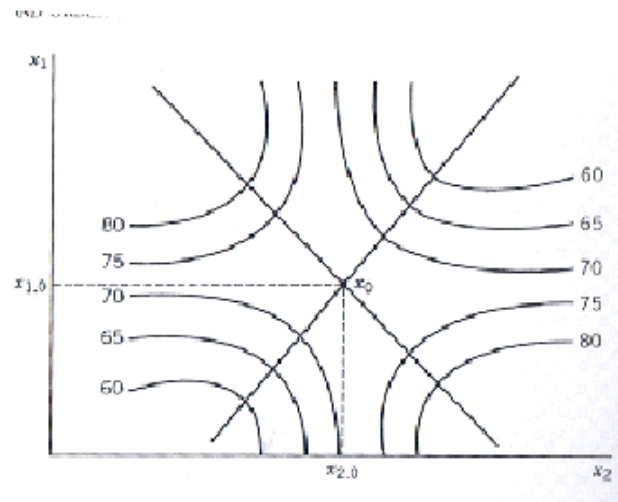


Figure 10 - Le cas d'un point minimax.

Les figures 8 et 9 montrent que les contours de réponse sont constants pour $m = 2$, dans le cas où $m = 3$ ces contours deviennent des surfaces de réponse. Dès que l'on a trois facteurs, il faut faire des coupures pour revenir au cas de deux facteurs. Dans ce cas, il est encore possible d'imaginer le type de surfaces de réponse rencontrées. Au delà de $m \geq 4$, il devient très difficile de savoir quelle est la forme de la surface de réponse. Une méthode d'aide à l'interprétation dans le cas où m est quelconque serait l'**analyse canonique** de l'équation de la surface de réponse. Cette analyse canonique est basée sur la simplification de la relation générale du modèle d'ordre 2 pour m facteurs quantitatifs

$$Y = \beta_0 + \sum_{i=1}^m \beta_i x_i + \sum_{i=1}^m \beta_{ii} x_i^2 + \sum_{i < j}^m \beta_{ij} x_i x_j + \epsilon, \quad (5.1)$$

en une forme canonique; ce qui permet de déterminer les coordonnées du point stationnaire du modèle, la valeur du modèle en ce point et sa nature (point maximum, minimum ou minimax) ainsi, que l'orientation des m axes principaux du modèle dans le domaine d'étude.

5.1.1 Coordonnées du point stationnaire

Ainsi, après estimation des paramètres par la méthode des moindres carrés, on obtient l'équation de surface de réponse qui s'écrit sous la forme matricielle suivante :

$$\hat{Y} = b_0 + x' \cdot b + x' \cdot \hat{B} \cdot x, \quad (5.2)$$

5.1.2 Nature du point stationnaire

a) Translation

Nous allons commencer par une translation de la fonction de réponse de l'origine ($x_1 = 0, x_2 = 0, \dots, x_m = 0$) au point stationnaire x_s qui sera la nouvelle origine. Posons

$$\begin{aligned} z &= x - x_s, \\ \widehat{Y} &= b_0 + x'.b + x'.\widehat{B}.x, \end{aligned} \quad (5.11)$$

cette transformation donne

$$\begin{aligned} \widehat{Y} &= b_0 + (z + x_s)'b + (z + x_s)'.\widehat{B}.(z + x_s), \\ &= b_0 + z'b + x'_s b + z'.\widehat{B}.z + x'_s.\widehat{B}.z + z'.\widehat{B}.x_s + x'_s.\widehat{B}.x_s, \\ &= [b_0 + x'_s b + x'_s.\widehat{B}.x_s] + z'b + z'.\widehat{B}.z + 2x'_s.\widehat{B}.z, \\ &= \widehat{Y}_s + z'.\widehat{B}.z, \end{aligned} \quad (5.12)$$

puisque $2x'_s.\widehat{B}.z = -z'b$. Cette équation représente la fonction de surface de réponse d'ordre 2, par rapport aux nouveaux systèmes d'axes z .

b) Rotation

La forme quadratique $z'.\widehat{B}.z$ peut être réduite en une expression ne contenant que des termes carrés purs en faisant une rotation. Soit P d'ordre $m \times m$ une transformation telle que $w = P'z$ et $P'P = I_m$.

Alors $z'.\widehat{B}.z = z'P\Lambda P'z = w'\Lambda w$ avec Λ diagonale et on a : $z'.\widehat{B}.z = \sum_{i=1}^m \lambda_i w_i^2$.

Les éléments de $\Lambda = (\lambda_i)_{i=1..m}$ sont les valeurs propres réelles de \widehat{B} , puisque \widehat{B} est une matrice réelle symétrique définie positive, et les colonnes de P sont les vecteurs propres associés d'où

$$\widehat{Y} = \widehat{Y}_s + w'P'\widehat{B}Pw = \widehat{Y}_s + w'\Lambda w \quad (5.13)$$

La transformation est une rotation des axes z_i pour former les axes w_i . Nous obtenons finalement l'équation canonique de la surface de réponse

$$\widehat{Y} = \widehat{Y}_s + \sum_{i=1}^m \lambda_i w_i^2. \quad (5.14)$$

c) Interprétation

L'équation (5.14) décrit la nature du point stationnaire. En effet,

- Si, pour $\forall i \in [1, \dots, m]$, $\lambda_i > 0$, le point stationnaire est un minimum, puisque tout déplacement à partir du point stationnaire dans n'importe quelle direction entrainera une augmentation de la réponse.
- Si, pour $\forall i \in [1, \dots, m]$, $\lambda_i < 0$, le point stationnaire est un maximum, puisque tout déplacement à partir du point stationnaire dans n'importe quelle direction entrainera une diminution de la réponse.
- Si, les valeurs propres prennent des signes différents, le point stationnaire est un point minimax.

Par ailleurs, si le domaine d'étude est cubique et possède des bornes égales -1 et $+1$, le point stationnaire est à l'intérieur du domaine si les coordonnées du point stationnaire sont, en valeur absolue, inférieures à $+1$.

Si, le domaine d'étude est sphérique, il faut que la distance d entre l'origine et le point stationnaire soit inférieure au rayon du domaine d'étude

$$d = OS = \left(\sum_{i=1}^m x_{si}^2 \right)^{\frac{1}{2}}. \quad (5.15)$$

En conclusion, si l'optimum est de nature recherché (minimum ou maximum) et s'il se trouve effectivement dans le domaine expérimental, l'analyse analytique permet de conclure. Cependant, dans le cas contraire c'est-à-dire, le point stationnaire n'est pas du type recherché ou il ne se trouve pas dans le domaine expérimental, la solution recherchée ne correspond pas à l'optimum global.

Une recherche locale doit alors être menée sur le domaine expérimental afin de trouver l'optimum local recherché. Pour cette recherche, on utilise la méthode dite "**Ridge Analysis**" qui consiste à se déplacer sur une (hyper) sphère de rayon ρ et trouver un optimum local. En fait, le problème revient à maximiser (ou minimiser selon notre objectif) l'équation

$$Y = b_0 + x'.b + x'.B.x, \quad (5.16)$$

sous la contrainte

$$\sum_{i=1}^m x_i^2 = \rho^2. \quad (5.17)$$

Cette démarche est répétée pour différentes valeurs de ρ , le résultat nous permet de représenter graphiquement la variation de la réponse et la variations des coordonnées du point optimal en

fonction de sa distance au centre du domaine. Soit donc à optimiser la fonction suivante

$$Q(x_1, \dots, x_m) = b_0 + x'.b + x'.B.x - \mu(x'x - \rho^2), \quad (5.18)$$

où μ représente un multiplicateur de Lagrange.

