

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Frères Mentouri - Constantine 1

Faculté des Sciences de la Nature et de la Vie
Département de Biologie Appliquée



جامعة الأخوة منتوري قسنطينة
UNIVERSITÉ DES FRÈRES
MENTOURI CONSTANTINE

Cours

Plans d'expériences



Réalisé par :

Dr. Mohamed Skander DAAS

Destiné aux étudiants du Master 1 Bio-industrie, Analyse et Contrôle

Année universitaire 2019/2020

Préface

Ce cours couvre les sujets de base en conception expérimentale et en analyse et est destiné aux ingénieurs de laboratoire et aux étudiants des cycles supérieurs et avancés. Les étudiants devraient avoir des connaissances de base sur les statistiques et quelques notions mathématiques telles que les tests statistiques, la régression linéaire etc.

Le but de ce cours est de montrer la puissance et l'utilité des plans d'expériences tout en expliquant comment planifier et analyser une étude expérimentale. C'est aussi une tentative de dissiper l'idée selon laquelle les PEX sont réservés aux personnes ayant une formation avancée en mathématiques. Il sera démontré que les PEX sont avant tout un outil logique qui peut être facilement saisi et appliqué, ne nécessitant que des compétences de base en mathématiques. Bien que les logiciels rendent les calculs plus simples et offrent une plus grande polyvalence, il est nécessaire de comprendre ce que le logiciel fait. À cette fin, aucun logiciel n'est utilisé avec ce texte, mais des calculatrices sont utilisées pour s'assurer que les bases sont apprises. L'intention est d'introduire les bases, de convaincre le lecteur de la puissance de cet outil, puis de recommander des ressources pour des études ultérieures.

Mots-clé :

plans d'expériences, factoriels, fractionnaires, surfaces de réponse, criblage, modélisation, optimisation.

Table des matières

1	Introduction	1
1.1	Qu'est-ce que les plans d'expériences (<i>PEX</i>)?	3
1.2	Termes et concepts communs des plans d'expériences	3
1.3	Cas d'utilisation des plans d'expériences	6
1.4	Objectifs d'utilisation des plans d'expériences	6
1.4.1	Objectif de comparaison d'alternatives	6
1.4.2	Objectif de criblage	7
1.4.3	Objectif de modélisation	7
1.5	Quelles sont les étapes des plans d'expériences?	8
1.6	Classes des plans d'expériences	9
2	Plans factoriels	11
2.1	Plans factoriels complets à deux niveaux	13
2.2	Matrice d'expériences	13
2.3	Effets globaux et effets moyens	15
2.4	Exemple applicatif	16
2.5	Représentation graphique des effets	17
2.5.1	Graphiques des effets principaux	17
2.5.2	Graphiques des effets d'interactions	18
2.6	Autres plans à deux niveaux	20
2.6.1	Les plans de <i>Koshal</i>	20
2.6.2	Les plans de <i>Rechtschaffner</i>	20
2.6.3	Les plans de <i>Plackett et Burmann</i>	20
2.7	Plans à plusieurs niveaux	21

2.7.1	Les tables de <i>Taguchi</i>	21
2.7.2	Carrés latins	21
2.7.3	Carrés de <i>Youden</i>	22
3	Signification des effets et validation du modèle	23
3.1	Test de signification des effets	25
3.1.1	Calcul de la variance des coefficients S_i^2	25
3.1.2	Détermination de la valeur critique t_{crit}	26
3.1.3	Réalisation du test et interprétation	26
3.1.4	Exemple applicatif	28
3.2	Régression linéaire et analyse du modèle	30
3.2.1	Régression linéaire	30
3.2.2	Test de validation du modèle	31
3.2.3	Quelques mesures statistiques d'évaluation du modèle	32
3.2.4	Exemple applicatif	35
4	Plans factoriels fractionnaires	37
4.1	Introduction	39
4.2	Plans factoriels fractionnaires	39
4.3	Notion d'alias et de contraste	40
4.3.1	Résolution	42
4.3.2	Génération d'alias	43
4.3.3	Contrastes h_i	44
4.4	Estimation des effets et des interactions a_i	44
4.5	Exemple applicatif	45
5	Plans pour surface de réponse	47
5.1	Introduction	49
5.2	Les plans composites centrés	50
5.3	Les plans de <i>Box-Behnken</i>	51
5.4	Les plans optimaux	52
6	Plans de mélange	55

6.1	Introduction	57
6.2	Types des plans de mélange	57
6.2.1	Plans de mélange en réseau	57
6.2.2	Plans de mélange centrés	59
6.2.3	Plans de mélange contraints (Sommets extrêmes)	61
6.3	Combinaison des plans de mélange avec d'autres plans	62
	Bibliographie	65

Table des figures

1	Introduction	1
1.1	Entrées et sorties d'une expérience	3
1.2	Exemple de facteurs, de niveaux et de réponses dans un plan d'expériences	3
1.3	Domaines de variation naturels / codés	5
1.4	Exemple graphique de prédiction de la réponse	7
1.5	Exemple graphique de l'optimisation de la réponse	8
2	Plans factoriels	11
2.1	Plans factoriels complets 2^2 et 2^3	13
2.2	Matrices d'expériences pour les plans factoriels complets 2^2 et 2^3	14
2.3	Matrice d'expériences avec réponses pour un plan factoriel complet 2^2	15
2.4	Matrice d'expériences avec effets pour un plan factoriel complet 2^2	16
2.5	Un exemple de deux domaines d'étude de deux facteurs	16
2.6	Matrice d'expériences avec 2^2 réponses de l'exemple étudié	17
2.7	Matrice d'expériences avec effets de l'exemple étudié	17
2.8	Domaines d'étude des facteurs T et D	18
2.9	Représentation graphique des effets principaux T et D	18
2.10	Représentation graphique de l'effet de l'interaction TD	19
2.11	Plan factoriel fractionnaire 2^{3-1}	19
2.12	Plan de Koshal pour trois facteurs	20
2.13	Plan Carrés latins pour 3 facteurs	21
3	Signification des effets et validation du modèle	23
3.1	Lecture d'une valeur t_{crit} depuis la table de <i>Student</i>	26

3.2	Plan factoriel fractionnaire 2^{3-1}	28
3.3	Matrice d'expériences avec effets de l'exemple étudié	28
3.4	Résultats des réponses estimées et des écarts au carré	29
3.5	Résultats de signification des effets	29
3.6	Résultats des réponses estimées et des écarts au carré	35
4	Plans fractionnaires	37
4.1	Plan factoriel fractionnaire 2^{3-1}	39
4.2	Plan factoriel fractionnaire 2^{5-2} à partir un plan factoriel complet 2^3	41
4.3	Matrices d'expériences avec contrastes pour un plan fractionnaire 2^{5-2}	43
4.4	Matrice d'expériences avec réponses pour le plan fractionnaire 2^{5-2}	45
4.5	Matrice d'expériences avec contrastes pour le plan fractionnaire 2^{5-2}	45
5	Plans pour surface de réponse	47
5.1	Plan composite centré pour 3 facteurs	50
5.2	Matrice d'expériences pour un plan composite centré pour 3 facteurs	51
5.3	Plan de <i>Box-Behnken</i> pour 3 facteurs	51
5.4	Matrice d'expériences pour un plan de <i>Box-Behnken</i> pour 3 facteurs	52
5.5	Matrice du plan optimal obtenu pour l'exemple étudié	53
6	Plans de mélange	55
6.1	Valeurs des composants dans un plan de mélange	57
6.2	Plan de mélange en réseau (3, 2)	58
6.3	Matrice d'expériences pour un plan de mélange en réseau (3, 2)	58
6.4	Plan de mélange en réseau (3, 2) augmenté	59
6.5	Matrice d'expériences pour un plan de mélange en réseau (3, 2) augmenté	59
6.6	Plan de mélange centré à 3 composants	60
6.7	Matrice d'expériences pour un plan de mélange centré à 3 composants	60
6.8	Plan de mélange centré à 3 composants (augmenté)	61
6.9	Matrice d'expériences pour un plan de mélange centré à 3 composants (augmenté)	61
6.10	Limites inférieures et supérieures dans un plan de mélange contraint	62
6.11	Un plan combiné de variables factorielles et de composants de mélange	63

Liste des tableaux

1 Introduction	1
1.1 Guide de sélection d'un plan	9
2 Plans factoriels	11
3 Signification des effets et validation du modèle	23
3.1 Table de <i>student</i>	27
3.2 Tableau d'analyse de la variance	32
3.3 Table de Fischer - Snedecor (Risque $\alpha = 0.05$)	34
3.4 Tableau d'analyse de la variance de l'exemple étudié	35
4 Plans fractionnaires	37
5 Plans pour surface de réponse	47
6 Plans de mélange	55

Chapitre 1

Introduction

1.1 Qu'est-ce que les plans d'expériences (*PEX*) ?

Les plans des expériences (*PEX*) est une branche de la statistique appliquée axée sur l'utilisation de la méthode scientifique pour la planification, la réalisation, l'analyse et l'interprétation de données issues d'essais ou d'expériences contrôlés. *PEX* est une méthodologie mathématique utilisée pour planifier et mener efficacement une étude scientifique qui modifie simultanément les variables d'entrée (X) appelées **Facteurs** afin de révéler leurs effets sur une variable de sortie (ou plusieurs) (Y) appelée **Réponse**.

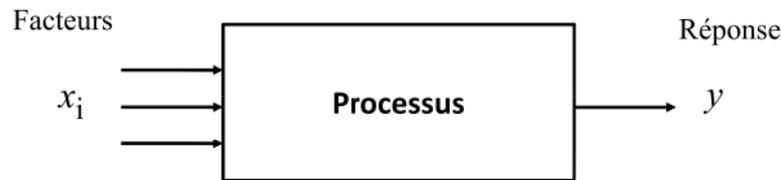


FIGURE 1.1: Entrées et sorties d'une expérience

En combinant plusieurs variables dans une étude au lieu de créer une étude pour chacune, le nombre d'essais requis sera considérablement réduit et une meilleure compréhension du processus est obtenue. Ceci est en contraste direct avec l'approche typique d'un facteur à la fois (*UFAF*), qui limite la compréhension et gaspille les données. En outre, les études *UFAF* ne peuvent pas assurer la détection des effets de combinaisons des facteurs (une condition qui sera ultérieurement définie comme une *interaction*).

