

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA

RECHERCHE SCIENTIFIQUE SAAD DAHLEB BLIDA

Faculté des sciences

Département de mathématiques



MEMOIRE DE MASTER

Spécialité : Recherche Opérationnelle

SYNTHESE SUR LA METHODOLOGIE DES PLANS D'EXPERIENCE : ETUDE, OPTIMALITE ET COMPARAISON

Présenté par

- AMEUR Assia
- DADDI Hadjira

Devant le jury composé de

F. HANNANE

Professeur, U. de Blida

Président

H.EL MOUSSAOUI

Maître assistant A

Promoteur

N.MESSAOUDI

Maître assistante A

Examinatrice

REMERCEMENT

Tout d'abord, nous remercions le maître de tout, DIEU le tout puissant qui nous a dirigés vers la bonne voie par sa compassion.

*Nos vifs remerciements accompagnés de toute notre gratitude vont tout d'abord à notre promoteur **Mr H. El MOUSSAOUI** pour nous avoir encadré, dirigé et conseiller tout long de ces moins de travail ainsi pour ces aide précieuse lors de la correction de ce mémoire et **Mr. F.HANNANE** pour les nombreux conseils et les informations qui nous a fournis sur les plans d'expériences.*

Nous n'oublierons pas ceux qui ont contribué de près ou de loin pour l'élaboration de ce modeste travail.

Nous remercions nos petites familles pour leur soutien morale et encouragement durant toute la formation.

En fin, on tient à remercier les membres de jury qui vont faire l'honneur d'apprécier notre travail.

DEDICACE

Je dédie ce modeste travail à mes très chers parents qui m'ont toujours soutenu pendant toute ma vie, je leur souhaite le bonheur et la bonne santé.

A toute ma famille frères sœurs oncles tantes cousins cousines.

A toute personne porte des sentiments pour moi.

Merci à tous ceux qui m'ont aidé près ou de loin.

Et a tous mes amis.

Et sans oublier mon binôme Hadjira que je remercie pour sa patience et je la souhaite la réussite à sa vie.

RESUME

La méthodologie de la recherche expérimentale (MRE) est un ensemble de méthode et mode de raisonnement destiné à tout expérimentateur désirant faire de la planification expérimentale. Elle a pour objet de lui permettre d'optimiser l'efficacité de sa recherche expérimentale quelque soit sa branche d'activité. Pour ce faire, elle va l'aider à exprimer au mieux son problème et lui proposer des stratégies expérimentales optimales en fonctions des objectifs qu'il s'est fixé et des moyennes dont il dispose.

Ce travail comporte deux objectifs : les algorithmes d'échange qui consiste à trouver la meilleure matrice selon les critères d'optimalité et une étude comparative des matrices d'expérience pour aider l'expérimentateur à prendre une décision.

الملخص

إن طريقة المنهج التجريبي هي مجموعة من الطرق و وسائل للتحليل يستعملها الباحثون من اجل تخطيط تجاربهم الهدف من الطريقة السماح للباحثين في تجويد فعالية تجاربهم مهما كان ميدان تخصصهم كما يساعد على تحليل المشكل لطرح إستراتيجية بحثة ذات جودة عالية بدلالة الأهداف المسطرة و الوسائل المتوفرة لهذه المنهجية. لهذا العمل هدفان : خوارزميات التبادل من اجل إيجاد أحسن مصفوفة و مقارنة المصفوفات لتوجيه اختيار المجرب.

LISTE DES ILLUSTRATIONS, GRAPHIQUE ET TABLEAUX

Figure 1.1 :	seule la variable x_i prend des niveaux différents, les autres variables étant fixée à des valeurs bien définies.	4
Figure 1.2 :	Niveaux des facteurs définissant des points expérimentaux dans l'espace expérimental.	5
Figure 2.1 :	Domaine d'étude d'un plan 2^2 .	11
Figure 2.2 :	Domaine expérimental d'un plan 2^3 .	12
Figure 2.3 :	Effet du facteur 1.	13
Figure 2.4 :	Plans composites centrés dans les faces.	32
Figure 2.5 :	Plans composite centrés pour l'étude de deux facteurs.	33
Figure 2.6 :	Plan de Box-Behnken pour 3 facteurs.	38
Figure 2.7 :	Plan de Doehlert pour l'étude de deux facteurs.	43
Figure 2.8 :	Emplacement des points expérimentaux du plan de Roquemore.	48
Figure 3.1 :	effet de facteur A .	56
Figure 3.2 :	effet de facteur B .	57
Figure 3.3 :	Hypothèse nulle.	59
Figure 3.4 :	Principe de l'analyse de la variance.	60
Figure 3.5 :	loi de F.	63
Figure 4.1 :	Représentation d'un mélange dans un système d'axes cartésiens.	74
Figure 4.2 :	Les compositions des mélanges à deux constituants peuvent être représentées par les points du segment de droite AB .	74
Figure 4.3 :	Représentation des mélanges à deux constituants sur un segment de droite.	75
Figure 4.4 :	Représentation de mélange M sur un segment de droite.	75
Figure 4.5 :	Représentation dans un système d'axes cartésiens d'un mélange de trois produits.	76
Figure 4.6 :	Représentation de l'ensemble des mélanges à trois constituants dans l'espace cartésien des trois produits A, B et C .	77
Figure 4.7 :	Représentation des mélanges à trois constituants à l'aide d'un triangle équilatéral.	78
Figure 4.8 :	Représentation des mélanges à quatre constituants par un tétraèdre régulier.	79
Figure 4.9 :	plan de mélanges ne comportant que les trois produits purs.	84
		85

Figure 4.10 :	plan de mélanges en réseaux (simplex lattice design) comportant les trois produits purs et les mélanges moitié-moitié. Ce plan est noté {3,2}.	
Figure 4.11 :	Plan de mélanges en réseaux comportant les trois produits purs et les mélanges 1/3-2/3 et 1/3-1/3-1/3. Ce plan est noté {3,3}.	85
Figure 4.12 :	Plan de mélanges centre comportant les trois produits purs, les mélanges.	86
Figure 4.13 :	plan de mélanges centré augmenté comportant les trois produits purs, les mélanges moitié-moitié, le mélange équiproportion des trois produits et les mélanges situés aux centres de gravité des simplex unitaires.	87
Figure 4.14 :	Les limites basses de tous les produits définissent deux zones : La zone interdite (zone ombrée) et la zone autorisée qui a la même forme géométrique que le domaine initial.	88
Figure 4.15 :	La somme des limites basses de tous les constituants égale à 0.	88
Figure 4.16 :	La somme des limites basses de tous les constituants.	88
Figure 4.17 :	La somme des limites basses de tous les constituants.	89
Figure 4.18 :	La somme des limites basses de tous les constituants.	89
Figure 4.19 :	Le triangle PQR représente le domaine d'étude restreint. Le point E est un mélange dont la composition peut être exprimée avec les corps purs A, B et C ou avec pseudo corps purs P, Q et R.	90
Figure 4.20 :	la teneur en produit A ne peut pas être supérieure à la limite haute U_1 et inférieure à la limite basse L_1 . Les proportions de ce produit dans les mélanges varient donc entre L_1 et U_1 . Le domaine d'étude est divisé en trois zones : deux zones interdites (zones ombrées) et zone autorisée (zone claire).	91
Figure 4.21 :	La somme des limites hautes de tous les constituants égale à 2.	92
Figure 4.22 :	La somme des limites hautes de tous les constituants supérieure à 1.	92
Figure 4.23 :	La somme des limites hautes de tous les constituants supérieure à 1.	92
Figure 4.24 :	la somme des limites hautes de tous les constituants égale à 1.	92
Figure 4.25 :	Les limites basses forment un triangle $A'B'C'$ qui doit être coupé par toutes les limites hautes pour que les limites soient compatibles. Ici L_3 n'est pas compatible avec les autres limites.	94
Figure 4.26 :	Les limites hautes forment un triangle $A'B'C'$ qui doit être coupé par toutes les limites basses pour que les limites soient compatibles. Ici L_1 n'est pas compatible avec les autres limites.	95

Figure 5.1 :	Ellipsoïde de confiance.	104
Tableau 1.1 :	Matrice d'expériences.	6
Tableau 2.1 :	Matrice d'expériences d'un plan 2^2 .	10
Tableau 2.2 :	Matrice d'expérience d'un plan 2^3 .	10
Tableau 2.3 :	Matrice des effets pour un plan 2^2 .	15
Tableau 2.4 :	Matrice des effets d'un plan 2^3 .	16
Tableau 2.5 :	Matrice des expériences (étude d'une émulsion de bitume).	18
Tableau 2.6 :	Comparaison des effets calculés.	18
Tableau 2.7 :	Matrice de calcul des contrastes pour les deux demi plans fractionnaires.	21
Tableau 2.8 :	Matrice d'expériences.	29
Tableau 2.9 :	Matrice d'expérience d'un plan composite pour deux facteurs.	33
Tableau 2.10 :	Plan de Box-Behnken pour trois facteurs.	39
Tableau 2.11 :	Plan de Box-Behnken pour quatre facteurs.	39
Tableau 2.12 :	Plan de Box-Behnken pour six facteurs.	42
Tableau 2.13 :	Plan de Box-Behnken pour sept facteurs.	42
Tableau 2.14 :	Matrice d'expériences d'un plan de Doehlert pour deux facteurs.	43
Tableau 2.15 :	Plans de Doehlert pour deux à quatre facteurs.	45
Tableau 2.16 :	Plans de Roquemore 311A pour trois facteurs.	47
Tableau 2.17 :	Plans de Roquemore 311B pour trois facteurs.	50
Tableau 2.18 :	Plans de Roquemore 416B pour quatre facteurs.	50
Tableau 2.19 :	Plan de Roquemore 416C pour quatre facteurs.	51
Tableau 3.1 :	Résultats Y pour 4 essais.	53
Tableau 3.2 :	Résultat Y pour 4 essais.	55
Tableau 3.3 :	Résultat $Y_{mesuré}$ et $Y_{estimé}$ pour 4 essais.	57
Tableau 3.4 :	Un plan d'expérience à un seul facteur.	61
Tableau 3.5 :	Tableau d'analyse de la variance.	64
Tableau 3.6 :	Tableau d'analyse de la variance.	68
Tableau 3.7 :	Plan d'expériences.	69
Tableau 3.8 :	Tableau d'analyse de la variance.	70
Tableau 4.1 :	Représentation en tableau d'un mélange à deux constituants.	80
Tableau 4.2 :	Représentation en tableau d'un mélange à trois constituants.	81
Tableau 6.1 :	Caractéristiques des matrices d'expériences du premier degré.	117

Tableau 6.2 :	Caractéristiques des matrices d'expériences du second degré à deux variables.	118
Tableau 6.3 :	Caractéristiques des matrices d'expériences du second degré à trois variables.	119
Tableau 6.4:	Caractéristiques des matrices d'expériences du second degré à quatre variables	120

TABLES DES MATIERE

<i>Résumé</i>	1
<i>Remerciements</i>	
<i>Dédicaces</i>	
<i>Table des matières</i>	
<i>Liste des illustrations graphiques et tableaux</i>	
<i>Liste des symboles</i>	
<i>Introduction Générale</i>	1
Chapitre 1 : Concepts fondamentaux sur la méthodologie de la recherche expérimentale.	3
1.1 introduction.	3
1.2 Limite des méthodes traditionnelles d'essais.	3
1.2.1 Etude d'un phénomène.	3
1.2.2 La méthode classique et ses inconvénients	3
1.3 La méthode des plans d'expériences.	4
1.4 Notion d'espace expérimental.	5
1.4.1 matrice d'expériences.	5
1.4.2 Coordonnées centrées réduites.	6
1.5 Notion de modélisation mathématique.	6
1.5.1 Intérêt de la représentation polynomiale.	7
1.5.2 Modèle de l'expérimentateur.	7
1.6 Système d'équations.	8
Chapitre 2 : Plans d'expérience classiques.	9
2.1 Plans de criblages.	9
2.1.1 Plans factoriels complets à deux niveaux.	9
2.1.1.1 Construction des plans factoriels complets à deux niveaux.	10
2.1.1.2 Plans factoriel à deux facteurs.	12
2.1.1.3 Plans factoriels à trois facteurs 2^3 .	15
2.1.1.4 Plans complets 2^k .	16
2.1.1.5 Optimalité des plans 2^k .	17
2.1.1.6 Avantages et inconvénients des plans factoriels complets à deux niveaux	17
2.1.2 Plans fractionnaires à deux niveaux.	18
2.1.2.1 Théorie des aliases.	18
2.1.2.2 Construction pratique des plans fractionnaires.	22
2.1.2.3 Notion de résolution.	25
2.1.2.4 Avantages et inconvénients.	25
2.1.3 Plan de Plackett et Burman.	26
2.1.3.1 Construction des plans de Plackett et Burman.	26
2.1.3.2 Modélisation mathématique.	29
2.2 Plans de modélisation.	30

2.2.1	Plans composites.	30
2.2.1.1	Construction des plans composites.	31
2.2.1.2	Les différentes classes de plans composites	32
2.2.1.3	Propriétés des plans composites.	34
2.2.1.4	Avantages et inconvénients.	37
2.2.2	Plans de Box-Behnken.	38
2.2.2.1	Construction des plans de Box-Behnken.	38
2.2.2.2	Propriété du plan de Box-Behnken.	40
2.2.2.3	Principaux plans de Box-Behnken.	41
2.2.2.4	Avantages et inconvénients.	42
2.2.3	Plans de Doehlert.	42
2.2.3.1	Plans de Doehlert pour deux facteurs.	43
2.2.3.2	Modèle mathématique postulé.	44
2.2.3.3	Plans de Doehlert pour k facteurs.	44
2.2.3.4	Avantage et inconvénients.	45
2.2.4	Plans de Roquemore.	46
2.2.4.1	Plans 311A de Roquemore.	46
2.2.4.2	Autres plans de Roquemore.	49
2.2.4.3	Avantage et inconvénients.	51
Chapitre 3 : Analyse de la variance dans les plans d'expérience.		52
3.1	Calcul des effets des facteurs.	52
3.1.1	Utilisation de la modélisation opérationnelle.	52
3.1.2	Modélisation opérationnelle à l'effet principale et calcul des résidus.	55
3.2	Analyse de la variance.	58
3.2.1	L'analyse de la variance, pourquoi ?	58
3.2.2	Cas où il n'y a qu'un seul facteur contrôlé.	58
3.2.2.1	L'hypothèse nulle.	58
3.2.3	Cas d'un facteur à k niveaux.	60
3.2.3.1	Variance globale.	61
3.2.3.2	Variance résiduelle.	62
3.2.3.3	La variance entre échantillons.	62
3.2.3.4	Equation fondamentale d'analyse de la variance.	62
3.2.3.5	Test de Snedecor.	63
3.2.3.6	Tableau d'analyse de la variance.	64
3.3	Cas des plans d'expériences sans répétition.	64
3.3.1	Analyse de la variance.	64
3.3.3.1	Décomposition de la variance totale.	65
3.3.3.1.1	Dispersion totale.	65
3.3.3.1.2	Dispersion expliquée par un facteur.	65
3.3.3.1.3	Variance d'une interaction.	66
3.3.3.1.4	Variance résiduelle.	67
3.3.3.1.5	Tableau d'analyse de la variance.	67

3.4 Cas des plans d'expérience avec répétitions.	69
3.4.1 Etude sur un exemple	69
3.4.2 Prévision du nombre de répétitions.	70
3.4.2.1 Hypothèses.	70
Chapitre 4 : Plans de mélanges.	72
4.1 Introduction.	72
4.2 La contrainte fondamentale des mélanges.	72
4.3 Représentation géométrique des mélanges.	73
4.3.1 Mélanges à deux constituants.	73
4.3.2 Lecture du diagramme des mélanges.	75
4.3.3 Mélange à trois constituants.	76
4.3.4 Lecture du diagramme des mélanges ternaires.	77
4.3.5 Mélanges à quatre constituants.	79
4.3.6 Mélanges à cinq constituants.	79
4.4 Représentation des mélanges par tableaux.	80
4.4.1 Mélange à deux constituants.	80
4.4.2 Mélanges à trois constituants.	80
4.4.3 Mélanges à m constituants.	81
4.5 Modèle mathématique pour l'interprétation des mélanges.	81
4.5.1 Modèle du premier degré.	81
4.5.2 Modèle du second degré.	82
4.5.3 Modèle du troisième degré.	82
4.6 Différent types de plans de mélanges.	83
4.6.1 Plan de type I.	83
4.6.1.1 Plans en réseaux.	83
4.6.1.2 Plan de mélanges centrés.	86
4.6.1.3 Plans de mélange centrés augmentés.	86
4.6.2 Plan de type II.	87
4.6.2.1 L'existence du domaine expérimental.	88
4.6.2.2 Calcul de la composition des mélanges expérimentaux.	90
4.6.3 Plan de type III.	90
4.6.3.1 L'existence du domaine expérimental.	91
4.6.3.2 Principe de la démarche.	93
4.6.3.2.1 Compatibilité des limites.	93
4.6.3.2.2 Calcul du nombre de sommets, arêtes et faces.	96
4.6.3.2.3 Algorithme XVERT de Snee & Marquardt.	97
Chapitre 5 : Plan D-optimaux et les algorithmes d'échanges.	99
5.1 Transmission des erreurs.	99
5.1.1 Précision sur les coefficients.	100
5.1.2 Précision sur les réponses prédites.	101

5.1.3	La fonction de variance de prédiction.	102
5.2	Critères d'optimalité.	102
5.2.1	Qualité de la représentation des coefficients.	103
5.2.2	Critère d'iso variance par rotation.	105
5.2.3	Critère de A-optimalité.	105
5.2.4	Critère de D-optimalité.	105
5.2.5	Critère de E-optimalité.	105
5.2.6	Critère de G-optimalité.	105
5.2.7	Critère de M-optimalité.	106
5.2.8	Critère d'orthogonalité.	106
5.2.9	Critère de presque orthogonalité.	107
5.3	Plans D-optimaux.	107
5.3.1	Principe des plans D-optimaux	107
5.3.2	Algorithme d'échange.	107
5.3.2.1	Procédure d'échange général.	108
5.3.2.2	Algorithme DETMAX.	109
5.3.2.3	Algorithme de Fedorov.	110
5.3.2.4	Algorithme de Fedorov modifiée.	112
5.3.2.5	Algorithme k-échange.	113
5.3.2.6	Algorithme kl-échange	115
Chapitre 6	<i>Comparaison des matrices d'expériences.</i>	117
6.1	Modèle polynomial du premier degré.	117
6.2	Modèle polynomial du second degré.	118
6.2.1	Matrices d'expérience à deux variables.	118
6.2.2	Matrices d'expérience à trois variables.	118
6.2.3	Matrices d'expérience à quatre variables	119
	Résumé	121

INTRODUCTION GENERALE

La méthode des plans d'expériences n'est pas une technique nouvelle, elle date en fait du début du siècle avec les travaux de Fisher en 1925 [1]. Les premiers utilisateurs de ces méthodes furent les agronomes qui ont vite compris l'intérêt des plans d'expériences et notamment la réduction du nombre d'essais lorsqu'on étudie de nombreux paramètres. Les plans d'expériences ont pris un essor considérable avec le développement de l'informatique et la puissance de calcul qui l'accompagne.

Les plans d'expériences devraient faire partie du bagage scientifique de tout expérimentateur. Ce n'est pas encore le cas et c'est pour cette raison que nous avons entrepris la rédaction du présent ouvrage. En effet, les plans d'expériences servent à optimiser l'organisation des essais. Cette organisation permet d'obtenir le maximum de renseignements avec le minimum d'expériences. Elle permet également d'obtenir la meilleure précision possible sur la modélisation des résultats. Cette méthode des plans d'expériences est basée sur des règles mathématiques strictes et elle exige une démarche rigoureuse de la part de l'expérimentateur.

Cette science de l'organisation des essais et de la modélisation des résultats expérimentaux nécessite une terminologie adaptée, précise et une mise en œuvre rigoureuse. C'est pourquoi nous commençons par un chapitre sur la terminologie et quelques notions de base sur les plans d'expériences. En fin des indications générales seront données sur les différents types de plans d'expériences et sur les usages que l'on peut en faire.

Dans notre travail, les principes généraux de construction des plans d'expériences sont présentés à partir de la notion d'espace expérimental. La représentation géométrique des points expérimentaux est très parlante mais elle est vite limitée lorsque la dimension de l'espace augmente. C'est pourquoi l'on utilise la représentation matricielle. A l'aide des deux représentations, géométrique et matricielle, les principaux plans d'expériences sont décrits : Plans factoriels complets à deux niveaux, plans fractionnaires, plan de Plackett et Burman, plans composites, plans de Box-Behnken, plans de Doehlert, plans de Roquemore et plan de mélange.

Ce mémoire comporte six chapitres. Dans le premier chapitre, nous allons expliquer brièvement les points essentiels de cette méthodologie de la recherche expérimentale. Le deuxième chapitre consacré à la description de sept plans classiques. Le chapitre trois est consacré à la description de l'analyse de la variance et les importantes notions des plans de

mélanges ont été décrite également au quatrième chapitre. Dans le cinquième chapitre nous avons abordés la notion de D-optimalité et les algorithmes d'échange utilisé dans la recherche des matrices D-optimales. En fin, des tables de comparaison entre les différents plan d'expériences ont été élaborées au chapitre six.

CHAPITRE 1

CONCEPTS FONDAMENTAUX SUR LA METHODOLOGIE DE LA
RECHERCHE EXPERIMENTALE**1.1 Introduction**

La méthodologie de la recherche expérimentale ou Planification des expériences fait partie des statistiques. Elle est utilisée par toute discipline où la réalisation d'expériences sur un système quelconque s'avère indispensable afin de vérifier des résultats théoriques, de tirer des informations pratiques ou d'optimiser le fonctionnement de celui-ci. Le système est considéré comme une enceinte fermée dans laquelle, seules sont prises en compte les entrées et les sorties, appelées respectivement facteurs et réponses. La constitution interne du système est totalement ignorée.

1.2 Limites des méthodes traditionnelles d'essais**1.2.1 Etude d'un phénomène**

L'étude d'un phénomène peut se schématiser de la manière suivante : l'expérimentateur s'intéresse à une grandeur, par exemple le rendement en blé d'une parcelle de terre, le prix de revient d'un produit chimique ou l'usure d'une pièce de moteur automobile. Cette grandeur dépendra d'un grand nombre de variables. Le rendement en blé sera fonction de la nature du terrain, de la qualité d'engrais incorporé, du climat, de l'exposition au soleil, etc... Sous une forme mathématique, on peut écrire que la grandeur d'intérêt y est une fonction de plusieurs variables x_i .

$$y = f(x_1, x_2, \dots, x_k) \quad (1.1)$$

L'étude du phénomène se ramène alors à la mesure de la grandeur en fonction des différentes valeurs que l'on peut donner aux variables. Nous allons décrire rapidement la méthode classique pour étudier la fonction (1.1).

1.2.2 La méthode classique

Traditionnellement, les essais sont effectués de manière séquentielle [8] en faisant varier les variables l'une après l'autre sans planification préalable de l'ensemble des essais à réaliser.

La méthode des plans d'expériences s'appuie sur deux notions essentielles, l'espace expérimental et la modélisation mathématique des grandeurs étudiées [9].

1.4 Notion d'espace expérimental

La fonction $y = f(x_1, x_2, \dots, x_k)$ peut être représentée dans un espace géométrique. Nous choisirons un repère cartésien, nous attribuerons le premier axe au premier facteur x_1 , le seconde au deuxième facteur x_2 , etc. Le nième axe au facteur x_n . L'espace ainsi défini est appelé l'espace expérimental. Pour réaliser une expérience, il faut fixer chacun des facteurs à une valeur bien précise. Cette valeur est appelée niveau du facteur. Lorsque l'on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes (borne inférieure ou niveau bas; borne supérieure ou niveau haut). L'ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut, s'appelle le domaine de variation du facteur. On a l'habitude de noter le niveau bas par -1 et le niveau haut par +1.

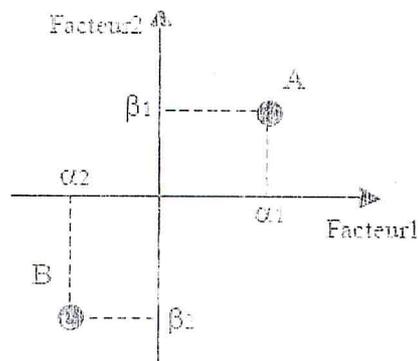


Figure 1.2: Niveaux des facteurs définissant des points expérimentaux dans l'espace expérimental

1.4.1 Matrice d'expériences

Chaque point expérimental peut être représenté soit géométriquement dans l'espace expérimental, soit à l'aide d'une matrice soit par un tableau donnant les niveaux. Les coordonnées du point expérimental sont les niveaux indiqués dans le tableau 1.1 qui porte le nom de matrice d'expériences.

Tableau 1.1: Matrice d'expériences

le nom de l'essai	Facteur1	Facteur2
A	α_1	β_1
B	α_1	β_1

Les matrices d'expériences représentent donc la disposition des points expérimentaux dans l'espace expérimental. Pour l'interprétation l'expérimentateur introduit une grandeur supplémentaire : la réponse. Il est alors amené à travailler dans un espace possédant une dimension en plus que l'espace expérimental.

1.4.2 Coordonnées centrées réduites

Lorsque l'on attribue la valeur -1 au niveau bas d'un facteur et la valeur $+1$ au niveau haut, on effectue deux modifications importantes, on déplace l'origine des mesures et on change l'unité de mesure. Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle variables centrées réduites, centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité. Le passage des variables d'origine z aux variables centrées réduite x , et inversement, est donné par la formule suivante:

$$x = \frac{z - z_0}{pas}$$

Où

$$z_0 = (\text{niveau bas} + \text{niveau haut})/2$$

$$pas = (\text{niveau haut} - \text{niveau bas})/2$$

L'intérêt des variables centrées réduites est de pouvoir présenter les plans d'expériences de la même manière quels que soient les domaines d'études retenus et quels que soient les facteurs. La théorie des plans d'expériences présente ainsi une grande généralité [10].

1.5 Notion de modélisation mathématique

Le modèle est une relation entre les facteurs x_1, x_2, \dots, x_k et la réponse que l'on désire étudier. La fonction $y = f(x_1, x_2, \dots, x_k)$ correspond à une surface de réponse (appelée hyper surface si $k > 3$) est trop générale et il est d'usage d'en prendre un développement limité de Taylor [9]. Si les dérivées du développement de Taylor peuvent être considérées comme constantes, le développement précédent prendra la forme d'un polynôme de degré plus ou moins élevé:

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \dots (1.2)$$

- y est la grandeur à laquelle s'intéresse l'expérimentateur
- x_i Représente le niveau attribué au facteur N° i
- a_0, a_i, a_{ij}, a_{ii} sont les coefficients du modèle

1.5.1 Intérêt de la représentation polynomiale

La représentation polynomiale de la réponse permet l'utilisation du calcul matriciel, contrairement aux autres formes de la modélisation, d'où le grand intérêt de la représentation polynomiale.

Les modèles utilisés par les autres disciplines font intervenir souvent des fonctions transcendentes (les exponentielles et les logarithmes). Les mathématiciens utilisent une représentation polynomiale (1.2).

❖ Avantages de la représentation polynomiale

- a_0 donne la valeur de la réponse au centre du domaine d'étude
- a_i (termes linéaires) représentent les effets des facteurs. On peut donc immédiatement faire une linéarisation de l'influence des facteurs sur la réponse, ce qui n'aurait pas été possible si la relation (1.1) comporte des fonctions transcendentes.
- Le modèle (1.2) est linéaire par rapport aux coefficients. On peut donc utiliser tout le calcul matriciel ce qui n'aurait pas été possible si la relation (1.1) comportait des fonctions transcendentes.
- On peut également déterminer toutes les interactions d'ordre 2, 3, ..., k ce qui n'aurait pas été possible si la fonction (1.1) comportait des fonctions transcendentes.
- On peut faire une optimisation c'est-à-dire connaître les valeurs exactes des facteurs donnant une réponse optimale.
- On peut trouver la valeur de la réponse en n'importe quel point intérieur au domaine expérimental sans réaliser l'expérience correspondante. Le modèle devient alors un modèle prédictif.