Annulons les dérivées partielles

$$\frac{\partial Q(x_1, \dots, x_m)}{\partial x_i} = b + 2.B.x - 2\mu x = 0, \quad (5.19)$$

on obtient

$$(B - \mu I_m)x = -\frac{1}{2}b. \quad (5.20)$$

La démarche comme décrite dans Mathieu et Phan-Tan-Luu (1978) par exemple, est la suivante

1. Choisir une valeur de μ ;
2. Résoudre l'équation : $(B - \mu I_m)x = -\frac{1}{2}b$;
3. Calculer $\rho = \sqrt{x'x}$;
4. Calculer la valeur de la réponse prévisionnelle en ce point.

Cette démarche est répétée jusqu'à ce que nous obtenions suffisamment de points pour pouvoir représenter graphiquement :

- L'évolution de la réponse maximale (ou minimale) prévisionnelle à une distance ρ du centre du domaine. Cette distance ρ variant de 0 (centre du domaine) à la limite du domaine expérimental d'intérêt.
- L'évolution des coordonnées du point optimal à une distance ρ du centre du domaine. Cette distance ρ variant de 0 (centre du domaine) à la limite du domaine expérimental d'intérêt.

Toutefois, l'obtention du point représentant le maximum (ou le minimum) est conditionnée par les valeurs que μ doit prendre. Le théorème suivant le précise,

Théorème 2 Si $\mu_1 > \lambda_i$ pour tout $i = 1 \dots m$, λ_i étant la $i^{\text{ème}}$ valeur propre de B , alors le point x_1 obtenu en résolvant l'équation $(B - \mu_1 I_m)x = -\frac{1}{2}b$ est un point pour lequel \hat{Y} atteint un maximum local pour ρ_1 (rayon associé à μ_1). Par contre si $\mu_1 < \lambda_i$ pour tout $i = 1 \dots m$, x_1 est un point pour lequel \hat{Y} atteint un minimum pour ρ_1 .

Illustration

L'illustration de ces techniques sera effectuée sur des cas réels que nous traitons dans le paragraphe qui suit :

5.2 Application

Parmi les plans étudiés nous avons choisis deux applications de plans de Plackett-Burman et les plans composites centrés vu le rôle qu'ils jouent dans les différents secteurs industriels comme, la biologie, la chimie, l'agroalimentaire, l'automobile et l'électronique, ... etc.

5.2.1 Les plans de Plackett-Burman

Problématique

Il s'agit de maximiser la production de la protéase acide par *Aspergillus niger*, qui est utilisée largement dans différentes industries : pharmaceutiques, alimentaires et chimique ...etc.

Le plan d'expérience

Une application de ce type de plans dans Mechakra-Maza (2002), a consisté à considérer sept facteurs (5 réels : corn-steep liquor (CSL), ammonium sulfate, déchets de dattes, urée et lactosérum et 2 facteurs erreurs). 8 essais ont été nécessaires pour modéliser la réponse. Les quantités ajoutées sont précisées dans le tableau suivant

Tableau 10 - Plan d'expérimentation (unités courantes)

Facteurs	Niveau bas	Niveau haut
X_1 : corn-steep liquor(CSL)	0	50 ml/l
X_2 : Ammonium sulfate	0	5 g/l
X_3 : Erreur	-	-
X_4 : Déchets de dattes	0	50 g/l
X_5 : Erreur	-	-
X_6 : Urée	0	5 g/l
X_7 : Lactosérum	0	50 ml/l

et le plan d'expérience associé est :

Tableau 11 - Plan d'expériences en unités codées

Expérience	Facteurs							Réponse
	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	Activité AP (mg/ml)
1	+	+	+	-	+	+	+	Y ₁ = 1832
2	-	+	+	+	-	+	-	Y ₂ = 418
3	-	-	+	+	+	-	+	Y ₃ = 437
4	+	-	-	+	+	+	-	Y ₄ = 1881
5	-	+	-	-	+	+	+	Y ₅ = 342
6	+	-	+	-	-	+	+	Y ₆ = 1748
7	+	+	-	+	-	-	+	Y ₇ = 1729
8	-	-	-	-	-	-	-	Y ₈ = 532

Analyse statistique et modélisation

Il s'agit de décrire le comportement de la réponse Y (la production), en fonction des variations des 7 facteurs

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7.$$

Les coefficients β_i sont ceux estimés par la méthode des moindres carrés. Les contributions dues aux facteurs erreurs X₃ et X₅ sont introduites dans le but de fournir une estimation de la variance résiduelles $\hat{\sigma}^2$ où $\hat{\sigma}^2 = \frac{8(\hat{\beta}_3^2 + \hat{\beta}_5^2)}{2}$.

Ainsi chaque estimateur $\hat{\beta}_i$ a pour variance $Var \hat{\beta}_i = \frac{\sigma^2}{8}$, et la statistique test t de student pour tester les coefficients est alors donnée par $\frac{\hat{\beta}_i}{\widehat{Var \hat{\beta}_i}} \sim T_2 \alpha$ (sous l'hypothèse $\beta_i = 0$) au seuil α donné.

Estimation et statistique des coefficients de régression

Tableau 12 - Estimation de coefficients de régression

	Coefficient β_i	Valeur du t test	Seuil de signification
X ₁	682.625	94.875	0.1%
X ₂	-34.625	-4.812	5%
X ₃	-6.125	-	-
X ₄	1.375	0.191	-
X ₅	8.125	-	-
X ₆	-17.625	-2.450	-
X ₇	50.875	-7.071	2%

donc le sous modèle de régression retenu au seuil 0.1% pour le rendement en l'activité protéasique (AP) est

$$Y = 1114.88 + 682.63X_1 + \varepsilon.$$

Résultat

La meilleure production de l'enzyme a été obtenue dans le milieu 4 en présence de CSL, d'urée et de dattes avec un résultat de 1881 u, d'activité protéasique mais les résultats statistiques ont permis de sélectionner uniquement le CSL comme facteur d'enrichissement du milieu. En effet, c'est le seul facteur ayant un coefficient de régression avec l'activité AP le plus élevé, de manière très significative (0.1%). Le coefficient de détermination (R^2), de l'ordre de 0.99 exprime la plausibilité du sous-modèle retenu.

5.2.2 Les plans composites centrés

Problématique

Il s'agit de maximiser la production d'une enzyme (α -Amylase) par la moisissure *Aspergillus oryzae*, qui est utilisé dans de multiples industries : alimentaire (pour améliorer la qualité d'un produit), pharmaceutique, textiles, ... etc.

Le plan d'expérience

En considérant 4 facteurs (amidon, extrait de levure, corn steep liquor et sels), en imposant une précision égale pour la réponse dans toutes les directions et en prenant en compte les effets interactifs et quadratiques des facteurs, 28 essais ont été effectués et sont consignés dans le tableau suivant :

Tableau 13 - Plan d'expériences en unités codées

N° expérience	A	B	C	D	Y
1	-1	-1	-1	-1	440
2	1	-1	-1	-1	530
3	-1	1	-1	-1	466
4	1	1	-1	-1	726
5	-1	-1	1	-1	580
6	1	-1	1	-1	580
7	-1	1	1	-1	360
8	1	1	1	-1	533
9	-1	-1	-1	1	200
10	1	-1	-1	1	360
11	-1	1	-1	1	546
12	1	1	-1	1	893
13	-1	-1	1	1	653
14	1	-1	1	1	666
15	-1	1	1	1	466
16	1	1	1	1	666
17	-2	0	0	0	533
18	2	0	2	0	1200
19	0	-2	0	0	466
20	0	2	0	0	733
21	0	0	-2	0	466
22	0	0	2	0	680
23	0	0	0	-2	333
24	0	0	0	2	413
25	0	0	0	0	333
26	0	0	0	0	333
27	0	0	0	0	333
28	0	0	0	0	332

Analyse statistique des résultats

Il s'agit de décrire le comportement de la réponse Y en fonction des variations des 4 facteurs retenus (qui sont continus et contrôlés de part leur nature). La réponse a été analysée par un modèle statistique linéaire du second degré,

$$Y = \beta_0 + \sum_{i=1}^{i=4} \beta_i X_i + \sum_{i=1}^{i=4} \beta_{ii} X_i^2 + \sum_{i < j} \beta_{ij} X_i X_j + \varepsilon.$$

Les coefficients β_i sont eux estimés par la méthode des moindres carrés.