1.2 Termes et concepts communs des plans d'expériences

Considérons le schéma suivant d'un processus de cuisson d'un gâteau (Figure 1.2). Il y a trois aspects dans un processus d'un plan d'expériences :

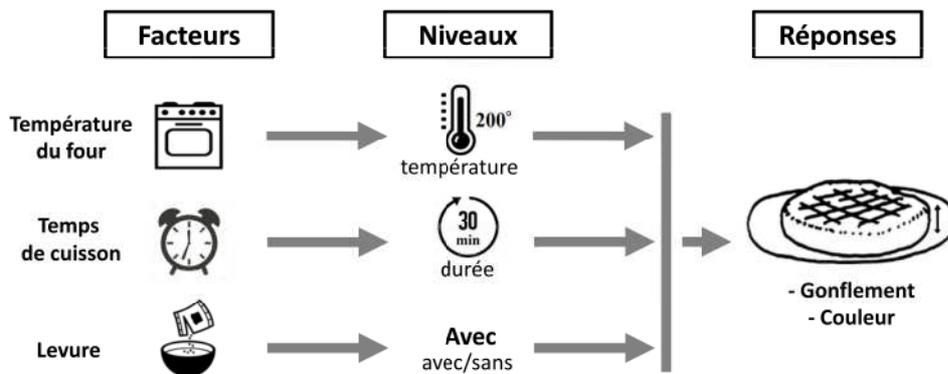


FIGURE 1.2: Exemple de facteurs, de niveaux et de réponses dans un plan d'expériences

- **Réponses** ou sorties de l'expérience : Dans le cas de la cuisson du gâteau, le goût, la consistance et l'apparence du gâteau sont des résultats mesurables potentiellement influencés par les facteurs et leurs niveaux respectifs. Les expérimentateurs souhaitent souvent éviter d'optimiser le processus pour une réponse au détriment d'une autre. Pour cette raison, les résultats importants sont mesurés et analysés afin de déterminer les facteurs et leurs niveaux qui fourniront le meilleur résultat global pour les caractéristiques essentielles de la qualité.
- **Facteurs** ou entrées pour le processus : Les facteurs peuvent être classés en variables contrôlables ou non contrôlables. Dans ce cas, les facteurs contrôlables sont les ingrédients du gâteau et le four dans lequel le gâteau est cuit. Les variables contrôlables seront appelées facteurs. Notez que la liste des ingrédients a été raccourcie pour cet exemple. De nombreux autres ingrédients pourraient avoir un impact significatif sur le résultat final (huile, eau, arôme, etc.). De même, il pourrait exister d'autres types de facteurs, tels que la méthode ou les outils de mixage, la séquence de mixage, etc. Les gens considèrent généralement un facteur de bruit comme un facteur incontrôlable qui provoque la variabilité dans des conditions de fonctionnement normales, mais dans certains cas ce facteur peut être contrôlé pendant l'expérience en utilisant le blocage et la randomisation.

Pour la température, le facteur est dit **quantitatif**, c'est un facteur qui a une valeur donnée. Par contre, si par exemple, nous voulons étudier l'effet du zeste de citron sur le goût du gâteau, en réalisant des expériences sans zeste de citron et des expériences avec zeste de citron, dans ce cas, le facteur « zeste de citron » est dit **qualitatif**, le niveau bas et le niveau haut correspondront aux deux modalités du facteur : -1 pour « sans » et +1 pour « avec ». Toutefois ce même facteur peut être quantitatif si nous étudions sa quantité au lieu de son absence/présence.

- **Niveaux et domaine d'étude** de chaque facteur : En plans d'expériences, lorsque nous étudions l'influence d'un facteur, en général, nous limitons sa variation entre deux bornes appelées respectivement : **niveau bas** (-1) et **niveau haut** (+1) qui définissent le domaine d'étude de ce facteur. Si nous considérons 5 niveaux de la température (140 160 180 200 220), son domaine d'étude sera [140, 220].

NB : Il est très important de bien choisir le domaine d'étude au début de chaque application des PEX, car les résultats trouvés et le modèle qui relie la réponse aux facteurs n'est valable que pour les valeurs appartenant au domaine d'étude fixé.

- **Interaction** : Les interactions se produisent lorsque l'impact d'un facteur dépend de la valeur d'un second facteur.
- **Blocage et réplication** : Le blocage est une technique expérimentale visant à éviter toute variation indésirable du processus d'entrée ou du processus expérimental. Par exemple, une expérience peut être menée avec le même équipement pour éviter toute variation de l'équipement. Les praticiens répliquent également des expériences, en effectuant la même combinaison plusieurs fois, afin d'obtenir une estimation de la quantité d'erreur aléatoire pouvant faire partie du processus.

- **Notion de modèle et de régression linéaire multiple** : La régression linéaire multiple est une méthode d'analyse de données quantitatives. Elle a pour but de mettre en évidence la liaison pouvant exister entre une variable dite expliquée (réponse), que nous la noterons y et plusieurs autres variables dites explicatives (facteurs) que nous les noterons x_1, x_2, \dots, x_k .

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

L'équation 1.2 montre un exemple d'un modèle quadratique.

$$y = \alpha_0 + \sum \alpha_i x_i + \sum \alpha_{ij} x_i x_j + \sum \alpha_{ii} x_i^2 + \epsilon \quad (1.2)$$

- **Variables codées** : L'utilisation des variables centrées réduites ou codées (-1, +1) présente l'intérêt de pouvoir généraliser la théorie des plans d'expériences quels que soient les facteurs ou les domaines d'études retenus. Utiliser les variables codées va permettre d'avoir pour chaque facteur le même domaine de variation [-1, +1] et de pouvoir ainsi comparer entre eux les effets des facteurs, et ce sont ces variables codées (sans dimensions et entre -1 et +1) qui sont utilisées dans l'équation du modèle obtenu. Exemple : nous choisissons pour le facteur température un domaine d'étude : [10°C, 40°C], nous affectons la valeur -1 à 10°C et +1 à 40°C et nous notons V_c la valeur codée et V_n la valeur naturelle. La figure 1.3 montre la correspondance proportionnelle entre les valeurs codées et naturelles.

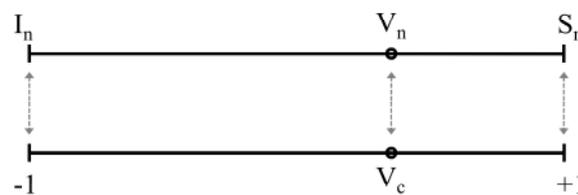


FIGURE 1.3: Domaines de variation naturels / codés

Où S_n et I_n sont respectivement la limite supérieure et inférieure du domaine d'étude de la variable température. Les variables centrées réduites sont sans unité. Une température de 20°C correspond à une variable centrée réduite de -0.33. Les formules de conversion sont :

$$V_c = \frac{2V_n - (S_n + I_n)}{(S_n - I_n)} \quad (1.3)$$

$$V_n = \frac{V_c(S_n - I_n) + (S_n + I_n)}{2} \quad (1.4)$$

- **Les tests d'hypothèses** : permettent de fournir une règle de décision. Une hypothèse offre deux possibilités : l'hypothèse nulle et l'hypothèse alternative. C'est l'hypothèse nulle qui est soumise au test et toute la démarche du test s'effectue en considérant cette hypothèse comme vraie. Les tests sont effectués à un niveau de signification basé sur une probabilité.

1.3 Cas d'utilisation des plans d'expériences

Un *PEX* permet à l'expérimentateur de manipuler plusieurs entrées pour déterminer leurs effets sur la sortie de l'expérience ou du processus.

- De nombreuses expériences maintiennent certains facteurs constants et de modifier le niveau d'une autre variable. L'utilisation de cette approche "un facteur à la fois" (OFAT) pour traiter les données est toutefois inefficace par rapport à l'évolution simultanée des niveaux des facteurs. En réalisant une expérience multi-factorielle, le *PEX* peut révéler des **interactions** critiques entre les facteurs qui sont souvent omises lors de la réalisation d'une expérience simple. Alors, les *PEX* peuvent révéler des problèmes cachés au cours du processus, ces problèmes cachés sont généralement associés aux interactions des différents facteurs.
- Les ingénieurs seront en mesure d'identifier les facteurs qui ont le plus d'impact sur le processus et ceux qui ont le moins d'influence sur les résultats du processus c-à-d **déterminer** et **quantifier** les effets clés (effets principaux et effets des interactions) dans un processus donné. Ce qui permet une éventuelle **réduction** du nombre de facteurs à considérer en **négligeant** certains parmi ces facteurs.
- En utilisant correctement les plans d'expériences, le **nombre d'essais** ou de tests peut être considérablement **réduit**. Un plan d'expériences robuste peut faire gagner du temps et du coût lors d'une étude expérimentale.
- Les *PEX* permettent aussi de **modéliser** la réponse en fonction des différents facteurs ce qui définit une relation mathématique entre la réponse et ces facteurs et de fournir des informations statistiques sur la qualité du modèle obtenu (degré de son prédictibilité, robustesse, adéquation, etc). Cette modélisation permet de répondre à des questions du genre :
 - Dans quels paramètres le processus offrirait-il des performances acceptables ? (Optimisation)
 - Quelle sera la réponse du processus sous une configuration donnée des facteurs ? (Prédiction)

1.4 Objectifs d'utilisation des plans d'expériences

1.4.1 Objectif de comparaison d'alternatives

Dans le cas de notre exemple de la cuisson du gâteau, nous pourrions vouloir comparer les résultats de deux différents types de farine. S'il s'avérait que la farine des différents fournisseurs avait un résultat non significatif, nous pourrions choisir le fournisseur le moins coûteux. Si le résultat était significatif, nous choisirions celui de la meilleure farine.

1.4.2 Objectif de criblage

Il existe souvent de nombreux facteurs possibles, dont certains peuvent être critiques et d'autres qui ne peuvent avoir que peu ou pas d'effet sur la réponse. En tant que but en soi, il est peut être souhaitable de **réduire le nombre de facteurs** à un ensemble relativement petit (2 à 5), de manière à concentrer l'attention sur le contrôle de ces facteurs. Les expériences de criblage sont un moyen efficace, avec un nombre minimal de tests, pour déterminer les facteurs importants. Nous pourrions poser une question : Quels sont les facteurs importants qui ont un impact sur la couleur du gâteau parmi : la température, la durée de la cuisson, la quantité du sucre et de la farine, le nombre d'œufs et l'utilisation du zeste de citron ?

1.4.3 Objectif de modélisation

La méthodologie de surface de réponse est un ensemble de techniques mathématiques et statistiques pour la construction de modèles mathématiques empiriques. L'exploitation de ces modèles peut répondre à plusieurs objectifs :

- **Objectif de prédiction** : Un modèle mathématique obtenu peut être utilisé pour prédire une réponse à un point quelconque dans les limites des domaines expérimentaux. La précision des réponses obtenues dépend du degré de capacité de prédiction du modèle.