1.5.2 Modèle de l'expérimentateur

Deux compléments doivent être apportés au modèle précédemment. Le premier est le manque d'ajustement. Cette expression traduit le fait que le modèle choisi par l'expérimentateur avant les essais est probablement un peu différent du modèle réel. Il y a un écart entre ces deux modèles. Cet écart est le manque d'ajustement. Le second est la prise en compte de la nature aléatoire de la réponse. En effet si on mesure plusieurs fois une réponse en un même point expérimental, on n'obtiendra pas exactement le même résultat. Il y a une dispersion des

résultats appelés erreur expérimentale. Ces deux écarts sont souvent réunis dans un seul écart, le résidu, noté e . Le modèle utilisée par l'expérimentateur s'écrit alors :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \dots + e$$

1.6 Système d'équations

Chaque point expérimental apporte une valeur de la réponse. Or cette représentation est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer. A la fin du plan d'expériences, on a un système de n équations (s'il y a n essais) à q inconnues (s'il y a q coefficients dans le modèle choisi a priori). Ce système s'écrit d'une manière simple en notation matricielle:

$$y = XA + e$$

Avec

y : vecteur des réponses

X : matrice de calcul, qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé

A : vecteurs des coefficients

e : vecteurs des résidus

Ce système ne peut pas, en général, être résolu simplement car le nombre d'équations est inférieur au nombre d'inconnues. En effet, il y a n équations et $q + n$ inconnues cette résolution ne peut pas être menée à bien que si l'on utilise une méthode de régression qui introduit q équations supplémentaires. La plupart du temps cette méthode est basée sur le critère d'optimisation des moindres carrés. On obtient les estimateurs les plus probables des coefficients à l'aide de la relation suivante :

$$\hat{a} = ({}^t X X)^{-1} {}^t X Y$$

Formule dans laquelle la matrice ${}^t X$ est la transposée de X .

Les plans factoriels complets à deux niveaux permettent aussi d'étudier k facteurs à raison de deux niveaux par facteur. Ils renferment toutes les combinaisons des k facteurs à leurs deux niveaux, soit 2^k combinaisons.

2.1.1.1 Construction des plans factoriels complets à deux niveaux

Matrice des essais

Pour étudier l'influence d'un facteur sur une réponse, il faut lui affecter au moins deux niveaux. Reprenons les notions du premier niveau est -1, le second par +1. Pour l'étudier deux facteurs 1 et 2, à raison de deux niveaux pour chaque facteur, il faut réaliser $2^2 = 4$ essais. Les conditions opératoires de chaque essai sont décrites dans le tableau 2.1.

Tableau 2.1 : Matrice d'expériences d'un plan 2^2

Essais N°	Facteur 1	Facteur 2
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1

Pour trois facteurs le nombre des essais à réaliser est huit essais. La matrice des essais ou d'expériences est donnée par le tableau 2.2.

Tableau 2.2 : Matrice d'expériences d'un plan 2^3

Essais N°	Facteur 1	Facteur 2	Facteur 3
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1

Pour k facteurs, le nombre d'essais est de 2^k d'où le nom de plan 2^k donné à cette famille de plan. La matrice des essais comporte k colonnes et 2^k lignes. Elle se construit simplement

- Colonne du 1^{er} facteur : alternance de -1 et +1
- Colonne du 2^{ième} facteur : alternance de -1 et +1 de 2 en 2
- Colonne du 3^{ième} facteur : alternance de -1 et +1 de 4 en 4
- Colonne du 4^{ième} facteur : alternance de -1 et +1 de 8 en 8, ... etc.

Domaine expérimental

Le domaine d'étude peut être représenté dans l'espace expérimental des facteurs à k dimensions. Lorsque $k = 2$, nous obtenons un carré. Les points expérimentaux sont situés aux quatre sommets du carré (Figure 2.1)

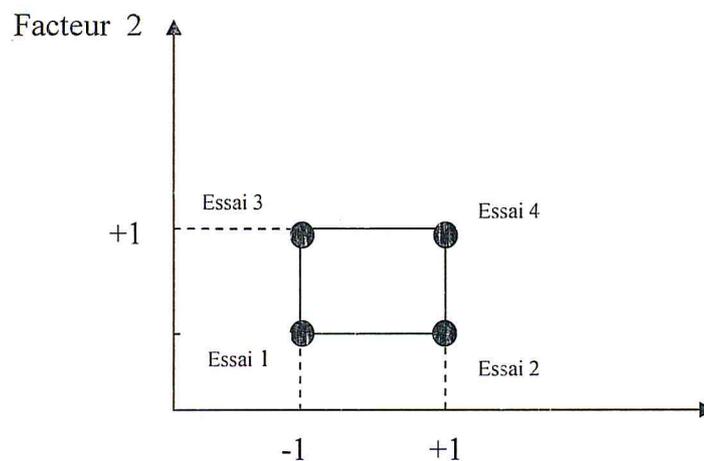


Figure 2.1 : Domaine d'étude d'un plan 2^2

Lorsque $k = 3$, le domaine expérimentale est représenté par un cube, dont les huit sommets représentent les huit essais de plan (Figure 2.2)

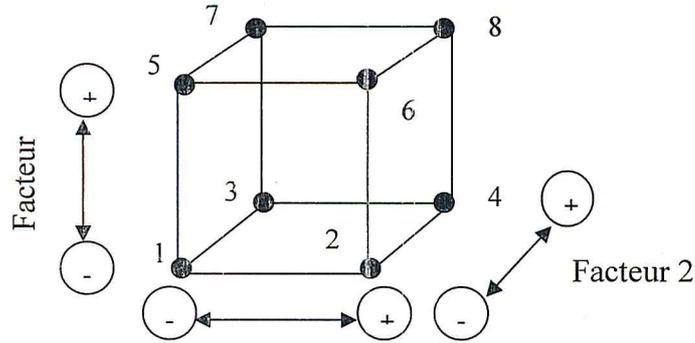


Figure 2.2 : Domaine expérimental d'un plan 2^3

Quand $K > 3$ la représentation géométrique du domaine n'est plus possible mais les points expérimentaux sont toujours aux sommets d'un hyper-cube à k dimensions.

2.1.1.2 Plans factoriel à deux facteurs

Pour deux facteurs, nous avons vu que le domaine d'étude est un carré. Par exemple la figure 2.1 représente un plan factoriel complet à deux facteurs. Le modèle mathématique postulé est un modèle de premier degré par rapport à chaque facteur. Le modèle est :

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 \quad (2.1)$$

y : la réponse

x_i : niveau attribué au facteur i

a_0 : valeur de la réponse au centre du domaine d'étude

a_1 (resp. a_2) : l'effet principal du facteur 1 (resp. facteur 2)

a_{12} : interaction entre les facteurs 1 et 2

Effets d'un facteur

L'expérimentateur ayant réalisé les essais est en possession de quatre valeurs de la réponse : y_1, y_2, y_3 et y_4 . Il a donc un système de quatre équations à quatre inconnues. Les inconnues étant les coefficients du modèle : a_0, a_1, a_2 et a_{12} . En remplaçant dans la relation (2.1) les x_i par leur valeur, on obtient :

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

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

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

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

La résolution de ce système donne :

$$a_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4) \quad (2.2)$$

$$a_1 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4) \quad (2.3)$$

$$a_2 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4) \quad (2.4)$$

$$a_{12} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4) \quad (2.5)$$

Signification de a_0

Le coefficient a_0 est la valeur de la réponse au centre du domaine d'étude. La formule (2.2) montre également que a_0 peut être considéré comme la moyenne des quatre réponses.

Signification de a_1

Plaçons-nous maintenant au niveau moyen du facteur 2. Pour cela, donnons la valeur zéro à x_2 . La relation (2.3) devient :

$$y = a_0 + a_1 x_1$$

Cette relation permet de tracer l'évolution de la réponse prédite dans un plan de coupe $x_2 = 0$ (figure 2.3). L'effet de facteur 1 (a_1) apparaît comme c'est la variation de la réponse lorsqu'on passe du niveau bas du facteur 1 au niveau haut de même facteur et lorsqu'on trace cet effet on aura :

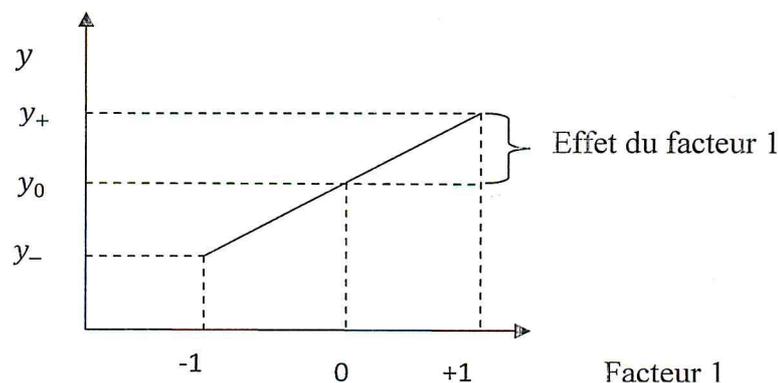


Figure 2.3 : Effet du facteur 1

Où y_+ (resp. y_-) est la valeur moyenne de la réponse au niveau haut (resp. bas) du facteur 1.

Signification de a_{12}

La relation (2.5) peut s'écrire :

$$Y = \frac{1}{2} \left(\frac{1}{2} (y_4 - y_3) - \frac{1}{2} (y_2 - y_1) \right) = \frac{1}{2} (y_+ - y_-) \quad (2.6)$$

L'interaction apparait comme étant la demi différence entre l'effet du facteur 1 au niveau haut du facteur 2 (effet noté y_+) et l'effet du facteur 2 (effet noté y_-). Elle traduit une variation de l'effet d'un facteur en fonction du niveau d'un autre facteur.

Matrice des effets

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

- **Calcul de la moyenne**

Le processus de calcul à adopter pour les effets peut s'appliquer en utilisant une colonne de signe + puisqu'il n'y a que ce signe dans la formule (2.2).

- **Calcul de l'interaction**

La suite des signes de la relation (2.5) est + - - +. Chacun des ces signes provient du x_1, x_2 figurant dans la relation (2.1). On peut retrouver cette suite de signes de la manière suivante : écrit en colonne, les signes correspondants à x_1 et à x_2 , puis en faisant le produit scalaire des éléments correspondants des colonnes des facteurs.

$$\begin{array}{ccc} x_1 & x_2 & x_1 x_2 \\ \left(\begin{array}{ccc} -1 & -1 & +1 \\ +1 & -1 & -1 \\ -1 & +1 & -1 \\ +1 & +1 & +1 \end{array} \right) \end{array}$$

Ayant la matrice d'expériences, il est facile de construire la matrice de effets (Tableau 2.3) en ajoutant une colonne de signe + pour la moyenne et en calculant celle de l'interaction comme précédemment.

Tableau 2.4 : Matrice des effets d'un plan 2^3

Essais N°	Moy	1	2	3	12	13	23	123	Réponses
1	+1	-1	-1	-1	+1	+1	+1	-1	38
2	+1	+1	-1	-1	-1	-1	+1	+1	37
3	+1	-1	+1	-1	-1	+1	-1	+1	26
4	+1	+1	+1	-1	+1	-1	-1	-1	24
5	+1	-1	-1	+1	+1	-1	-1	+1	30
6	+1	+1	-1	+1	-1	+1	-1	-1	28
7	+1	-1	+1	+1	-1	-1	+1	-1	19
8	+1	+1	+1	+1	+1	+1	+1	+1	16
Effets	27,25	-1	-6	-1	0,25	0,25	0,25	0	

Seuls les effets 2 et 3 sont significatifs.

Conclusions de l'étude

La concentration de l'acide gras est pratiquement sans influence sur la stabilité de l'émulsion. Par contre, la dilution de l'acide chlorhydrique est un facteur important à effet négatif. La nature du bitume est également importante, la meilleure stabilité sera obtenue : avec le bitume B, il n'y a aucune interaction significative.

Remarque

L'effet est négatif quand la réponse diminue lorsque le facteur correspondant passe du niveau -1 au niveau +1.

2.1.1.4 Plans complets 2^k

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

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j \quad (2.7)$$

On peut écrire sous forme matricielle :

$$Y = XA \quad (2.8)$$

Dans le cas des plans factoriels à deux niveaux, la matrice X qui est orthogonale est appelée matrice d'Hadamard. Elle possède l'importante propriété : $X'X = NI$ où N le nombre d'essais et I la matrice d'unité. Si l'on reporte cette expression dans la formule (2.1) on obtient :

$$\hat{A} = (X'X)^{-1} X'Y = (NI)^{-1} X'Y \quad (2.9)$$

D'où

$$\hat{A} = \frac{1}{N} X'Y$$

2.1.1.5 Optimalité des plans 2^k

Tous les effets et les interactions sont estimés d'une manière indépendante, avec la variance minimale $var(y)/N$ [5].

2.1.1.6 Avantages et inconvénients des plans factoriels complets à deux niveaux

Les avantages [8] des plans factoriels sont nombreux et nous n'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.
- L'interprétation des résultats est à la portée de tout expérimentateur et ne demande pas des connaissances approfondies en statistique.
- 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.

L'inconvénient de ces plans est que le nombre des essais à réaliser devient rapidement important. De plus, on ne peut calculer que les effets principaux et les interactions donc on ne peut pas obtenir de modèles du second degré et plus.

2.1.2 Plans fractionnaires à deux niveaux

Les plans fractionnaires sont appelés plans 2^{k-p} . Le 2 signifie que chaque facteur prend 2 niveaux, le k indique que k facteurs sont étudiés, le p signifie qu'il y a p facteurs supplémentaires par rapport au nombre de facteurs du plan complet. L'avantage des plans fractionnaires est évident : la charge expérimentale est divisée par 2^p car $2^{k-p} = 2^k / 2^p$.

2.1.2.1 Théorie des alias

Reprenons le dernier exemple. L'expérimentateur a réalisé huit essais. Aurait-il pu en réaliser moins, par exemple, la moitié ? Supposons qu'il ne réalise que la moitié des essais soit $2^{3-1} = 2^2 = 4$ essais. La matrice d'expériences se présente selon le tableau 2.5 et les réponses sont, bien sûr les mêmes que celles de l'exemple précédent.

Tableau 2.5 : Matrice des expériences (étude d'une émulsion de bitume)

Essais N°	1	2	12	y
1	-1	-1	1	$y_1 = 30$
2	1	-1	-1	$y_2 = 37$
3	-1	1	-1	$y_3 = 26$
4	1	1	1	$y_4 = 16$

On peut comparer les effets calculés à partir de ce plan fractionnaire avec le plan complet (Tableau 2.6).

Tableau 2.6 : Comparaison des effets calculés

Effets	Plan complet	Plan fractionnaire
Moyenne	27,25	27,25
1	-1	0,75
2	-6	-6,25
3	-4	-4,25

Ces résultats sont tout à fait comparables à ceux qui ont été obtenus avec le plan complet de huit essais. Il semble donc que, pour un effort moindre on puisse obtenir les mêmes résultats. Cela paraît surprenant et n'a-t-il pas fallu rien donner en échange des quatre essais ? Quel est donc le prix à payer pour faire moins d'essais pour faire

apparaître la réponse ? Reprenons la valeur de l'effet 3 et de l'interaction 12 calculée avec le plan complet :

$$a_3 = \frac{1}{8}(-y_1 - y_2 - y_3 - y_4 + y_5 + y_6 + y_7 + y_8)$$

$$a_{12} = \frac{1}{8}(+y_1 - y_2 - y_3 + y_4 + y_5 - y_6 - y_7 + y_8)$$

Et additionnant :

$$a_3 + a_{12} = \frac{1}{4}(-y_2 - y_3 + y_5 + y_8)$$

On retrouve la quantité :

$$c_3 = a_3 + a_{12} = -4,25$$

C'est-à-dire que c_3 est égal à l'effet principal a_3 augmenté de l'interaction a_{12} . On dit que a_1 et a_{12} sont **aliasés**. La quantité c_3 peut être appelée **alias** ou **contraste** ou simplement effet.

On verra de même que :

$$c_1 = a_1 + a_{23} = -0,75$$

$$c_2 = a_2 + a_{13} = 6,25$$

Le modèle du plan fractionnaire 2^{3-1} s'écrit :

$$y = c_0 + c_1x_1 + c_2x_2 + c_3x_3$$

Les hypothèses d'interprétation

Tous les plans fractionnaires posent le même problème d'interprétation des résultats. Comme on n'effectue pas toutes les expériences du plan complet, on ne peut pas obtenir la valeur de toutes les interactions. Il faut créer soi-même des informations supplémentaires de remplacement. Ces informations supplémentaires doivent être réalistes et compatibles avec l'étude menée. Elles sont introduites sous forme d'hypothèses et elles demandent à être vérifiées avant la conclusion de l'étude. Les hypothèses de travail les plus souvent retenues sont les suivantes :

- **Hypothèse 1**

Les interactions d'ordre 3 ou d'ordre plus élevé sont considérées comme négligeables. On élimine ainsi un grand nombre d'inconnues. Mais cette hypothèse peut parfois être mise en défaut.

- **Hypothèse 2**

Si un contraste est nul, cela peut signifier :

-Que les effets et les interactions aliasés sont tous nuls. C'est l'hypothèse est la plus probable et c'est celle que retiendrons sous le nom d'hypothèse 2.

- Que les effets et les interactions aliasés se compensent. Cette hypothèse est

Probable et nous ne la retiendrons pas.

- **Hypothèse 3**

Si deux contrastes sont faibles, on supposera que leur interaction l'est aussi et si un contraste est faible et l'autre fort, on supposera que leur interaction est faible.

- **Hypothèse 4**

Si deux contrastes sont forts, on se méfiera de leur interaction qui peut l'être également.

Les hypothèses présentées ici sont très souvent vérifiées mais, il arrive parfois qu'elles soient mises en défaut. Il est toujours possible d'en adopter d'autres en fonction du problème traité et des risques encourus. Pour une bonne analyse des résultats il est prudent de toujours bien préciser les hypothèses d'interprétation que l'on a retenues [7].

Calcul des contrastes

Reprenons la matrice des effets d'un plan 2^3 où l'ordre des essais a été choisi pour faire ressortir 2 plans 2^2 pour les facteurs 1 et 2. Le plan 2^3 a donc été divisé en 2 demi-plans (Tableau 2.7)

Tableau 2.7 : Matrice de calcul des contrastes pour les deux demi-plans fractionnaires

Essais N°	I	1	2	3	12	13	23	123
5	1	-1	-1	1	1	-1	-1	1
2	1	1	-1	-1	-1	-1	1	1
3	1	-1	1	-1	-1	1	-1	1
8	1	1	1	1	1	1	1	1
1	1	-1	-1	-1	1	1	1	-1
6	1	-1	1	1	-1	1	-1	-1
7	1	1	1	1	-1	-1	1	-1
4	1	1	-1	-1	1	-1	-1	-1

1^{ère} demi-plan

2^{ème} demi-plan

Considérons le demi-plan supérieur. En notation de Box, on constate que 3 et 12 sont égaux puisqu'ils ont la même suite de signes + - - +, on peut écrire :

$$3 = 12$$

Or, on avait démontré

$$c_3 = a_3 + a_{12}$$

Donc $3 = 12$ est équivalent à $c_3 = a_3 + a_{12}$. Cette relation d'équivalence est valable dans les deux sens et elle constitue la base de la théorie des alias. On montre de même que :

$$1 = 23 \text{ est équivalent à } c_1 = a_1 + a_{23}$$

$$2 = 13 \text{ est équivalent à } c_2 = a_2 + a_{13}$$

On peut trouver ces relations toujours dans le demi plan supérieur, les deux colonnes de signe + permettent d'écrire :

$$I = 123$$

En multipliant successivement cette relation appelée le générateur des alias par 1, 2 et 3 :

$$1. I = 1.123 = 1^223 = 23$$

$$2. I = 2.123 = 12^23 = 13$$

$$3. I = 3.123 = 123^2 = 12$$

2.1.2.2 Construction pratique des plans fractionnaires

La construction pratique des plans fractionnaires est basée sur :

- On choisit un plan complet et l'on écrit sa matrice de calcul des effets. On appelle Cette matrice le plan de base.
- Dans ce plan de base, on choisit p colonnes de signes correspondants à p interactions les plus élevées possibles (pour qu'elle ait le plus de chance d'être faible) pour étudier les facteurs supplémentaires.

Pour illustrer cette construction des plans fractionnaires prenons comme plan de base 2^3 , la matrice des effets comprend 3 interactions d'ordre 2 et une d'ordre 3. Le plan complet permet d'étudier 3 facteurs sur la colonne 1, 2 et 3.

Essais I 1 2 3 12 13 23 123

1	}	1	-1	-1	-1	1	1	1	-1
2		1	1	-1	-1	-1	-1	1	1
3		1	-1	1	-1	-1	1	-1	1
4		1	1	1	-1	1	-1	-1	-1
5		1	-1	-1	1	1	-1	-1	1
6		1	1	-1	1	-1	1	-1	-1
7		1	-1	1	1	-1	-1	1	-1
8		1	1	1	1	1	1	1	1

Si nous voulons étudier 4 facteurs, nous conservons les trois premières colonnes de signes pour les 3 premiers facteurs et nous choisirons la colonne d'une interaction pour le quatrième facteur. Pour fixer les idées, prenons la colonne de l'interaction 12 et écrivons que l'on attribue au facteur supplémentaire 4 les niveaux définis par les signes de cette interaction :

$$4 = 12$$

D'où le générateur d'aliases

$$4.4 = 4.12$$

En multipliant ce générateur successivement par 1, 2, 3 et 4, on obtient :

$$1 = 24 \text{ est équivalent à } c_1 = a_1 + a_{24}$$

$$2 = 14 \text{ est équivalent à } c_2 = a_2 + a_{14}$$

$$4 = 12 \text{ est équivalent à } c_4 = a_4 + a_{12}$$

$$3 = 1234 \text{ est équivalent à } c_3 = a_3 + a_{1234}$$

On aurait pu aliaser le facteur 4 sur une autre interaction, on aurait eut d'autres valeurs des contrastes. Il est tout à fait possible d'étudier deux facteurs supplémentaires, en utilisant un plan 2^{5-2} . On peut choisir comme aliases par exemple :

$$4 = 12$$

$$5 = 13$$

D'où les générateurs d'aliases indépendants :

$$I = 124$$

$$I = 135$$

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

$$I. I = 124.135$$

$$I. I = 2345$$

Deux facteurs supplémentaires introduisent donc un groupe d'aliases ou GGA comportant quatre termes.

$$I = 124 = 135 = 2345$$

On utilise ce GGA pour savoir comment les facteurs et les interactions sont aliasés dans les contrastes que l'on calcule avec ce plan fractionnaire. Par exemple pour déterminer le contraste c_1 on multipliant tous les termes du GGA par la colonne 1 :

$$I. I = 1.124 = 1.135 = 1.2345$$

$$1 = 24 = 25 = 12345 \text{ est équivalent à } c_1 = a_1 + a_{24} + a_{35} + a_{12345}$$

Pour un plan de base 2^4 , la matrice des effets comprend six interactions d'ordre 2, quatre d'ordre 3, et une d'ordre 4. Le plan complet permet d'étudier quatre facteurs 1, 2, 3 et 4. Pour étudier 8 facteurs, le plan complet contient 256 effets et interactions, le plan $2^{8-4} = 2^4$ ne permettra de calculer que 16 contrastes. Il faut quatre interactions pour étudier les quatre facteurs supplémentaires. On peut choisir comme alias par exemple :

$$5 = 123$$

$$6 = 124$$

$$7 = 134$$

$$8 = 1234$$

Les quatre générateurs d'alias indépendants sont :

$$I = 1235 = 1246 = 1347 = 12348$$

Les générateurs dépendants se calculent à partir des générateurs indépendants en les multipliant 2 à 2, 3 à 3, et 4 à 4.

Multiplication 2 à 2 :

$$1235.1246 = 3456$$

$$1235.1347 = 2457$$

$$1235.12348 = 458$$

$$1246.1347 = 2367$$

$$1246.12348 = 368$$

$$1347.12348 = 278$$

Multiplication 3 à 3 :

$$1235.1246.1347 = 1567$$

$$1235.1246.12348 = 12568$$

$$1235.1347.12348 = 13578$$

$$1246.12348.1347 = 14678$$

Multiplication 4 à 4 :

$$1235.1246.1347.12348 = 2345678$$

Le groupe de générateurs d'aliases (GGA) sera donc :

$$I = 278 = 368 = 458 = 1235 = 1246 = 1347 = 3456 = 2457 = 2367 = 12348 = 1567 = \\ 12568 = 13578 = 14678 = 2345678$$

Le calcul complet est présenté ici pour le facteur 1 :

$$c_1 = a_1 + a_{235} + a_{246} + a_{347} + a_{567} + a_{1278} + a_{1368} + a_{1458} + a_{13456} \\ + a_{12457} + a_{12367} + a_{2568} + a_{2348} + a_{3578} + a_{4678} + a_{12345678}$$

De même si l'on néglige les interactions d'ordre supérieur à trois on peut écrire:

$$\begin{array}{lll} c_1 = a_1, & c_6 = a_6 + a_{38}, & c_{14} = a_{14} + a_{26} + a_{37}, \\ c_2 = a_2 + a_{78}, & c_7 = a_7 + a_{28}, & c_{23} = a_{23} + a_{15} + a_{67}, \\ c_3 = a_3 + a_{68}, & c_8 = a_8 + a_{45} + a_{36} + a_{27}, & c_{24} = a_{24} + a_{16} + a_{57}, \\ c_4 = a_4 + a_{58}, & c_{12} = a_{12} + a_{35} + a_{46}, & c_{34} = a_{34} + a_{17} + a_{56}, \\ c_5 = a_5 + a_{48}, & c_{13} = a_{13} + a_{25} + a_{47}, & c_{234} = a_{234} + a_{18}, \end{array}$$

2.1.2.3 Notion de résolution

Nous avons vu que dans un plan fractionnaire les effets principaux sont aliasés avec des interactions d'ordres différents. La résolution X d'un plan est égale à 1 plus la valeur de l'ordre de l'interaction d'ordre le moins élevé. Par exemple si un facteur principal est aliasé avec des interactions d'ordre 2, 3 et 4, la résolution sera $1+2 = 3$. On écrit souvent la résolution en chiffres romains. On écrit résolution III.

2.1.2.4 Avantages et inconvénients

Les avantages des plans fractionnaires sont les même que ceux des plans factoriels complets à deux niveaux. Mais, en plus ils permettent d'échapper au principal inconvénient de ces derniers : le grand nombre d'essais. Par exemple on commence par un plan 2^{6-3} et s'il s'avère nécessaire de faire des essais complémentaires pour lever les ambiguïtés, on poursuivra les expériences en exécutant un deuxième plan 2^{6-3} qui, ajouté

au premier donnera deux plans 2^{6-3} soit $2 \cdot 2^{6-3}$ ou 2^{6-3+1} que l'on écrira plus simplement 2^{6-2} .

Le plus important serait peut-être que l'expérimentateur doit faire un effort pour apprendre à maîtriser l'emploi des plans factoriels fractionnaires.

2.1.3 Plan de Plackett et Burman

2.1.3.1 Construction des plans de Plackett et Burman

La structure des alias de ces plans est très différente de celle des plans complets 2^k ou des plans fractionnaires 2^{k-p} . D'un point de vue pratique, cette structure est intéressante car elle minimise l'influence des interactions lorsqu'on calcule les contrastes correspondant aux effets principaux. C'est pourquoi nous ferons une distinction dans les plans basés sur des matrices d'Hadamard. Cette distinction ne suivra pas les faits historiques. En effet, Plackett et Burman ont publié tous les plans basés sur des matrices d'Hadamard depuis 2 jusqu'à 100 lignes. Malgré tout nous ferons la distinction suivant :

- Nous appellerons plans factoriels complets ou fractionnaires tous les plans Basés sur des matrices d'Hadamard et dont le nombre de lignes est à 2^k .
- Nous appellerons plans Plackett et Burman tous les plans basés sur des matrices d'Hadamard et dont le nombre de lignes n'est pas égal à 2^k .