Estimation et et statistique des coefficient de régression pour Y

Estimated Regression Coefficients for Y					
Term	Coef	SE Coef	T	P	
Constant	308,000	44,07	6,989	0,000	
A	107,375	17,99	5,968	0,000	
B	49,208	17,99	2,735	0,017	
C	32,125	17,99	1,786	0,097	
D	16,458	17,99	0,915	0,377	
A*A	129,385	17,99	7,192	0,000	
B*B	62,635	17,99	3,481	0,004	
C*C	56,010	17,99	3,113	0,008	
D*D	6,010	17,99	0,334	0,744	
A*B	44,812	22,03	2,034	0,063	
A*C	-29,437	22,03	-1,336	0,204	
A*D	12,313	22,03	0,559	0,586	
B*C	-97,187	22,03	-4,411	0,001	
B*D	46,063	22,03	2,090	0,057	
C*D	35,063	22,03	1,591	0,136	
S = 88,14 R-Sq = 91,5% R-Sq(adj) = 82,3%					

A l'aide d'outils classique de la régression linéaire multiple, on peut juger de la qualité de l'ajustement du modèle. En effet, le coefficient de détermination R^2 de l'ordre de 0.92 permet de confirmer la qualité de la modélisation de la réponse.

Une analyse de la variance a permis de tester les effets estimés des 4 facteurs ; De plus, l'erreur pure, estimée à partir des 4 répétitions du point central, a permis d'évaluer l'erreur expérimentale et de réaliser un test du manque d'ajustement.

$F_{obs} = 3.74$ étant non significatif indique que le modèle explique parfaitement les variations de l'activité en fonction des facteurs retenus

Analysis of Variance for Y						
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	14	1087015	1087015	77644	10,00	0,000
Linear	4	366090	366090	91522	11,78	0,000
Square	4	467759	467759	116940	15,05	0,000
Interaction	6	253166	253166	42194	5,43	0,005
Residual Error	13	100986	100986	7768		
Lack-of-Fit	10	93486	93486	9349	3,74	0,153
Pure Error	3	7500	7500	2500		
Total	27	1188001				

L'importance des effets de facteurs est illustrée par la graphe de Pareto,

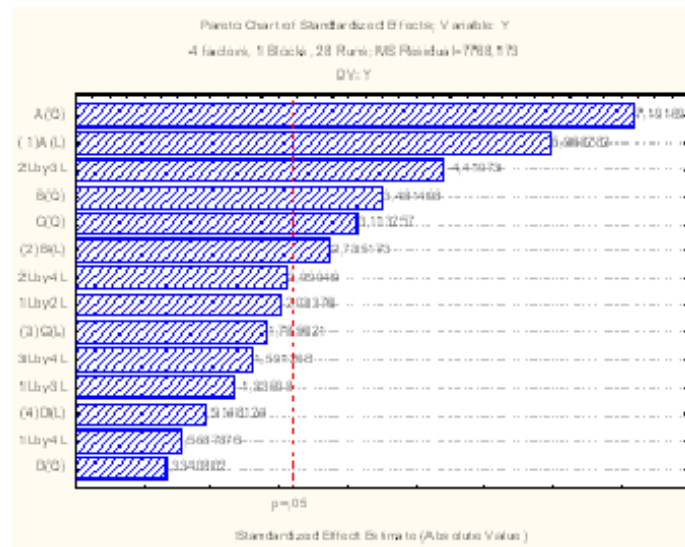


Figure 11 - Graphe de Pareto illustrant l'importance des effets de facteurs

Optimisation de la réponse

Représentation graphique Pour déterminer les valeurs des facteurs qui optimisent la réponse Y, nous avons tenté une représentation graphique par des courbes d'iso-réponses. Les tracés suivants illustrent quelques unes d'entre-elles.