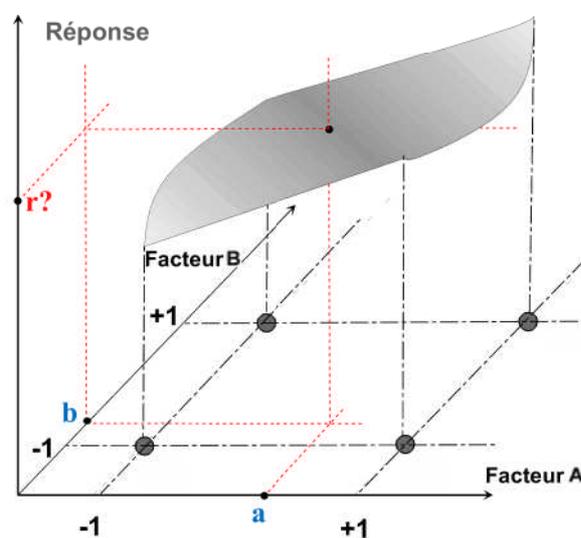


FIGURE 1.4: Exemple graphique de prédiction de la réponse

- **Objectif d'optimisation** : Une optimisation est effectuée pour déterminer les valeurs des entrées du processus à utiliser pour obtenir la sortie du processus souhaitée. Les objectifs d'optimisation habituels peuvent être de maximiser le rendement d'un processus, de minimiser le temps de traitement nécessaire à la fabrication d'un produit ou d'atteindre une spécification du produit cible.

Nous pourrions poser une question : Comment rendre le gâteau aussi humide que possible sans le désintégrer ?

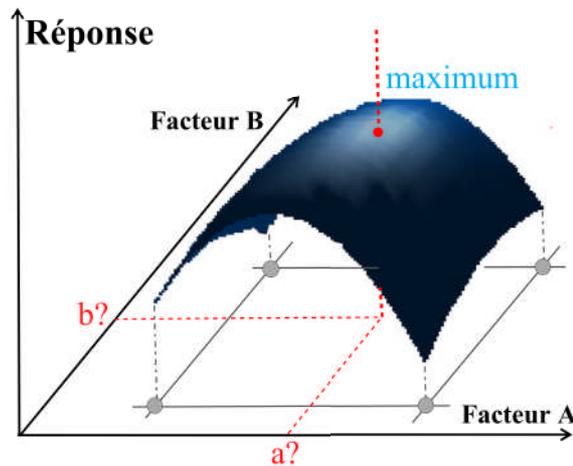


FIGURE 1.5: Exemple graphique de l'optimisation de la réponse

- **Objectif d'optimisation multi-réponse** (Équilibrer les compromis) : L'objectif est de déterminer les paramètres des facteurs permettant d'optimiser simultanément (un compromis) plusieurs réponses.

1.5 Quelles sont les étapes des plans d'expériences ?

Obtenir de bons résultats par un *PEX* implique les sept étapes suivantes :

- **Poser des objectifs** : Les objectifs d'une expérience sont mieux déterminés par une discussion en équipe. Tous les objectifs doivent être écrits, même ceux qui ne semblent pas très intéressants. Le groupe devrait discuter des objectifs clés et des objectifs "Bons mais pas vraiment nécessaires". La priorisation des objectifs vous aide à choisir la direction à prendre en ce qui concerne la sélection des facteurs, les réponses et un plan particulier.
- **Sélectionner les variables du processus et leurs plages** : Les variables d'un processus incluent à la fois les entrées et les sorties - c'est-à-dire les **facteurs** et les **réponses**. La sélection de ces variables s'effectue mieux en équipe. L'équipe devrait
 - Inclure tous les facteurs importants (selon le jugement de l'ingénieur).
 - Déterminer une plage d'étude pour chaque facteur.
 - Être prudent en choisissant les niveaux bas et haut des facteurs.
 - Vérifier les réglages des facteurs pour des combinaisons non pratiques ou impossibles.
 - Inclure toutes les réponses pertinentes.

- **Choisir un plan expérimental** : Le choix d'un plan expérimental dépend des objectifs de l'expérience, du nombre de facteurs à étudier et de la quantité de ressources disponibles. La table 1.1 montre un simple guide exemple de sélection d'un plan.

TABLE 1.1: Guide de sélection d'un plan

Nombre de facteurs	Objectif comparatif	Objectif de dépistage	Surface de réponse
1	Plan totalement aléatoire à 1 facteur	–	–
2 - 4	Plan de bloc aléatoire	Factoriel complet ou fractionnaire	Central composite ou Box-Behnken
5 ou plus	Plan de bloc aléatoire	Factoriel fractionnaire ou Plackett-Burman	Dépistage en premier pour réduire le nombre de facteurs

- **Exécuter le plan** : Les expériences du plan sont réalisées et les résultats des essais (réponses) sont rassemblés.
- **Vérifier que les données sont cohérentes avec les hypothèses expérimentales** : Dans tous les modèles, nous formulons des hypothèses et nous exigeons également que certaines conditions soient approximativement remplies à des fins d'estimation. Ceux-ci sont :
 - Les systèmes de mesure sont-ils capables de répondre à toutes vos questions ?
 - Vos réponses sont-elles susceptibles d'être bien approchées par de simples modèles polynomiaux ?
 - Les résidus (la différence entre les prédictions du modèle et les observations réelles) se sont-ils bien comportés ?
- **Analyser et interpréter les résultats** : En supposant l'existence d'un modèle de départ que nous souhaitons l'adapter à nos données expérimentales et que l'expérience ait été conçue correctement pour notre objectif, la plupart des progiciels *PEX* analyseront ces données et peuvent fournir plusieurs statistiques numériques ainsi que graphiques.
- **Utiliser / présenter les résultats** :

1.6 Classes des plans d'expériences

Il existe trois grandes familles de plans d'expériences :

- **Les plans de criblages** : dont l'objectif est de découvrir les facteurs les plus influents sur une réponse donnée en un minimum d'expériences.

- **Les plans pour surface de réponse** : dont l'objectif est de trouver une relation mathématique (modèle) qui lie les réponses mesurées aux variables associées aux facteurs soit via une démarche mathématique analytique ou purement matricielle. Ce modèle peut être aussi utilisé à des fins d'optimisation du processus étudié.
- **Les plans de mélange** : dont l'objectif est le même que la deuxième famille mais où les facteurs ne sont pas indépendants et sont contraints. Comme exemple de contrainte, la somme des fractions molaires d'un mélange doit être égale à 1.

Chapitre 2

Plans factoriels

Dans les plans factoriels, plusieurs facteurs sont étudiés simultanément pendant l'étude. Des facteurs qualitatifs et/ou quantitatifs peuvent être pris en compte. L'objectif de ces plans est d'identifier les facteurs qui ont un effet significatif sur la réponse, ainsi que d'étudier l'effet des interactions (en fonction du plan d'expériences utilisé). Des prédictions peuvent également être effectuées lorsque des facteurs quantitatifs sont présents, mais des précautions doivent être prises car certains plans sont très limités dans le choix du modèle prédictif. Par exemple, dans les plans à deux niveaux, seule une relation linéaire entre la réponse et les facteurs peut être utilisée, ce qui peut ne pas être réaliste.

2.1 Plans factoriels complets à deux niveaux

Ce sont des plans factoriels où le nombre de niveaux pour chaque facteur est limité à deux (-1 et +1). Le fait de limiter les niveaux à deux et de mener une expérience factorielle complète réduit le nombre de traitements et permet d'étudier tous les facteurs et toutes leurs interactions. La figure 2.1 montre graphiquement deux plans factoriels complets pour deux et trois facteurs.

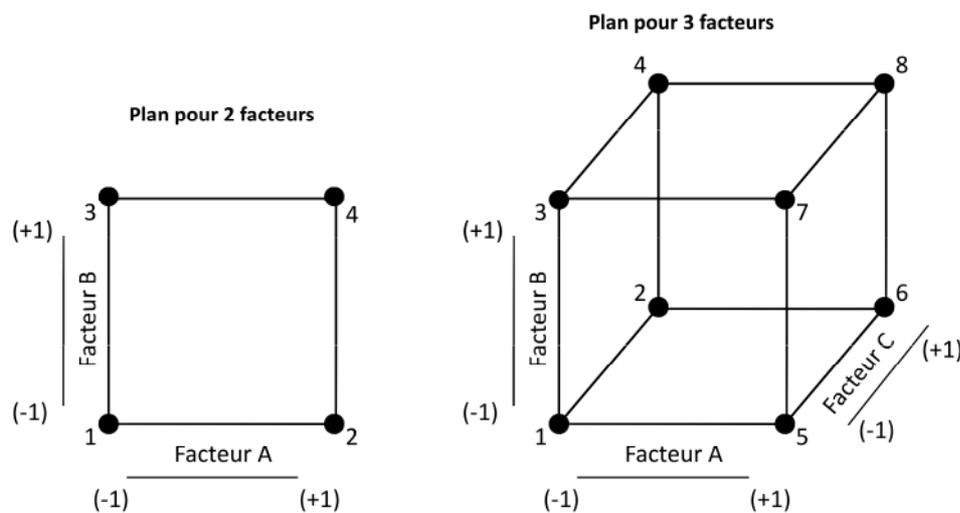


FIGURE 2.1: Plans factoriels complets 2^2 et 2^3

Si tous les facteurs sont quantitatifs, les données de ces expériences peuvent être utilisées à des fins de prédiction, à condition qu'un modèle linéaire soit approprié pour modéliser la réponse (puisque seuls deux niveaux sont utilisés, la courbure ne peut pas être modélisée).

2.2 Matrice d'expériences

La matrice d'expériences montre toutes les combinaisons possibles des niveaux haut et bas pour chaque facteur d'entrée. Ces niveaux haut et bas peuvent être codés +1 et -1. Par exemple, une expérience à 2 facteurs nécessitera 4 essais.

Un plan pour lequel nous avons k facteurs est appelé un plan 2^k . Le nombre d'expériences à réaliser sera donc 2^k expériences. Ce nombre devient rapidement très important. Par exemple pour seulement 7 facteurs, il faudrait 128 expériences. Pour diminuer le nombre des essais en conservant la possibilité d'étudier tous les facteurs, les plans factoriels fractionnaires à deux niveaux ont été proposés (voir chapitre 4).