Ainsi un plan de essais basé sur une matrice d'Hadamard sera appelé un plan de Plackett et Burman, tandis qu'un plan basé sur les même matrices mais ayant 32 essais sera appelé plan factoriel.

Examinons maintenant en détail la méthode de construction des matrices d'expériences d'Hadamard : Ces matrices se construisent selon un algorithme qui donne la 1^{ère} ligne de la matrice d'expériences, les autres lignes étant générées par permutation circulaire de cette 1^{ère} ligne et la dernière ligne étant une ligne ne comportant que des signes(-) [7].

▪ $N=4$

+ + -

▪ $N=8$

+++--

▪ $N=12$

++-++++-

▪ $N=16$

++++-+-+--

▪ $N=24$

++++-+-+--+-+--

▪ $N=32$

---+-+--+--+--++++-+-+--+

Pour illustrer la construction des plans de Plackett et Burman prenons par exemple $N=8$, la matrice d'expériences est :

Essais	1	2	3	4	5	6	7
1	+1	+1	+1	-1	+1	-1	-1
2	-1	+1	+1	+1	-1	+1	-1
3	-1	-1	+1	+1	+1	-1	+1
4	+1	-1	-1	+1	+1	+1	-1
5	-1	+1	-1	-1	+1	+1	+1
6	+1	-1	+1	-1	-1	+1	+1
7	+1	+1	-1	+1	-1	-1	+1
8	-1	-1	-1	-1	-1	-1	-1

Cette matrice est une matrice d'Hadamard. Nous allons la transformer pour la présenter selon la disposition classique des plans factoriels. On sait qu'il est possible d'inverser les signes + et - sans modifier les résultats de l'expérimentation. Effectuons cette transformation :

Essais	1	2	3	4	5	6	7
1	-1	-1	-1	+1	-1	+1	+1
2	+1	-1	-1	-1	+1	-1	+1
3	+1	+1	-1	-1	-1	+1	-1
4	-1	+1	+1	-1	-1	-1	+1
5	+1	-1	+1	+1	-1	-1	-1
6	-1	+1	-1	+1	+1	-1	-1
7	-1	-1	+1	-1	+1	+1	-1
8	+1	+1	+1	+1	+1	+1	+1

Ordonnons maintenant ces lignes pour retrouver la séquence des signes des colonnes des matrices des effets des plans 2^3 .

	1	2	3	13	123	12	23	I
	-1	-1	-1	+1	-1	+1	+1	+1
	+1	-1	-1	-1	+1	-1	+1	+1
	-1	+1	-1	+1	+1	-1	-1	+1
	+1	+1	-1	-1	-1	+1	-1	+1
	-1	-1	+1	-1	+1	+1	-1	+1
	+1	-1	+1	+1	-1	-1	-1	+1
	-1	+1	+1	-1	-1	-1	+1	+1
	+1	+1	+1	+1	+1	+1	+1	+1

Nous retrouvons, au déplacement près de quelques colonnes, un plans 2^3 et comme Plackett et Burman n'envisagent pas d'interaction, il s'agit d'un plan factoriel fractionnaire 2^{7-4} ou les quatre facteurs supplémentaires sont confondus avec les quatre interactions. On peut écrire : un Plackett et Burman $N=8$ est un plan factoriel

Si nous voulons étudier 11 facteurs, il faudra donc au minimum 12 expériences, la 1^{ère} ligne de cette matrice est : ++ - +++ - - - + -

A partir de cette 1^{ère} ligne, les lignes suivantes sont générées par permutation circulaire (vers la droite) et nous obtenons la matrice d'expériences suivante, comportant 11 colonnes et 12 lignes (tableau 2.8).

Tableau 2.8 : Matrice d'expériences

Essais N°	1	2	3	4	5	6	7	8	9	10	11
1	+1	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1
2	-1	+1	+1	-1	+1	+1	+1	-1	-1	-1	+1
3	+1	-1	+1	+1	-1	+1	+1	+1	-1	-1	-1
4	-1	+1	-1	+1	+1	-1	+1	+1	+1	-1	-1
5	-1	-1	+1	-1	+1	+1	-1	+1	+1	+1	-1
6	-1	-1	-1	+1	-1	+1	+1	-1	+1	+1	+1
7	+1	-1	-1	-1	+1	-1	+1	+1	-1	+1	+1
8	+1	+1	-1	-1	-1	+1	-1	+1	+1	-1	+1
9	+1	+1	+1	-1	-1	-1	+1	-1	+1	+1	-1
10	-1	+1	+1	+1	-1	-1	-1	+1	-1	+1	+1
11	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1	+1
12	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

2.1.3.2 Modélisation mathématique

Les plans de Plackett et Burman utilisent comme modèle mathématique des polynômes du premier degré sans interaction, tel que :

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n$$

L'interprétation des plans Plackett et Burman est analogue à celle des plans factoriels fractionnaires. Il est donc indispensable de connaître la théorie des aliases si l'on veut exploiter correctement les résultats d'essais. Malheureusement ces plans n'ont pas été présentés, dans le passé, comme des plans factoriels fractionnaires donc comme des plans délicats à interpréter, mais comme un moyen d'étude facile à utiliser et ne posant pas de problème. D'où un certain nombre d'échecs entraînant le rejet global des plans d'expériences par certains expérimentateurs. L'utilisation de ces plans a également développé de mauvaises habitudes comme par exemple celle qui consiste à utiliser les

colonnes non attribuées à des facteurs pour estimer l'erreur expérimentale. Cette pratique est tout-à-fait à déconseiller car il y a risque de prendre une interaction entre deux facteurs comme erreur expérimentale. Cette très mauvaise habitude est malheureusement trop fréquente et même conseillée dans certains manuels ?

Il y a pourtant un plan intéressant dans les plans de Plackett et Burman c'est celui à douze essais, matrice d'Hadamard (12,12) qui permet d'étudier onze facteurs en douze essais. Ce plan correspond à un plan 2^{11-7} mais ne réclame que douze essais contre seize pour le plan fractionnaire. Il peut être très utile si l'économie de quatre expériences fait gagner beaucoup de temps ou d'argent. La difficulté réside dans l'interprétation car les confusions entre effets principaux et interactions n'est pas simple [12].

Le plan factoriel. Il peut être soit complet fractionnaire à deux niveaux par facteurs.

2.2 Plans de modélisation

2.2.1 Plans composites

Les plans composites permettent de calculer un modèle polynomial du second degré. On parle de modèle quadratique. Ils sont composés d'une première partie qui est un plan fractionnaire ou complet, puis d'un ou plusieurs essais au centre du domaine d'étude et enfin d'essais supplémentaires permettant le calcul du modèle quadratique.

Les **plans centraux composites** sont des plans intéressants dans la mesure où ils permettant de faire une étude d'un système de manière séquentielle.

L'utilisateur peut réaliser les premiers essais du plan factoriel et calculer un modèle linéaire. Dans le cas où la précision du modèle par rapport au procédé réel n'est pas suffisante, les essais complémentaires sont réalisés et un modèle quadratique est calculé.

Le plan composite est constitué de trois parties distinctes, autorisant une étude séquentielle :

1. Le plan factoriel. Il peut être soit complet fractionnaire à deux niveaux par facteurs.
2. Le plan en étoile. Les points expérimentaux sont situés à la même distance du centre du domaine d'étude.
3. Les points au centre du domaine.

2.2.1.1 Construction des plans composites

Revenons en détail sur les trois éléments des plans composites :

- **Plan factoriel**

La partie factorielle est constituée de points d'expériences situés aux sommets de carrés, de cubes ou d'hyper cubes suivant le nombre de facteurs étudiés, et si possible des points de contrôle situés au centre du domaine d'étude.

- **Plan en étoile**

Le choix de l'emplacement des points en étoile est lié aux conditions expérimentales et à des considérations théoriques basées sur les critères d'optimalité. Les conditions expérimentales dépendent de la facilité ou de l'impossibilité d'atteindre certaines zones du domaine d'étude. Par exemple, si l'on a choisi, pour la partie factorielle, un domaine d'étude de 80 °C à 99 °C pour l'étude d'une solution aqueuse, on ne pourra guère élargir le domaine vers les hautes températures lors du choix des points en étoile. Dans ce cas, on pourra choisir un plan central composite à faces centrées. Il convient néanmoins de toujours réfléchir à ce problème d'extension de domaine au moment de la préparation du plan d'expériences pour ne pas se faire surprendre.

Les considérations théoriques seront examinées en détail dans les paragraphes suivants et permettront de trouver la valeur de α .

- **Points au centre**

On contrôle la validité et la stabilité des plans factoriels et des plans en étoile en

Ajoutant des points d'expériences situés, le plus souvent, au centre du domaine d'étude.

Ces points sont importants et ils ont plusieurs rôles :

- ✓ Ils servent à tester la validité du modèle du premier degré,
- ✓ Ils servent à s'assurer qu'il n'y a pas de glissement (stabilité) entre les deux séries d'essais,
- ✓ Ils permettent d'obtenir une estimation de l'erreur expérimentale,
- ✓ Ils diminuent l'erreur de prédiction près du point central,
- ✓ Ils interviennent dans le calcul de α .

Le nombre total N d'essais à réaliser est la somme de chacun des parties qui composent le plan :

- Les essais du plan factoriel, soit n_f ,
- Les essais du plan en étoile, soit n_a ,
- Les essais au centre, soit n_0 .

Le nombre total n d'essais d'un plan composite est donné par la relation :

$$N = n_f + n_a + n_0$$

2.2.1.2 Les différentes classes de plans composites

On précise parfois deux types de différentes classes de plans composites :

- ❖ **Plans composites centrés dans les faces** : ces plans peuvent être considérés comme un cas particulier des plans composites centrés, la charge expérimentale est la même, chaque facteur est étudié sur 3 niveaux seulement (figure 2. 4).

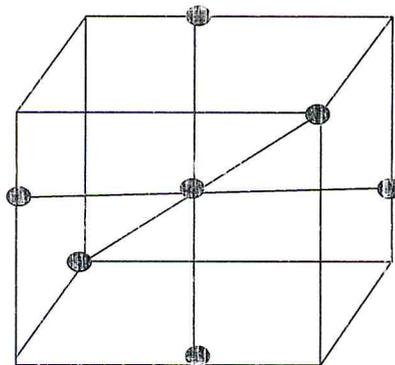


Figure 2.4 : Plans composites centrés dans les faces

- ❖ **Plans composites centrés** : Si les points en étoile sont à l'extérieur du domaine cubique (figure 2. 5).

Le nombre de niveaux d'un plan composite est de cinq par facteur : le point central, les deux niveaux du plan factoriel et les deux niveaux des points en étoile.

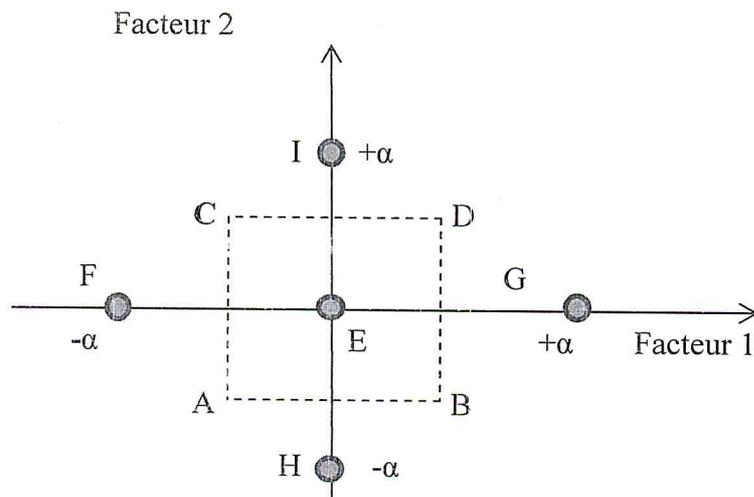


Figure 2.5 : Plans composite centrés pour l'étude de deux facteurs

- Les points A , B , C et D sont les points expérimentaux d'un plan factoriel 2^2 .
- Le point E est le point central. Ce point peut avoir été répliqué une ou plusieurs fois,
- Les points F , G , H et I sont les points d'un plan en étoile.

La matrice d'expériences correspondante à ce plan est donnée par le tableau 2.9. La distance des points en étoile au centre du domaine est notée par la lettre grecque α .

Tableau 2.9 : Matrice d'expériences d'un plan composite pour deux facteurs.

Essais n°	Facteur 1	Facteur 2
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1
5	$-\alpha$	0
6	$+\alpha$	0
7	0	$-\alpha$
8	0	$+\alpha$
9 à 11	0	0

2.2.1.3 Propriétés des plans composites

➤ Modèle mathématique postulé

Est un modèle du second degré dans lequel on conserve le coefficient constant. Les termes du premier degré, les termes rectangles et les termes du second degré. Il n'y a pas d'interaction d'ordre 3 ou plus élevé.

- Pour deux facteurs le modèle s'écrit :

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2$$

- Pour trois facteurs le modèle s'écrit :

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + e$$

➤ Matrice de calcul

La matrice du calcul est notée X . Elle est caractérisée par p colonnes et n lignes.

Pour deux facteurs la matrice de calcul X est une matrice de dimensions (11×6) puisqu'il y a 11 expériences et 6 coefficients dans le modèle postulé :

$$X = \begin{pmatrix} 1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -\alpha & 0 & 0 & \alpha^2 & 0 \\ 1 & \alpha & 0 & 0 & \alpha^2 & 0 \\ 1 & 0 & -\alpha & 0 & 0 & \alpha^2 \\ 1 & 0 & \alpha & 0 & 0 & \alpha^2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

➤ Matrice d'information

La matrice d'information est définie à partir du produit matriciel entre la transposée de la matrice du calcul, notée $'X$. Il s'agit d'une matrice carrée contenant p ligne et p colonnes. Le déterminant de cette matrice est un indicateur algébrique qui traduit la qualité d'une matrice d'expériences de taille n .

La matrice d'information est calculée à partir de la matrice de calcul, elle égale tXX .

Pour 2 facteurs on a :

$${}^tXX = \begin{pmatrix} 11 & 0 & 0 & 0 & 4+2\alpha^2 & 4+2\alpha^2 \\ 0 & 4+2\alpha^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4+2\alpha^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 4+2\alpha^2 & 0 & 0 & 0 & 4+2\alpha^4 & 4 \\ 4+2\alpha^2 & 0 & 0 & 0 & 4 & 4+2\alpha^4 \end{pmatrix}$$

Où tX matrice transposée de la matrice de calcul.

D'une manière générale, la matrice d'information peut s'écrire :

$${}^tXX = \begin{pmatrix} N & 0 & 0 & 0 & n_f+2\alpha^2 & n_f+2\alpha^2 \\ 0 & n_f+2\alpha^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & n_f+2\alpha^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & n_f & 0 & 0 \\ n_f+2\alpha^2 & 0 & 0 & 0 & n_f+2\alpha^4 & n_f \\ n_f+2\alpha^2 & 0 & 0 & 0 & n_f & n_f+2\alpha^4 \end{pmatrix}$$

➤ Critères d'optimalité

Nous n'envisageons ici que le cas où tous les points axiaux sont à la même distance (en grandeurs centrées réduites) du centre du domaine d'étude. Suivant les critères d'optimalité choisis, la valeur de α ne sera pas la même.

• Iso variance par rotation

Les éléments de la matrice d'information doivent satisfaire la relation [9] :

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

$$\Leftrightarrow 2n_f = 2\alpha^4$$

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

Dans le cas d'un plan composite pour deux facteurs ayant ses points en étoile situés à $\alpha = (4)^{1/4} = 1.414$ du centre, la matrice d'information est la suivante :

$$({}^tXX)^{-1} = \begin{pmatrix} 1 & & & & & \\ & 8 & & & & \\ & & 8 & & & \\ & & & 4 & & \\ 8 & & & & 12 & 4 \\ 8 & & & & 4 & 12 \end{pmatrix}$$

Cette matrice respecte bien le critère d'iso variance par rotation : $3 \times 4 = 12$. Dans ces conditions, l'erreur de prédiction des réponses calculées est la même pour les points situés à la même distance du point central. Le critère d'iso variance par rotation est respecté quand tous les points du plan sont à la même distance du point central, c'est-à-dire sur un cercle, une sphère ou une hyper sphère.

- **Presque orthogonalité**

C'est un autre critère d'optimalité très utilisé. Au lieu de l'isovariance par rotation, on pourrait vouloir respecter le critère d'orthogonalité. Il faudrait une disposition des points qui entraîne une matrice $'XX$ diagonale. Cela n'est pas possible car on ne peut pas annuler les éléments correspondants aux termes constants et aux termes carrés. Rappelons qu'il faut que la sous matrice, obtenue en éliminant la première ligne et la première colonne de $({}^tXX)^{-1}$ doit être diagonale. On démontre que cela est possible si [11].

$$\alpha = \left(\frac{n_f (\sqrt{N} - \sqrt{n_f})^2}{4} \right)^{1/4}$$

Pour deux facteurs on a $\alpha=1.21$. Ecrivons la matrice d'information correspondante.

$${}^tXX = \begin{pmatrix} 12 & 0 & 0 & 0 & 6,928 & 6,928 \\ 0 & 6,928 & 0 & 0 & 0 & 0 \\ 0 & 0 & 6,928 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 6,928 & 0 & 1 & 0 & 8,287 & 4 \\ 6,928 & 0 & 0 & 0 & 4 & 8,287 \end{pmatrix}$$

La matrice de dispersion correspondante est bien presque orthogonale puisque si l'on élimine la première ligne et la première colonne, on obtient une matrice diagonale.

$$({}^tXX)^{-1} = \begin{pmatrix} 0,239 & 0 & 0 & 0 & -0,135 & -0,135 \\ 0 & 0,144 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,144 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,25 & 0 & 0 \\ -0,135 & 0 & 1 & 0 & 0,233 & 4 \\ -0,135 & 0 & 0 & 0 & 0 & 0,233 \end{pmatrix}$$

➤ **Ecart type des coefficients du modèle prédictif**

La variance de chacun des coefficients du modèle postulé est donnée par la formule.

$$\text{var}(\hat{a}_i) = \text{Diag var}(\hat{A}) = \sigma^2 \text{Diag}(X^T X)^{-1}$$

En prenant les racines carrées des éléments diagonaux de la matrice de variance covariance, on obtient les écarts types des coefficients.

$$\sigma(\hat{a}_0) = 0,500 \sigma$$

$$\sigma(\hat{a}_0) = 0,489 \sigma$$

$$\sigma(\hat{a}_1) = 0,353 \sigma$$

$$\sigma(\hat{a}_1) = 0,380 \sigma$$

$$\sigma(\hat{a}_2) = 0,353 \sigma$$

$$\sigma(\hat{a}_2) = 0,380 \sigma$$

$$\sigma(\hat{a}_{12}) = 0,500 \sigma$$

$$\sigma(\hat{a}_{12}) = 0,500 \sigma$$

$$\sigma(\hat{a}_{11}) = 0,395 \sigma$$

$$\sigma(\hat{a}_{11}) = 0,483 \sigma$$

$$\sigma(\hat{a}_{22}) = 0,395 \sigma$$

$$\sigma(\hat{a}_{22}) = 0,483 \sigma$$

2.2.1.4 Avantages et inconvénients

Les plans composites permettent de mener un travail progressif et rationnel. L'expérimentateur peut réaliser un premier plan factoriel. Ce plan lui indiquera si le domaine retenu doit être conservé et quelle sera l'influence des différents facteurs. S'il existe encore des ambiguïtés à la fin de l'analyse de ce premier plan factoriel, l'expérimentateur pourra faire un plan complémentaire. Ayant défini le bon domaine et connaissant les facteurs influents, il doit s'assurer que le modèle du premier degré est valide. S'il l'est, il peut s'arrêter, s'il ne l'est pas, il doit passer à un modèle du second degré. C'est là que les points expérimentaux supplémentaires du plan en étoile sont précieux.

2.2.2 Plans de Box – Behnken

Box et Behnken ont proposé en 1960 [3] ces plans qui permettent d'établir des modèles du second degré. En effet tous les facteurs ont trois niveaux, -1, 0 et +1. Ces plans sont faciles à mettre en œuvre et possèdent la propriété de séquentialité.

2.2.2.1 Construction des plans de Box-Behnken

Pour 3 facteurs nous avons un cube pour 4 et 5 facteurs nous aurons des hyper cubes à 4 et 5 dimensions (lorsque le nombre de facteurs dépasse 3, nous parlerons de 4- cube, 5-cube, ..., k-cube). On place les points expérimentaux non pas aux sommets du cube ou l'hyper cube, mais au milieu des arêtes, au centre des faces (carrés) ou au centre des cubes.

- Pour 3 facteurs, on a un cube. Le cube possède 12 arêtes et donc le plan à 12 essais. On ajoute 3 essais au centre du domaine d'étude. On obtient au total 15 essais (figure 2.6)

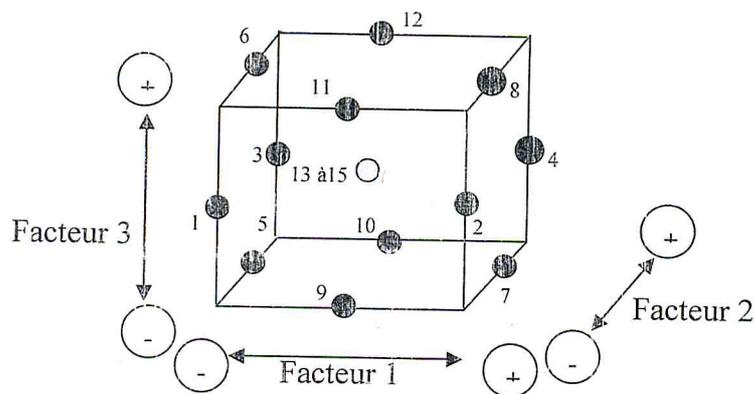


Figure 2.6 : Plan de Box-Behnken pour 3 facteurs

La construction du plan suit le principe suivant : deux facteurs décrivent un carré (4 essais d'un plan 2^2) et les coordonnées correspondantes du troisième facteur valent zéro. Le tableau 2.7 indique ces essais.

Tableau 2.10 : Plan de Box-Behnken pour trois facteurs

Essais	Facteur 1	Facteur 2	Facteur 3
1	-1	-1	0
2	1	-1	0
3	-1	1	0
4	1	1	0
5	-1	0	-1
6	1	0	-1
7	-1	0	1
8	1	0	1
9	0	-1	-1
10	0	1	-1
11	0	-1	1
12	0	1	1
13 à 15	0	0	0

- Pour 4 facteurs, le plan est bâti sur un hyper cube à 4 dimensions. Les points expérimentaux sont situés au centre des 24 carrés. La construction du plan est la suivante : deux facteurs décrivent un carré (4 essais d'un plan 2^2), le troisième et le quatrième facteur ont leurs coordonnées égales à zéro. On pourrait ajouter trois points au centre du domaine d'étude (Tableau 2.11)

Tableau 2.11 : Plan de Box-Behnken pour quatre facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4
1 à 4	± 1	± 1	0	0
5 à 8	± 1	0	± 1	0
9 à 12	± 1	0	0	± 1
13 à 16	0	± 1	± 1	0
17 à 20	0	± 1	0	± 1
21 à 24	0	0	± 1	± 1
25 à 27	0	0	0	0

Dans ce tableau les niveaux des facteurs sont indiqués par ± 1 . Cela signifie que ce facteur prend successivement les niveaux $+1$ et -1 et que toutes les combinaisons sont réalisées.

2.2.2.2 Propriété du plan de Box-Behnken

- **Modèle mathématique postulé**

Le modèle utilisé est un modèle du second degré avec les interactions d'ordre deux pour 3 facteurs le modèle s'écrit :

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

- **Matrice de calcul**

La matrice de calcul est construite à partir du plan d'expériences et du modèle postulé. Pour trois facteurs la matrice est donnée par :

$${}^tXX = \begin{pmatrix} 1 & -1 & -1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 & -1 & 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & 0 & -1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & -1 & 0 & -1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & -1 & 0 & -1 & 0 & 1 & 0 & 1 \\ 1 & -1 & 0 & 1 & 0 & -1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & -1 & -1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & 1 & 1 \\ 1 & 0 & -1 & 1 & 0 & 0 & -1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- **Matrice d'information**

Cette matrice est calculée à partir de la matrice de calcul (tXX). Pour trois facteurs on a :

Tableau 2.12 : Plan de Box-Behnken pour six facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4	Facteur 5	Facteur 6
1 à 8	± 1	± 1	0	± 1	0	0
9 à 16	0	± 1	± 1	0	± 1	0
17 à 24	0	0	± 1	± 1	0	± 1
25 à 32	± 1	0	0	± 1	± 1	0
33 à 40	0	± 1	0	0	± 1	± 1
41 à 48	± 1	0	± 1	0	0	± 1
49 à 52	0	0	0	0	0	0

Tableau 2.13 : Plan de Box-Behnken pour sept facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4	Facteur 5	Facteur 6	Facteur 7
1 à 8	0	0	0	± 1	± 1	± 1	0
9 à 16	± 1	0	0	0	0	± 1	± 1
17 à 24	0	± 1	0	0	± 1	0	± 1
25 à 32	± 1	± 1	0	± 1	0	0	0
33 à 40	0	0	± 1	± 1	0	0	± 1
41 à 48	± 1	0	± 1	0	± 1	0	0
49 à 52	0	± 1	± 1	0	0	± 1	0
57 à 62	0	0	0	0	0	0	0

2.2.2.4 Avantages et inconvénients

Les plans de Box-Behnken sont faciles à construire, et permettent de réaliser un modèle du deuxième degré. Mais dès que le nombre de facteurs à étudier devient grand ces plans obligent à faire beaucoup d'essais.

2.2.3 Plans de Doehlert

La caractéristique principale des plans de Doehlert [12,1] est d'avoir une répartition uniforme des points expérimentaux dans l'espace expérimental. C'est une démarche différente de celle que nous avons rencontrée jusqu'ici. On ne cherche pas le meilleur emplacement des points expérimentaux pour un modèle donné. On répartit simplement les points de manière régulière sans se poser de question sur le critère d'optimalité.

2.2.3.1 Plans de Doehlert pour deux facteurs

Pour deux facteurs les points expérimentaux forment un hexagone régulier (Figure 2.7). Tous les points sont à la même distance du centre du domaine d'étude.

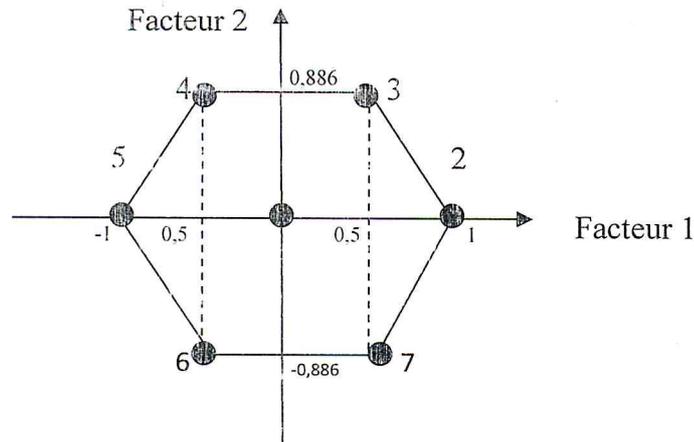


Figure 2.7 : Plan de Doehlert pour l'étude de deux facteurs

La matrice d'expériences correspondante à ce plan est donnée par le tableau 2.10.

Tableau 2.14 : Matrice d'expériences d'un plan de Doehlert pour 2 facteurs

Essais N°	Facteur 1	Facteur 2
1	0	0
2	1	0
3	0,5	0,866
4	-0,5	0,866
5	-1	0
6	-0,5	-0,866
7	0,5	-0,866

La disposition des points expérimentaux selon la figure 2.7 conduit à cinq niveaux pour le premier facteur x_1 et trois niveaux pour le deuxième facteur x_2 . Tous les points du plan de Doehlert sont sur un cercle de rayon unité (en grandeurs centrées réduites). Le domaine défini par les plans Doehlert est un domaine sphérique, un cercle pour deux facteurs, une sphère pour 3 facteurs et une hyper sphère pour plus de 3 facteurs.