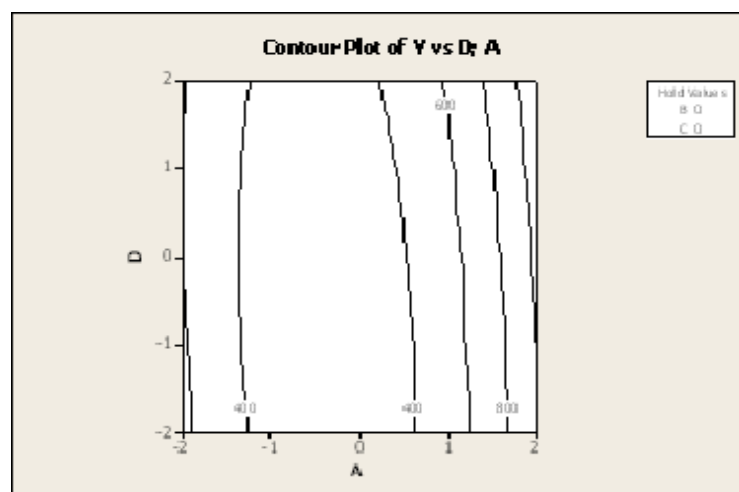


Figure 12 - Graphe contour de surface de réponse

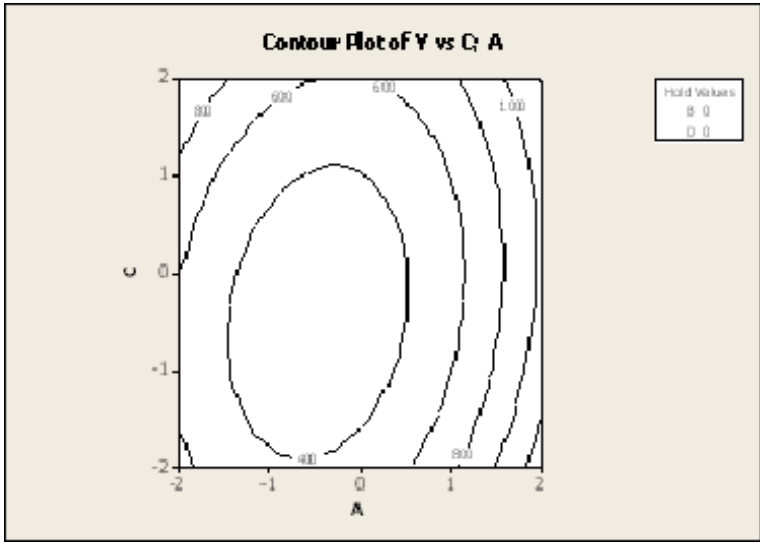


Figure 13 - Graphe contour de surface de réponse

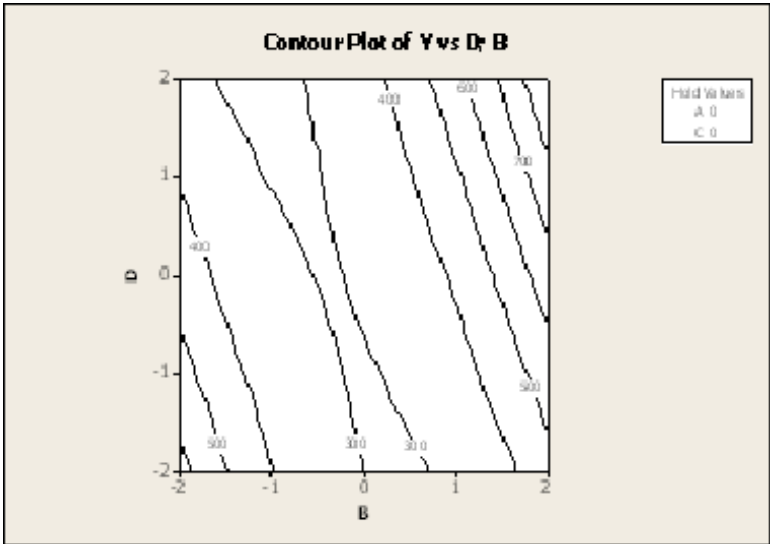


Figure 14 - Graphe contour de surface de réponse

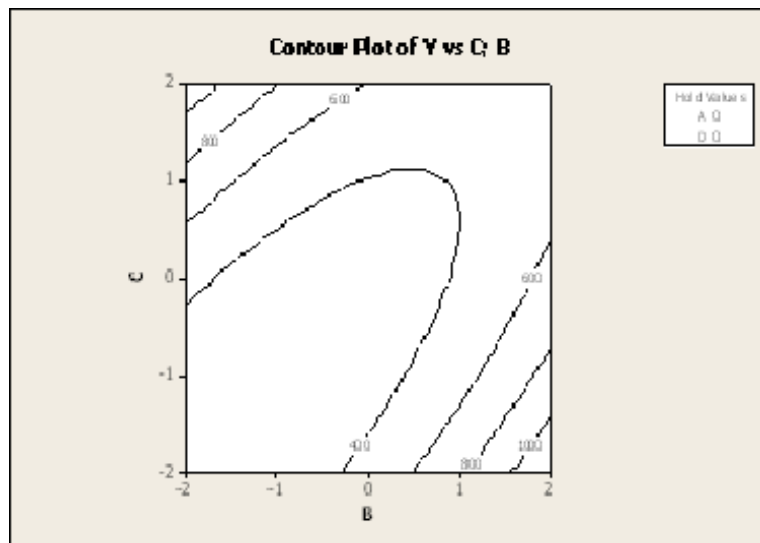


Figure 15 - Graphe contour de surface de réponse

Les profils de ces courbes totalement différents ne permettent donc pas de détecter facilement l'optimum recherché ni sa nature et alors le recours à une méthode analytique s'est imposé.

Analyse canonique de la surface de réponse

La décomposition canonique est donnée par

$$\hat{Y} = 279.59 + 147.85w_1^2 + 90.29w_2^2 + 36.31w_3^2 - 20.42w_4^2,$$

du polynôme décrivant la réponse estimée, a permis de confirmer que le point stationnaire (valeur codée)

$$X_s = \begin{pmatrix} +0.7688 \\ -0.0744 \\ +0.1283 \\ +1.8617 \end{pmatrix},$$

est un point minimax.

L'étude a été poursuivie en utilisant la procédure de la "ridge analysis". Pour notre étude et à la périphérie de la région expérimentale, (fig16 et fig17) les essais réalisés en condition optimale ont permis l'obtention d'une activité α -amylasique de 1166 U ($\mu\text{moles}/\text{min}/\text{ml}$), soit une augmentation

de 111% de la production de l'enzyme recherchée.

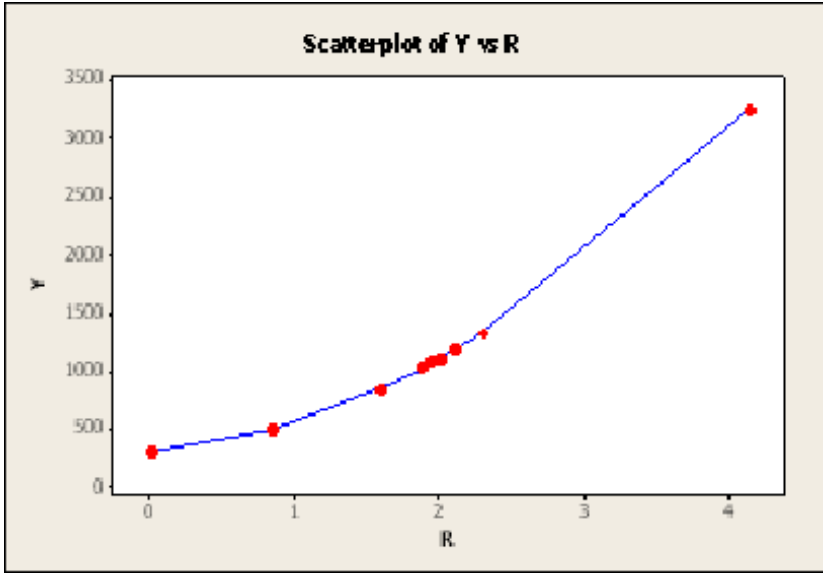


Figure 16 -

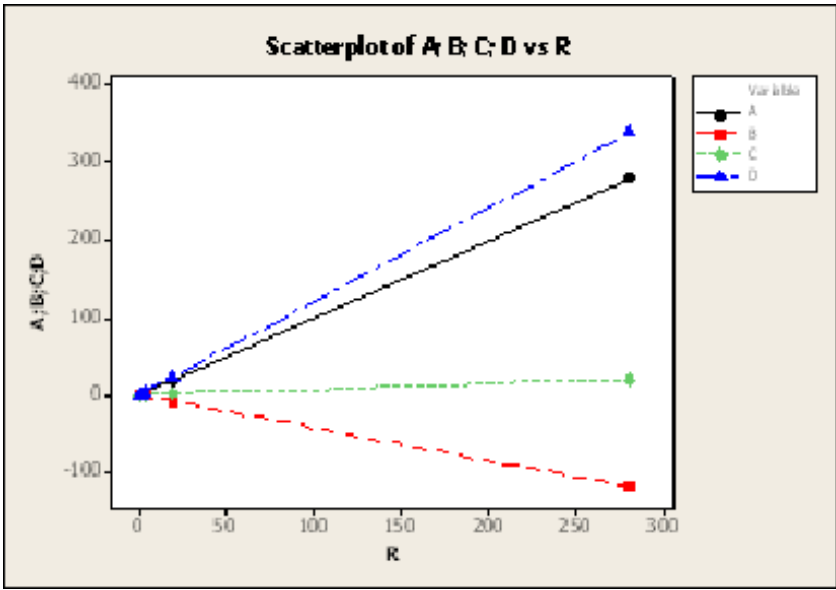


Figure 17 -

Conclusion

Au cours de ce travail, nous nous sommes intéressés aux plans d'expériences associés aux surfaces de réponse d'ordre 1. De part leur construction facile et propriétés d'optimalités ils se révèlent très intéressants pour une sélection rapide et efficace des facteurs les plus influents sur une réponse mesurée. Toutefois, les niveaux des facteurs de ces plans au nombre de deux, ne nous permettent pas de détecter une éventuelle interaction ou effet de courbure de la réponse étudiée. De ce fait, l'introduction des plans d'expériences d'ordre 2 s'avère donc nécessaire. Ces derniers possèdent non seulement des propriétés d'optimalités assez intéressantes mais aussi ils permettent de déterminer l'ensemble de paramètres expérimentaux qui produisent une valeur optimale de la réponse. Si le modèle est validé, l'équation de surface de réponse peut-être utilisée pour prédire la réponse étudiée à n'importe quel point du domaine expérimental.

Notons que le domaine expérimental associé à ces plans, est supposé avoir une forme suffisamment régulière (cubique ou sphérique). Si la forme du domaine d'intérêt est quelconque, les plans que nous venons d'étudier ne peuvent être appliqués ou du moins ils perdent de leur qualité : c'est le cas où une ou plusieurs expériences se révèlent dangereuses à réaliser, par exemple. La conception d'autres plans s'impose donc naturellement.

Concernant les facteurs, ceux utilisés sont considérés comme non liés ; qu'en serait -il dans un problème de formulation, par exemple, où les facteurs sont liés par une certaine relation ?