Une matrice d'expériences peut se construire à la main en suivant l'algorithme de *Yates* ; pour une colonne c nous alternons : une série de 2^{c-1} (-1) et de 2^{c-1} (+1). La figure 2.2 montre les matrices d'expériences pour deux plans factoriels complets 2^2 et 2^3 .

Exp	x_1	x_2
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1

Exp	x_1	x_2	x_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

FIGURE 2.2: Matrices d'expériences pour les plans factoriels complets 2^2 et 2^3

Le modèle mathématique postulé est un modèle du premier degré par rapport à chaque facteur. L'équation 2.1 représente le modèle postulé sans interactions.

$$y = a_0 + \sum_{i=1}^k a_i x_i + \epsilon \quad (2.1)$$

Où a_0 représente l'effet théorique et les a_i représentent les effets principaux des différents facteurs.

Le modèle avec interactions d'ordre 2 prend en considération les interaction entre chaque facteur et un autre. Les effets des interactions sont quantifiés et représentés par les coefficients a_{ij} de l'équation 2.2.

$$y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i < j} \sum_{i < j} a_{ij} x_i x_j + \epsilon \quad (2.2)$$

Le modèle générique à k facteurs avec toutes les interactions est dit **modèle complet**, il peut être décrit par l'équation 2.3.

$$y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i < j} \sum_{i < j} a_{ij} x_i x_j + \dots + \sum \dots \sum_{i < j < \dots < k} a_{ij \dots k} x_i x_j \dots x_k + \epsilon \quad (2.3)$$

2.3 Effets globaux et effets moyens

Soit la **matrice d'expériences avec réponses** (y_i) pour un plan factoriel complet 2^2 pour deux facteurs nommés x_1 et x_2 .

Exp	x_1	x_2	Réponse (y_i)
1	-1	-1	y_1
2	+1	-1	y_2
3	-1	+1	y_3
4	+1	+1	y_4

FIGURE 2.3: Matrice d'expériences avec réponses pour un plan factoriel complet 2^2

Les effets principaux des facteurs ainsi que leurs interactions peuvent être quantifié :

Les effets moyens d'un facteur aux niveau -1 et +1 correspondent aux moyennes des réponses pour chaque niveau. les effets moyens du facteur x_1 sont donnés par les équations 2.4 et 2.5.

$$a_1^- = \frac{y_1 + y_3}{2} \quad (2.4)$$

$$a_1^+ = \frac{y_2 + y_4}{2} \quad (2.5)$$

Idem pour le facteur x_2 , les effets moyens sont donnés par les équations 2.6 et 2.7.

$$a_2^- = \frac{y_1 + y_2}{2} \quad (2.6)$$

$$a_2^+ = \frac{y_3 + y_4}{2} \quad (2.7)$$

Effet global d'un facteur : L'effet global d'un facteur est défini comme la variation moyenne de la réponse en passant du niveau bas du facteur à son niveau haut. Dans un **plan factoriel équilibré à deux niveaux**, l'effet estimé d'un facteur est la variation moyenne de la réponse entre ses deux niveau (sachant que la réponse dans un niveau est représentée par son effet moyen).

$$a_1 = \frac{a_1^+ - a_1^-}{2} = \frac{-y_1 + y_2 - y_3 + y_4}{4} \quad (2.8)$$

$$a_2 = \frac{a_2^+ - a_2^-}{2} = \frac{-y_1 - y_2 + y_3 + y_4}{4} \quad (2.9)$$

D'une manière générale, quand le modèle choisi est un polynôme, les coefficients des termes du premier degré sont les effets des facteurs.

Réponse théorique : est la moyenne des réponses observées aux niveaux -1 et +1.

$$a_0 = \frac{+y_1 + y_2 + y_3 + y_4}{4} \quad (2.10)$$

Effets des interactions : Pour calculer l'effet d'une interaction entre plusieurs variables x_1, x_2, \dots, x_k , nous ajoutons à la **matrice des effets** une colonne, que nous la baptisons $x_i x_j \dots x_k$, et que nous l'obtenons en faisant le produit "ligne à ligne" des colonnes des variables. L'effet moyen et l'effet global d'une interaction peuvent être calculés de la même manière que celle de l'effet global d'un facteur et son effet moyen. La matrice des effets principaux et des effets des interactions (pour deux facteurs nous avons qu'une seule interaction) pour un plan factoriel complet pour deux facteurs est montré dans la figure 2.4.

Exp	moy	x_1	x_2	$x_1 x_2$	Réponse (y_i)
1	+1	-1	-1	+1	y_1
2	+1	+1	-1	-1	y_2
3	+1	-1	+1	-1	y_3
4	+1	+1	+1	+1	y_4
Effets a_i	a_0	a_1	a_2	a_{12}	

FIGURE 2.4: Matrice d'expériences avec effets pour un plan factoriel complet 2^2

Le coefficient a_{12} mesure donc la variation de l'effet du facteur x_1 quand le niveau du facteur x_2 est modifié. Nous pouvons aussi montrer que le même coefficient a_{12} mesure également la variation de l'effet du facteur x_2 quand le niveau du facteur x_1 est, lui aussi, modifié.

Donc, le modèle mathématique linéaire peut s'écrire par l'équation 2.11

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + \epsilon \quad (2.11)$$

2.4 Exemple applicatif

Considérons le processus de cuisson d'un gâteau dont nous voulons étudier l'effet de la température (T) et la durée (D) de la cuisson sur l'épaisseur (E) du gâteau. L'expérimentateur décide d'utiliser un plan d'expériences factoriel complet 2^2 . Il définit le domaine d'étude des deux paramètres comme suit (figure 2.5) :

	Niveau bas : -1	Niveau haut : +1
Température : T	150°	200°
Durée : D	15 min	25 min

FIGURE 2.5: Un exemple de deux domaines d'étude de deux facteurs

Les résultats mesurés de l'épaisseur (E) pour chaque expérience réalisée, sont donnés dans la figure 2.6.

Exp	T	D	Epaisseur E_i (mm)
1	-1	-1	26
2	+1	-1	22
3	-1	+1	52
4	+1	+1	24

FIGURE 2.6: Matrice d'expériences avec 2^2 réponses de l'exemple étudié

Pour déterminer les coefficients du modèle complet, nous calculons l'effet théorique, les effets principaux et l'effet d'interaction en remplissant la matrice des effets (voir figure 2.7).

Exp	moy	T	D	TD	Epaisseur E_i (mm)
1	+1	-1	-1	+1	26
2	+1	+1	-1	-1	22
3	+1	-1	+1	-1	52
4	+1	+1	+1	+1	24
Effets a_i	$a_0 = 31$	$a_1 = -8$	$a_2 = 7$	$a_{12} = -6$	

FIGURE 2.7: Matrice d'expériences avec effets de l'exemple étudié

Le modèle obtenu est donné par l'équation 2.12.

$$E = 31 - 8T + 7D - 6TD + \epsilon \quad (2.12)$$

2.5 Représentation graphique des effets

La représentation graphique permet une interprétation facile (visuelle) de l'effet des facteurs principaux et des interactions sur la réponse.

2.5.1 Graphiques des effets principaux

Puisqu'il est souvent difficile d'obtenir des informations en regardant une équation mathématique, les graphiques des effets principaux sont presque toujours importants. La représentation graphique des effets principaux peut être tracée à partir du tableau des réponses moyennes des effets principaux. Elle permet de :

- Montrer quels sont les facteurs qui ont un effet significatif et ceux qui ont un effet nul ou négligeable. La grandeur d'un effet est définie par la grandeur de l'angle entre l'axe des abscisses et le

tracé du graphique.

- Montrer si les facteurs ont un effets négatif ou positif. Cela dépend du signe de l'angle ; si l'angle est positif cela signifie que l'effet du facteur est positif. Si l'angle est négatif cela signifie que l'effet du facteur est négatif.
- Comparer les effets des différents facteurs. Cette comparaison est effectuée en fonction de la grandeur d'angle des différents facteurs (quelque soit son signe).

	T	D
Niveau bas : -1	$a_T^- = 23$	$a_D^- = 38$
Niveau haut : +1	$a_T^+ = 39$	$a_D^+ = 24$

FIGURE 2.8: Domaines d'étude des facteurs T et D

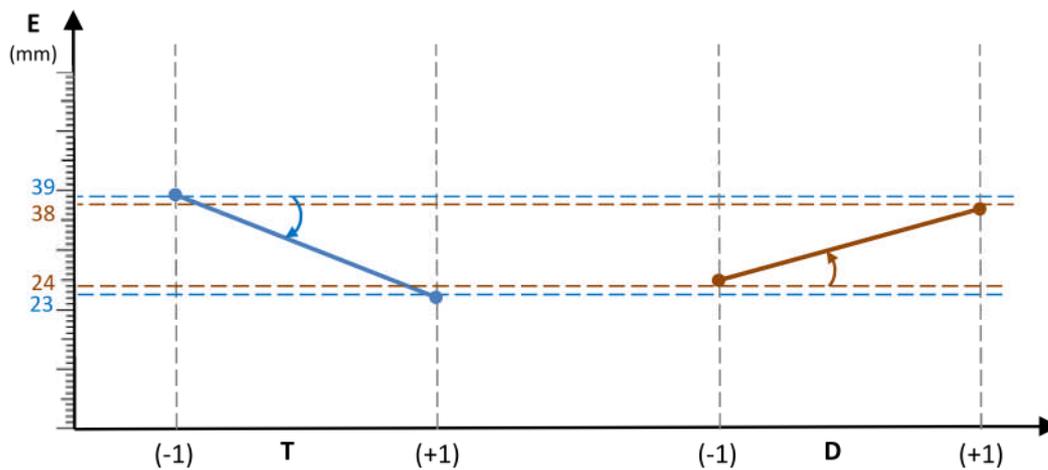


FIGURE 2.9: Représentation graphique des effets principaux T et D

En observant les deux graphiques des effets de la température et de la durée de cuisson dans la figure 2.9, nous constatons que les deux facteurs ont des effets globaux significatifs sur l'épaisseur du gâteau. Or, la température a un effet global négatif (en utilisant une température élevée de cuisson, donnera un gâteau moins épais et vice-versa), par contre la durée montre un effet global positif (en augmentant la durée de cuisson, nous obtiendrons un gâteau plus épais et vice-versa). Nous constatons aussi que l'effet global de la température sur l'épaisseur du gâteau est légèrement plus grand que celui de la durée de cuisson.

2.5.2 Graphiques des effets d'interactions

Lorsque des interactions significatives existent entre les facteurs expérimentaux, les diagrammes d'effets principaux ne racontent pas toute l'histoire sur les facteurs qui s'interagissent et peuvent même être trompeurs. Dans de tels cas, les graphiques d'interactions doivent être produits

pour chaque paire de facteurs. La représentation graphique de l'effet d'une interaction peut être tracée à partir du tableau des réponses moyennes des interactions. Moins les lignes sont parallèles, plus une interaction significative est probable.