2.2.3.2 Modèle mathématique postulé

Le modèle mathématique postulé des plans de Doehlert est, en général, un modèle du second degré avec interactions d'ordre 2. Pour 2 facteurs, on a :

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2$$

2.2.3.3 Plan de Doehlert pour k facteurs

Les plans de Doehlert sont utilisables pour un nombre quelconque de facteurs. Nous indiquons sur un seul tableau les plans Doehlert jusqu'à 4 facteurs. Le plan à 4 facteurs comprend les essais de 1 à 21. Le tableau 2.15 montre bien que le premier plan de Doehlert (essais 1 à 7) est au niveau zéro du troisième facteur. On peut donc traiter deux facteurs en réalisant les essais 1 à 7. Puis, si cela est nécessaire effectuer les essais 8 à 13 pour étudier le troisième facteur. Il en sera de même pour le quatrième facteur. Cette possibilité de conduite séquentielle d'une étude est fort appréciable dans certains cas. Nous n'avons signalé qu'un seul point central, mais rien n'empêche, d'en réaliser plusieurs.

Tableau 2.15 : Plans de Doehlert pour deux à quatre facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4
1	0	0	0	0
2	1	0	0	0
3	0,5	0,866	0	0
4	-1	0,866	0	0
5	-0,5	0	0	0
6	0,5	-0,866	0	0
7	0	-0,866	0	0
8	0,5	0,289	0,816	0
9	-0,5	0,289	0,816	0
10	0	-0,577	0,816	0
11	0,5	-0,289	-0,816	0
12	-0,5	-0,289	-0,816	0
13	0	0,577	-0,816	0
14	0,5	0,289	0,204	0,791
15	-0,5	0,289	0,204	0,791
16	0	-0,577	0,204	0,791
17	0	0	-0,612	0,791
18	0,5	-0,289	-0,204	-0,791
19	-0,5	-0,289	-0,204	-0,791
20	0	0,577	-0,204	-0,791
21	0	0	0,612	-0,781
22	0	0	0	0

2.2.3.4 Avantage et inconvénients

La multiplicité des niveaux dans les plans de Doehlert ne facilite pas la construction du plan d'expériences. Lorsqu'on passe à quatre facteurs et plus il devient difficile de définir même les niveaux de chacun des facteurs. Il faut alors avoir recours à des matrices préparées à l'avance. Les plans de Doehlert permettent d'obtenir un modèle du second degré avec un minimum d'essais. Si on les compare aux plans composites on constate un

léger gain pour deux facteurs (sept essais) alors qu'il faut neuf essais avec un plan composite pour explorer le domaine expérimental et calculer les coefficients du modèle. Dans le cas où les essais sont longs et coûteux, l'économie de deux expériences peut être avantageuse.

2.2.4 Plans de Roquemore

Les plans d'expériences proposés par Roquemore en 1976 [20] sont conçus pour des surfaces du second degré. Ils essayent de respecter à la fois le critère d'isovariance par rotation et le critère de presque orthogonalité. C'est la raison pour laquelle on les appelle les plans hybrides. Il est impossible de respecter simultanément ces deux critères. Ces plans sont presque saturés. Ils sont donc utiles lorsqu'on veut faire le moins d'essais possible. Ces plans existent pour trois, quatre et six facteurs et portent des noms particuliers. Ces noms sont constitués de trois indicateurs :

1. Le nombre de facteurs
2. Le nombre d'expériences
3. Une lettre pour différencier deux plans ayant le même nombre d'expériences et construits pour traiter le même nombre de facteurs

Pour trois facteurs, il y a deux plans de Roquemore, le 311A et le 311B. Ils permettent d'étudier 3 facteurs en 11 essais. Pour quatre facteurs il y a deux plans de Roquemore le 416A et le 416B. Pour 6 facteurs, il y a un plan de Roquemore, le 628A. Nous étudierons en détail le plan 311A et nous indiquerons simplement les autres plans sans les analyser.

2.2.4.1 Plan 311A de Roquemore

Comme son nom l'indique, ce plan permet d'étudier 3 facteurs en 11 essais.

➤ Matrice d'expériences

Les plans expérimentaux sont situés comme l'indique le tableau 2.11.

Tableau 2.16 : Plans de Roquemore 311A pour trois facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 2
1	0	0	2
2	0	0	-2
3	-1,414	-1,414	1
4	1,414	-1,414	1
5	-1,414	1,414	1
6	1,414	1,414	1
7	-2	0	-1
8	2	0	-1
9	0	-2	-1
10	0	2	-1
11	0	0	0

➤ **Emplacement des points expérimentaux**

L'examen du tableau 2.16 permet de distinguer les éléments suivants :

- Un plan 2^2 est constitué de quatre points situés aux sommets d'un carré (essais 3, 4, 5 et 6). Ces essais sont au niveau +1 pour le facteur 3.
- Quatre points situés aux sommets d'un carré (essais 7, 8, 9 et 10) décalé de 45° par rapport au carré précédent. Ces points sont au niveau -1 du facteur 3.
- Trois points situés sur un axe passant par les centres des deux carrés précédents, le premier point est situé au niveau 0 du facteur 3, c'est le point central. Les deux autres points sont respectivement situés au niveau +2 et -2 pour le troisième facteur (figure 2. 8).

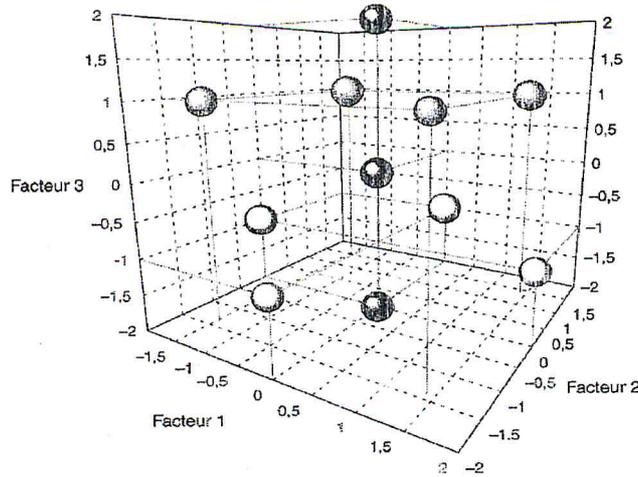


Figure 2.8 : Emplacement des points expérimentaux du plan de Roquemore

➤ **Modèle mathématique**

Le modèle mathématique postulé des plans de Roquemore est un modèle du second degré avec interactions d'ordre deux. Pour trois facteurs on a :

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

➤ **Matrice de calcul**

Pour le plan de Roquemore 311A, la matrice de calcul X est une matrice (10,10) puisqu'il y a 10 coefficients dans le modèle postulé et que l'on effectue 11 essais. Ce plan est donc presque saturé.

➤ **Matrice d'information**

La matrice d'information est une matrice (10,10)

$${}'XX = \begin{pmatrix} 11 & & & & & & & & & & 16 & 16 & 16 \\ & 16 & & & & & & & & & & & \\ & & 16 & & & & & & & & & & \\ & & & 16 & & & & & & & & & \\ & & & & 16 & & & & & & & & \\ & & & & & 16 & & & & & & & \\ & & & & & & 16 & & & & & & \\ & & & & & & & 16 & & & & & \\ 16 & & & & & & & & 48 & 16 & 16 & & \\ 16 & & & & & & & & 16 & 48 & 16 & & \\ 16 & & & & & & & & 16 & 16 & 40 & & \end{pmatrix}$$

L'isovariance par rotation est respectée pour les facteurs 1 et 2 car $48=3 \times 16$. Elle ne l'est pas pour le facteur 3 ($40 \neq 3 \times 16$).

➤ **Matrice de dispersion**

La matrice de dispersion est également une matrice (10,10). Elle va nous servir à calculer les variances des coefficients.

$$\begin{pmatrix} 1 & & & & & & -0,188 & -0,188 & -0,25 \\ & 0,0625 & & & & & & & \\ & & 0,0625 & & & & & & \\ & & & 0,0625 & & & & & \\ & & & & 0,0625 & & & & \\ & & & & & 0,0625 & & & \\ & & & & & & 0,0625 & & \\ -0,188 & & & & & & & 0,061 & 0,029 & 0,029 \\ -0,188 & & & & & & & & 0,029 & 0,061 & 0,039 \\ -0,25 & & & & & & & & & 0,029 & 0,039 & 0,094 \end{pmatrix}$$

➤ **Critère d'optimalité**

L'examen de la matrice d'information montre que le plan de Roquemore 311A pour 3 facteurs respecte le critère d'isovariance par rotation pour 2 facteurs et ne le respecte pas pour le troisième. L'examen de la matrice de dispersion montre que le plan de Roquemore 311A ne respecte pas le critère de presque orthogonalité. Mais les valeurs des éléments hors diagonale principale sont faibles.

2.2.4.2 Autres plans de Roquemore

Les principaux plans hybrides de Roquemore concernent les plans à trois ou quatre facteurs. Ils existent aussi pour six facteurs. Nous indiquons ces divers plans dans les tableaux suivants : 2.12, 2.13. Les propriétés de ces plans s'obtiennent de la même manière que le plan 311A.

Tableau 2.17 : Plans de Roquemore 311B pour trois facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3
1	0	0	2,449
2	0	0	-2,449
3	-0,751	-2,106	-1
4	0,751	-2,106	1
5	-0,751	2,106	1
6	0,751	2,106	-1
7	-2,106	-0,751	1
8	2,106	-0,751	-1
9	-2,106	0,751	-1
10	2,106	0,751	1
11	0	0	0

Tableau 2.18 : Plans de Roquemore 416B pour quatre facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4
1	0	0	0	1,732
2	0	0	0	-0,269
3	-1	-1	-1	0,605
4	1	-1	-1	0,605
5	-1	1	-1	0,605
6	1	1	-1	0,605
7	-1	-1	1	0,605
8	1	-1	1	0,605
9	-1	1	1	0,605
10	1	1	1	0,605
11	-1,518	0	0	-1,050
12	1,518	0	0	-1,050
13	0	-1,518	0	-1,050
14	0	1,518	0	-1,050
15	0	0	-1,518	-1,050
16	0	0	1,518	-1,050

Tableau 2.19 : Plan de Roquemore 416C pour quatre facteurs

Essais N°	Facteur 1	Facteur 2	Facteur 3	Facteur 4
1	0	0	0	1,765
2	-1	-1	-1	0,568
3	1	-1	-1	0,568
4	-1	1	-1	0,568
5	1	1	-1	0,568
6	-1	-1	1	0,568
7	1	-1	1	0,568
8	-1	1	1	0,568
9	1	1	1	0,568
10	-1,470	0	0	-1,050
11	1,470	0	0	-1,050
12	0	-1,470	0	-1,050
13	0	1,470	0	-1,050
14	0	0	-1,470	-1,050
15	0	0	1,470	-1,050
16	0	0	0	0

2.2.4.3 Avantages et inconvénients

Avec ces plans on peut étudier 4 facteurs en 16 essais seulement et réaliser un modèle du deuxième degré. Si on les compare aux plans de Doehlert, il faut réaliser 21 essais. L'inconvénient de ces plans est qu'ils existent que pour 3, 4 et 6 facteurs seulement.

CHAPITRE 3

ANALYSE DE LA VARIANCE DANS LES PLANS D EXPERIENCES

L'ingénieur doit très souvent répondre à une question d'importance considérable, qui est : est ce que tel ou tel facteur influence le résultat de mesure? Plus précisément, de telles questions peuvent être : est-ce que telle impureté affecte les caractéristiques mécaniques de tel matériau? Ou bien : est ce que tel paramètre, qu'on a de bonnes raisons physiques d'estimer important, influence réellement les résultats de mesure ?

L'outil pour ce genre d'étude est l'analyse de la variance qui nous permettra de déterminer à partir de quel seuil, un effet peut être considéré comme significatif .

3.1 Calcul des effets des facteurs**3.1.1 Utilisation de la modélisation opérationnelle**

Cette modélisation permet de représenter en clairs tous les effets des facteurs à chaque niveau. On peut estimer le plus simplement possible toutes les valeurs y en fonction de la position de facteur. On détermine aisément de ce fait l'optimum pour y qui peut être soit un minimum soit un maximum .

La modélisation va conduire à calculer l'effet de chaque facteur, par rapport à la moyenne générale de l'ensemble des résultats. Elle se représente sous forme d'une formule composée de la valeur moyenne et des différents facteurs (ou interactions) affectés de leurs niveaux respectifs. Si on veut par exemple étudier deux facteurs A et B, l'écriture symbolique du modèle est la suivante :

$$y_{estimée} = I + (E_{A1}E_{A2})(A) + (E_{B1}E_{B2})(B) \dots \dots \dots (3.1)$$

Avec :

I : Est la moyenne générale

E_{A1} : Est l'effet de A au niveau 1

E_{A2} : Est l'effet de A au niveau 2

E_{B2} : Est l'effet de B au niveau 1

E_{B2} : Est l'effet de B au niveau 2

Qui sont calculés de la manière suivants

E_{A1} =(moyenne des résultats au niveau 1)-moyenne générale

E_{A2} =(moyenne des résultats au niveau 2)-moyenne générale

E_{B1} =(moyenne des résultats au niveau 1)-moyenne générale

E_{B2} =(moyenne des résultats au niveau 2)-moyenne générale

Le modèle 1, est une représentation simple de données, permet au moyen de la méthode d'analyse de la variance de tester statiquement l'influence éventuelle de chaque facteur et interaction sur la caractéristique étudiée.

Expliquant cette modélisation à travers l'exemple suivant :

Prenons par exemple l'étude de la consommation d'un moteur (M_1) par rapport à 2 facteurs A et B :

Soit un moteur M_1 , deux facteurs à étudier A (carburants) et B (système d'injection) les deux facteurs A et B prend deux niveaux A_1 et A_2 pour le facteur A, et B_1 , B_2

Pour le facteur B, et la réponse y est la consommation en grammes. Les résultats du plan d'expériences réalisé sont présentés dans le tableau 3.1 (Il s'agit d'un plan 2^2)

Tableau 3.1 : Résultats Y pour 4 essais

N ^o	A	B	Y
1	-1	-1	$Y_1=100$
2	-1	1	$Y_2=110$
3	1	-1	$Y_3=95$
4	1	1	$Y_4=105$

Les combinaisons possibles sont : A_1B_1 , A_1B_2 , A_2B_1 , A_2B_2 pour chaque combinaison l'essai réalisé donne une résultat Y :

L'effet moyen est :

$$I = \frac{(Y_1 + Y_2 + Y_3 + Y_4)}{4} = \frac{100 + 110 + 95 + 105}{4} = 102.5$$

L'effet du facteur A est obtenu par la formule :

$$\text{l'effet de A} = (Y_3 + Y_4)/2 - (Y_1 + Y_2)/2$$

= niveau le plus haut - le niveau le plus bas

$$\text{Donc l'effet de A} = (95 + 105)/2 - (100 + 110)/2$$

$$= 100 - 105 = -5$$

$$\text{Et l'effet de B} = (Y_2 + Y_4)/2 - (Y_1 + Y_3)/2$$

$$= (110 + 105)/2 - (100 + 95)/2$$

$$= 107.5 - 97.5 = 10$$

L'effet de A au niveau 1

$$E_{A_1} = \frac{(100 + 110)}{2} - 102.5 = +2.5$$

L'effet de A au niveau 2

$$E_{A_2} = \frac{(95 + 105)}{2} - 102.5 = -2.5$$

Effet totale de A = (+2.5 à -2.5)

(décroissant du + vers le -)

L'effet de B au niveau 1

$$E_{B_1} = \frac{(100 + 95)}{2} - 102.5 = -5$$

L'effet de B au niveau 2

$$E_{B_2} = \frac{(110 + 105)}{2} - 102.5 = +5$$

Effet totale de B = (-5 à +5)

(Croissant du - vers le +)

Si on considère que l'optimum est un minimum, le résultat final est obtenu pour $A=2$ et $B=1$, c'est-à-dire on fixe le facteur A au niveau 2, et le facteur B au niveau 1.

Alors

$$Y_{min}=102.5-2.5-5 = 95$$

La modélisation opérationnelle est utilisée systématiquement dans tous les cas que les facteurs soient qualitatifs, la modélisation choisie est universelle car elle permet de calculer aisément la position des facteurs optimisent la réponse y .

3.1.2 Modélisation opérationnelle à l'effet principal et calcul des résidus

Cette modélisation permet de représenter en clair tous les effets des facteurs à chaque niveau, on peut estimer les plus simplement possibles toutes valeurs de Y en fonction de la position des facteurs, prenons par exemple l'étude de consommation d'un moteur M_2 par rapport à deux facteurs A et B, cas d'un plan 2^2 avec les facteurs A (carburant) et B (système d'injection) et leurs niveaux 1 et 2.

Mesurons la consommation Y du moteur M_2 sur une distance donnée à partir des deux facteurs sans modifier tout autre facteur (tableau 3.2), on a quatre combinaisons possibles :

$$A_1B_1, A_1B_2, A_2B_1, A_2B_2$$

Tableau 3.2 : résultat Y pour 4 essais

N°	A	B	Y
1	-1	-1	40
2	1	-1	39
3	-1	1	31
4	1	1	36

Calcul des effets de A et B :

$$\text{L'effet de A} = (39 + 36)/2 - (40 + 31)/2 = 37.5 - 35.5 = +2$$

$$\text{L'effet de B} = (31 + 36)/2 - (40 + 39)/2 = 33.5 - 39.5 = -6$$

La modélisation va conduire à calculer l'effet de chaque facteur, par rapport à la moyenne générale de l'ensemble des résultats, c'est-à-dire calcule des effets au niveau bas et au niveau haut pour chaque facteur

L'écriture symbolique du modèle est la suivante :

$$Y_{\text{estimé}} = I + A + B$$

Avec I : représente la moyenne générale :

$$I = (40 + 39 + 31 + 36)/4 = 36.5$$

L'effet de A au niveau 1

$$E_{A_1} = 35.5 - 36.5 = -1$$

L'effet de A au niveau 2

$$E_{A_2} = 37.5 - 36.5 = +1$$

L'effet de B au niveau 1

$$E_{B_1} = 39.5 - 36.5 = +3$$

L'effet de B au niveau 2

$$E_{B_2} = 33.5 - 36.5 = -3$$

Représentation graphique des effets

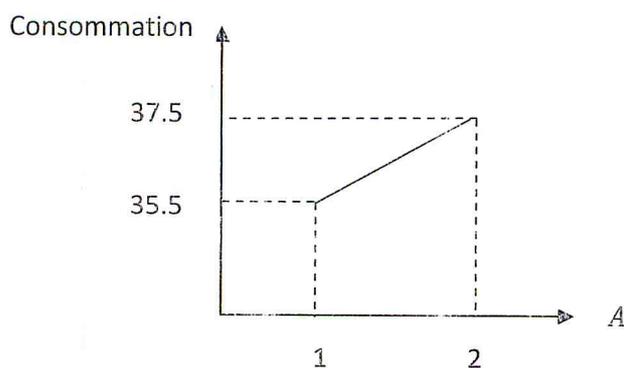


Figure 3.1 : effet de facteur A

Nous constatons que la consommation augmente lorsqu'on fixe le carburant au niveau 2

Nous permet d'aborder la notion d'hypothèse nulle.

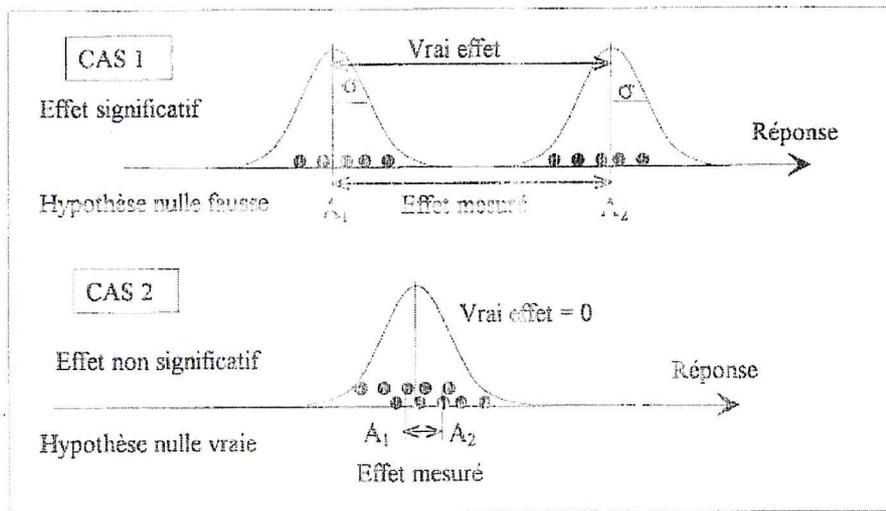


Figure 3.3 : Hypothèse nulle

Niveau -1. Il est probable que toutes ces expériences ne donneront pas le même résultat. On observera une dispersion des résultats autour d'une moyenne. On réalise de même plusieurs expériences lorsque A est au niveau +1.

Notations :

A_1 : Moyenne des résultats lorsque $A = -1$

A_2 : Moyenne des résultats lorsque $A = +1$

σ : Ecart-type des écarts observés autour de la position moyenne.

Dans le premier cas de la figure 3.3, il est évident que l'effet du facteur A est significatif. L'écart observé lorsqu'on passe de niveau -1 au niveau +1 est significativement différent de la dispersion résiduelle. Dans le deuxième cas, l'écart observé n'est pas le fait d'un effet du facteur A . Cet écart provient d'un simple écart dû l'échantillonnage.

On peut schématiser le principe de l'analyse de la variance par la figure 3.2 dans le cas d'un facteur significatif.

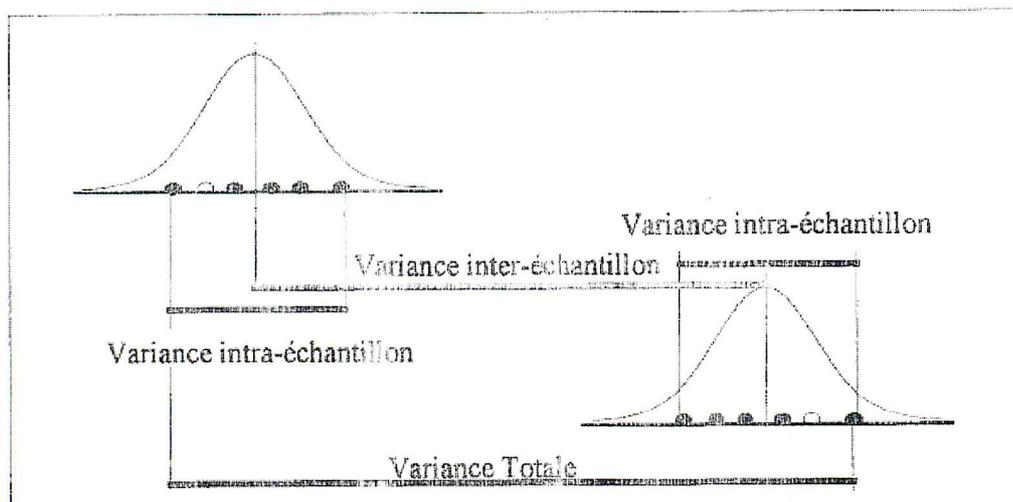


Figure 3.4 : Principe de l'analyse de la variance

La variance totale sur l'ensemble des essais se décompose en deux variances :

- La variance inter-échantillon (V_A) due à l'effet du facteur.
- La variance intra-échantillon (V_R) due à la dispersion sur la réponse.
- Si le facteur n'est pas significatif, la variance inter-échantillon sera proportionnellement supérieure à la variance intra-échantillon. Pour le tester, on établit le rapport : $F = V_A/V_R$.

L'hypothèse nulle consiste à supposer que l'effet de A est nul, le rapport F ne doit donc pas dépasser une certaine limite (F_{limite}).

On dira que A est significatif si F est supérieure à (F_{limite}) fixé par la table de Snedecor.

3.2.3 Cas d'un facteur à k niveaux

Considérons maintenant l'expérience suivante, on réalise k échantillons (nombre de niveaux) de n observations. Cela revient à faire un plan d'expériences à un seul facteur.

Les résultats peuvent être regroupés dans le tableau suivant :

Tableau 3.4 : Un plan d'expérience à un seul facteur

N°	Niveau de A	Réponse 1		Réponse j		Réponse n	Moyenne	Effet
1	A_1							
i	A_i			y_{ij}			\bar{y}_i	a_{Ai}
k	A_k							
moyenne générale							\bar{y}	

Notation :

- y_{ij} : valeur de la $j^{\text{ième}}$ réponse de niveau i
- \bar{y}_i : moyenne des réponses de la ligne i
- \bar{y} : moyenne de l'ensemble des réponses
- N : nombre total de réponse = $k \cdot n$.
- a_{Ai} : effet du facteur A au niveau i

Dans une ligne, le facteur A est toujours au même niveau. Mais, il existe néanmoins une variation de la réponse due aux facteurs non contrôlés.

On peut calculer :

- Pour chaque niveau du facteur A son effet $a_{Ai} = \bar{y}_i - \bar{y}$.
- Pour chaque essai le résidu $r_{ij} = y_{ij} - \bar{y}_i$ qui représente les variations non expliquées par le facteur A .

3.2.3.1 Variance globale

La variance globale est la variance observée en considérant l'ensemble des N essais. Elle se calcule par la formule classique de la variance :

$$V_t = \frac{\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2}{N - 1} = \frac{Q_T}{v_T}$$

Le nombre de degrés de liberté étant égal :

$$v_T = N - 1$$

Cette variance globale se décompose en deux variances :

- La variance résiduelle à l'intérieur des échantillons (V_R)
- La variance entre les échantillons (V_A)

3.2.3.2 Variance résiduelle

La variance à l'intérieur des échantillons (à l'intérieur des lignes dans notre exemple), est l'effet des facteurs non contrôlés. On la notera Variance résiduelle V_R .

$$V_R = \frac{\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2}{k \cdot (n - 1)} = \frac{Q_R}{v_R} = \frac{\sum r_{ij}^2}{V_R}$$

Le nombre de degrés de liberté par ligne étant de $n - 1$, le nombre de ligne étant k , le nombre de degrés de liberté est donc le $k \cdot (n - 1)$.

3.2.3.3 La variance entre échantillons

La variance entre échantillons (entre les lignes) provient du facteur contrôlé. En effet, si le facteur A n'est pas influent, les moyennes de toutes les lignes seront très proches, la variance entre échantillons sera du même ordre de grandeur que la variance résiduelle. En fait, on pourrait montrer que dans ce cas, V_A est une estimation de V_R . Dans le cas contraire, la variance V_A deviendra importante, elle sera significativement différente de la variance résiduelle.

$$V_A = \frac{n \sum_{i=1}^k (\bar{y}_i - \bar{y})^2}{p - 1} = \frac{Q_A}{v_A}$$

Le nombre de degré de liberté v_A est égal au nombre de lignes moins une (nombre de niveaux moins un).

3.2.3.4 Equation fondamentale d'analyse de la variance

On peut montrer que la somme des carrés de la totalité des écarts à la moyenne peut se décomposer de la sorte :

$$Q_T = Q_A + Q_R$$

De même on peut noter que :

$$v_T = v_A + v_R = N - 1$$

Pour vérifier si le facteur A n'a pas influence significative (hypothèse nulle), nous allons comparer les deux estimations de variances V_A et V_R . Pour cela, il faut tester si V_A est du même ordre de grandeur que V_R ou s'il est significativement différent. Nous réaliserons cette comparaison en utilisant le test de Snedecor.

3.2.3.5 Test de Snedecor

Le test de Snedecor est un test qui permet de comparer deux variances. Il est donc parfaitement adapté à notre problème puisque nous cherchons à comparer les variances V_A et V_R . Le rapport $\frac{V_A}{V_R}$ suit une loi de Fisher (loi de Snedecor Figure 3.5) à ν_A et ν_R degrés de liberté.