Enfin, si les surfaces de réponse qui ont été considérées lors de notre modeste travail, sont au plus d'ordre 2, c'est parce que le nombre d'essais augmente rapidement en fonction du degré du polynôme et du nombre de facteurs contrôlés. Au delà de cet ordre, une modélisation par une fonction non linéaire, pourrait être envisageable, moyennant le développement d'autres critères d'optimalité des plans.

Annexe 1

Procédure d'optimisation par la méthode du simplexe

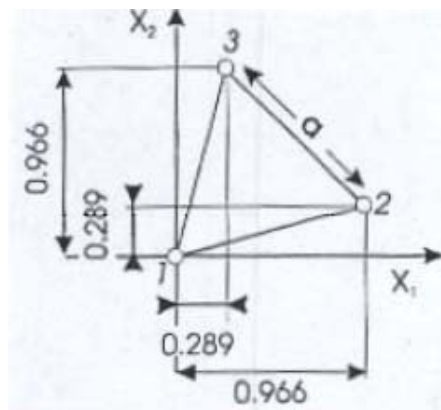
Définition *La méthode de simplexe est une méthode d'optimisation rapide, ponctuelle, séquentielle et monotone. Elle est utilisée quand on n'a qu'une seule variable expliquée ou réponse Y et ne postule aucune forme de modèle mathématique.*

Le but de la méthode du simplexe est de forcer le simplexe à s'écartier des zones où la réponse est mauvaise, pour se rapprocher de celles où elle est bonne. Betteridge en a décrit élégamment la philosophie sous-jacente en disant que c'est comme "gravir une montagne sans carte, mais avec l'objectif d'atteindre le sommet". Après la démarche qui consiste à effectuer l'expérience initiale dans des conditions expérimentales correspondant aux coordonnées des sommets du simplexe et à partir du simplexe de départ, une série de règles pour l'optimisation est utilisée pour organiser une séquence d'expériences.

Construction du simplexe initial le premier problème dans l'optimisation par la méthode du simplexe est le choix du plan initial pour la construire du plan d'expérience résultant. Différents choix sont possibles selon l'orientation de celui-ci.

1. *Un sommet est situé à l'origine des axes*

Représentation graphique



Le plan d'expériences général de ce simplexe est

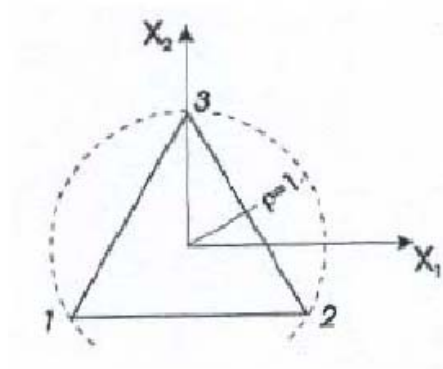
N° point	x_1	x_2	x_3	...	x_{m-1}	x_m
1	0	0	0	...	0	0
2	p	q	q	...	q	q
3	q	p	q	...	q	q
\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
m	q	q	q	...	p	q
$m+1$	q	q	q	...	q	p

où

$$p = \frac{i-1 + \sqrt{i+1}}{i\sqrt{2}}, q = \frac{\sqrt{i+1}-1}{i\sqrt{2}}, i = 2, \dots, m.$$

2. Le centre du plan est situé au centre du simplexe

Représentation graphique



Le plan d'expérience général de ce simplexe est

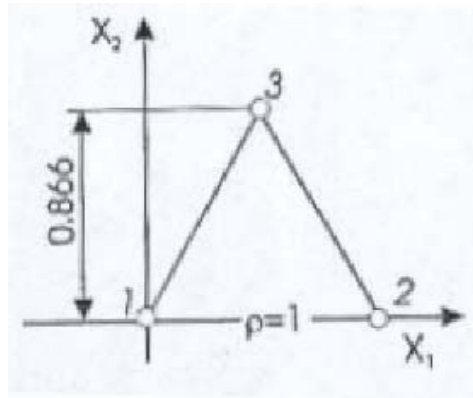
N° point	x_1	x_2	x_3	x_4	...	x_m
1	$-a_1$	$-a_2$	$-a_3$	$-a_4$...	$-a_m$
2	a_1	$-a_2$	$-a_3$	$-a_4$...	$-a_m$
3	0	$2a_2$	$-a_3$	$-a_4$...	$-a_m$
4	0	0	$3a_3$	$-a_4$...	$-a_m$
5	0	0	0	$4a_4$...	$-a_m$
\vdots	\vdots	\vdots	\vdots	\vdots	...	\vdots
$m+1$	0	0	0	0	...	ma_m

avec

$$a_i = \sqrt{\frac{1}{2i(i+1)}}; i = 1, \dots, m.$$

3. Le simplexe de départ est orienté de telle façon qu'un côté soit parallèle à l'axe du facteur qui semble le plus important.

Représentation graphique



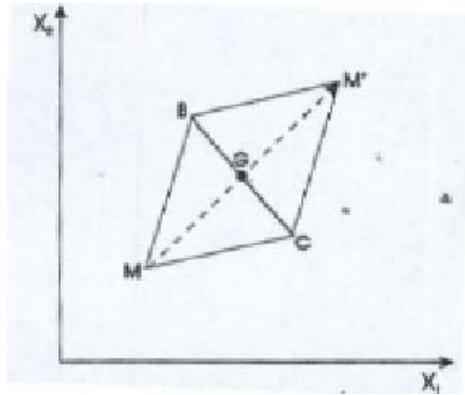
Le plan d'expérience de ce simplexe est

$N^{\circ} \text{ point}$	x_1	x_2	x_3	...	x_{m-1}	x_m
1	0	0	0	...	0	0
2	1	0	0	...	0	0
3	0.5	0.866	0	...	0	0
4	0.5	0.289	0.816	...	0	0
\vdots	\vdots	\vdots	\vdots	...	\vdots	\vdots
m	0.5	0.289	0.204	...	$\sqrt{\frac{m}{2(m-1)}}$	0
$m+1$	0.5	0.289	0.204	...	$\frac{1}{\sqrt{2(m-1)m}}$	$\sqrt{\frac{m+1}{2m}}$

Détermination des coordonnées du point réfléchi M'

Géométriquement, on reflète le simplexe à travers le côté opposé au point le plus faible.

- Dans le cas de deux facteurs



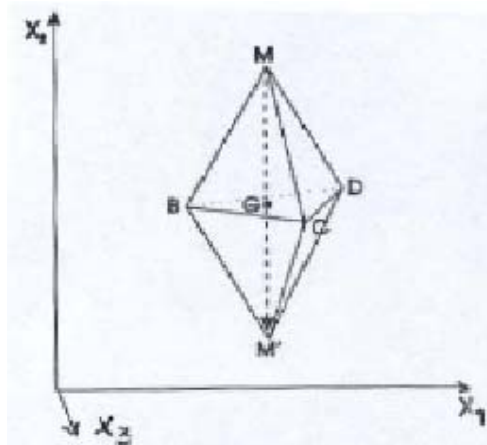
Le point réfléchi M' est le point symétrique du point le plus mauvais M par rapport à G le centre de gravité de B et C. Ainsi :

$$x_{iG} = (x_{iM} + x_{iM'})/2 = (x_{iB} + x_{iC})/2,$$

et donc

$$x_{iM'} = 2x_{iG} - x_{iM} = 2[(x_{iB} + x_{iC})/2] - x_{iM}$$

- Dans le cas de trois facteurs



on a

$$x_{iG} = (x_{iB} + x_{iC} + x_{iD})/3,$$

donc

$$x_{iM'} = 2[(x_{iB} + x_{iC} + x_{iD})/3] - x_{iM}.$$

On peut avoir le cas de plus de trois facteurs, mais de tels simplexes ne pourront pas être représentés en trois dimensions. D'une manière générale, les coordonnées du point G centre de gravité est donnée par la formule