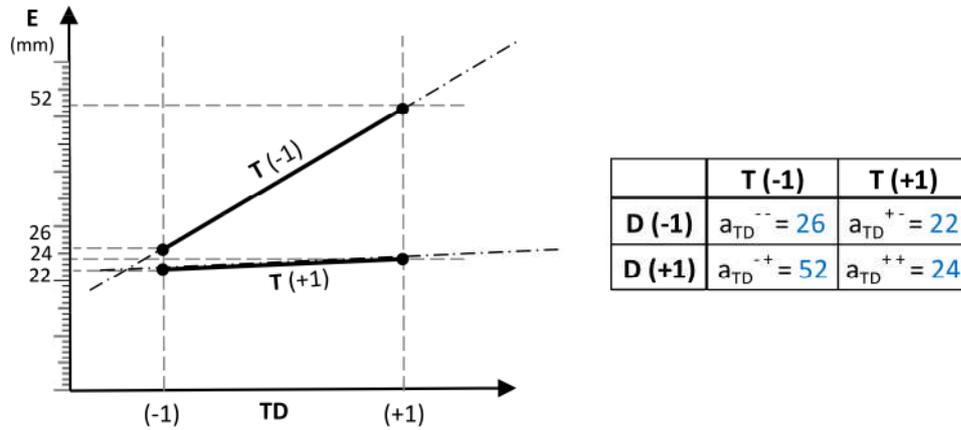


FIGURE 2.10: Représentation graphique de l'effet de l'interaction TD

La figure 2.10 représente les données des réponses moyennes du facteur T pour les deux niveaux du facteur D. Les deux lignes du graphique ne sont pas parallèles et montre un effet important de l'interaction entre T et D.

2.5.2.1 Plans factoriels fractionnaires à deux niveaux 2^{k-q}

Il s'agit d'une catégorie spéciale des plans à deux niveaux où toutes les combinaisons des niveaux des facteurs ne sont pas prises en compte et l'expérimentateur peut choisir les combinaisons à exclure. Sur la base des combinaisons exclues, certaines interactions ne peuvent pas être déterminées. La figure 2.11 représente un plan factoriel fractionnaire pour trois facteurs. En augmentant q de 1, le nombre d'expériences et encore réduit à la moitié.

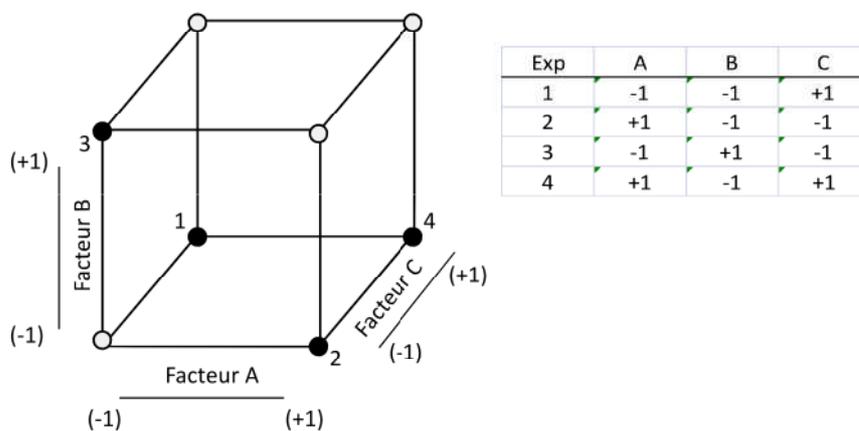


FIGURE 2.11: Plan factoriel fractionnaire 2^{3-1}

2.6 Autres plans à deux niveaux

2.6.1 Les plans de *Koshal*

Les plans de *Koshal* permettent uniquement de déterminer les effets principaux des facteurs, les interactions ne sont pas prises en considération. L'équation 2.13 montre leur modèle mathématique.

$$Y = a_0 + \sum_{i=1}^k a_i x_i + \epsilon \quad (2.13)$$

Les plans de *Koshal* offrent l'avantage de donner directement l'effet des facteurs. Ils forment une partie d'un plan factoriel qui est toujours extensible pour obtenir un plan complet ou fractionnaire. La Figure 2.12 montre un plan de *Koshal* pour 3 facteurs.

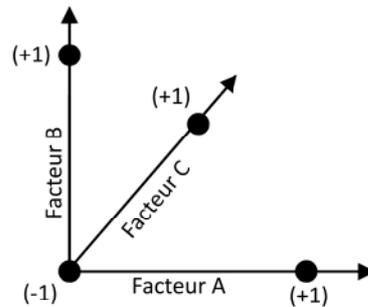


FIGURE 2.12: Plan de Koshal pour trois facteurs

2.6.2 Les plans de *Rechtschaffner*

Les plans de *Rechtschaffner* sont des plans factoriels fractionnaires qui considèrent les effets principaux des facteurs et leurs interactions d'ordre deux. Les interactions d'ordre supérieur à 2 sont supposées négligées. L'équation 2.14 montre le modèle mathématique adopté.

$$Y = a_0 + \sum_{i=1}^k a_i x_i + \sum_{i < j}^k \sum_{i < j} a_{ij} x_i x_j + \epsilon \quad (2.14)$$

Un plan fractionnaire de résolution III peut être choisi pour obtenir un plan de *Rechtschaffner*.

2.6.3 Les plans de *Plackett et Burmann*

Il s'agit d'une catégorie spéciale des plans factoriels fractionnaires à deux niveaux, où seulement quelques essais spécifiquement choisis sont effectués pour étudier uniquement les effets principaux (c'est-à-dire aucune interaction). Les plans de *Plackett et Burman* sont le plus souvent saturés.

L'équation 2.15 montre le modèle mathématique sans interactions analogue à celui des plans de *Ko-shal*.

$$Y = a_0 + \sum_{i=1}^k a_i x_i + \epsilon \quad (2.15)$$

2.7 Plans à plusieurs niveaux

2.7.1 Les tables de *Taguchi*

Les plans de tables orthogonales de *Taguchi* est un type de plans factoriels fractionnaires général. Il s'agit d'un plan orthogonal très fractionné basé sur une matrice proposée par *M. Genichi Taguchi* permettant de considérer un sous-ensemble sélectionné de combinaisons de plusieurs facteurs à plusieurs niveaux. Les tables orthogonales de *Taguchi* sont équilibrées pour garantir que tous les niveaux de tous les facteurs sont pris en compte de manière égale. Pour cette raison, les facteurs peuvent être évalués indépendamment les uns des autres malgré la fractionalité du plan. Dans le plan de *Taguchi*, seuls les effets principaux et les interactions à deux facteurs sont pris en compte, et les interactions d'ordre supérieur sont supposées inexistantes.

2.7.2 Carrés latins

Les carrés latins (le terme Carré Latin a été utilisé pour la première fois par Euler, 1782) sont utilisés lorsque le nombre des niveaux des facteurs est supérieur à deux et lorsque il n'existe pas d'interactions (ou qu'elles sont négligeables) entre les facteurs. Un carré latin pour 3 facteurs est illustré dans la figure 2.13. Ce type de plans est souvent utilisé pour les variables discrètes et son modèle mathématique est souvent un modèle sans interactions.

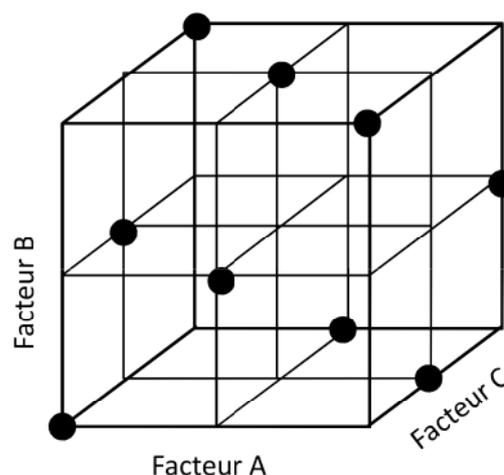


FIGURE 2.13: Plan Carrés latins pour 3 facteurs

2.7.3 Carrés de *Youden*

Ce type de plans considère deux variables discrètes et prend plus de quatre niveaux. Ils ont servi également de base à l'établissement des plans pour simulations numériques. Ce type de plans est similaire à celui des carrés latins. Le nombre des essais est réduit en retirant des points au plan complet.

Chapitre 3

Signification des effets et validation du modèle

3.1 Test de signification des effets

L'influence des facteurs et de leurs interactions est interprétée par les coefficients du modèle postulé. Il faut donc trouver une valeur étalon (t_{crit}) pour la prise de décision si l'effet d'un facteur ou d'une interaction est important ou non (Un effet sera dit significatif s'il est, pour un risque donné, significativement différent de 0). Le test de Student a pour but de fournir une règle de décision. La valeur à tester t_i sera le rapport de la valeur du coefficient a_i sur la valeur de son écart-type S_i (équation 3.1) :

$$t_i = \frac{a_i}{S_i} \quad (3.1)$$

3.1.1 Calcul de la variance des coefficients S_i^2

En statistique la formule qui détermine la variance des coefficients S_i^2 en fonction de la variance des écarts S^2 est donnée par l'équation 3.2.

$$S_i^2 = kS^2 \quad (3.2)$$

Où la constante k dépend du modèle postulé et de la matrice d'expériences. Généralement k est très long à calculer (Des logiciels spécialisés possèdent les algorithmes pour faire ce calcul). Mais dans le cas des plans factoriels la relation est plus simple devient (3.3) :

$$S_i^2 = \frac{1}{n}S^2 \implies S_i = \sqrt{\frac{S^2}{n}} \quad (3.3)$$

Les calculs statistiques qui permettent de déterminer l'écart type S_i font intervenir la variance des écarts (les différences entre les valeurs expérimentales y_i et les valeurs estimées (prédites) par le modèle \hat{y}_i^2) selon l'équation 3.4.

$$S^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p} \quad (3.4)$$

Où n est le nombre d'expériences réalisées et p est le nombre de coefficients du modèle (constante incluse).

Remarque : Pour pouvoir conduire les calculs statistiques il est clair que $(n - p)$ doit être différent de 0. Pour cela, en pratique il est nécessaire de négliger un ou plusieurs termes (en général l'interaction (les interactions) d'ordre plus élevé pour que p soit différent de n . Si un plan complet 2^3 est considéré, ça donne 8 expériences et 8 termes du modèle (a_0 , 3 effets principaux, 3 effets d'interactions d'ordre deux et 1 effet d'interaction d'ordre trois). Pour permettre le test statistique, l'interaction d'ordre trois peut être négligée ce qui donne un modèle réduit avec 7 termes ($p = 7$).