Cette courbe donne la répartition du rapport entre les variances de deux échantillons issus de populations à variances identiques.

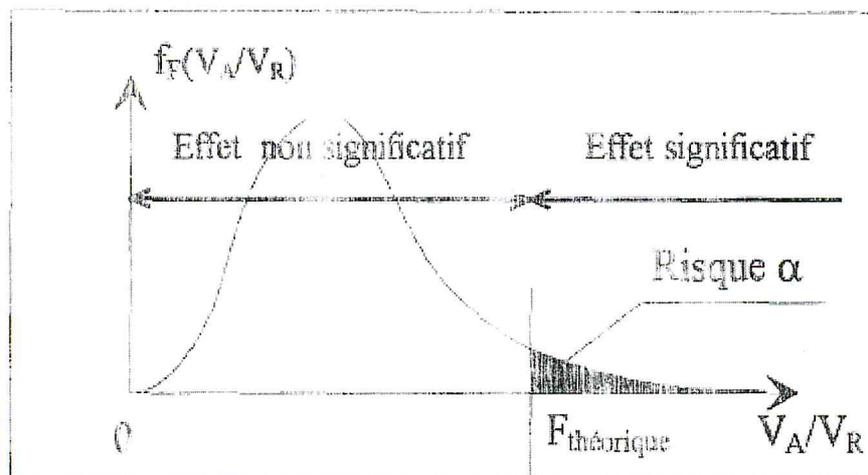


Figure 3.5 : loi de F

En fait dans cette courbe, seule la partie supérieure nous intéresse, car nous cherchons à montrer que V_A est très supérieure à V_R . Le rapport sera donc toujours supérieur à 1 si l'effet de A est significatif.

La courbe nous montre que dans le cadre de l'hypothèse nulle, le rapport peut être égale à l'infini. Cependant, la probabilité est très faible. Pour conclure sur un effet, il faut donc prendre le risque de conclure que cet effet soit significatif alors que nous sommes dans le cadre de l'hypothèse nulle. Nous le noterons risque α . Il est généralement fixé à 5%.

En fonction du risque choisi, la loi de F déterminera un seuil de refus de l'hypothèse nulle que l'on appellera $F_{théorique}$.

Le test consiste donc à comparer le rapport $\frac{V_A}{V_R}$ avec la valeur $F_{théorique}$ que nous trouvons dans le tableau de la loi de F en fonction :

- De v_A et v_R degré de liberté du facteur étudié et de la résiduelle
- Du risque choisit

La valeur $F_{théorique}$ se lit dans le tableau de la loi de Snedecor fournie en annexe. Nous choisissons la colonne correspondant au nombre de degré de liberté de la variance V_A (v_A) et la ligne correspondant à v_R . Les deux tableaux correspondent au risque de 5% .

3.2.2.6 Tableau d'analyse de la variance

Afin de simplifier les calculs et la présentation des résultats, on regroupe les calculs dans un tableau (tableau 3.5) qu'on appelle « tableau d'analyse de la variance »

Tableau 3.5 : Tableau d'analyse de la variance

variation	somme des carrés Q	Degré de liberté	Variance	F_{exp}	$F_{théo}$	Signification
Entre échantillon	$n \sum_{i=1}^k (\bar{y}_i - \bar{y})^2$	$v_A = k - 1$	$V_A = \frac{n \sum_{i=1}^k (\bar{y}_i - \bar{y})^2}{k-1}$	$\frac{V_A}{V_R}$	Pour $v_A = k - 1$	Si $F_{exp} > F_{théo}$
Intérieur échantillon	$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 = \sum_{ij} r_{ij}^2$	$v_R = N - k$	$V_R = \frac{\sum_{ij} r_{ij}^2}{N - k}$		$v_R = N - k$	
Totale	$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y})^2$	$N - 1$				

Avec n : nombre de fois ou le facteur 4 est répété sur un niveau.

3.3 Cas des plans d'expériences sans répétition

Un plan sans répétition est un plan d'expériences qui n'est réalisé qu'une seule fois. La réponse, correspondant à un essai, est donc obtenue à partir d'une seule expérience.

3.3.1 Analyse de la variance

Pour chaque effet de facteur et pour chaque interaction, nous pouvons calculer la variance de ceux-ci. L'analyse de la variance consiste à comparer ces variances à la variance résiduelle grâce au test de Snedecor.

3.3.1.1 Décomposition de la variance totale

3.3.1.1.1 Dispersion totale

Les réponses obtenues au cours des N essais d'un plan d'expériences présentent une dispersion autour de leur moyenne générale \bar{y} caractérisée par la somme des carrés totaux

$$Q_T = \sum_{i=1}^N (y_i - \bar{y})^2 = \sum_{i=1}^N y_i^2 - N\bar{y}^2.$$

Cette dispersion a deux causes principales :

- Les changements de niveaux des facteurs du plan.
- Les erreurs aléatoires.

Nous pouvons donc calculer la variance totale par la formule suivante :

$$V_T = \frac{Q_T}{v_T}$$

Où v_T représente le degré de liberté totale, on le calcule en appliquant la relation suivante :

$$v_T = N - 1$$

3.3.1.1.2 Dispersion expliquée par un facteur

Considérons un facteur A à k niveaux. La dispersion expliquée par A est appelée somme des carrés de A et désignée par Q_A . Elle est égale à la dispersion des k moyennes A_i de A autour de la moyenne générale \bar{y} , multipliée par l'effectif n_A de ces moyennes (n_A est supposé constant et vaut N/k).

$$Q_A = n_A \sum_{i=1}^k (A_i - \bar{y})^2$$

Où

A_i c'est la moyenne des résultats obtenus pour les essais réalisés avec A au niveau i .

CHAPITRE 4

PLANS DE MELANGES

4.1 Introduction

Dans les plans d'expérience classiques, les facteurs doivent être indépendants cela signifie que l'on peut choisir en toute liberté le niveau d'un facteur quelque soient les niveaux déjà attribués aux autres facteurs.

Les plans de mélanges sont également souvent caractérisés par de nombreuses contraintes qui peuvent peser sur le choix des proportions des constituants.

Les plans de mélanges ont d'abord été étudiés par des anglo-saxon (Claringbold, Cornell, Crozier, Marquardt, Quenuille, Scheffé, Snee, etc..) Ce sont eux qui ont introduit la terminologie propre à cette science des plans de mélanges. Nous nous inspirerons de leurs travaux et nous adopterons une terminologie proche de la leur à chaque fois que cela sera possible. Afin de ne pas compliquer la tâche des expérimentateurs français qui seront appelés à lire ou à écrire sur ce sujet, nous évitons, chaque fois que cela sera possible, les transpositions terminologiques inutiles. Le vocabulaire non spécifique aux mélanges est le même que celui des plans d'expériences classiques. La méthodologie de mise en forme de l'étude et la conduite des essais sont également tout à fait comparables à celle des plans d'expériences classiques.

Il faut bien se rendre compte qu'il y a plan de mélanges lorsque la réponse étudiée dépend des proportions des constituants du mélange et non pas des quantités de mélange utilisé.

Nous commencerons par examiner le problème de la non-indépendance des facteurs qui est à la base de la distinction entre les plans de mélanges et les plans d'expériences classiques. La non-indépendance des facteurs est exprimée par la contrainte fondamentale des mélanges [16].

4.2 La contrainte fondamentale des mélanges

Soit un mélange ayant n constituant. Le premier constituant représente un certain pourcentage de mélange. Ainsi, chaque constituant participe pour une certaine part au mélange total. Mais l'ensemble des constituants du mélange forme un tout et la somme de

leur de leurs teneurs est égale à 100%. Lorsque l'on a défini les proportions des $n - 1$ premiers constituants, la proportion du dernier ne peut plus être choisie, elle est déjà déterminée. Si l'on note par x_i la teneur en constituant i , la somme des teneurs de tous les constituants du mélange satisfait à la relation :

$$\sum_{i=1}^n x_i = 100\%$$

La teneur de chaque constituant reste comprise entre 0 et 100%. Lorsqu'on augmente la teneur de l'un des constituants, la teneur des autres est automatiquement diminuée pour que la somme de toutes les teneurs reste égale à 100%. Si au lieu d'utiliser les pourcentages, on ramène la somme des teneurs des différents constituants à l'unité on écrira :

$$\sum_{i=1}^n x_i = 1$$

Cette relation s'appelle la **contrainte fondamentale des mélanges**. C'est à cause de cette contrainte que les mélanges doivent être étudiés à part. Les représentations géométriques des plans de mélanges sont différentes des représentations utilisées pour les plans d'expériences classiques et les modèles mathématiques sont, eux profondément modifiés.

4.3 Représentation géométrique des mélanges

Comme pour les plans d'expériences, on pourrait utiliser les axes cartésiens. Le premier axe serait attribué au premier constituant ou premier facteur pour reprendre la terminologie des plans d'expériences. Le second axe, orthogonal au premier, serait attribué au second constituant ou second facteur, et ainsi de suite. Mais, la contrainte des mélanges va modifier cette représentation. En effet, les teneurs x_i ne sont plus indépendantes et il faut en tenir compte. Voyons d'abord un mélange à deux constituants.

4.3.1 Mélange à deux constituants

Représentation cartésienne, l'axe O_{x_1} étant orthogonal à l'axe O_{x_2} . Les axes sont gradués en proportions variant de 0 à 1. Un mélange quelconque contenant x_a de A et x_b de B est représenté par un point situé à l'intersection des coordonnées x_a et x_b (figure 4.1). Ce point qui représente un mélange est appelé soit **point de composition**, soit **point de mélange**, ou tout simplement point [13].

La contrainte des mélanges introduit une relation entre x_a et x_b :

$$x_a + x_b = 1$$

Que l'on peut écrire :

$$x_b = 1 - x_a$$

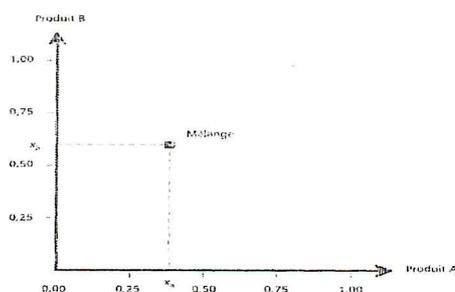


Figure 4.1 : Représentation d'un mélange dans un système d'axes cartésiens

Cette dernière relation n'exprime que les points de coordonnées x_a et x_b sont sur une droite de pente -1, coupant l'axe du produit A au point d'abscisse 1 (point A sur la figure 4.2) et l'axe du produit B au point d'ordonnée 1 (point B sur la figure 4.2). Les compositions coordonnées x_a et x_b variant entre 0 et 1, il n'y a que le segment AB qui soit utile. Toutes les compositions possibles des mélanges des deux produits A et B sont représentées par les points de ce segment de droite.

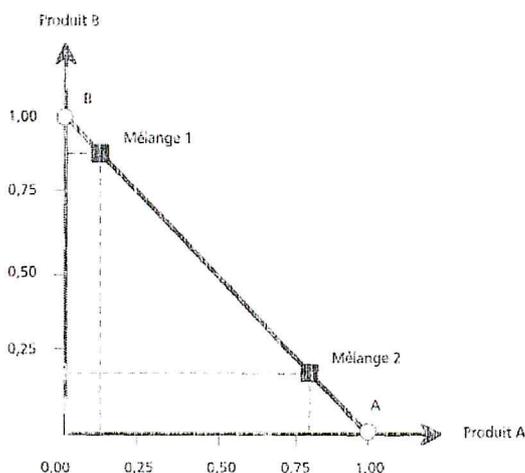


Figure 4.2 : Les compositions des mélanges à deux constituants peuvent être représentées par les points du segment de droite AB

On peut donc ne conserver que ce segment de droite et abandonner les axes O_{x_1} et O_{x_2} . Le produit pur A est représenté par l'une des extrémités du segment AB . Le produit B est représenté par l'autre extrémité. Ce segment porte une double graduation (figure 4.3), celle des teneurs du premier produit et celle du second produit. La lecture de la composition des mélanges sur ce segment demande un certain apprentissage. Elle se lit à la fois de droite à gauche et de gauche à droite.

4.3.2 Lecture du diagramme des mélanges

Le segment de la figure 4.3 représente les mélanges des deux constituants A et B . Le produit A est à droite et le produit B est à gauche. La graduation inférieure est celle du produit A . Elle varie de 0 à gauche à 1 à droite. La graduation supérieure est celle du produit B . Elle varie de 0 à droite à 1 à gauche.

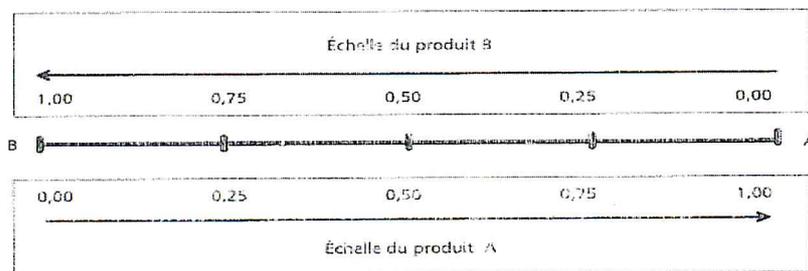


Figure 4.3 : Représentation des mélanges à deux constituants sur un segment de droite

Les échelles se lisent, l'une dans un sens, l'autre dans l'autre sens. La somme des teneurs, en un point de l'échelle, est égale à l'unité (figure 4.4)

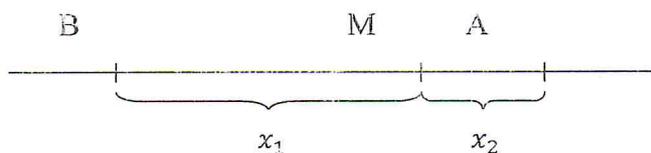


Figure 4.4 : Représentation de mélange M sur un segment de droite.

Tout point M du segment AB est représentatif d'un mélange binaire :

-en A le constituant A est seul dans le mélange.

-en B le constituant B est seul dans le mélange.

-en M , x_1 , la proportion de A , et x_2 , la proportion de B , sont telles que :

$$x_1 = \frac{MB}{AB} \quad \text{et} \quad X_2 = \frac{MA}{AB}$$

4.3.3 Mélange à trois constituants

Trois axes orthogonaux entre eux représentent chacun les teneurs d'un produit dans le mélange (figure 3.5). le point ayant pour coordonnée x_a, x_b et x_c représente le mélange ayant une teneur x_a du produit A, une teneur x_b du produit B et une teneur x_c du produit C.

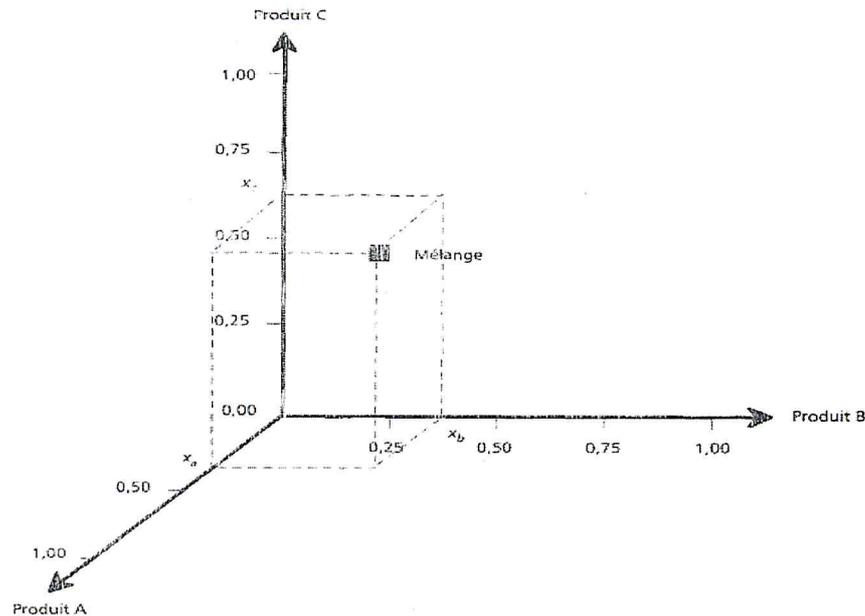


Figure 4.5 : Représentation dans un système d'axes cartésiens d'un mélange de trois produits

La contrainte des mélanges introduit la relation :

$$x_a + x_b + x_c = 1$$

Qui exprime que les points de coordonnées x_a, x_b et x_c sont sur un plan passant par les trois points d'abscisse 1 sur les axes de coordonnées. Les teneurs x_a, x_b et x_c varient entre 0 et 1, par conséquent le domaine des mélanges se limite à un triangle équilatéral ayant pour sommets les trois points d'abscisses 1 (figure 4.6). C'est ce triangle que l'on utilise pour représenter les mélanges à trois composants. Là aussi, il faut s'habituer à lire la composition des mélanges ternaires sur ce diagramme triangulaire.

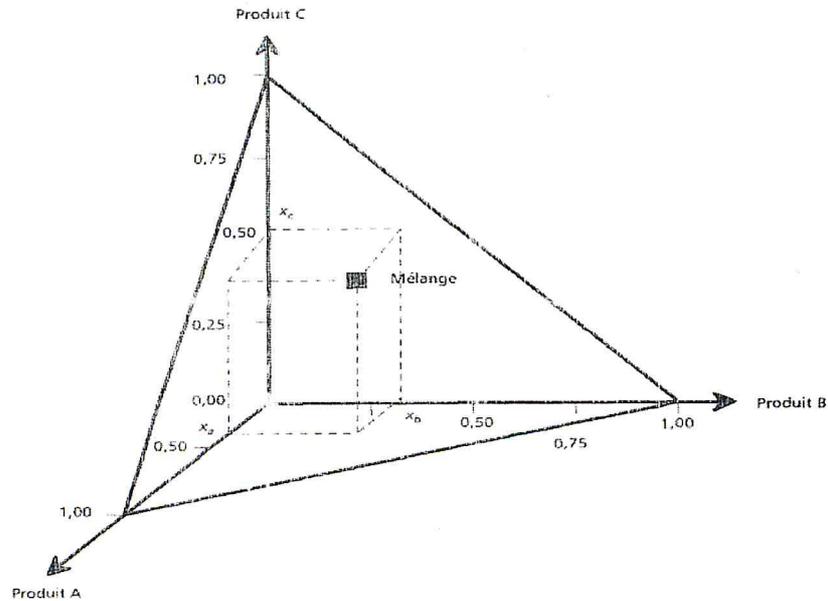


Figure 4.6 : Représentation de l'ensemble des mélanges à trois constituants dans l'espace cartésien des trois produits A , B et C

4.3.4 Lecture du diagramme des mélanges ternaires

Les produits purs sont aux sommets du triangle équilatéral. Les mélanges binaires sont représentés par les côtés du triangle. Par exemple le côté gauche AB du triangle (figure 3.7) représente les compositions des mélanges des seuls produits A et B , et ces mélanges ne contiennent pas de produit C .

Chaque côté du triangle équilatéral est gradué de 0 à 1. On n'a laissé que l'échelle d'un seul produit sur chaque côté. L'échelle de l'autre produit se déduit par complément à 1. Le côté AB est gradué pour le produit A : 1 est au point A et 0 au point B . C'est l'échelle de produit A . Le côté BC est gradué pour le produit B : 1 est au point B et 0 est au point C . Le côté AC est gradué pour le produit C : 1 est au point C et 0 est au point A .

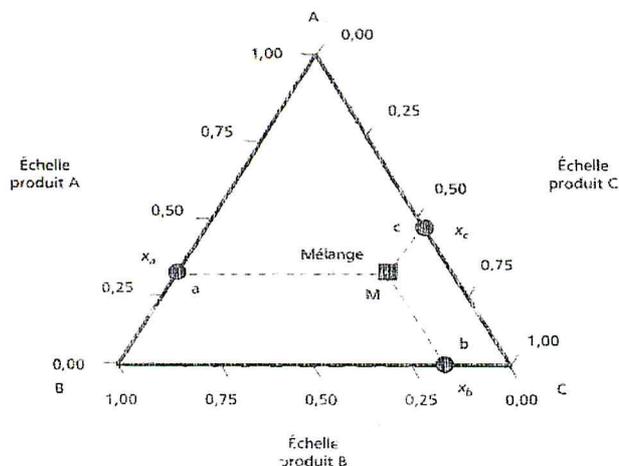


Figure 4.7 : Représentation des mélanges à trois constituants à l'aide d'un triangle équilatéral

Un point de la surface intérieure du triangle équilatéral représente un mélange ternaire. Les teneurs de chaque produit se lisent sur les côtés du triangle. La teneur en produit A se lit sur le côté AB qui est l'échelle du produit A. On projette le point M sur AB parallèlement à BC (côté opposé à A). De même la teneur en B se lit sur l'échelle de B (côté BC) en projetant le point M sur BC parallèlement à AC (côté opposé à B). Enfin la teneur en C se lit sur l'échelle de C (côté AC) en projetant le point M sur AC parallèlement à AB (côté opposé à C).

Les teneurs respectives de trois produits dans le mélange sont données par les relations :

$$bM = Ba = \text{teneur en A dans le mélange } M$$

$$cM = Cb = \text{teneur en B dans le mélange } M$$

$$aM = Ac = \text{teneur en C dans le mélange } M$$

Les propriétés géométriques du triangle équilatéral assurent le respect de la contrainte fondamentale des mélanges :

$$Ac + Ba + Cb = AB = AC = BC = 1$$

Remarque :

Trois points définissent une figure géométrique à deux dimensions, qui est appelée «simplex» par les Anglo-Saxons. Ici, le triangle ABC (deux dimensions) peut être considéré comme un simplex (trois dimensions). De même, une figure géométrique à trois

dimensions définir par quatre points, non dans un même plan, est appelée un simplexe. D'une manière générale, la figure géométrique à k dimensions définie par $k + 1$ points sera appelée un simplexe dans la suite de l'ouvrage. Quand les points du simplexe sont équidistants les uns des autres, le simplexe est dit régulier. Par exemple, les sommets d'un triangle équilatéral forment un simplexe régulier.

4.3.5 Mélanges à quatre constituants

Les quatre produits purs sont aux sommets d'un tétraèdre régulier (figure 4.8). Le tétraèdre est un simplexe puisqu'il est défini par quatre points dans un espace à trois dimensions.

Les mélanges binaires sont représentés par les cotés du tétraèdre. Les mélanges ternaires sont représentés par les faces du tétraèdre qui sont des triangles équilatéraux. Les mélanges quaternaires sont représentés par les points du volume intérieur du tétraèdre. On obtient les compositions d'un mélange en projetant le point représentatif du mélange sur les faces et sur les cotés du tétraèdre.

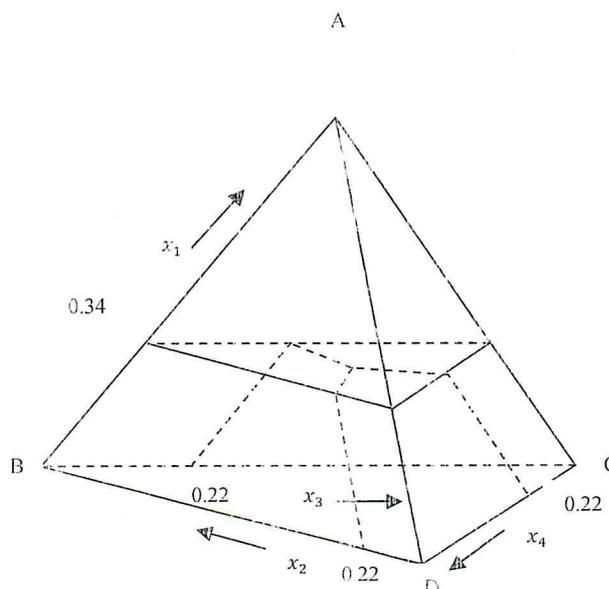


Figure 4.8 : Représentation des mélanges à quatre constituants par un tétraèdre régulier

Afin de ne pas commettre d'erreur dans la lecture des compositions, on fera attention aux conventions qui ont été adoptées pour graduer les cotés du tétraèdre.

4.3.6 Mélange à cinq constituants

La représentation géométrique n'est plus possible car il faudrait dessiner un volume régulier à quatre dimensions. Nous allons donc recourir à une autre représentation pour illustrer tous les mélanges quel que soit le nombre de constituants : les tableaux.

4.4 Représentation des mélanges par des tableaux

Comme pour les plans d'expériences, on associe à la représentation géométrique une représentation en tableau. Cette représentation permet de représenter tous les mélanges quel que soit le nombre de constituants. Contrairement à la représentation géométrique, il n'y a pas de limite et l'on peut représenter un mélange à m constituants, m étant aussi grand que l'on désire.

4.4.1 Mélange à deux constituants

Un mélange à deux constituants est représenté par un tableau (tableau 4.1) qui comprend trois colonnes :

- la première colonne contient le nom des différents mélanges. Nous avons adopté une numérotation selon l'ordre des nombres entiers. On pourrait également prendre des lettres ou des noms quelconques.
- La seconde colonne contient les proportions du produit A.
- La troisième colonne contient les proportions du produit B.

Le tableau comprend autant de lignes qu'il y a de mélanges à étudier. Ici nous avons indiqué trois mélanges. La somme des proportions d'une ligne quelconque est toujours égale à l'unité puisqu'il s'agit d'un mélange donné. Cette propriété est la conséquence de la contrainte fondamentale des mélanges.

Tableau 4.1 : Représentation en tableau d'un mélange à deux constituants

N° du mélange	Produit A	Produit B
1	$X_{a,1}$	$X_{b,1}$
2	$X_{a,2}$	$X_{b,1}$
3	$X_{a,3}$	$X_{b,1}$

4.4.2 Mélanges à trois constituants

Le tableau (4.2) représente un mélange ternaire. Il possède une colonne de plus que le tableau (4.1) puisqu'il y a un constituant supplémentaire. Là aussi, la somme des proportions de chaque ligne est égale à 1 pour toutes les lignes.

Tableau 4.2 : Représentation en tableau d'un mélange à trois constituants

N° du mélange	Produit A	Produit B	Produit C
1	$X_{a,1}$	$X_{b,1}$	$X_{c,1}$
2	$X_{a,2}$	$X_{b,2}$	$X_{c,2}$
3	$X_{a,3}$	$X_{b,3}$	$X_{c,3}$

4.4.3 Mélange à m constituants

Un tableau représentant une série de mélanges à m constituants possède $m + 1$ colonnes et autant de lignes que nécessaire. La somme des proportions des constituants d'un mélange est toujours égale à 1. Dès que $m \geq 5$, la représentation en tableau est la seule possible.

4.5 Modèle mathématique pour l'interprétation des mélanges

4.5.1 Modèle du premier degré

Parmi les modèles mathématiques possibles, l'un des plus simples est celui du premier degré. On suppose que les variations de la réponse sont proportionnelles aux teneurs des constituants du mélange. On pourrait adopter, en premier lieu, le même modèle que celui des plans factoriels classiques dont les facteurs sont indépendants.

Dans le cas particulier des mélanges, nous avons toujours la relation [16] :

$$\sum_i x_i = 1$$

Qui traduit que la somme des compositions de tous les constituants du mélange représente la totalité de ce mélange. De ce fait, le modèle du 1^{er} degré s'écrit :

$$y = a_1 x_1 + a_2 x_2 + \dots + a_k x_k$$

Un modèle de mélange du 1^{er} degré n'a pas de constante. Vérifions le modèle linéaire avec constante :

$$y = b_0(x_1 + x_2 + \dots + x_k) + b_1 x_1 + b_2 x_2 + \dots + b_k x_k$$

Qui est bien de la forme annoncée un peu plus haut avec :

Il est plus courant d'utiliser le modèle simplifié qui ne comprend que les termes du premier degré, les termes rectangles et le produit des trois constituants. Le modèle simplifié, appelé modèle cubique restreint, est le suivant :

$$y = \sum_i a_i x_i + \sum_{i < j} a_{ij} x_i x_j + \sum_{i < j < k} a_{ijk} x_i x_j x_k$$

Et le nombre de coefficients du modèle à déterminer est égal : $\frac{k(k^2+5)}{6}$

4.6 Différent types de plans de mélanges

Les plans de mélanges permettent de choisir les mélanges à tester de façon à avoir les meilleures estimations des coefficients a_i, a_{ij}, \dots , ce sont des plans optimaux. Nous distinguons différents types de plans selon les contraintes aux quelles sont soumises les compositions x_i [15].