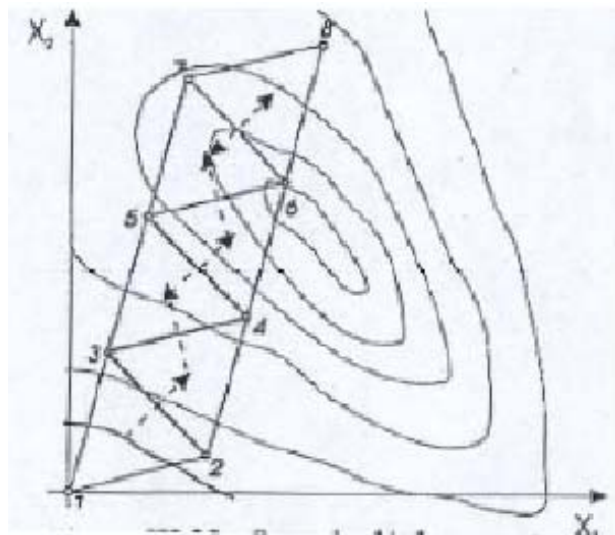
$$x_{iG} = 1/m \sum_{\substack{1 \leq Z \leq m+1 \\ Z \neq M}} x_{iZ},$$

où x_{iZ} la $i^{\text{ème}}$ coordonnée de point Z, d'où

$$x_{iM'} = 2/m \sum_{\substack{1 \leq Z \leq m+1 \\ Z \neq M}} x_{iZ} - x_{iM}.$$

Quand arrête-on la prospection de la surface de réponse ?

Dans la recherche de la meilleure réponse et afin de s'éloigner du plus mauvais point, il est nécessaire d'appliquer une série de règles :



1. Se déplacer d'un pas après chaque réponse.
2. Dans le simplexe de départ le point pour lequel la réponse est la plus mauvaise est remplacé par son symétrique par rapport au centre de gravité des m points restants.
3. Si le point obtenu par symétrie correspond à la réponse la plus mauvaise du nouveau simplexe, former à sa place un autre simplexe en remplaçant l'avant-dernier par son symétrique.
4. Si un nouveau sommet dépasse une contrainte imposée, le rejeter et reprendre les règles 2 et 3.

5. Lorsqu'un sommet du simplexe où la meilleure réponse a été obtenue après M itérations, M étant défini par la relation

$$M = 1,65m + 0,05m^2,$$

(m = nombre de facteurs), on arrête la recherche et on peut le considérer comme optimum.

Remarque

- *Une recherche plus précise de l'optimum peut être conduite à partir du point stationnaire en diminuant la taille du simplexe (de moitié ou du quart.).*
- *Si l'erreur expérimentale est trop grande par rapport à la variation supposée de la réponse, l'apparition d'une erreur de biais sera évitée en remplaçant toutes les anciennes observations du dernier simplexe par de nouvelles, chaque fois que $2(m + 1)$ expériences auront été faites.*

Annexe 2

Les matrices d'Hadamard

Définition Une matrice H carrée d'ordre n est dite matrice d'Hadamard si : $\forall (i, j) \in [1, \dots, n]$: $H_{(i,j)} \in \{-1, 1\}$ et telle que :

$$H^T H = nI_n,$$

avec I_n matrice identité de dimension n .

La condition d'existence d'une matrice d'Hadamard est que $n = 0$ ou 2 modulo(4). Il faudra donc mener un nombre d'essais multiple de 4 pour pouvoir utiliser une telle matrice.

Propriétés La matrice d'Hadamard est une matrice qui possède des propriétés remarquables, notamment :

1. Elle est carrée.
2. Elle possède un nombre de lignes (ou de colonnes) multiple de quatre.
3. Ses éléments sont uniquement des $+1$ ou des -1 .
4. Le produit scalaire de deux colonnes (ou de deux lignes) est toujours égal à zéro. (donc elle est orthogonales)

$$H^T H = nI_n.$$

5.

$$|\det H| = n^{\frac{1}{2}n}.$$

6.

$$H^T H = H H^T.$$

7. Si H' une matrice formée en échangeant deux lignes ou deux colonnes de la matrice H d'Hadamard, alors H' est une matrice d'Hadamard.
5. Si H' une matrice formée d'une matrice H en remplaçant quelques lignes $H_{i,*}$ par $\overline{H_{i,*}}$ ou colonnes $H_{*,i}$ par $\overline{H_{*,i}}$ alors H' est une matrice Hadamard si et seulement si H est une matrice d'Hadamard. (\overline{H} la matrice complémentaire de H , $H_{i,*}$ dénotera les $i^{\text{ème}}$ ligne de H).
8. Il n'existe aucune matrice d'Hadamard $n \times n$ pour $n \notin \{1, 2, \dots, 4k : k \in \mathbb{N}\}$.
9. Si H' et H'' deux matrices d'Hadamard alors le produit de Kronecker $H = H' \cdot H''$ est une matrice d'Hadamard.
10. Une matrice d'Hadamard H est régulière si la somme de chaque ligne dans Z est constante.

Annexe3

Analyse canonique

Posons $\lambda_1, \lambda_2, \dots, \lambda_i, \dots, \lambda_m$ les valeurs propres de la matrice symétrique \widehat{B} , et $w_1, w_2, \dots, w_i, \dots, w_m$ les vecteurs propres associés. On a

$$\widehat{B}w_i = w_i\lambda_i ; i = 1, 2, \dots, m, \quad (1)$$

où $w_i'w_i = 1$, et si P est la matrice d'ordre $m \times m$ dont les colonnes sont les vecteurs propres normalisés associés aux valeurs propres de \widehat{B} , l'équation (1) s'écrit sous la forme

$$\widehat{B}P = P\Lambda, \quad (2)$$

où Λ est la matrice diagonale dont les éléments sont les valeurs propres de \widehat{B} .

Multiplions l'équation (2) par $P' (= P^{-1})$, il vient

$$P'\widehat{B}P = P'P\Lambda = \Lambda. \quad (3)$$

Puisque $P'P = I$, l'équation (5.2)(Cf. chap 5) s'écrit sous la forme

$$\begin{aligned} \widehat{Y} &= b_0 + x'.I.b + x'.I.\widehat{B}.I.x, \\ \widehat{Y} &= b_0 + x'.P'P.b + x'.P'P.\widehat{B}.P'P.x, \\ \widehat{Y} &= b_0 + (Px)'(Pb) + (Px)'PBP'(Px) \end{aligned} \quad (4)$$

Posons : $X = Px$ et $\theta = Pb$ (ou $x = P'X$ et $b = P'\theta$) alors

$$\widehat{Y} = b_0 + X'\theta + X'\Lambda X, \quad (5)$$

$$\widehat{Y} = b_0 + \theta_1 X_1 + \dots + \theta_m X_m + \lambda_1 X_1^2 + \dots + \lambda_m X_m^2. \quad (6)$$

Cette forme canonique : $X'\Lambda X$ est réduite a une expression ne contenant que les termes quadratiques purs, correspond à une rotation des axes initiaux Par dérivation de l'équation (6), nous déduisons les coordonnées du point stationnaire (X_{1s}, \dots, X_{ms})

$$X_{is} = \frac{-\theta_i}{2\lambda_i}, \quad i = 1, \dots, m. \quad (7)$$

La réponse en ce point stationnaire (S) est alors

$$Y_s = b_0 - \frac{1}{4}\theta'.\Lambda^{-1}.\theta. \quad (8)$$

Les quantités importantes à considérer dans cette forme canonique sont

- Les signes et l'ordre de grandeur relative des valeurs propres λ_i qui détermineront la nature du point stationnaire,
- Les coordonnées X_{is} du point stationnaire,
- La nature de la transformation $X = P'x$.