3.1.2 Détermination de la valeur critique t_{crit}

Afin de pouvoir tester la signification d'un effet avec un risque donné α , le test de Student est utilisé ; le rapport t_i est comparé à une valeur critique t_{crit} pour un risque α et un degré de liberté $ddl = n - p$. Cette valeur critique peut être directement lue à partir de la table de *Student* (voir table 3.1).

$$t_{crit} = v(\alpha, ddl) \quad (3.5)$$

Par exemple la valeur critique de student t_{crit} pour un modèle dont $p = 7$, $n = 8$ et pour un risque $\alpha = 0.05$ correspond à la cellule de l'intersection de la première ligne ($n - p = 1$) et la sixième colonne ($\alpha = 0.05$) dans la table de student 3.1.

ddl/α	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
1	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6
2	0.816	1.080	1.386	1.886	2.920	4.303	6.965	9.925	14.09	22.33	31.60

FIGURE 3.1: Lecture d'une valeur t_{crit} depuis la table de *Student*

3.1.3 Réalisation du test et interprétation

L'hypothèse selon laquelle l'effet a_i est nul s'appelle l'hypothèse nulle et est notée H_0 . N'importe quelle autre hypothèse qui diffère de l'hypothèse H_0 s'appelle l'hypothèse alternative et est notée H_1 ($\alpha\%$ de confiance pour que l'effet soit significativement différent de zéro). Nous aurons :

- H_0 : l'effet a_i est nul c-à-d a_i n'est pas significatif.
- H_1 : l'effet a_i n'est pas nul c-à-d a_i est significatif.

Après avoir calculé les rapports t_i , ces valeurs sont comparées avec la valeur critique t_{crit} pour déterminer la signification des effets ou leur non signification.

- Si $|t_i| > t_{crit}$, l'hypothèse H_0 est rejetée (H_1 est acceptée) c-à-d l'effet a_i est significatif.
- Sinon ($|t_i| \leq t_{crit}$), l'hypothèse H_0 est acceptée (H_1 est rejetée) c-à-d l'effet a_i n'est pas significatif.

À partir du t de Student, l'évaluation de la probabilité peut révéler que le coefficient est peu significatif (ou nul) ou non. Cette probabilité est la p_{valeur} .

- Si la p_{valeur} est proche de 0 (probabilité nulle), le coefficient est influent car il n'est pas nul.
- Si la p_{valeur} est proche de 1, le coefficient est proche de 0 et donc négligeable. Si la p_{valeur} possède une valeur intermédiaire, le coefficient est peut-être légèrement significatif ou non significatif.

TABLE 3.1: Table de *student*

dll/α	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.005	0.002	0.001
1	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	127.3	318.3	636.6
2	0.816	1.080	1.386	1.886	2.920	4.303	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	2.871	3.174	3.390
120	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

3.1.4 Exemple applicatif

En prenant l'exemple précédent de la cuisson d'un gâteau ou nous considérons l'ajout ou le non ajout de la levure comme un facteur supplémentaire. Nous voulons déterminer *quels sont les facteurs et les interactions qui ont un effet significatif et ceux qui ont un effet négligeable*. L'expérimentateur décide d'utiliser un plan d'expériences factoriel complet 2^3 . Il définit le domaine d'étude des trois paramètres comme suit (figure 3.2) :

	T	D	L
Niveau bas : -1	$a_T^- = 23$	$a_D^- = 38$	sans
Niveau haut : +1	$a_T^+ = 39$	$a_D^+ = 24$	avec

FIGURE 3.2: Plan factoriel fractionnaire 2^{3-1}

Les résultats mesurés de l'épaisseur (E_i) pour chaque expérience réalisée ainsi que les coefficients calculés sont donnés dans la figure 3.3.

Exp (i)	moy	T	D	L	TD	TL	DL	TDL	E_i^{exp}
1	1	-1	-1	-1	1	1	1	-1	66,82
2	1	1	-1	-1	-1	-1	1	1	45,22
3	1	-1	1	-1	-1	1	-1	1	69,22
4	1	1	1	-1	1	-1	-1	-1	38,48
5	1	-1	-1	1	1	-1	-1	1	66,6
6	1	1	-1	1	-1	1	-1	-1	74,82
7	1	-1	1	1	-1	-1	1	-1	74,2
8	1	1	1	1	1	1	1	1	74,28
	$a_0=63,705$	$a_T=-5,505$	$a_D=0,34$	$a_L=8,77$	$a_{TD}=-2,16$	$a_{TL}=7,58$	$a_{DL}=1,425$	$a_{TDL}=-0,125$	

FIGURE 3.3: Matrice d'expériences avec effets de l'exemple étudié

L'équation 3.6 montre le modèle mathématique complet :

$$\hat{E} = 63,705 - 5,505T + 0,34D + 8,77L - 2,16TD + 7,58TL + 1,425DL + 0,125TDL \quad (3.6)$$

- Pour répondre à la question déterminant la signification des différents effets, nous utilisons le test de *Student*.

Afin de pouvoir conduire le test statistique de *Student*, nous choisissons de supprimer le dernier terme du modèle complet (l'interaction d'ordre le plus élevé) pour éviter le cas où le dénominateur $n - p = 0$. Le nouveau modèle à considérer est 3.7

$$\hat{E} = 63,705 - 5,505T + 0,34D + 8,77L - 2,16TD + 7,58TL + 1,425DL \quad (3.7)$$

- Pour calculer S^2 nous calculons les réponses estimées \hat{E}_i par le modèle 3.7 et les résidus dans le tableau 3.4.

Exp (i)	E_i	\hat{E}_i	$(E_i - \hat{E}_i)^2$
1	66,82	66,945	0,015625
2	45,22	45,095	0,015625
3	69,22	69,095	0,015625
4	38,48	38,605	0,015625
5	66,6	66,475	0,015625
6	74,82	74,945	0,015625
7	74,2	74,325	0,015625
8	74,28	74,155	0,015625
		$\Sigma(E_i - \hat{E}_i)^2$	0,125

FIGURE 3.4: Résultats des réponses estimées et des écarts au carré

- Nous aurons : $S^2 = \frac{\sum_{i=1}^8 (y_i - \hat{y}_i)^2}{n-p} = \frac{0.125}{8-7} = 0.125$
- Les valeurs des S_i seront :
 $S_0 = S_T = S_D = S_L = S_{TD} = S_{TL} = S_{DL} = \sqrt{\frac{S^2}{n}} = \sqrt{\frac{0.125}{8}} = 0.125$
- La valeur de t_{crit} pour un risque $\alpha = 0.05$ et un degré de liberté $ddl = n - p = 8 - 7$ est prélevée de la table de Student 3.1 $t_{crit} = v(0.05, 1) = 12, 71$
- Finalement nous calculons les valeurs $t_i = \frac{a_i}{S_i}$ et nous comparons leurs valeurs absolues à t_{crit} dans le tableau 3.5 pour établir une décision concernant la signification des effets des facteurs et de leurs interactions.

Effet	a_i	s_i	t_i	$ t_i \geq t_{crit} ?$	a_i est significatif ?
a_0	63,705	0,125	509,64	Oui	Sinificatif
a_T	-5,505	0,125	44,04	Oui	Sinificatif
a_D	0,34	0,125	2,72	Non	Non Sinificatif
a_L	8,77	0,125	70,16	Oui	Sinificatif
a_{TD}	-2,16	0,125	17,28	Oui	Sinificatif
a_{TL}	7,58	0,125	60,64	Oui	Sinificatif
a_{DL}	1,425	0,125	11,4	Non	Non Sinificatif

FIGURE 3.5: Résultats de signification des effets

- D’après la figure 3.5, la durée de cuisson D et l’interaction DL ont un effet négligeable (considérés comme = 0) c-à-d qu’il y a moins de 5% de risque que leurs effets ne soient pas nuls. Quant aux effets de la température, de la levure et les interactions TD et TL sont significatifs c-à-d qu’il ont plus de 5% de risque que ces effets ne soient pas nuls (considérés comme $\neq 0$).

3.2 Régression linéaire et analyse du modèle

3.2.1 Régression linéaire

Supposons que nous avons un modèle polynomial sous la forme générique suivante (3.8) :

$$y = a_0 + \sum_{i=1}^k a_i x_i \quad (3.8)$$

Ce polynôme peut être utilisé pour représenter plusieurs formes de modèles (linéaires, quadratiques, etc). Par exemple, pour écrire le modèle $y = 3 + 2.1z_1 + 4z_2 + 2z_1z_2 + z_2^2$ sous la forme 3.8 nous posons $x_1 = z_1$, $x_2 = z_2$, $x_3 = z_1z_2$ et $x_4 = z_2^2$. Notez que, bien qu'un modèle ne soit pas linéaire, le modèle de régression est toujours linéaire car il est linéaire dans les paramètres.

À la fin du plan d'expériences, un système de n équations (s'il y a n essais) et à p inconnues (s'il y a p coefficients dans le modèle choisi) peut être écrit d'une manière simple en notation matricielle :

$$y = Xa + \epsilon \quad (3.9)$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdot & \cdot & x_{1p} \\ 1 & x_{21} & x_{22} & \cdot & \cdot & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n1} & x_{n2} & \cdot & \cdot & x_{np} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \cdot \\ \cdot \\ a_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \cdot \\ \cdot \\ \epsilon_n \end{bmatrix} \quad (3.10)$$

y : est le vecteur des réponses

X : est la matrice de calcul des coefficients ou matrice du modèle qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé

a : est le vecteur des coefficients

ϵ : est le vecteur des écarts

Ce système ne peut pas être résolu simplement car en général le nombre d'équations est inférieur au nombre d'inconnues (n équations et pn inconnues). Cette résolution peut être menée en utilisant une méthode de régression. Souvent cette méthode est basée sur la méthode d'optimisation des moindres carrés. Les estimations des coefficients obtenues sont les éléments du vecteur \hat{a} .

$$\hat{a} = (X'X)^{-1}X'y \quad (3.11)$$

Où X' est la transposée de la matrice X .

Pour calculer directement les valeurs des coefficients, il existe plusieurs logiciels qui exécutent ce calcul.