4.6.1 Plan de type I

Les plans de mélanges de type I supposent que les produits purs ont la propriété d'usage recherchée et qu'il n'y a aucune contrainte qui pèse sur eux. L'usage est de classer les principaux types de plans de mélange suivant l'emplacement des points représentatifs des compositions. Ces points sont repartis dans le domaine d'étude selon des dispositions particulières dont nous allons passer en revue les principales :

- Les plans en réseaux.
- Les plans de mélanges centrés.
- Les plans de mélanges centrés augmentés.

4.6.1.1 Plans en réseaux

Les plans en réseaux (simplex lattice design dans la littérature anglo-saxonne) sont les premiers plans de mélanges qui ont été introduits par Schéffe dans les années 1958-1965. Les points de composition sont régulièrement répartis dans l'espace d'étude. Pour les mélanges à trois constituants, le plus simple des plans de mélanges en réseaux ne contient que les trois produits purs. Les points représentatifs de ces produits sont aux sommets du triangle équilatéral (figure 4.9). Les coordonnées de ces points sont 0 et 1.

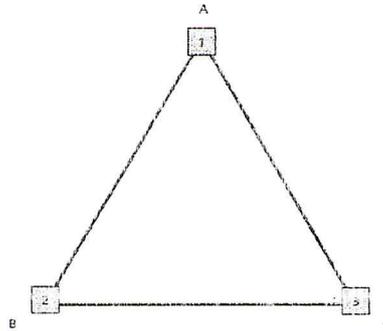


Figure 4.9 : plan de mélanges ne comportant que les trois produits purs

Les plans qui ne font intervenir que les produits purs peuvent titrer enrichi par les points de composition moyenne, c'est-à-dire par les mélanges contenant 50% de chacun des produits. Ces points de composition moyenne sont situés au milieu des côtes du triangle. Leurs coordonnées sont 0 et $1/2$. On remarquera sur la figure 4.9 que ces points de composition moyenne se répartissent de manière symétrique par rapport aux côtés du triangle. Ils forment un réseau régulier, d'où le nom de ces plans.

Ce réseau est caractérisé par la répartition des points dans l'espace expérimental, c'est-à-dire par leurs coordonnées. La longueur unité du côté a été divisée par deux, le réseau a un pas de $1/2$. Si l'on veut mettre un réseau de points plus serré. Il faut diviser la longueur unité des côtes par trois afin d'obtenir des points régulièrement répartis. Le pas de ce nouveau réseau est $1/3$. Les coordonnées des points sont 0, $1/3$, et $2/3$. On peut continuer à construire ainsi des réseaux de plus en plus serrés en divisant l'unité par m . le pas du réseau est $1/m$ et les coordonnées des points expérimentaux sont 0, $1/m$, $2/m$, $3/m$ jusqu'à $m/m = 1$. Si k est un nombre de la suite des entiers : 1, 2, 3, ...etc. les coordonnées des points des plans en réseaux s'écrivent : $\frac{k}{m}$

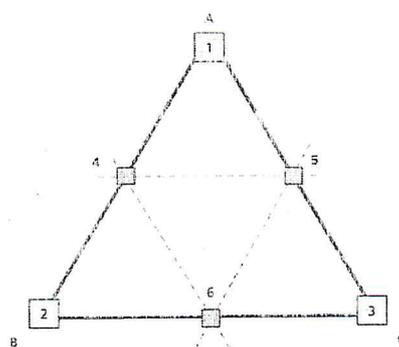


Figure 4.10 : plan de mélanges en réseaux (simplex lattice design) comportant les trois produits purs et les mélanges moitié-moitié. Ce plan est noté $\{3,2\}$

Pour nommer ces plans, on adopte la convention suivante : on utilise deux chiffres, le premier correspond au diviseur utilisé pour établir le pas du réseau. Ces deux chiffres sont séparés par une virgule et mis entre deux accolades. Ici le plan de la figure 4.9 est un plan $\{3,2\}$, soit 3 constituants et un pas de $1/2$. Le plan de la figure 3.10 est un plan $\{3,3\}$, soit 3 constituants et un pas de $1/3$. Si un plan de mélanges fait intervenir 5 constituants et que l'expérimentateur décide un pas de $1/4$, il s'agit d'un $\{5,4\}$.

Cette notation permet de calculer très rapidement le nombre de mélanges à préparer et à étudier connaissant le nombre de constituants du mélange et le nombre m . Pour un plan $\{q, m\}$. Le nombre total de mélange différents est égal à :

$$C_{q+m-1}^m = \frac{(q+m-1)!}{(m)!(q-1)!}$$

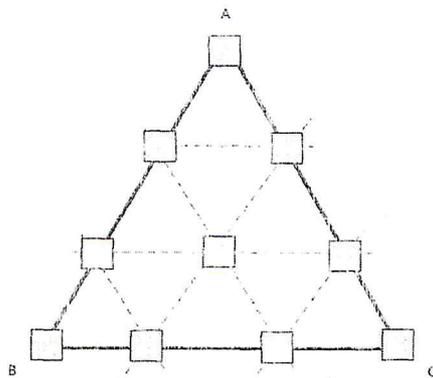


Figure 4.11 : Plan de mélanges en réseaux comportant les trois produits purs et les mélanges $1/3$ - $2/3$ et $1/3$ - $1/3$ - $1/3$. Ce plan est noté $\{3,3\}$.

Le plan de mélange en réseaux $\{3,3\}$ de la figure 4.11 possède 10 mélanges différents à étudier, pour $m = 3$, $q = 3$ on a :

$$C_{q+m-1}^m = \frac{(q+m-1)!}{(m)!(q-1)!} = \frac{(3+3-1)!}{(3)!(3-1)!} = 10$$

4.6.1.2 Plan de mélanges centrés

Les plans de mélanges centrés (simplex-centroid design dans la littérature anglo-saxonne) se distinguent des plans en réseaux par la présence systématique d'un point central, c'est-à-dire d'une composition contenant autant de chacun des constituants du mélange. Le plan centre permettant d'étudier trois constituants comprend :

- Les produits purs.
- Les mélanges moitié-moitié de produits purs.
- Le mélange contenant un tiers de chaque produit pur (point central).

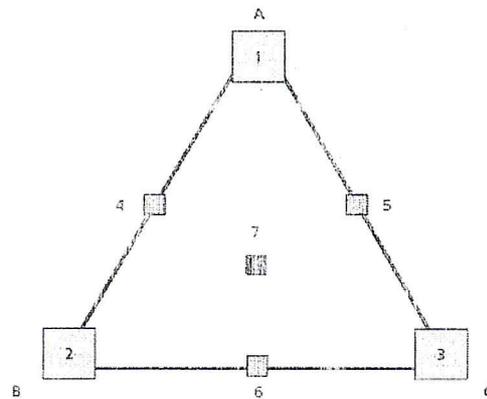


Figure 4.12 : Plan de mélanges centré comportant les trois produits purs, les mélanges moitié-moitié et le mélange équiproportion des trois produits purs

Le nombre N de mélanges à étudier pour des plans de mélanges centré à q constituants est donné par la formule :

$$N = 2^q - 1 = 8 - 1 = 7$$

4.6.1.3 Plans de mélanges centrés augmentés

Les plans de mélanges centrés augmentés sont des plans de mélanges centrés auxquels on ajoute les centres de gravité des simplex unitaires. Pour le plan de mélanges centré à trois composants comprend quatre simplex unitaires (figure 4.13). Le centre de gravité de l'un des simplex unitaires est déjà occupé (point noir, numéro 7). Il reste donc à ajouter trois points (points blancs numéro 8, 9 et 10) au centre des trois autres simplex unitaires.

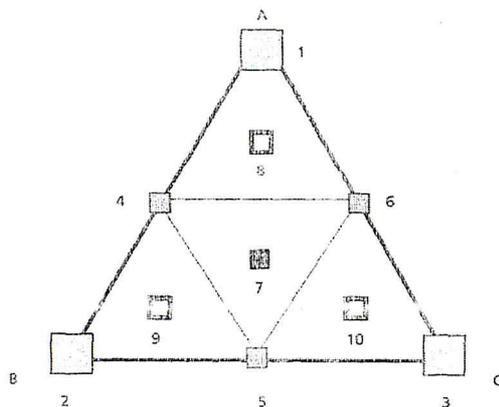


Figure 4.13 : plan de mélanges centré augmenté comportant les trois produits purs, les mélanges moitié-moitié, le mélange équiproportion des trois produits et les mélanges situés aux centre de gravité des simplex unitaires.

4.6.2 Plan de type II

Les plans de mélanges de type II sont tel que : $x_i \geq l_i$ c'est-à-dire les teneurs des constituants ne peuvent pas descendre au-dessous d'une certaine limite l_i . Donc le domaine des compositions est alors divisé en deux zones : la zone interdite au-dessous de la limite l_i et la zone autorisée au-dessus de cette limite. Cette dernière conserve la forme du domaine initiale, un triangle équilatéral s'il y a trois constituants, un tétraèdre régulier s'il y a quatre constituants et un hyper polyèdre régulier s'il y a plus de quatre constituants. Le domaine des mélanges possibles (figure 4.14) découle des relations suivantes :

$$\sum_i x_i = 1$$

$$x_1 \geq L_1$$

$$x_2 \geq L_2$$

$$x_3 \geq L_3$$

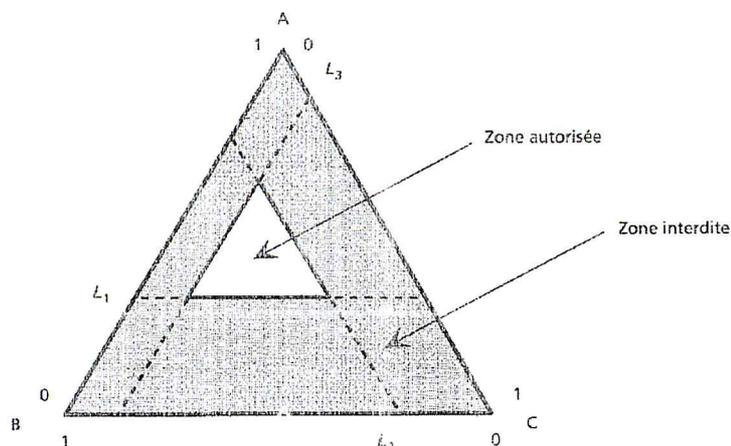


Figure 4.14 : les limites basses de tous les produits définissent deux zones :

La zone interdite (zone ombrée) et la zone autorisée qui a la même forme géométrique que le domaine initiale.

4.6.2.1 L'existence du domaine expérimental

On considère un mélange ne comportant que deux constituants A et B. On suppose que la limite basse du produit A est égale à L_A et que celle du produit B est égale à L_B . On envisage les différents cas suivants [13].

1. Les limites basses L_A et L_B sont chacune égale à 0. Toutes les proportions de A et de B sont alors possibles dans le mélange de ces deux produits (figure 4.15). Dans ce cas, on a : $L_A + L_B = 0$



Figure 4.15 : La somme des limites basses de tous les constituants égale à 0.

2. La limite basse L_A est plus grand que 0. Les teneurs en A varient alors de L_A à 1 et celles de B de zéro à $1 - L_A$ (figure 4.16). Dans ce cas, on a :

$$L_A + L_B < 1$$



Figure 4.16 : La somme des limites basses de tous les constituants.

3. Les deux limites basses L_A et L_B sont plus grandes que 0. Les proportions de A varient de L_A à $1 - L_B$ et celles de B de L_B à $1 - L_A$ (figure 4.17). Dans ce cas, on a :

$$L_A + L_B < 1$$



Figure 4.17 : La somme des limites basses de tous les constituants.

4. Les deux limites basses vérifient la relation $L_A + L_B = 1$. Dans ce cas, il n'y a plus qu'un seul mélange possible. La concentration en A est L_A et celle de B est L_B (figure 4.18)



Figure 4.18 : La somme des limites basses de tous les constituants.

5. Les deux limites basses vérifient la relation $L_A + L_B > 1$, il n'y a plus de mélanges possible. Cette analyse permet de conclure que, dans le cas d'un mélange binaire, les mélanges sont possibles si les limites basses respectent la relation : $L_A + L_B \leq 1$

Dans le cas d'un mélange ternaire (produits A, B et C), si les limites basses des deux constituants vérifient $L_A + L_B < 1$, les mélanges de A et B sont possibles.

Les proportions du troisième constituant sont alors comprises entre zéro et $1 - (L_A + L_B)$. Si la limite basse de C atteint la valeur de $1 - L_A - L_B$, il n'y a plus de mélange possible. Il faut donc que :

$$L_A + L_B + L_C \leq 1$$

D'une manière générale, pour que des mélanges soient possibles entre k constituants, les limites basses doivent respecter la relation suivante :

$$\sum L_i \leq 1$$

4.6.2.2 Calcul de la composition des mélanges expérimentaux

Soit un premier constituant A du mélange dont la teneur varie de 1 (produit pur) à L_1 (limite inférieure), un deuxième constituant B dont la teneur varie de 1 à L_2 et un troisième constituant c dont la teneur varie de 1 à L_3 . Les points A, B et C représentent les

corps purs (s'il y a plus de trois constituants, il faut étendre les raisonnements au tétraèdre puis aux hyperpolyèdres). Dans le cas de trois constituants, le domaine d'étude est ramené au triangle PQR (figure 4.19). Les points P, Q et R représentent des mélanges particuliers que l'on appelle pseudo corps purs.

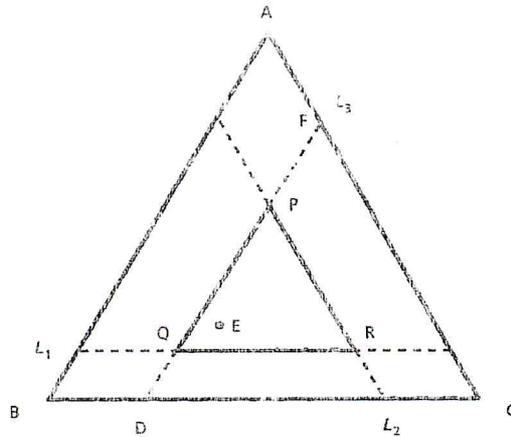


Figure 4.19 : Le triangle PQR représente le domaine d'étude restreint. Le point E est un mélange dont la composition peut être exprimée avec les corps purs A, B et C ou avec pseudo corps purs P, Q et R.

La composition d'un mélange E du domaine peut s'exprimer par rapport aux :

-vrais corps purs A, B, C... désignons par $x_1, x_2, x_3, \dots, x_k$ les valeurs correspondantes.

Les x_i représentent la composition réelle du mélange

-pseudo corps purs $A'B'C', \dots$ soient $x'_1, x'_2, x'_3, \dots, x'_k$ les valeurs correspondantes.

La relation qui lie les x'_i aux x_i est :

$$x'_i = \frac{x_i - l_i}{1 - \sum l_i}$$

4.6.3 Plan de type III

Les plans de type III sont tel quels proportions du produit A varient entre une valeur limite basse L_1 et une valeur limite haute U_1 . Le domaine des compositions est divisé en trois zones : une première zone interdite L_1 (zone ombrée inférieure sur la figure 4.20), une zone autorisée (zone claire sur la figure 4.20) et une deuxième zone interdite par la limite U_1 (zone ombrée supérieure sur la figure 4.20). ces deux limites ne conservent pas la forme du

domaine initial. Le produit B peut, lui aussi, avoir une limite haute U_2 et une limite basse L_2 . Les autres produits peuvent avoir, de même, des limites hautes et des limites basses. L'ensemble de ces contraintes va fortement modifier la manière d'aborder la construction des plans de mélanges.

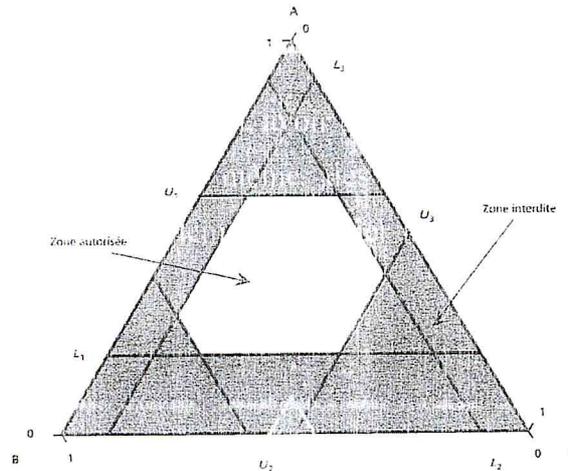


Figure 4.20 : la teneur en produit A ne peut pas être supérieure à la limite haute U_1 et inférieure à la limite basse L_1 . Les proportions de ce produit dans les mélanges varient donc entre L_1 et U_1 . Le domaine d'étude est divisé en trois zones : deux zones interdites (zones ombrées) et zone autorisée (zone claire).

4.6.3.1 L'existence du domaine expérimental

Nous avons vu conséquences d'une limite basse imposée (type II) à un ou plusieurs constituants d'un mélange. Nous allons examiner maintenant les conséquences d'une limitation des proportions de l'un ou de plusieurs constituants vers les hautes valeurs.

On considère un mélange ne comportant que deux constituants, A et B. on suppose que la limite haute du produit A est égale à U_A et que celle du produit B est égale à U_B . On envisage les différents cas suivants :

- Les limites hautes U_A et U_B sont chacune égale à 1. Toutes les proportions de A et de B sont alors possibles dans le mélange de ces deux produits (figure 4.21). dans ce cas, on a :



Figure 4.21 : La somme des limites hautes de tous les constituants égale à 2.

- La limite haute est plus petite que 1. Les teneurs en A varient alors de 0 à U_A et celles de B de $1 - U_A$ à 1 (figure 4.22). dans ce cas, on a : $U_A + U_B = 1$.



Figure 4.22 : La somme des limites hautes de tous les constituants supérieure à 1.

- Les deux limites hautes U_A et U_B sont plus petites que 1. Les proportions de A varient de $1 - U_B$ à U_A et celles de B de $1 - U_A$ à U_B (figure 4.23). dans ce cas, on a :



Figure 4.23 : La somme des limites hautes de tous les constituants supérieure à 1.

- Les deux limites hautes vérifient la relation $U_A + U_B = 1$. Dans ce cas, il n'y a plus qu'un seul mélange possible. La concentration en A est U_A et celle de B est U_B (figure 4.24).



Figure 4.24 : la somme des limites hautes de tous les constituants égale à 1.

- Les deux limites hautes vérifient la relation $U_A + U_B < 1$, il n'y a plus de mélange possible. Cette analyse permet de conclure que, dans le cas d'un mélange binaire, les mélanges sont possibles si les limites hautes respectent la relation :

$$U_A + U_B \geq 1$$

Remarque

Dans le cas d'un mélange tertiaire (produits A, B et C), si limites hautes des deux premier constituants vérifient $U_A + U_B > 1$, les mélanges de A et B sont possibles les proportions du troisième constituant sont alors comprises entre $1 - (U_A + U_B)$ et 1. si la limite haute de atteint la valeur de $1 - (U_A + U_B)$ il n'y a plus qu'un seul

mélange possible. Si la limite haute de C est inférieure à cette valeur, il n'y a plus de mélange possible. Il faut donc que :

$$(U_A + U_B + U_C) \geq 1$$

D'une manière générale, pour que des mélanges soient possibles entre k constituants, les limites hautes respectent la relation suivante :

$$\sum U_i > 1$$

Conclusion

On conclure que l'existence du domaine d'étude il doit vérifier les relations suivantes :

- La somme des limites basses inférieures à un $\sum L_i \leq 1$
- La somme des limites hautes supérieure à un $\sum U_i > 1$

4.6.3.2 Principe de la démarche

4.6.3.2.1 Compatibilité des limites

Définition :

On appelle étendue relative du produit i la différence entre la limite haute U_i et la limite basse L_i . Si l'on note l'étendue Δ_i , on a :

$$\Delta_i = U_i - L_i$$

Contraintes sur les limites hautes

Soient trois produits A, B et C dont on impose les limites basses L_1 , L_2 et L_3 . Ces limites forment un triangle $A'B'C'$ (figure 4.25) dont l'intérieur et le pourtour représentent le domaine d'étude restreint.

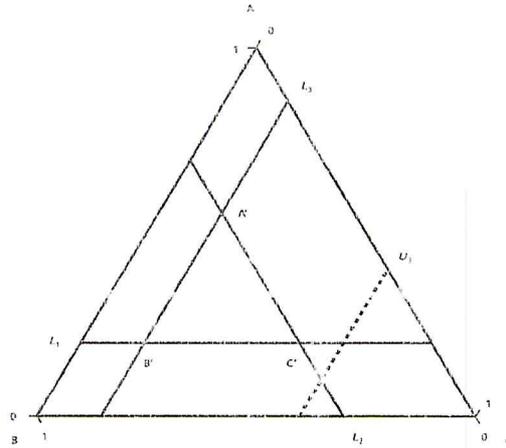


Figure 4.25 : Les limites basses forment un triangle $A'B'C'$ qui doit être coupé par toutes les limites hautes pour que les limites soient compatibles.

Ici L_3 n'est pas compatible avec les autres limites

On sait que ce domaine est toujours à l'intérieur du domaine initial et qu'il lui est semblable. Introduisons des limites hautes U_1 , U_2 et U_3 pour chacun des produits. Si ces limites hautes coupent le triangle $A'B'C'$, on dira que les limites sont compatibles. Par contre, si l'une au moins de ces limites hautes ne coupe pas le triangle $A'B'C'$, on dira que les limites sont incompatibles. On cherche la condition pour que, les limites basses étant données, les limites hautes soient compatibles avec les limites basses.

On considère maintenant la limite basse L_i est parfaitement définie. Pour que les limites soient compatibles, il faut que la limite haute U_i coupe le triangle $A'B'C'$. Il faut donc la limite de haute U_i ne peut dépasser la quantité $1 - \sum_{j \neq i} L_j$.

En retranchant L_i aux deux membres de l'inégalité :

$$U_i \leq 1 - \sum_{j \neq i} L_j$$

Nous obtenons :

$$U_i - L_i \leq 1 - \sum_j L_j$$

Donc la limite haute U_i est compatible avec la limite basse L_i si :

$$\Delta_i \leq 1 - \sum_j L_j$$

Contraintes sur les limites basses :

On définit les limites hautes U_1 , U_2 et U_3 et on désire que les limites basses L_1 , L_2 et L_3 soient compatibles avec les limites hautes. Les limites basses doivent couper le domaine défini par les limites hautes.

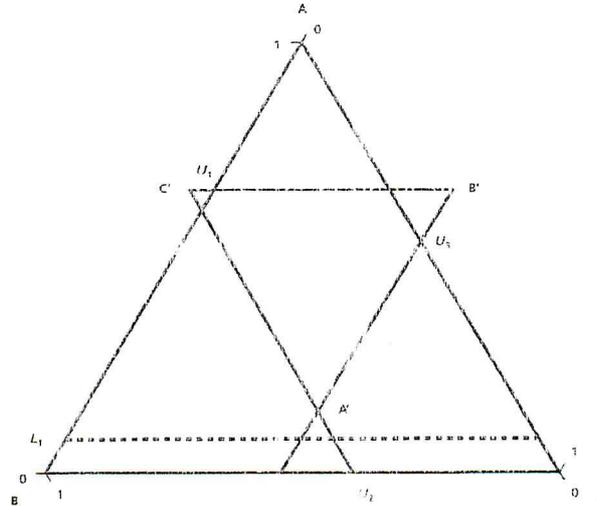


Figure 4.26 : Les limites hautes forment un triangle $A'B'C'$ qui doit être coupé par toutes les limites basses pour que les limites soient compatibles.

Ici L_1 n'est pas compatible avec les autres limites

La limite inférieure d'un constituant i ne peut être inférieure à $1 - \sum_{j \neq i} U_j$. En retranchant U_1 aux deux membres de l'inégalité :

$$L_i \geq 1 - \sum_{j \neq i} U_j$$

Il vient :

$$L_i - U_i \leq 1 - \sum_j U_j$$

Donc la limite basse L_i est compatible avec la limite haute U_i .

Si :

$$\Delta_i \leq \sum_j U_j - 1$$

4.6.3.2.2 Calcul du nombre de sommets, arêtes et faces

Théorème 1 :

Le nombre de sommet du domaine d'étude d'un mélange ayant k constituants est donné par la formule suivante :

$$N_{\text{sommets}} = k + \sum_{r=1}^{r=k} [(k - 2r)\Gamma_{(r)} - (r - 1)\Lambda_{(r)}]$$

Formule dans laquelle :

- r est un nombre entier variant de 1 à k .
- $\Gamma_{(r)}$ est la somme des k termes $\Gamma_{(1)}, \Gamma_{(2)}, \Gamma_{(3)}, \dots, \Gamma_{(k)}$, que l'on calcule de la manière suivante.
- $\Gamma_{(1)}$ est le nombre étendues qui sont inférieures à la plus petite des deux expressions, $\sum U_i - 1$ ou $1 - \sum L_i$. Cette plus petite expression sera conservée pour la suite des calculs.
- $\Gamma_{(2)}$ est calculé de la manière suivante : on fait la somme de deux étendues. Si cette somme est inférieure à $\sum U_i - 1$ (ou $1 - \sum L_i$ suivant les cas), on la prend en compte. Si cette somme est égale ou supérieur, on ne la prend pas en compte. On examine ainsi toutes les combinaisons possibles de sommes de deux étendues. $\Gamma_{(2)}$ est égal au nombre des sommes prises en compte.
- $\Gamma_{(3)}$ est calculé de la même manière que $\Gamma_{(2)}$, mais en considérant des sommes de trois étendues. Toutes les combinaisons de trois étendues sont examinées.
- $\Gamma_{(4)}$ et les suivants jusqu'à $\Gamma_{(k)}$ sont calculés de la même manière que précédemment en examinant les sommes de 4, 5, ..., k étendues.
- $\Lambda_{(r)}$ est calculé de la même manière que $\Gamma_{(r)}$, mais en ne retenant que les étendues ou la somme des est égale exactement à :

$$\sum U_i - 1$$

(ou $1 - \sum L_i$ suivant les cas).

Théorème 2 :

Le nombre d'arêtes du domaine d'étude d'un mélange ayant k constituants est donné par la formule :

$$N_{arêtes} = C_k^{k-2} + \sum_{r=1}^{r=k-2} L_{(r)} C_{k-r}^{k-r-2} - \sum_{r=2}^{r=k} [\Gamma_{(r)} + \Lambda_{(r)}] C_r^{r-2}$$

Formule dans laquelle :

- C_k^p est le nombre de combinaisons de k objets p à p .
- Les autres termes ont la même signification que précédemment.

Théorème 3 :

Lorsque le nombre de constituants dépasse trois, on peut utiliser une formule pour calculer le nombre de faces de l'hyper polyèdre. Cette formule est semblable à celles qui ont été indiquées pour les sommets et les arêtes. Le nombre de faces est donné par :

$$N_{faces} = C_k^{k-3} + \sum_{r=1}^{r=k-3} L_{(r)} C_{k-r}^{k-r-3} - \sum_{r=3}^{r=k} [\Gamma_{(r)} + \Lambda_{(r)}] C_r^{r-3}$$

Pour mener à bien l'expérimentation, il ne suffit pas de connaître le nombre de sommets, d'arêtes et de faces du domaine d'étude. Il faut encore positionner ces points d'expériences dans l'espace ou l'hyperespace expérimental. Jusqu'à trois constituants, ce problème est facilement résolu par un graphique. Pour quatre constituants, on peut encore essayer de débrouiller avec des schémas. Au-delà, cela devient pratiquement impossible et des algorithmes spécifiques ont été développés pour effectuer ces calculs. Parmi ceux-ci on peut citer l'algorithme XVERT de Snee et Marquardt.