Listes des figures

Figure 1 - Les plans équiradiaux pour 2 facteurs	31
Figure 2 - Plan simplexe pour deux facteurs	32
Figure 3 - Graphes linéaires de Taguchi	38
Figure 4 - Plan composite pour 2 facteurs	41
Figure 5 - Plan de Box-Behnken pour 3 facteurs	49
Figure 6 - Plan de Doehlert pour 2 facteurs	53
Figure 7 - Plan de Hoke pour 3 facteurs	56
Figure 8 - Le cas d'un point maximum	59
Figure 9 - Le cas d'un point minimum	59
Figure 10 - Le cas d'un point minimax	60
Figure 11 - Graphe de Pareto illustrant l'importance des effets de facteurs	71
Figure 12 - Graphe contour de surface de réponse	71
Figure 13 - Graphe contour de surface de réponse	72
Figure 14 - Graphe contour de surface de réponse	72
Figure 15 - Graphe contour de surface de réponse	73

Liste des tableaux

Tableau 1 - Table d'ANOVA	22
Tableau 2 - Résolution - Propriété	27
Tableau 3 - Nombre maximum de facteurs étudiés sur un plan de base	27
Tableau 4 - Les plans équiradiaux pour 2 facteurs	30
Tableau 5 - Table des signes plus et moins pour les plans de Plackett et Burman	34
Tableau 6 - La table de Taguchi $L_{27}(3^{13})$	36
Tableau 7 - Valeurs α suivant le nombre de points au centre pour satisfaire la propriété de presque orthogonalité	44
Tableau 8 - Récapitulatif des paramètres des plans bloqués orthogonalement	47
Tableau 9 - Nombre d'expériences nécessaires pour la réalisation d'un plan BB	50
Tableau 10 - Plan d'expérimentation (unités courantes)	65
Tableau 11 - Plan d'expériences en unités codées	66
Tableau 12 - Estimation de coefficients de régression	67
Tableau 13 - Plan d'expériences en unités codées	69

Liste des symboles et notations

	Description
β	Vecteur des coefficients théoriques (paramètres)
$\widehat{\beta}$	Vecteur des coefficients estimés pour la modélisation de la réponse
β_i	Effets moyens relatifs aux facteurs x_i
β_{ij}	Interactions d'ordre un entre les effets
σ^2	Variance expérimentale
$\widehat{\sigma}^2$	Variance résiduelle
ϵ	Le vecteur des n résidus
CMa	Carré moyen factoriel
CMr	Carré moyen résiduel
A	Effet principal du facteur A
$A \circ B$	Effet interaction de A et B
m	Nombre de facteurs
p	Nombre de modalités du facteur
n	Nombre d'expérience
SCE	Somme des carrés des écarts
X	Matrice du modèle ou matrice des effets
$X'X$	Matrice d'information
$(X'X)^{-1}$	Matrice de dispersion
Y	Vecteur des réponses expérimentales, mesurées
\widehat{Y}	Vecteur des réponses estimées
\mathcal{U}	L'ensemble des n unités expérimentales
\mathcal{E}	Domaine expérimental
d_x	La fonction de variance
VIF_j	Le facteur d'inflation
P	Transformation orthogonale
$O(\mathcal{R}^m)$	Le groupe orthogonal de \mathcal{R}^m
M	Matrice des moments
1_ν	Fonction indicatrice à ν dimension

\mathcal{F}_p^m	L'espace vectoriel
$\mathcal{P}_{m-1}(\mathcal{F}_p)$	L'espace projectif
$\text{Card}\mathcal{P}_{m-1}(\mathcal{F}_p)$	Le cardinal de l'ensemble $\mathcal{P}_{m-1}(\mathcal{F}_p)$
ξ_α	L'espace de contrastes
$\beta_\alpha(l)$	La partition de \mathcal{F} orthogonal à α
mod	Modulo
α	direction
\mathcal{G}	Le sous-groupe intrabloc
R	Rayon
O	Matrice orthogonale

Notation	Lecture
A	Matrice (ou vecteur)
A_m	Matrice carrée d'ordre m
$A_{n \times p}$	Matrice A à n lignes et p colonnes
A'	Transposée de la matrice A
A^{-1}	Inverse de la matrice A
\hat{A}	Matrice (ou vecteur) estimé

REFERENCES

- Anderson, V. L. and Mclean, R. A. (1974). Design of Experiments : A Realistic Approach, CRC-Press, New-York. basel.
- Anderson, M. J. and Whitcomb, P. J. (2005). RSM Simplified Optimizing Processes Using Response Surface Methods for Design of Experiments, Productivity Press.
- Baasel, W. D. (1990). Preliminary chemical engineering plant design, Springer.
- Baumert, L. Golomb, S. W. and Hall, M. (1962). Discovery of an Hadamard matrix of order 92. Amer. Math. soc. Bull., 68, 237-238.
- Box, G. E. P. (1952). Multi-factor designs of first order, Biometrika 39 (1), 1-45.
- Box, G. E. P. (1988). Signal to noise ratios. Performance criteria and transformations (with discussion), Technometrics, 30, p. 140.
- Box, G. E. P. and Behnken, D. W. (1960). Some New Three-Level Designs for the Study of Quantitative Variables, Technometrics, 2, 455-475.
- Box, G. E. P. and Draper, N. R. (1987). Empirical model-building and response surface, John Wiley, New York.
- Box, G. E. P. and Hunter, J. S. (1957). Multifactor experimental designs for exploring response surfaces, Ann. Math. Stat., 195-241.
- Box, G. E. P. , Hunter, W. G. and Hunter. J. S. (1978). Statistics for Experimenters : An introduction to design, data analysis and model building, Wiley, New York.
- Box, G. E. P. and Wilson, K. B. (1951). On the Experimental Attainment of Optimum Conditions, J. R. Statist. Soc., Ser. B, 13 (1), 1-45.
- Bruns, R. E. Scarminio, I. S. and Neto, B. B. (2006). Statistical Design-chemometrics, Elsevier Chemometrics.
- Cochran, W.G. (1939). Long-term agricultural experiments (with discussion), J. R. Statist. Soc., Suppl. 6 (2), 104-148.
- Cohen, V. (1989). Introduction aux plan d'expériences, Revue statistique appliquée, 37 (2), 17-45.
- Coursol, J. (1980). Technique statistique des modèles linéaires, Cimpa, Nice.
- Dagnelie, P. (2000). La planification des expériences : Choix des traitements et dispositif expérimental, journal de la société Française de statistique. 141 (1-2), 5-29.
- Dagnelie, P. (2003). Principes d'expérimentation : Planification des expériences et analyse de leurs résultats, Gembloux, Press agronomiques, édition Electronique.
- Dean, A. and Voss, D. (1999). Design and Analysis of Experiments, Springer-Verlag New-York, Inc.
- Djekrif-Dakhmouche, S. , Gheribi-Aoulmi, Z. , Meraihi, Z. and Bennamoune, L. (2006). Application of statistical design to the optimization of culture medium for α -amylase production by *Aspergillus niger* ATCC 16404 grown on orange waste powder, journal of food engineering, 73 (2), 190-197.