3.2.2 Test de validation du modèle

Dans les cas où il y a plus de deux échantillons de test, une ANOVA est utilisée pour déterminer s'il existe des différences statistiquement significatives entre les moyennes des échantillons (Dans les cas de deux échantillons, le test t suffit à vérifier s'il existe des différences statistiquement significatives entre les moyennes des échantillons). La procédure de test implique une analyse de variance (ANOVA) et la réalisation du test F (*Fisher-Snedecor*) qui teste la signification de la régression dans sa globalité (il teste la nullité de tous les coefficients en même temps). Il ne permet donc pas de préjuger la signification particulière des coefficients pris isolément. C'est ce que fait le test de *Student* qui teste un à un la signification des coefficients. Il s'agit de tester les hypothèses :

$$H_0 : a_1 = a_2 = \dots = a_p = 0$$

$$H_1 : \text{il existe au moins un } a_i \neq 0.$$

La variance de la variable à expliquer (ou totale) se décompose en somme de la variance expliquée par le modèle et de la variance résiduelle.

$$STCE = SCEL + SCER \quad (3.12)$$

- **Somme totale des carrés des écarts (STCE)** : est la somme totale des carrés corrigés de y .

$$STCE = \sum_{i=1}^n (y_i - \bar{y})^2 \quad (3.13)$$

- **Somme des carrés des écarts dues à la liaison (SCEL)** : est la somme des carrés expliquée par le modèle

$$SCEL = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 \quad (3.14)$$

- **Somme des carrés des écarts des résidus (SCER)** : est la somme des carrés des résidus.

$$SCER = \sum_{i=1}^n \hat{e}_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (3.15)$$

Notons que : y_i sont les réponses observées lors de la réalisation des expériences, \bar{y} est la moyenne des réponses et \hat{y}_i sont les réponses estimées à l'aide du modèle. De plus des "carrés moyens" sont définis qui sont le quotient d'une somme de carrés par son degré de liberté (*CML*, *CMR* et *CMT*).

SCEL aura $p - 1$ degrés de liberté (p est le nombre de coefficients estimés du modèle).

SCER aura $n - p$ degrés de liberté (où n est le nombre d'expériences réalisées).

STCE aura $n - 1$ degrés de liberté.

Le tableau dit analyse de la variance se présente sous la forme suivante :

TABLE 3.2: Tableau d'analyse de la variance

Variation due à	Somme des carrés	ddl	Carré moyen	F
Liaison	$SCEL$	$p - 1$	$\frac{SCEL}{p-1} = CML$	$F_{cal} = \frac{CML}{CMR}$
Résidus	$SCER$	$n - p$	$\frac{SCER}{n-p} = CMR$	
Totale	$STCE$	$n - 1$	$\frac{STCE}{n-1} = CMT$	

Le test F permet alors de comparer le F_{cal} que nous avons calculé dans le tableau précédent avec le $F_{crit}(ddl_1, ddl_2, \alpha)$ lu à partir de la table de Fisher-Snedecor (Tableau 3.3) avec les degrés de liberté $ddl_1 = p - 1$, $ddl_2 = n - p$ et avec un risque fixé à l'avance α . Le critère est de rejeter H_0 au seuil α si $F_{cal} > F_{crit}$.

- Accepter H_0 implique que l'on conclut qu'il n'y a pas de relation globale entre x_i et y . Ceci peut signifier que :

- Le modèle utilisé n'est pas adéquat.
- La variation des x_i influent peu ou pas sur la variation de y .

- Au contraire, rejeter H_0 implique que nous concluons que la variation des x_i influe la variation de y .

Autrement dit :

- Si $F_{cal} > F_{crit}$ le modèle de la régression linéaire est considéré comme *valide*.
- Sinon, le modèle est considéré comme *non valide*.

3.2.3 Quelques mesures statistiques d'évaluation du modèle

En effectuant une analyse de la variance du modèle de régression, plusieurs mesures statistiques peuvent être obtenues (R^2 , R_{ajuste}^2 , $R_{prédit}^2$ et la *Precision – adequate*). Ces mesures permettent de tester la qualité du modèle.

3.2.3.1 Coefficient de détermination R^2

Le coefficient de détermination R^2 montre dans quelle mesure un modèle de fonction correspond aux données. Plus R^2 est proche de 1, meilleur est la correspondance. R^2 est donc une mesure de la qualité du modèle, il est toujours entre 0 et 1. S'il est égal à 1, le modèle permet de retrouver

les valeurs des réponses mesurées.

$$R^2 = 1 - \frac{SCER}{STCE} = \frac{SCEL}{STCE} \quad (3.16)$$

Attention, le R^2 ne révèle pas tout sur la qualité du modèle. R^2 doit être considéré comme une donnée descriptive, intéressante en soi, et pratique pour comparer des modèles sur les mêmes données, mais il ne peut être considéré comme une note absolue. Dans la pratique, il est difficile d'indiquer la valeur d'un bon R^2 car les valeurs varient beaucoup d'une discipline à l'autre et du processus étudié.

3.2.3.2 R au carré ajusté R_{ajuste}^2

Le R_{ajuste}^2 est une version modifiée du R^2 qui a été ajustée en fonction du nombre de prédicteurs du modèle. Le R_{ajuste}^2 peut être négatif, mais ce n'est généralement pas le cas. Il est toujours inférieur à R^2 .

$$R_{ajuste}^2 = 1 - \frac{CMR}{CMT} = 1 - \frac{(n-1)(1-R^2)}{n-p} \quad (3.17)$$

3.2.3.3 Somme des carrés des erreurs résiduelles prédites $PRESS$

$PRESS$ (Predicted Residual Error Sum of Squares) est une mesure de l'adaptation du modèle à chaque point du plan. Le $PRESS$ est calculé en prédisant d'abord où chaque point doit se trouver à partir d'un modèle contenant tous les autres points, à l'exception du point en question. Les résidus au carré (différence entre les valeurs réelles et prédites) sont ensuite additionnés. $PRESS$ est utile pour déterminer la valeur de $R_{prédit}^2$.

$$PRESS = \sum_{i=1}^n (y_i - \hat{y}_{-i})^2 \quad (3.18)$$

3.2.3.4 R au carré prédit $R_{prédit}^2$

Le $R_{prédit}^2$ indique dans quelle mesure un modèle de régression prédit les réponses à de nouvelles observations.

$$R_{prédit}^2 = 1 - \frac{PRESS}{STCE} \quad (3.19)$$

Le $R_{prédit}^2$ est calculé en supprimant systématiquement chaque observation de l'ensemble de données, en estimant l'équation de régression et en déterminant dans quelle mesure le modèle prédit l'observation supprimée. Comme le R_{ajuste}^2 , le $R_{prédit}^2$ peut être négatif et il est toujours inférieur à R^2 .

Même si vous ne prévoyez pas d'utiliser le modèle pour les prévisions, le $R^2_{\text{prédit}}$ fournit néanmoins des informations cruciales.

3.2.3.5 Précision adéquate Pre_{adeq}

C'est un rapport signal sur bruit. Il compare la plage des valeurs prédites aux points du plan à l'erreur de prédiction moyenne. Des ratios supérieurs à 4 indiquent une discrimination de modèle adéquate.

$$Pre_{adeq} = \frac{\max(\hat{y}) - \min(\hat{y})}{\sqrt{\frac{p*CMR}{n}}} \tag{3.20}$$

TABLE 3.3: Table de Fischer - Snedecor (Risque $\alpha = 0.05$)

n_1	1	2	3	4	5	6	7	8	9	10	12	15	20
n_2													
1	161	199,5	215,7	224,6	230,2	234	236,8	239	240,5	241,9	243,9	245,9	248
2	18,5	19	19,16	19,25	19,3	19,33	19,35	19,4	19,38	19,4	19,41	19,43	19,45
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,7	8,66
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6	5,96	5,91	5,86	5,8
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,1	4,06	4	3,94	3,87
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44
8	5,32	4,46	4,07	3,84	3,69	3,58	3,5	3,44	3,39	3,35	3,28	3,22	3,15
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94
10	4,96	4,1	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77
11	4,84	3,98	3,59	3,36	3,2	3,09	3,01	2,95	2,9	2,85	2,79	2,72	2,65
12	4,75	3,89	3,49	3,26	3,11	3	2,91	2,85	2,8	2,75	2,69	2,62	2,54
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,6	2,53	2,46
14	4,6	3,74	3,34	3,11	2,96	2,85	2,76	2,7	2,65	2,6	2,53	2,46	2,39
15	4,54	3,68	3,29	3,06	2,9	2,79	2,71	2,64	2,59	2,54	2,48	2,4	2,33
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28
17	4,45	3,59	3,2	2,96	2,81	2,7	2,61	2,55	2,49	2,45	2,38	2,31	2,23
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19
19	4,38	3,52	3,13	2,9	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16
20	4,35	3,49	3,1	2,87	2,71	2,6	2,51	2,45	2,39	2,35	2,28	2,2	2,12
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,1
22	4,3	3,44	3,05	2,82	2,66	2,55	2,46	2,4	2,34	2,3	2,23	2,15	2,07
23	4,28	3,42	3,03	2,8	2,64	2,53	2,44	2,37	2,32	2,27	2,2	2,13	2,05
24	4,26	3,4	3,01	2,78	2,62	2,51	2,42	2,36	2,3	2,25	2,18	2,11	2,03
25	4,24	3,39	2,99	2,76	2,6	2,49	2,4	2,34	2,28	2,24	2,16	2,09	2,01
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,2	2,13	2,06	1,97
28	4,2	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96
29	4,18	3,33	2,93	2,7	2,55	2,43	2,35	2,28	2,22	2,18	2,1	2,03	1,94
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2	1,92	1,84
60	4	3,15	2,76	2,53	2,37	2,25	2,17	2,1	2,04	1,99	1,92	1,84	1,75
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	1,66
infini	3,84	3	2,6	2,37	2,21	2,1	2,01	1,94	1,88	1,83	1,75	1,67	1,57

3.2.4 Exemple applicatif

En prenant l'exemple précédent (voir section 3.1.4) de la cuisson d'un gâteau où nous avons obtenu le modèle mathématique suivant :

$$\widehat{E} = 63,705 - 5,505T + 0,34D + 8,77L - 2,16TD + 7,58TL + 1,425DL \quad (3.21)$$

Nous voulons étudier la validité du modèle obtenu avec un risque ($\alpha = 0.05$) ainsi que sa qualité (sa correspondance aux données mesurées, sa capacité de prédiction et sa précision adéquate) en effectuant une analyse de la variance (*ANOVA*) et en calculant les différentes mesures statistiques.