4.6.3.2.3 Algorithme XVERT de Snee&Marquardt [18]**Détermination des sommets à partir des limites :**

- Classer les constituants en ordre croissant de domaine (Δ_i).
- Construire avec les $k - 1$ premiers constituants un plan factoriel complet à 2 niveaux (L_i et U_i) et compléter le dernier constituant pour que $\sum x_i = 1$.

- Ramener si nécessaire la x_i du dernier constituant à sa U_i et augmenter de la différence la valeur du niveau d'un des autres constituants chaque fois que cela est possible (nouveaux sommets).

Détermination des arêtes et de leur centre

- Recherche les sommets ayant la même frontière p ,
 - $p = k - r - 1$ (dimension $r = 1$ pour les arêtes, $r = 2$ pour les faces,...)
 - Combinaisons de $p = 2$ caractéristiques communes.
- Ne retenir que les combinaisons de sommets avec les niveaux avec composition égal à L_i ou U_i .
- Calculer la composition du centre de l'arête en faisant la moyenne de compositions des sommets correspondants.

Détermination des faces et de leur centre

- Recherche les faces ayant la même frontière p .
 - $p = k - r - 1$ (dimension $r = 1$ pour les faces...).
 - Combinaisons de $p = 1$ caractéristique commune.
- Ne retenir que les combinaisons de sommets avec les niveaux avec composition égal à L_i ou U_i .
- Calculer la composition du centre de la face en faisant la moyenne des compositions des sommets correspondants.

CHAPITRE 5

PLANS D-OPTIMAUX ET LES ALGORITHMES D'ÉCHANGES

Les plans D-optimaux sont souvent utilisés lorsque le domaine expérimental n'est pas entièrement accessible. Il existe des contraintes qui interdisent l'accès à certaines régions du domaine d'étude. Ces plans permettent d'explorer la plus grande partie possible du domaine expérimental malgré les restrictions dues aux contraintes.

Les plans D-optimaux sont aussi souvent utilisés lorsque l'on veut diminuer le nombre des essais d'un plan classique. On impose un nombre d'essais et l'algorithme de calcul des plans D-optimaux conserve, pour un modèle donné, les meilleurs points d'expériences.

La recherche de ces meilleurs points expériences n'est possible que si l'on possède un logiciel dans lequel est intégré un algorithme spécifique de calcul des plans D-optimaux. Les deux algorithmes les plus connus sont celui de Mitchell et celui de Fedorov. Il s'agit d'algorithmes basés sur une procédure d'échange.

Les résultats de l'algorithme représentent une bonne solution au problème posé, mais on n'est jamais sûr qu'il s'agisse de la meilleure solution. En effet, le logiciel calcule le déterminant de la matrice tXX pour des ensembles de points expérimentaux et cherche la valeur plus élevée [11].

Dans ce chapitre, nous présenterons les critères d'optimalités, ensuite, nous discuterons des plans D-optimaux.

5.1 Transmission des erreurs

Les réponses mesurées par l'expérimentateur sont entachées d'une erreur expérimentale, ce sont donc des grandeurs aléatoires. Cette erreur peut être estimée par l'écart-type de la population et déterminée sur un échantillon. Les niveaux des facteurs et le modèle mathématique choisi a priori sont supposés n'introduire aucune erreur et ne sont pas considérés comme des variables aléatoires. La matrice X n'est donc pas une grandeur aléatoire et elle n'introduit aucune erreur. Seule la matrice Y est une matrice dont les éléments sont des variables aléatoire. Les variations des éléments de cette matrice entraînent donc des variations sur les coefficients du modèle mathématique puis qu'il existe la relation :

$$\hat{A} = ({}^tXX)^{-1} {}^tXY$$

Les réponses prédites comme des grandeurs aléatoires puis qu'il existe la relation :

$$\hat{Y} = X\hat{A}$$

Il résulte de cette courte analyse que les erreurs sur les coefficients du modèle et sur les réponses prédites ne seront pas les mêmes selon la valeur des erreurs sur les réponses mesurées, le modèle choisi à priori et l'emplacement des points expérimentaux dans le domaine d'étude. Nous allons examiner en détail les liens qui existent entre ces différentes erreurs.

5.1.1 Précision sur les coefficients

Lorsque les hypothèses de la régression linéaire sont respectées et lorsque les variances des réponses mesurées sont issues d'une même population pour tout le domaine d'étude, les statisticiens ont démontré la relation suivante :

$$\text{var}(\hat{A}) = \sigma^2 ({}^tXX)^{-1} \quad (5.1)$$

Où σ^2 est la variance des résidus. Cette variance regroupe celle des réponses mesurées et celle du manque d'ajustement du modèle. Nous appellerons cette variance globale soit variance des résidus, soit variance de régression. Dans la relation (5.1), $\text{var}(\hat{A})$ est la matrice des variances-covariances des coefficients. La formule précédente, pour l'usage que nous voulons en faire, devient alors

$$\text{Diag var}(\hat{A}) = \sigma^2 \text{Diag} ({}^tXX)^{-1}$$

C'est la relation des variances des coefficients. Cette relation est très importante. En effet, elle montre qu'il ya trois composantes qui induisent des erreurs sur les coefficients. Ces composantes sont :

- L'erreur commise sur les réponses à cause du terme σ^2 , les erreurs commises sur les réponses mesurées se transmettent donc aux coefficients du modèle mathématique.
- L'emplacement des points expérimentaux. La position des points expérimentaux, dans le domaine d'étude génère des erreurs plus ou moins élevées sur les coefficients du modèle mathématique. C'est la matrice X qui est responsable de ce transfert. Cela signifie qu'un expérimentateur qui conduit des expériences avec

beaucoup de soin et obtient des résultats d'essai très précis peut quand même avoir des modèles mathématiques médiocres s'il a mal disposé les points expérimentaux dans le domaine expérimental.

- Le modèle mathématique choisi a priori. Le choix initial du modèle mathématique engendre des écarts d'ajustement plus ou moins élevés sur les coefficients du modèle. Les responsables de ce transfert sont σ^2 et la matrice X , cela signifie que l'on doit s'attendre à de grosses erreurs sur certains coefficients lorsque le modèle mathématique choisi a priori est très éloigné du modèle réel.

5.1.2 Précision sur les réponses prédites

La méthode des moindres carrés permet, grâce au jeu des coefficients qui a été calculé, d'obtenir les réponses dans tout le domaine d'étude, c'est-à-dire que l'on peut calculer les réponses pour tous les points du domaine d'étude même si aucune expérience n'y a été réalisée. Soit un point p du domaine d'étude, défini par ses coordonnées, la réponse prédite en ce point est donnée par :

$$\hat{y}_p = {}^t x_p \hat{A}$$

Si p est un des points expérimentaux du plan d'expériences, la matrice-vecteur ${}^t x_p$ est une ligne de la matrice de calcul X . La matrice-vecteur x_p sera appelée le vecteur modélisé du point p , \hat{A} est la matrice des coefficients calculés avec la méthode des moindres carrés.

L'incertitude sur les coefficients se transmet à la réponse calculée :

$$\text{var}(\hat{y}_p) = \text{var}({}^t x_p \hat{A})$$

Dans cette relation la matrice ligne ${}^t x_p$ dépend des coordonnées d'un point du domaine d'étude et il a été admis par hypothèse que les coordonnées des points expérimentaux étaient parfaitement connues et n'introduisaient pas d'erreurs. On peut donc sortir le vecteur modélisé du point p des parenthèses :

$$\text{var}(\hat{y}_p) = {}^t x_p \text{var}(\hat{A}) x_p$$

Dans cette expression la variance de \hat{A} est connue et l'on sait qu'elle est égale à :

$$\text{var}(\hat{A}) = \sigma^2 ({}^t X X)^{-1}$$

La variance de la réponse calculée au point p est donc :

$$\text{var}(\hat{y}_p) = {}^t x_p \sigma^2 ({}^t X X)^{-1} x_p$$

Cette relation permet de calculer l'incertitude sur la réponse calculée au point p . On constate que cette erreur sur la réponse calculée (ou réponse prédite) dépend de quatre grandeurs :

- l'erreur expérimentale sur les réponses mesurées

- la position du point p dans le domaine d'étude
- l'ensemble des points qui ont été utilisés pour établir les coefficients du modèle, c'est-à-dire le plan d'expériences lui-même
- le modèle postulé choisi pour interpréter les résultats (par la matrice de calcul des coefficients et la variance des résidus)

5.1.3 La fonction de variance de prédiction

L'erreur commise sur les réponses mesurées dépend de la nature de l'expérimentation, du degré de précision de la technologie employée, du soin, de l'habileté de l'expérimentateur, et de bien d'autres facteurs dont l'expérimentateur est responsable. Ces facteurs ne dépendent pas de la théorie des plans d'expériences mais de la pratique expérimentale [9]. Pour séparer cette partie expérimentale de celle qui dépend de la théorie, on introduit la fonction de variance de prédiction $d^2(\hat{y}_p)$:

$$d^2(\hat{y}_p) = {}^t x_p ({}^t X X)^{-1} x_p$$

On peut tracer les courbes d'égalité d'erreur de prédiction dans le domaine d'étude. Il est intéressant de constater que la fonction d'erreur de prédiction ne dépend pas des résultats des expériences c'est-à-dire des valeurs des réponses mesurées. Elle dépend essentiellement de l'emplacement des points expérimentaux dans le domaine d'étude et du modèle postulé. On peut donc savoir, avant de commencer l'expérimentation, comment la précision des réponses prédites sera affectée par le choix de l'emplacement des points expérimentaux et par celui du modèle. En faisant apparaître les écart-types et la fonction de la variance de prédiction, la relation peut s'écrire :

$$\text{var}(\hat{y}_p) = \sigma^2 {}^t x_p ({}^t X X)^{-1} x_p = \sigma^2 d^2(\hat{y}_p)$$

Si la fonction d'erreur de prédiction est plus petite que 1, on a une meilleure précision sur les réponses prédites que sur les réponses mesurées. Si la fonction d'erreur de prédiction est plus grande que 1, c'est l'inverse. En général, on s'arrange pour que l'erreur de prédiction ne soit pas plus élevée que l'erreur expérimentale. On évite donc de faire des prévisions dans les régions où la fonction d'erreur de prédiction est supérieure à l'unité.

5.2 Critères d'optimalité

Les erreurs sur les réponses prédites sont liées à la matrice X . Comme cette matrice est connue, il est possible d'évaluer la qualité de l'expérimentation avant le début des expériences. La matrice X dépend du modèle mathématique postulé et de l'emplacement des

points expérimentaux. Un bon plan d'expériences sera un plan permettant d'obtenir la plus faible erreur sur les réponses prédites. La règle que l'on peut suivre est d'obtenir une erreur sur les réponses prédites de l'ordre de grandeur de l'erreur sur les réponses mesurées. Selon les critères choisis, l'emplacement des points expérimentaux pourra varier d'un plan à l'autre. Il y a plusieurs critères d'optimalité. Il y a ceux qui s'intéressent à la répartition de la variance dans le domaine d'étude. C'est par exemple le critère d'iso variance par rotation. Il y a les critères d'optimalité qui ont pour objectif d'obtenir un modèle mathématique de bonne qualité. Ces derniers critères s'intéressent à la précision des coefficients du modèle [11].

5.2.1 Qualité de la représentation des coefficients

Plusieurs auteurs ont montré que la limite du domaine de confiance des coefficients est donnée par la relation:

$$(A - \hat{A})(X^T X)(A - \hat{A}) = qs^2 F_{\alpha}(p, v) \quad (3.2)$$

Où

s^2 est une estimation de la variance expérimentale

F est la statistique correspond au test de Fisher

v est le niveau de confiance choisi

X est la matrice d'expérience

\hat{A} est le vecteur calculé

q est le nombre de degrés de liberté associé à s^2

Ce domaine est un hyper ellipsoïde dans l'espace des coefficients, centré sur le vecteur calculé \hat{A} . Avec un risque α , on peut dire que les valeurs vraies des coefficients sont dans cet hyper ellipsoïde. Pour deux coefficients, l'hyper ellipsoïde est représenté graphiquement par :

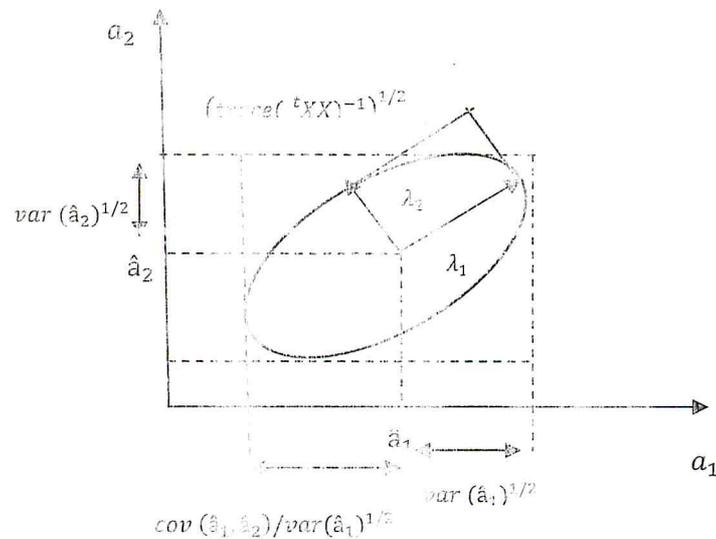


Figure 5.1 : Ellipsoïde de confiance

Où (\hat{a}_1, \hat{a}_2) sont les deux coefficients trouvés par la régression multilinéaire, (λ_1, λ_2) les deux valeurs propres de l'ellipsoïde. Cet hyper ellipsoïde représente les performances de l'ensemble des observations pour un modèle donné. Il peut être décrit par 3 caractérisations :

A. Volume

Le volume de l'ellipsoïde est égal à $\pi(\lambda_1\lambda_2)^{\frac{1}{2}} = \pi(\det({}^tXX)^{-1})^{\frac{1}{2}}$. Ce volume est lié au déterminant de la matrice de dispersion $(X^tX)^{-1}$. Plus le volume sera petit, plus l'hyper ellipsoïde tend vers un point et plus on se rapprochera de la solution vraie A .

B. Forme

Lorsque la forme est très allongée, alors il existe une grande disparité sur la précision des coefficients. Si $\lambda_1 = \lambda_2$ l'ellipsoïde devient une hyper sphère, et donc tous les coefficients sont déterminés avec la même précision.

C. Orientation

Si les axes principaux de l'hyper ellipsoïde sont parallèles aux axes des coefficients alors, la valeur calculée d'un coefficient sera indépendante des valeurs calculées des autres coefficients.

D'après la formule (5.2), l'hyper ellipsoïde dépend de la matrice de $({}^tXX)$, donc les critères sont fondés sur les matrices d'information $({}^tXX)$ et de dispersion $({}^tXX)^{-1}$.

5.2.2 Critère d'iso variance par rotation

Le critère d'isovariance par rotation est respecté si l'erreur de prédiction est la même pour des points situés à égale distance du centre du domaine d'étude.

5.2.3 Critère de A-optimalité

Le critère de A-optimalité vise à minimiser la somme des variances des coefficients du modèle. Les éléments diagonaux de la matrice de dispersion $({}^tXX)^{-1}$ étant proportionnels aux variances des coefficients, on minimisera donc la trace de $({}^tXX)^{-1}$. Le calcul d'un plan A-optimal est numériquement coûteux.

5.2.4 Critère de D-optimalité

Vise à minimiser la variance globale de tous les coefficients du modèle. Cette variance est proportionnelle au déterminant de la matrice de dispersion. A lieu de minimiser le déterminant de $({}^tXX)^{-1}$, on maximise le déterminant de tXX , évitant ainsi l'opération lourde d'inversion de matrice. C'est ce critère qui est le plus fréquemment utilisé et qui est implémenté dans les logiciels spécialisés.

5.2.5 Critère de E-optimalité

Une matrice d'expériences est dite E-optimal, si elle conduit à la valeur propre maximale de $({}^tXX)^{-1}$ la plus faible possible.

5.2.6 Critère de G-optimalité

Le critère-G prend en compte la plus grande valeur, sur tout le domaine expérimental, de la fonction de variance $d^2(\hat{y}_p)$ engendré par la matrice X.

$$d_{max} = \text{Max} (d^2(\hat{y}_p))$$

La meilleure matrice d'expériences vis-à-vis de ce critère étant celle qui a la plus petite valeur.

Efficacité-G

L'efficacité-G d'une matrice d'expériences est donnée par la formule suivante :

$$Eff - G = \frac{p}{Nd_{max}} 100$$

Une matrice d'expériences est proche de la matrice d'expériences optimal G , si la valeur de l'efficacité- G avoisine 100 %.

5.2.7 Critère de M-optimalité

Le critère M permet de tenir compte de la qualité d'information apportée par l'expérience. Ce critère est indépendant du nombre des essais du plan d'expériences.

On appelle matrice des moments M , la matrice définie par :

$$M = \frac{({}^t X X)}{N}$$

Grace à ce critère on peut comparer deux matrices d'expériences, n'ayant pas forcément le même nombre d'expériences. Soient M_1 et M_2 deux matrices des moments associées à deux matrices d'expériences constituées respectivement de N_1 et N_2 expériences.

$$M_1 = \frac{({}^t X_1 X_1)}{N_1}$$

$$M_2 = \frac{({}^t X_2 X_2)}{N_2}$$

Nous dirons que le premier plan est plus efficace que le deuxième plan vis-à-vis du critère M si :

$$|M_1| > |M_2|$$

5.2.8 Critère d'orthogonalité

Une matrice d'expériences est orthogonale si elle permet d'obtenir des estimations des coefficients indépendants. Cela se caractérise par des axes de l'ellipsoïde parallèles aux axes des coefficients. Cette propriété est obtenue quand $({}^t X X)^{-1}$ (ou ${}^t X X$) est diagonale, donc quand les covariances des coefficients sont nulles

5.2.9 Critère de presque orthogonalité

Avec les modèles du second degré, il est impossible d'obtenir une matrice de dispersion $({}^tXX)^{-1}$ parfaitement diagonale. Il reste toujours un élément qui associe le terme constant du modèle aux termes du second degré. C'est pour quoi on a introduit le critère de presque orthogonalité. Si on exclut la première ligne et la première colonne de $({}^tXX)^{-1}$, on obtient une sous-matrice ne contenant plus les termes constants. Si cette sous-matrice est diagonale, on dit qu'il y a presque orthogonalité.

5.3 Plans D-optimaux

L'inconvénient majeur des plans à sommets extrêmes et des plans dérivés est leur grand nombre d'essais. Les méthodes des plans D-optimaux sont souvent utilisées lorsque l'on veut diminuer le nombre des essais d'un plan. On désire, en général, un modèle mathématique fournissant des réponses calculées aussi précises que possible. L'objectif est donc de trouver le meilleur compromis : établir un modèle mathématique capable de fournir de bonnes prévisions et effectuer le minimum d'essais [22].

5.3.1 Principe des plans D-optimaux

Un plan $Pest$ dit D-optimal pour résoudre un problème donné, s'il minimise le critère

$$\text{critère } - D = \left| \frac{{}^tX^tX}{n} \right|^{-1}$$

Ce qui équivaut à minimiser le déterminant de la matrice de variance-covariance des paramètres du modèle.

Le critère-D est une mesure globale de la précision des paramètres du modèle et sa minimisation est plus spécialement utile quand le but de l'expérimentation est d'analyser les effets des facteurs sur la réponse (modélisation).

Si les domaines individuels des facteurs sont limités à l'intervalle $[-1,1]$, le critère D est toujours plus grand ou égal à 1[1].

5.3.2 Algorithme d'échange

Précédemment on a décrit les principaux et critères fondamentaux pour la construction des plans D-optimaux, mais il n'explique pas le processus de sélection. Basée sur la

complexité des conceptions D-optimales et le grand nombre de combinaisons possibles des expériences, des algorithmes sont utilisés pour le processus de sélection. Le présent chapitre traite les algorithmes d'échange.

Un algorithme d'échange consiste à sélectionner les matrices X^* optimales par échange d'un ou plusieurs points du plan de départ généré et à répéter cette procédure d'échanges jusqu'à trouver la meilleure matrice selon les critères D-optimale. Les algorithmes peuvent être divisés en deux groupes : groupe-1 algorithme sert à ajouter et supprimer les points d'une manière séquentielle, et un algorithme de groupe-2 consiste à réaliser un échange par un ajout et une suppression simultanée de points (Meyer & Nachtsheim 1995). Nous allons maintenant expliquer et évaluer les cinq différents algorithmes qui sont universellement applicable, mais différent dans le temps de calcul utilisé, la qualité et l'efficacité des modèles générés.

5.3.2.1 Procédure d'échange général

Afin de commencer le processus de sélection, nous devons créer un plan de démarrage avec n expériences X_n . Le but de l'algorithme d'échange est de supprimer ou ajouter des points à cette matrice et de déterminer l'effet de cette modification. Soit $d(x_x)$, la fonction de la variance de prédiction au point x .

$$d(x_x) = x_x^t * (X_n^t X_n)^{-1} * x_x \quad (5.3)$$

L'ajout d'une nouvelle expérience x_j à la matrice X_n , va créer une nouvelle matrice X_{n+1} , selon Aguitar et al (1995), la relation entre X_n et X_{n+1} est donnée par.

$$(X_{n+1}^t X_{n+1}) = (X_n^t X_n) + (x_j * x_j^t) \quad (5.4)$$

En outre, l'influence de l'échange peut être utilisée pour mettre à jour le déterminant de la nouvelle matrice sans le recalculer. Dans ce cas, le déterminant augmente proportionnel à la variance de prédiction $d(x_j)$ du point ajouté x_j .

$$|X_{n+1}^t X_{n+1}| = |X_n^t X_n| * (1 + d(x_j)) \quad (5.5)$$

La suppression d'une expérience x_i de la matrice de départ X_n repose sur le même principe on obtient ainsi:

$$(X_{n-1}^t X_{n-1}) = (X_n^t X_n) - (x_i * x_i^t) \quad (5.6)$$

$$|X_{n-1}^t X_{n-1}| = |X_n^t X_n| * (1 - d(x_i)) \quad (5.7)$$

5.3.2.2 Algorithme DETMAX

L'algorithme DETMAX a été publié par Mitchell (1974) et est un exemple typique de l'algorithme de catégorie groupe-1. Avec une matrice de départ aléatoire de n points, l'algorithme cherche à améliorer le déterminant de la matrice d'information en ajoutant ou en supprimant un point. L'expérience d'ajouter x_j est celui avec la plus grande variance de prédiction $d(x_j)$. Comme l'équation (5.5) montre, cette expérience conduit à l'augmentation du déterminant. Le résultat d'un tel procédé d'échange est un déterminant qui est supérieure ou égale au précédent.

Jusqu'à présent, cet algorithme serait l'algorithme de Wynn-Michelle de 1972, mais Mitchell (1974) a modifié cette approche pour obtenir une plus grande flexibilité et des excursions ont permis au design. Dans ce cas, une excursion signifie que le plan de ($n + 1$ - point) ne doit pas être réduit immédiatement à un plan de n -points, mais peut devenir un modèle à $n + 2$ points. Par conséquent, le remplacement de plus d'un point dans le plan de départ peut être possible au cours d'une itération. La limite d'excursions est définie par Michell (1974) pour être $k = 6$. En considérant un plan avec les n meilleurs points actuels, l'algorithme ajoute ou supprime jusqu'à k points jusqu'à ce que la limite de déplacement est atteints. La taille du plan crée varie de $n - k$ à $n + k$. Si aucune amélioration dans le déterminant n'est trouvée au cours de cette excursion, tous les plans créés sont enregistrés dans une liste F qui contient les plans d'échec. Cet ensemble F est utilisé pour la prochaine excursion où l'on considère deux règles différentes définies par Michell (1974, p.204):
 D le plan actuelle, à tout moment au cours d'une excursion. les règles à utiliser sont les suivantes:

- (i) Si le nombre de points dans D est supérieur à n , il faut soustraire un point si D n'est pas dans F et ajouter un point autrement.
- (ii) Si le nombre de points dans D est inférieur à n , ajouter un point si D n'est pas dans F et soustraire une autre point.

L'algorithme suivant (5.1) montre l'utilisation de ces règles d'excursions et la visualisation de flux de l'algorithme avec une notation de programmation simple et abstraite.

Générer un plan aléatoire avec un nombre de points désiré ;

tant que La limite de déplacement n'est pas atteinte **faire**

```

si le nombre de candidats est égal au nombre de points souhaitées alors
  | l'ajout ou la supprimer un point aléatoirement;
sinon si le nombre de candidats est plus grand que le nombre de points souhaité alors
  | si le nouveaux plan n'est pas à l'intérieur de l'ensemble de plans d'échec alors
  |   | supprimer le candidat dont la variance de prédiction est la plus petite;
  |   sinon
  |     | ajouter le candidat dont sa variance de prédiction est plus grande ;
sinon si le nombre de candidats est inférieur au nombre de pistes souhaitées alors
  | si le nouveaux plan n'est pas à l'intérieur de l'ensemble de plans d'échec alors
  |   | ajouter le candidat dont sa variance de prédiction est plus grande ;
  |   sinon
  |     | supprimer le candidat dont la variance de prédiction est la plus petite;
si aucune amélioration sur le déterminant constants alors
  | sauvegarder le plan dans la liste des plans en échec ;
sinon
  | liste claire de designs l'échec ;
  
```

Algorithme 5.1 : Algorithme de DETMAX (Mitchell (1974))

Remarque

Si l'ensemble F est vide, l'algorithme ajoute simplement et soustrais des points, tel que décrit dans le premier paragraphe. Considérant les plans échec des itérations précédentes, l'algorithme procède toujours dans le sens d'un plan à n points dans F , qui a déjà conduit à l'échec. Puis il change de direction (Mitchell1974). Si une excursion améliore le plan, les plans d'échec dans F sont supprimés et un nouveau départ est effectué (d'Aguiar et al.1995, Mitchell 1974).

5.3.2.3 Algorithme de Fedorov

Fedorov (1972) est un algorithme d'échange simultané qui garde toujours la taille n du plan désirée sans excursions. Après la génération d'un plan de départ aléatoire, l'algorithme sélectionne un point x_i du plan doit être éliminé par un point x_j de l'ensemble candidat. La procédure d'ajout et la suppression d'un point est fait en une seule étape et peut être notée comme une véritable d'échange. L'effet d'un tel échange peut être démontré par l'utilisation de la matrice d'information. En référence à l'équation (5.4) et (5.6), un échange est une addition et une soustraction simultanée d'un point.

$$(X_{nouveau}^t X_{nouveau}) = (X_{ancien}^t X_{ancien}) - (x_i * x_i^t) + (x_j * x_j^t) \quad (5.8)$$

Contrairement à l'échange général de la section (5.3.2.1), Fedorov (1972) a examiné l'interaction entre la fonction de variance des deux candidats pour le calcul du nouveau déterminant. Au lieu d'ajouter et de supprimer la variance de prédiction des deux points selon l'équation (5.5) et (5.7), il définit une fonction qui modifie le déterminant de la matrice de la manière suivante:

$$|X'_{nouveau} X_{nouveau}| = |X'_{ancien} X_{ancien}| * (1 + \Delta(x_i, x_j)) \quad (5.9)$$

Le calcul de la valeur Δ pour un couple de x_i et x_j utilise la variance de la prédiction des deux points et une fonction de variance combinée appelée $d(x_i, x_j)$.

$$\Delta(x_i, x_j) = d(x_j) - [d(x_i)d(x_j) - (d(x_i, x_j))^2] - d(x_i) \quad (5.10)$$

$$d(x_i, x_j) = x_i^t * (X_n^t X_n)^{-1} * x_j = x_j^t * (X_n^t X_n)^{-1} * x_i \quad (5.11)$$

L'idée de base de l'algorithme de Fedorov est de calculer la valeur de Δ pour tous les couples possibles (x_i, x_j) et sélectionner l'une avec la plus grande valeur. Le point x_i est pris à partir de sélectionné, et x_j peut être soit pris à partir des points restants ou à partir de l'ensemble de tous les candidats. Considérant l'ensemble des points restants qu'on appelle une recherche épuisante qui évite les points double dans le plan. Avec une recherche non épuisante, la sélection d'une expérience deux fois est possible. Comme l'équation (5.9) montre, les points avec la plus haute Δ -valeur augmentent la valeur du déterminant. De plus si plus un couple avec la même Δ -valeur est trouvé, l'algorithme choisit un parmi eux d'une manière au hasard. Tandis que si un couple avec un effet Δ -valeur positif est trouvé, l'algorithme échange les points et met à jour la matrice d'information et la matrice de dispersion. Parfois, on se trouve devant un couple qui augmente le déterminant mais pas d'une manière significative. Pour cela, Fedorov a défini une valeur seuil ϵ_{fed} , si le maximum Δ -valeur est plus petite que ϵ_{fed} alors on arrête l'échange. En générale, on prend comme ϵ_{fed} , la valeur 10^{-6} (Aguitar et al). Le schéma général de l'algorithme est donné dans la liste (5.2).