- Doehlert, D. H. (1970). Uniform shell designs, *Appl Stat*, 231–9.
- Doehlert, D. H. and Klee, V. L. (1972). Experimental designs through level reduction of a D-dimensional cuboctahedron, *Discrete Math*, vol 2, 309-334.
- Drain, D. (1997). *Handbook of Experimental Methods for Process Improvement Experimental design*, CRC-Press.
- Draper, R. Gaffke N. and Pukelsheim, F. (1993). Rotability of variance surface and moments matrices, *Journal of Statistical Planning and Inference*, 36, 347-356.
- Droesbeke, J. J. , Fine, J. et Saporta, G. (1997). *Plans d'expériences : Applications à l'entreprise*, Editions Technip.
- Duby, C. (1989). La méthode Taguchi : Valeur et Limite, *revue statistique appliquée*, XXXVII (2), 7-16.
- Einax, J. W. , Zwanziger, H. W. and S. Geib (1997). *Chemometrics in Environmental Analysis*, Wiley-VCH.
- Feneuille, D. , Mathieu, D. et Phan-Tan-Lun, R. (1978). *Méthodologie de la Recherche Expérimentale : Introduction, outils mathématiques, études des surfaces de réponse, matrices de mélanges, méthodes du simplex*, Université d'Aix-Marseille 3, Publication du LMRE.
- Finney, D. J. (1945). The fractional replication of factorial experiments, *Ann. Eugenics* 12 (4), 291-301.
- Finney, D. J. (1946). Recent developments in the design of field experiments, III. Fractional replication, *J. Agric. Sci.* 36 (3), 184-191.
- Fisher, R. A. (1925). *Statistical methods for research workers*, Edinburgh, Oliver and Boyd, 239 p.
- Fisher, R. A. (1926). The arrangement of field experiments, *J. Ministry. Agric.* 33, 503-513.
- Fisher, R. A. (1935). *The design of experiments*, Oliver and Boyd, Londres.
- Giesbrecht, F. G. and Gumpertz, M. L. (2004). *Planning, construction and statistical analysis of comparative experiments*, Willey-IEEE.
- Goupy, J. L. (1990). Etude comparative de divers plans d'expériences, *revue statistique appliquée*, 38 (4), 5-44.
- Goupy, J. (1999). *Plans d'expériences pour surfaces de réponse*, Dunod, Paris.
- Goupy, J. (2005). *Pratiquer les plans d'expériences*, Dunod, Paris.
- Goupy, J. (2006). les plans d'expériences , *revue Modulad*, n° 34, 37-79.
- Goupy, J. (2006). *Introductions aux plans d'expériences*, Dunod, Paris.
- Goupy, J. et Creighton L. (2006). *Introduction aux plans d'expériences*. Paris, Dunod, 336 p.
- Goupy, J. *Plans d'expériences PE 230 P1-26*.
- Govert, B. *La planification expérimentale vue par Taguchi*. Université Catholique de Louvain.
- Govert, B. *Plans pour le criblage de facteurs*, Université Catholique de Louvain.

- Govert, B. Plans d'expériences pour l'estimation de surface de réponse, Université Catholique de Louvain.
- Goos, P. (2002). The optimal design of blocked and split-plot experiments, Springer.
- Hoke, A. T. (1974). Economical second-order designs based on irregular fractions of the 3^m factorial, *Technometrics*, 16 (3), 375-384.
- Hunter, J. S. (1985). Statistical design applied to product design, *J. qual. technol.*, 17, 201-221.
- Joblet, E. et Delarminat, E. (1989). Méthodologie Taguchi appliquée à une étude de surface de réponse, *Revue statistique appliquée*, 37 (2), 103-116.
- Kackar, N. (1985). Off-line quality control, parameter design and the Taguchi method, *J. Quality Technology* 17, 176-188.
- Khuri, A. and Comell, J. (1987). Response surface : Designs and Analyses, Marcel Dekker, New-York.
- Khuri, A. L. and Cornell, J. A. (1996). Response surface, Designs and Analyses, CRC Press.
- Kiefer, J. (1959). Optimum experimental designs (with discussion). *J. R. Stat. Soc., Ser. B*, 21 (2), 272-319.
- Kobilinsky, A. (2001). Les Plans Fractionnaires, laboratoire de Biométrie.
- Lamure, A. Methodologie des plans d'expériences.
- Lazic, Z. R. (2006). Design of Experiments in Chemical Engineering : A practical Guide, Wiley-VCH.
- Lewis, G. A., Mathieu, D. and Phan-Tan-Luu, R. (1999). Pharmaceutical experimental design, Inc., New York. Basel.
- Legendre, A. M. (1805). Nouvelles Méthodes pour la Détermination des Orbites des Comètes. Courcier, Paris.
- Louvet, F. et Delplanque, L. (2005). Les plans d'expériences par la méthode Taguchi, Expérimentique.
- Luangpaiboon, P. (2001). Proposed finishing strategies based on experimental design for process optimisation, *Thammasat Int. J. Sc. Tech.*, 6 (3).
- McLean, R. A. and Anderson, V. L. (1984). Applied Factorial and Fractional Designs, Marcel Dekker, New-York.
- McKetta, J. J. et Cuningham, W. A. (1989). Encyclopedia of chemical Processing and Design, CRC-Press.
- Mechakra-Maza, A. , Gheribi-Aoulmi, Z. , Meraihi, Z. , Bousseboua, H. (2002). Utilisation de la planification expérimentale pour l'optimisation de la production de la protéase acide par *Aspergillus niger*. Science and technologie, numéro spécial-D, 59-63.
- Meilgaard, M. , Civille, G. V. and Carr, B. T. (1999). Sensory evaluation techniques, CRC Press.

- Mendham, J. (2006). *Analyse chimique quantitative de Vogel*, Mendham. Denney. Barnes. Thomas, Boeck université.
- Muhammad, A. S. Deborah, L. I. and Kowalski, B. R. (1986). *Chamometrics*, Wiley.IEEE.
- Myers, R. H. (1971). *Response Surface Methodology*, Boston :Allyn and Bacon.
- Myres, R. H. and Montgomery, D. C. (1995). *Response Surface Methodology : Process and Product Optimisation Using Designed Experiments*, New York Wiley.
- Nielloud, F. and Marti-Mestres, G. (2000). *Pharmaceutical Emulsion and Suspensions*, CRC Press, Marsel Dekker, Inc. New-York. Basel.
- Nguyen, N. K. ,Borkowski, J. J. (2008). New 3-level response surface designs constructed from incomplete block designs, *journal of statistical planing and inference* 138, 294-305.
- Park, S. H. (1996). *Robust Design and Analyis for Quality Engineering* , Springer.
- Pilet, M. (1998). Construire facilement des plans de résolution IV à partir des tables de Taguchi, *Revue statistique appliquée*, 45 (4), 85-100.
- Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments, *Biometrika*, 33 (4), 305–325.
- Rekab, K. and Shaikh, M. (2005). *Statistical Design of experiments with Engineering applications*, CRC Press.
- Saporta, G. (2006). *Prababilités, analyse des données statistique*, Edition TECHNIP.
- Scheffé, H. (1958). Experiments with mixtures, *J. R. Stat. Soc., Ser. B*, 20 (2), 344-360.
- Taguchi, G. (1959). Linear Graphs for Orthogonal Arrays and their Applications to Experimental Designs with the Aid of Various Techniques, *Rep. Statist. Appl. Res.* 6 (1), 1-43.
- Taguchi, G. (1960). Tables of Orthogonal Arrays and Linear Graphs, *Rep. Statist. Appl. Res.* (6) 5, 176-227.
- Taguchi, G. (1987). *System of experimental design : engineering methods to optimize quality and minimize costs* (2 vol.). White Plains, Unipub/Kraus,552+656p.
- Tomassone, R. *Méthodologie des plans d'expériences*, Biométrie, Parfums Christien Dior. (Séminaire LVMH Aport de la Statistique dans le Recherche et le Contrôle).
- Tressler, E. (2004). *A survy of the Hadamard conjecture (thèse)*, Blacksburg, Virginia.
- Ulmo, J. et Pflugfelder, B. (1979). Méthode d'exploration de la surface de réponse pour l'optimisation d'une variable dépendante dans le cas de facteurs contrôlés quantitatifs, *revue statistique appliquée*, 27 (4), 23-36.
- Vuchkov, I. N. and Boyadjieva, L. N. (2001). *Quality Improvement with Design of Experiments*, Springer.
- William, P. (1998). *Meredith John, Statistical Design and Analysis of Experiments*, Siam.
- Yates, F. (1935). Complex experiments (with discussion), *J. R. Statist. Soc., Ser. B*, 2 (2), 181-247.
- www.itl.nist.gov