Générer un plan aléatoire avec un nombre de points désiré n :

tant que des couples avec Δ positive sont trouvés alors

pour $i = 1$ à n faire

calculer la variance de la prédiction $d(x_i)$ pour le point x_i de plan;

pour $j = 1$ à N faire

calculer la variance de prédiction $d(x_j)$ pour le point d'appui x_j ;

calculer la fonction de variance $d(x_i, x_j)$ pour le couple (x_i, x_j) ;

calculer la fonction delta $\Delta(x_i, x_j)$ pour le couple ;

vérifier si le delta est le maximum et enregistrer le couple (x_i, x_j) ;

si le maximum des deltas est positif alors

si plus qu'un couple avec le même delta maximal alors

sélectionner aléatoirement un couple ;

change le point sélectionné x_i avec x_j ;

mettre à jour la matrice d'informations et la matrice de dispersion ;

réinitialiser delta maximal ;

Algorithme 5.2 : Algorithme de Fedorov (1972).

5.3.2.4 Algorithme de Fedorov modifiée

Cook & Nachtsheim (1980) font une comparaison de différents algorithmes pour construire des plans D-optimal exacts et ils ont inventé un algorithme propre en fonction de l'algorithme de Fedorov 1972. L'algorithme de Fedorov détermine la valeur de Δ pour tous les couples d'échange possibles pendant une itération mais n'utilise qu'une seule valeur pour effectuer un échange. Ce calcul est coûteux lors de son utilisation. Avec la version modifiée par Cook & Nachtsheim (1980), chaque itération de l'algorithme de Fedorov est décomposés en $[n]$ étapes, une pour chaque Point d'appui dans le plan au début de l'itération. Avec une matrice de conception d'un ordre aléatoire, l'algorithme commence avec le premier point d'appui x_i et calcule les Δ -valeur pour tous les couples possibles avec ce point d'appui. Après avoir trouvé le meilleur échange de ce point, le plan est mis à jour et le point d'appui prochain est proposé pour un échange. En d'autres termes, une itération de l'algorithme de Fedorov est modifiée pour échanger jusqu'à des points de design n si le déterminant augmenterait (Cook & Nachtsheim 1980, Atkinson & Donev 1998). Ce comportement est visualisé dans la liste (5.3).

La différence entre les deux approches peut être montrée par l'exemple suivant : Considérant un plan souhaitée avec 5 points et un ensemble de candidat $N = 20$ expériences, $n * N = 100$ couples possibles peuvent être proposées pour un échange. Chaque itérations de l'algorithme de Fedorov on calcul ces 100 valeur Δ pour tous les couples possibles et on

utilise une seule d'entre eux pour réaliser l'échange. En revanche, la version modifiée de l'algorithme commence par le premier point de plan et ne calcule que les 20 Δ -valeur pour les couples possibles, y compris ce point. Après cela, un échange est effectué et la matrice de dispersion doit être mise à jour. L'algorithme se poursuit avec le point de design suivant et calcule à nouveau 20 valeurs du Δ . En général, à la fois l'algorithme calcule les 100 valeurs au cours d'une itération, la version d'échange calcule jusqu'à 3 points au cours de cette procédure.

Une étude réalisée par Cook & Nachtsheim(1980) montre que l'approche modifiée est peut être deux fois plus vite que l'algorithme de Fedorov.

Générer un plan aléatoire avec un nombre de points désiré n :

tant que des couples de points avec un Δ positive peuvent être trouvés alors

pour $i = 1$ à n faire

 calculer la variance de prédiction $\hat{\sigma}^2(x_i)$ pour le point x_i ;

pour $j = 1$ à N faire

 calculer la variance de prédiction $\hat{\sigma}^2(x_j)$ pour le point x_j ,

 calculer la fonction de variance $d(x_i, x_j)$ pour le couple (x_i, x_j) ;

 calculer la fonction delta $\Delta(x_i, x_j)$ pour le couple ;

 sélectionner le couple avec un delta maximal;

si le maximum des deltas est positif alors

si il ya plusieurs couples avec la même valeur de delta maximal alors

 sélectionner aléatoirement un couple ;

 changer le point sélectionné x_i avec x_j ;

 mettre à jour la matrice d'informations et la matrice de dispersion ;

 réinitialiser delta maximal ;

Algorithme 3.3: Algorithme de Fedorov modifié par Cook & Nachtsheim(1980)

5.3.2.5 Algorithme k-échange

Au cours d'une comparaison de l'algorithme de Fedorov (1972) et l'algorithme de Mitchell (1974) DETMAX, Johnson & Nachtsheim (1983) les points sélectionnés par l'algorithme de suppression de Fedorov sont normalement pas les uns avec la plus faible variance de prédiction, mais la fréquence des rangs de points supprimés pourrait être caractérisée comme biaisé en faveur des rangs de la variance inférieures. Simplement dit, au lieu de considérer tous les candidats ou seulement l'un avec la plus petite variance de

prédiction, un ensemble de k points ayant la plus faible variance devrait être sélectionné. Similaire à l'algorithme de Fedorov modifié, une itération est maintenant décomposée en k étapes. À l'intérieur de chacune de ces k étapes les valeurs Δ -valeurs sont calculées et le couple correspondant serait échangé si le déterminant augmente (Johnson & Nachtsheim 1983)

La similitude avec les algorithmes précédemment décrits peuvent être reconnue si nous définissons k -valeur. Avec $k = 1$ l'algorithme est similaire à l'algorithme de Wynn-Mitchell de la section (5.3.2.2), et avec $k = n$, il devient l'algorithme de Fedorov modifié (Cook & Nachtsheim 1980). La sélection des k -valeur est difficile et dans la plupart des cas dépendent de problème proposé. Une valeur commune, qui est également suggéré par Johnson & Nachtsheim (1983), est $k = 3$ ou $k = 4$. Meyer & Nachtsheim (1995) plus tard ont conseillé de sélectionner la valeur tel que:

$$k \leq \frac{n}{4}$$

La liste (5.4) montre la structure générale de l'algorithme k -échange qui réduit la liste des points de plan en sélectionnant uniquement les k points avec la plus faible variance de prédiction.

Générer un plan aléatoire avec un nombre de points désiré n ;

tant que des couples de points avec un Δ positive peuvent être trouvés alors

calculer la variance de prédiction $d(x_i)$ pour tous les points de plan ;

sélectionner k points de plan avec la variance de prédiction le plus faible;

pour $i = 1$ à n faire

calculer la variance de prédiction $d(x_i)$ pour le point x_i ;

pour $j = 1$ à N faire

calculer la variance de prédiction $d(x_j)$ pour le point x_j ;

calculer la fonction de variance $\bar{d}(x_i, x_j)$ pour le couple (x_i, x_j) ;

calculer la fonction delta $\Delta(x_i, x_j)$ pour le couple i ;

sélectionner le couple avec un delta maximal;

si le maximum des deltas est positif alors

si il ya plusieurs couples avec la même valeur de delta maximal alors

 sélectionner aléatoirement un couple ;

changer le point sélectionné x_i avec x_j ;

mettre à jour la matrice d'informations et la matrice de dispersion ;

réinitialiser delta maximal ;

Algorithme 5.4: algorithme k -échange (Johnson & Nachtsheim (1983)).

Considérons l'exemple avec $n = 5$ et $N = 20$, le k -échange accélère le processus de sélection en réduisant le nombre complet pour le calcul de Δ . Similaire à la procédure de Fedorov modifiée, chaque itération de l'algorithme de base Fedorov est décomposée en Δ -calculs pour chaque point de plan. Mais en fonction de k -valeur défini, l'algorithme tient compte que certains points de plan. Avec $k = 3$ l'algorithme k -échange calcule trois fois les 20 Δ -valeur et effectue jusqu'à trois échange.

Cette approche a une plus grande efficacité de calcul par rapport à la procédure de Fedorov modifié. En particulier pour les données de taille n grande, la définition de k -valeur a un impact. La qualité du plan n'est pas toujours aussi bonne que celle de l'algorithme de Fedorov normale, mais en ce qui concerne le temps de calcul nécessaire, l'algorithme donne de bons résultats.

5.3.2.6 Algorithme kl -échange

Le kl -échange est également une modification de l'algorithme de Fedorov et donc un type d'algorithme groupe-2. Développé par Atkinson & Donev (1989), cet algorithme permet de réduire la liste des points d'appui et d'échange avant de calculer la valeur Δ pour tous les couples possibles. L'utilisation du k point ayant la plus faible variance de prédiction est similaire à la procédure de k -échange. En plus de cela, nous ne considérons que les l candidats avec la plus grande variance de prédiction entre les points d'appui. En définissant $k = n$ et $l = N$, cet algorithme est la procédure de Fedorov. A l'initialisation normale de l'algorithme avec $k < n$ et $l < N$ réduit la quantité de calculer Δ et d'accélérer l'algorithme. Comme l'algorithme de Fedorov, le kl -échange sélectionne les points dont la valeur Δ est la plus grande et effectue un seul échange. Même si le nom kl -échange apparaît comme une modification de k -échange, nous devons observer que le kl -procédure s'appuie sur l'algorithme de Fedorov et n'effectue pas l'échange multiple au cours d'une itération. Seule la notion de choisir le point k de la plus faible variance de prédiction est similaire à l'algorithme de k -échange.

Atkinson & Donev (1989) ont défini deux modifications de cet algorithme. La première est semblable à l'idée de base de l'algorithme Fedorov modifié par Cook & Nachtsheim (1980) et réalise plus d'une échange au cours de chaque itération. Avec la sélection de $l = 1$, cela a changé kl -échange et devient soit l'algorithme de Fedorov modifié avec $k = n$ ou k -échange avec $k < n$. La seconde modification modifie les critères de sélection pour les points k et l .

Au lieu d'un choix dépend de la variance, l'algorithme sélectionne de façon aléatoire parmi les points de plan et les points d'appui (Atkinson & Donev 1989). Dans la liste (5.5), la version de base est montrée sans aucune modification.

Générer un plan aléatoire avec un nombre de points désiré n ;

```

tant que des couples de points avec un  $\Delta$  positive peuvent être trouvés alors
  calculer la variance de prédiction  $d(x_i)$  pour tous les points de plan ;
  sélectionner  $k$  points de plan avec la variance de prédiction le plus faible;
  calculer la variance de prédiction  $d(x_j)$  pour le point  $x_j$ ;
  sélectionner  $l$  points de plan avec la variance de prédiction le plus grande;
  pour  $i = 1$  à  $k$  faire
    calculer la variance de prédiction  $d(x_i)$  pour le point  $x_i$ ;
    pour  $j = 1$  à  $l$  faire
      calculer la variance de prédiction  $d(x_j)$  pour le point  $x_j$ ;
      calculer la fonction de variance  $d(x_i, x_j)$  pour le couple  $(x_i, x_j)$ ;
      calculer la fonction delta  $\Delta(x_i, x_j)$  pour le couple ;
      vérifier si le delta est le maximum et enregistrer le couple  $(x_i, x_j)$ ;
    si le maximum des deltas est positif alors
      si il ya plusieurs couples avec la même valeur de delta maximal alors
        sélectionner aléatoirement un couple ;
      changer le point sélectionné  $x_j$  avec  $x_i$  ;
      mettre à jour la matrice d'informations et la matrice de dispersion ;
      réinitialiser delta maximal ;
  
```

Algorithme 5.5 : Algorithme kl-échange (atkinson & Donev (1989))

Par rapport à l'algorithme de Fedorov, le kl-échange réduit la quantité de calcul de Δ entre chaque itération. Pour notre exemple à $n = 5$, $N = 20$ et $k = l = 3$ conduit au calcul de seulement $k * l = 3 * 3 = 9$ valeurs de Δ . En comparaison avec les 100 couples de l'algorithme de Fedorov, cette approche kl accélère le calcul d'une manière efficace. Certes, nous devons observer que la quantité d'itérations effectuées dans kl-échange peut être plus élevée parce que tous les couples sont pris en considération pour un échange. Mais en général, l'algorithme kl-échange crée des modèles très rapides et dans la plupart des cas donnent des résultats acceptables compte tenu de la D-optimalité.

CHAPITRE 6

COMPARAISON DES MATRICES D'EXPERIENCES

Le but de ce chapitre est de comparer les performances des matrices d'expériences pour tirer des indications qui puissent guider l'expérimentateur dans son choix.

Comparaison des qualités des matrices d'expériences

Dans les différents tableaux qui regroupent les caractéristiques des matrices d'expériences, nous avons reporté, pour chaque matrice d'expérience: le nombre d'expériences ('nb. exp'), le nombre de niveaux pris par chaque variable ('nb. niveaux'), le déterminant de la matrice de dispersion $|(\text{}^tXX)^{-1}|$ (critère optimal-D), la trace de $\text{}^tXX$ (critère A), le déterminant de la matrice des moments noté $|M|$ (critère M), la valeur maximale de la fonction de variance de prédiction d_{max} et l'efficacité-G (critère G).

6.1 Modèle polynomial du premier degré

Un modèle polynomial du premier à deux variables comporte quatre coefficients à estimer ($p=4$). Les valeurs des critères de qualité calculées pour chacune des matrices d'expériences sont regroupées dans le tableau (6.1).

Tableau 6.1 : Caractéristiques des matrices d'expériences du premier degré.

Nom	Plan factoriel 2^2	Plackett et Burman (N=4)	Plan factoriel 2^3	Plackett et Burman (N=8)	Plackett et Burman (N=12)
Nb. exp	4	4	8	8	12
Nb. niveaux	2; 2	2	2; 2	2	2
$ (\text{}^tXX)^{-1} $	$3.9 \cdot 10^{-3}$	$15.6 \cdot 10^{-3}$	$5.96 \cdot 10^{-8}$	$4.77 \cdot 10^{-7}$	$1.35 \cdot 10^{-12}$
Trace $ (\text{}^tXX)^{-1} $	1	0.75	1	0.88	0.92
$ M $	1	1	1	1	1
d_{max}	1	0.75	1	0.88	0.92
Eff-G(%)	100	100	100	99	99

6.2 Modèle polynomial du second degré

6.2.1 Matrices d'expérience à deux variables

Un modèle polynomial du second degré à deux variables comporte six coefficients à estimer ($p=6$). Les valeurs des critères de qualité, calculées pour chacune des matrices d'expériences, sont regroupées dans le tableau 6.2.

Tableau 6.2 : Caractéristiques des matrices d'expériences du second degré à deux variables.

Nom	Plan factoriel 3^2	Composite ($\alpha = 1.414$)	Doehlert
Nb. exp	9	11	7
Nb. niveaux	3 ; 3	5 ; 5	5 ; 3
$ (\text{}^tXX)^{-1} $	$1.92 \cdot 10^{-4}$	$1.02 \cdot 10^{-5}$	$3.2 \cdot 10^{-2}$
Trace $ (\text{}^tXX)^{-1} $	2.1389	1.19	6.0001
$ M $	$9.754 \cdot 10^{-3}$	$5.5 \cdot 10^{-2}$	$2.52 \cdot 10^{-4}$
d_{max}	0.81	0.63	1
Eff-G(%)	82.30	86.58	85.71

D'après le tableau 6.2, si on compare les propriétés des trois plans : plan factoriel complet à trois niveaux, plan composite et plan de Doehlert.

- Pour le plan composite le déterminant de $(\text{}^tXX)^{-1}$ est petit par rapport aux autres plans. Suivant le critère de D-optimalité le plan composite est meilleur. Il en est de même pour la trace de $(\text{}^tXX)^{-1}$ et donc le critère A est mieux respecté.
- On voit que le déterminant de la matrice des moments pour le plan composite est plus grand que celui des deux autres plans. Suivant le critère M le plan composite est plus efficace.
- De même les valeurs de d_{max} et $Eff - G(\%)$ montrent que le plan composite est toujours meilleur.

6.2.2 Matrices d'expérience à trois variables

Le nombre de coefficients à estimer est de dix ($p=10$), les caractéristiques des matrices d'expérience à trois variables sont rassemblées dans le tableau 6.3.

Tableau 6.3 : Caractéristiques des matrices d'expériences du second degré à trois variables.

Nom	Plan factoriel 3^3	Plan Box-Behnken	Plan composite ($\alpha = 0.9$)	Plan Requemore 311A	Plan Requemore 311B
Nb. exp	27	15	15	11	11
Nb. niveaux	3 : 3 : 3	3 : 3 : 3	5 : 5 : 5	5 : 5 : 5	5 : 5 : 5
$ ({}^tXX)^{-1} $	$1.702 \cdot 10^{-11}$	$3.97 \cdot 10^{-8}$	$7.44 \cdot 10^{-11}$	$2.01 \cdot 10^{-11}$	$2.16 \cdot 10^{-11}$
Trace $ ({}^tXX)^{-1} $	1.18	2.27	2.08	12	11.8
$ M $	$2.85 \cdot 10^{-4}$	$4.364 \cdot 10^{-5}$	$2.33 \cdot 10^{-2}$	$1.4 \cdot 10^{-4}$	$1.75 \cdot 10^{-4}$
d_{max}	0.51	0.73	0.99	1.16	1
Eff-G(%)	72.73	91.32	67.34	78.62	90.90

- Les deux matrices suivantes, les matrices Requemore 311A et Requemore 311B, la seconde est meilleure à tous points de vue, et les caractéristiques de cette matrice est excellentes ($d_{max}=1$; Eff-G=90%).
- Pour le plan factoriel complet à trois niveaux le déterminant de $({}^tXX)^{-1}$ est le plus petit par rapport les autres plans (le critère-D est satisfait). De même pour la trace de $({}^tXX)^{-1}$ est plus petite et donc le critère A est mieux respecté.
- On voit que la valeur de d_{max} est la meilleur pour les plan factoriel complet à trois niveaux aussi.
- D'après le critère M on voit que la matrice d'expérience du plan composite pour trois facteurs est plus efficace que les autres plans.
- Et pour Eff - G(%) le plan Box-Behnken est meilleur.

6.2.3 Matrices d'expérience à quatre variables

Les matrices d'expérience à quatre variables regroupées dans le tableau 6.4 permettent l'estimation des quinze coefficients du modèle polynomial du second degré correspondant ($p=15$).

Tableau 6.4: Caractéristiques des matrices d'expériences du second degré à quatre variables.

Nom	Roquemore 416B	Roquemore 416C	Composite orthogonale ($\alpha = 1.21$)	Composite retatable ($\alpha = 1.414$)	Box- Behnken
Nb. exp	16	16	25	25	25
Nb. niveaux	5 ; 5 ; 5 ; 4	5 ; 5 ; 5 ; 4	5 ; 5 ; 5 ; 5	5 ; 5 ; 5 ; 5	3 ; 3 ; 3 ; 3
Trace $ ({}^tXX)^{-1} $	18.80	18.49	15.22	13.33	13.33
$ M $	$7.7 \cdot 10^{-9}$	$8 \cdot 10^{-8}$	$1.5 \cdot 10^{-8}$	$9.3 \cdot 10^{-8}$	$9.3 \cdot 10^{-8}$
d_{max}	1.27	1.21	1.78	1	1
Eff-G(%)	74.03	77.49	33.74	60	60

Les matrices 416B et 416C ont des caractéristiques très proches, sont satisfaisantes, avec une légère préférence pour la 416C.

Parmi les matrices d'expériences à vingt-cinq points, les matrices composite retatable et de Box-Behnken, bien que différentes, ont les mêmes caractéristiques. Ces deux matrices ont un déterminant de matrice M et fonction de variance maximale sont très satisfaisantes ($|M|=9.3 \cdot 10^{-8}$; $d_{max}=1$). Ici encore, le choix de la matrice dépend des contraintes sur le nombre de niveaux des variables. En revanche, la matrice d'expérience composite orthogonale est de qualité médiocre (Trace $|({}^tXX)^{-1}|=15.22$; $|M|=1.5 \cdot 10^{-8}$; $d_{max}=1.78$; Eff-G(%)=33.74).

CONCLUSION GENERALE

La méthodologie de la Recherche Expérimentale, propose une démarche à l'expérimentateur prêt à prendre le temps de la réflexion pour engager une stratégie optimale qui doit lui donner avec la plus grande précision possible les réponses aux questions posées.

Aussi, la planification expérimentale devient une méthode, voire une philosophie, pour aborder un problème. La méthodologie intervient pour formaliser les limites déterminées par l'expérimentateur (qui doit rester le seul maître des décisions), et pour vérifier si la décision finale est bien en accord avec les objectifs fixés.

Ce travail est le résultat d'une synthèse où de nombreux outils existants interviennent dans le but d'aider l'expérimentateur à diriger lui-même et méthodiquement sa recherche. Pour cela, nous avons voulu tenir compte du choix de la démarche à l'interprétation des résultats, des critères mathématiques et des idées subjectives des expérimentateurs. Nous avons préféré volontairement modérer « légèrement » l'importance de critères mathématiques stricts qui permettent d'obtenir des résultats rigoureux, pour introduire les objectifs personnels de l'expérimentateur et ainsi proposer des solutions correspondant un peu plus à la réalité des objectifs qui parfois restent difficilement modélisables pour être introduits dans les algorithmes de recherche.

APPENDICE A

LISTE DES SYMBOLES ET ABREVIATION

a_0	: Constant d'un modèle
a_i	: Coefficients des termes de premier degré
$a_{i,i}$: Coefficients des termes carrés
$a_{i,j}$: Coefficients des termes rectangles
\hat{a}_i	: Coefficients de la variables x_i dans un modèle déterminé pa la méthode des moindres carrés
A	: Vecteur des coefficients du modèle
\hat{A}	: Vecteur des coefficients déterminés par la méthode des moindres carrés
α	: Distance au centre des points en étoile dans un plan composite, niveau attribué
β	: Distance au centre des points expérimentaux pour le $2^{ième}$ plan proposé
c_i	Contraste du facteur i
$d(\hat{y}_u)$: Fonction d'erreur de prédiction
$d^2(\hat{y}_u)$: Fonction de la variance de prédiction
ddl	: Degré de liberté
Δ	: L'écart d'ajustement
e_i	: Ecart résiduel
ε_i	: Valeur vraie de l'écart e_i
e	: Vecteur des écarts
E	: Esperance mathématique
$Eff - G$: Efficacité vis-à-vis du critère G
f	: Fonction
F	: F de Fisher-Snedecor
h	: Le nombre de composantes à retenir dans la régression PLS
h_i	: Le $i^{ième}$ élément diagonal de la matrice $X({}^tXX)^{-1} {}^tX$
I	: Matrice unité, générateur d'aliases, moyenne dans un tableau des effets
k	: Nombre de facteurs
n_0	: Nombre d'essais au centre du domaine d'étude
n_f	: Nombre d'essais du plan factoriel
n_α	: Nombre d'essais du plan étoile
N	: Nombre d'essais d'un plan d'expériences

Bibliographie

1. B.GOVARETS « Plans optimaux » article, Institut de statique, université catholique de Louvain.
2. B.Govarets 'Plans d'expériences pour traiter des problèmes de mélange' article, Institut de statistique, université catholique de Zouvain.
3. Box, G.E.P. Behnken, D.W., 'Some new three level designs for the study of quantitative variables' Technometrics, Vol. 2, P.455-475 (1960).
4. Dohler, D. H., 'Uniform shell design' Appl. Stat, n° 19, P.231 (1970).
5. El Mossawi Hicham ' Proposition de plans d'expériences dans la méthodologie de la recherche expérimental : Etude et programmation' Thèse de magister, université Blida (2004).
6. Franquart, P, 'Optimisation multicritère et méthodologie de la recherche expérimental' Thèse en sciences, Université d'Aix-Marseille (1992).
7. Goupy, J., 'Plans d'expériences', Techniques de l'Ingénieur, traité analyse et caractérisation et caractérisation, P. 230, P. 1-20 (1997).
8. Goupy. J 'La méthode des plans d'expériences', DUNOD Paris, P. 303 (1998).
9. Goupy, J 'Modélisation par les plans d'expériences', P.409 DUNOD Paris (1999).
10. Goupy, J. 'Introduction aux plans d'expériences' DUNOD Paris, P.290 (2001).
11. Goupy, J. 'Plans d'expériences pour surface de réponse', P.341, 409, DUNOD Paris (1999).
12. Goupy, J. 'Etude comparative de divers plans d'expériences' Rue de statistique appliquée : Vol. 38, n°4, P. 5-44 (1990).
13. Goupy, J. 'Plans d'expériences-Les mélanges_Goupy', DUNOD Paris (1999).
14. Gill, L., 'Plans d'expériences en gestions industrielle', Lavoisier, Paris (2003).
15. J.Antony, 'Design of experiments for Engineer and scientists', (2003).
16. HANNAN, F, 'Plans d'expériences pour mélanges à deux niveaux et facteurs externes' Thèse de Doctorat en mathématique appliquées.
17. Mathieu, D., Phan-Tan-Lun, R., Sergent, M., 'Méthodologie de la recherche expérimentale' ZPRAI université d'Aix-Merseille (1998).
18. Maurice, P., 'Les plans d'expériences par la méthode Taguchi' rue thénard, Paris (1997).

- P : Nombre de facteur supplémentaires pour un plan fractionnaire
 P_h : Vecteur des pentes des droites des moindres carrés passant par l'origine (Régression PLS)
 q : Nombre de coefficients d'un modèle mathématique
 Q^2 : Coefficient de la prédiction
 s^2 : Estimation de variance
 σ_y : Ecarte-type de l'erreur sur les réponses mesurées
 σ^2 : Variance des écarts
 $\sigma(\hat{\alpha}_i)$: Ecart-type des coefficients déterminés par la méthode des moindres carrés
 $var(\hat{A})$: Variance des coefficients estimés par la méthode des moindres carrés
 $var(\hat{y})$: Variance de l'erreur sur les réponses mesurées
 x_i : Niveau attribué aux différents facteurs et exprimé en valeurs centrées réduite, variable i
 x_{ij} : Élément de la ligne i et de la colonne j de la matrice X
 x_u : Vecteur modélisé du point u
 X : Matrice de calcul
 tX : Matrice transposée de X
 $({}^tXX)$: Matrice d'information
 $({}^tXX)^{-1}$: Matrice de dispersion
 y_i : Réponse de l'essai n° i
 y : Vecteur des réponses observées
 \hat{y}_i : Réponse prédite de l'essais i
 \hat{y} : Vecteur des réponses prédites
 y_+ : Moyenne des réponses au niveau haut d'un facteur
 y_- : Moyenne des réponses au niveau bas d'un facteur
 Y : Matrice des réponses observées
 z : Niveau attribué au facteur exprimé en unité courante
 z_0 : Valeur centrale de l'intervalle $[-1, +1]$ exprimé en unité courante

19. Peissik, A., 'Methodologie de la recherche expérimentale : propriétés et caractérisations des matrices d'expériences pour les méthodes polynomiaux du seconde degré Thèse en sciences, Université d'Aix-Marseille (1995).
20. Roquemore, K.G, 'Hybrid designs for quadratic Response surfaces', Techno metrics, Vol. 18, n°4, P.419-423 (1976).
21. Schimmerling, P., Sisson, J. C., Zaidi, A ; 'Pratique des plans d'expériences', Edition Tec Doc, Paris, P. 526 (1998).
22. Stephane GAZUT 'Conception et mise en œuvre de nouvelle méthodes d'élaboration de plans d'expériences pour l'apprentissage de modèles.
23. Yuzhu, H 'Program for Doehlert matrix design of experiments', Trends in Analytical Chemistry, Vol. 8, n°4 (1989).