Exp (i)	E_i	\widehat{E}_i	$(E_i - \widehat{E}_i)^2$	$(\widehat{E}_i - \bar{E})^2$
1	66,82	66,945	0,015625	10,4976
2	45,22	45,095	0,015625	346,3321
3	69,22	69,095	0,015625	29,0521
4	38,48	38,605	0,015625	630,01
5	66,6	66,475	0,015625	7,6729
6	74,82	74,945	0,015625	126,3376
7	74,2	74,325	0,015625	112,7844
8	74,28	74,155	0,015625	109,2025
	63,705		0,125	1371,8892
	\bar{E}		$\Sigma(E_i - \widehat{E}_i)^2$	$\Sigma(\widehat{E}_i - \bar{E})^2$

FIGURE 3.6: Résultats des réponses estimées et des écarts au carré

$$SCEL = \sum_{i=1}^n (\widehat{E}_i - \bar{E})^2 = 1371,8892$$

$$CML = \frac{SCEL}{p-1} = \frac{1371,8892}{7-1} = 228,6482$$

$$SCER = \sum_{i=1}^n (E_i - \widehat{E}_i)^2 = 0.125$$

$$CMR = \frac{SCER}{n-p} = \frac{0.125}{8-7} = 0.125$$

$$STCE = SCEL + SCER = 1372,0142$$

$$CMT = \frac{STCE}{n-1} = \frac{1372,0142}{8-1} = 196,0020..$$

TABLE 3.4: Tableau d'analyse de la variance de l'exemple étudié

Variation	Somme des carrés	ddl	Carré moyen	F
Liaison	1371,8892	7 - 1	228,6482	$F_{cal} = \frac{CML}{CMR} = 1829.1856$
Résidus	0.125	8 - 7	0.125	
Totale	1372,0142	8 - 1	196,0020	

Pour tester la validité du modèle obtenu, F_{cal} est comparé à F_{crit} . F_{crit} est obtenu de la table de *Fisher* pour les paramètres $\alpha = 0.05$, $dll1 = p - 1$ et $dll2 = n - p$ ($F_{crit}(0.05, 6, 1) = 234$). $F_{cal} > F_{crit} \Rightarrow$ le modèle de régression est valide.

$R^2 = \frac{2791.75}{2794.875} = \frac{SCEL}{STCE} = 0,99990889$ et $R_{ajuste}^2 = 1 - \frac{CMR}{CMT} = 1 - \frac{0.125}{196,002028} = 0,999362$. Nous avons les valeurs de R^2 et R_{ajuste}^2 sont très proches de 1 ce qui prouve que la qualité du modèle est très bonne en ce qui concerne sa correspondance aux données observées.

$$Pre_{adeq} = \frac{\max(\hat{y}) - \min(\hat{y})}{\sqrt{\frac{p * CMR}{n}}} = \frac{\max(66.945, 45.095, 69.095, 38.605, 66.475, 74.945, 74.325, 74.155) - \min(66.945, 45.095, 69.095, 38.605, 66.475, 74.945, 74.325, 74.155)}{\sqrt{\frac{7 * 0.125}{8}}} = 109.882 . Pre_{adeq} > 4 \Rightarrow \text{le modèle a une bonne précision adéquate.}$$

Le calcul de la valeur *PRESS* est long ; nous donnons directement sa valeur $PRESS = 8$ (En utilisant un logiciel nous pouvons directement obtenir toutes les statistiques).

$R_{predit}^2 = 1 - \frac{PRESS}{STCE} = 1 - \frac{8}{1372,0142} = 0.9942$. La valeur de R_{predit}^2 est très proche de 1 \Rightarrow le modèle a une grande capacité de prédire de nouvelles observations.

Dans l'étude du modèle ci-dessus nous avons négligé que le terme *TDL*. Or, d'après l'étude de la signification des effet effectuée dans la section 3.1.4 nous avons trouvé que les effets de *D* et *TDL* sont négligeables. Si nous négligeons ces deux terme et nous étudions la qualité du nouveau modèle simplifié (équation 3.7), nous aurons des résultats qui reflètent aussi la bonne qualité de ce modèle simplifié.

Chapitre 4

Plans factoriels fractionnaires

4.1 Introduction

La visualisation d'une expérience comportant plus de trois facteurs est plus complexe et difficile à décrire, car les dimensions requises pour la visualisation sont égales au nombre de facteurs, mais le principe des expériences en angle est le même, quel que soit le nombre de facteurs. Pour les études où quatre facteurs ou plus présentent un intérêt, comme dans un test de robustesse ou une étude de dépistage, il est assez courant d'utiliser des plans factoriels fractionnaires.

4.2 Plans factoriels fractionnaires

Comme son nom l'indique, le plan factoriel fractionnaire est une fraction d'un plan factoriel complet. Un plan factoriel fractionnaire est construit de manière à pouvoir encore identifier les principaux effets sans acquérir les informations détaillées fournies par un plan factoriel complet. Les expériences sont sélectionnées en utilisant une sélection symétrique des angles, des diagonales et des diagonales opposées.

La demi-fraction d'un plan à trois facteurs nécessite quatre essais. Le plan comprend deux ensembles de quatre essais chacun, exactement équivalents sur le plan mathématique, où chaque ensemble peut être choisi (4.1).

Un plan factoriel complet est facilement complété à l'aide de l'ensemble des autres paramètres des facteurs. Le plan est équilibré, ce qui signifie que chaque facteur est exécuté avec le même nombre de fois pour chaque niveau.

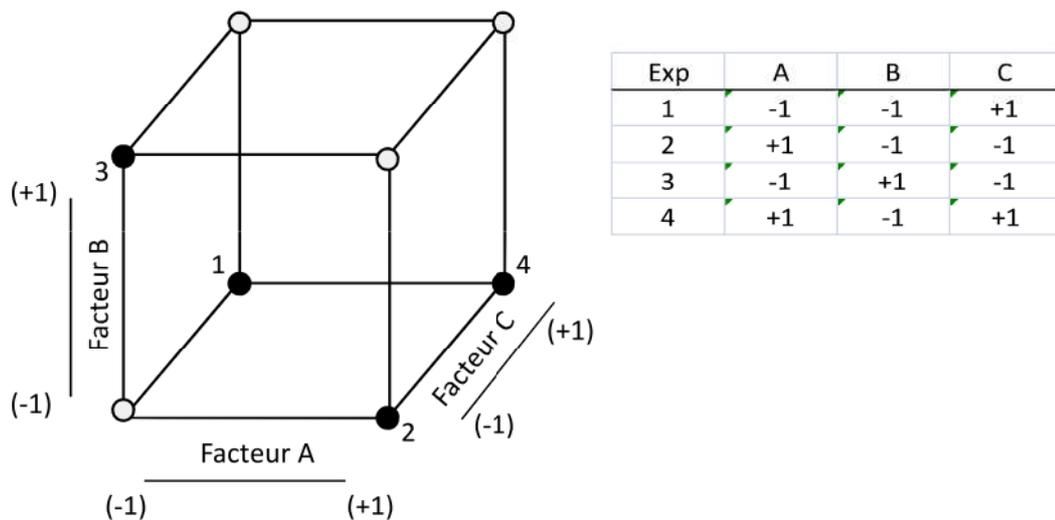


FIGURE 4.1: Plan factoriel fractionnaire 2³⁻¹

En général, un plan factoriel fractionnaire peut être décrit par $N = 2^{k-p}$ où :

- N est le nombre d'expériences.
- k est le nombre de facteurs à étudier.
- p est la taille de la fraction ($1 \Rightarrow 1/2, 2 \Rightarrow 1/4, 3 \Rightarrow 1/8, \text{etc.}$).

Par conséquent, $N = 2^{4-1}$ signifie que quatre facteurs seront examinés en $2^3 = 8$ exécutions.

La force d'un plan factoriel fractionnaire réside dans le fait qu'il permet de sélectionner de nombreux facteurs en utilisant relativement peu d'expériences. Cependant, l'inconvénient de ce type de plan est que les effets sont confondus. En fonction de la résolution (*Res*), le plan factoriel fractionnaire prend en charge les effets des facteurs principaux et d'interaction.

Les plans fractionnaires sont généralement utilisés en tant que plans de criblage pour déterminer quels sont les facteurs les plus influents sans forcément étudier les interactions d'ordre 2. C'est souvent le cas si le nombre de facteurs est très élevé.

4.3 Notion d'alias et de contraste

Un plan factoriel fractionnaire est généré à partir d'un plan factoriel complet en choisissant une structure d'alias. La structure d'alias détermine quels effets sont confondus. Par exemple, un plan factoriel fractionnaire 2^{5-2} avec les 5 facteurs peut être généré en utilisant un plan factoriel complet 2^3 impliquant trois facteurs (disons A, B et C), puis en choisissant de confondre les deux facteurs restants D et E avec les interactions générées par $D = A * B$ et $E = A * C$. Ces deux expressions sont appelées les **générateurs d'alias**.

Ainsi, par exemple, lorsque le plan est exécuté et que l'expérimentateur estime les effets pour le facteur D , ce qui est réellement estimé est une combinaison de l'effet principal de D et de l'interaction à deux facteurs impliquant A et B .

Une caractéristique importante d'un plan fractionnaire est la **relation de définition**, qui donne à l'ensemble des colonnes d'interaction une valeur égale dans la matrice du plan à une colonne de signes plus, notée I .

Pour l'exemple ci-dessous, puisque $D = AB$ et $E = AC$, alors ABD et ACE sont deux colonnes de signes plus, et par conséquent $BDCE$ aussi. Dans ce cas, la relation de définition du plan factoriel fractionnaire est :

$$I = ABD = ACE = BCDE \quad (4.1)$$

La relation de définition permet de déterminer le motif d'alias du plan.

Exp	I	A	B	C	D	E	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE	ABC	ABD	ABE	ACD	ACE	ADE	BCD	BCE	BDE	CDE	ABCD	ABCE	ABDE	ACDE	BCDE	ABCDE	
25	+	-	-	-	+	+	+	+	-	-	+	-	-	-	-	+	-	+	+	+	+	-	+	+	-	-	-	-	+	+	-	-	
2	+	+	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	+	
19	+	-	+	-	-	+	-	+	+	-	-	-	+	+	-	-	+	+	-	-	+	+	+	-	-	+	-	+	+	-	-	+	
12	+	+	+	-	+	-	+	-	+	-	-	+	-	-	+	-	-	+	-	-	+	+	-	+	-	+	-	+	-	+	+	+	
13	+	-	-	+	+	-	+	-	-	+	-	-	+	+	-	-	+	+	-	-	+	+	-	+	+	-	+	-	-	+	+	-	
22	+	+	-	+	-	+	-	+	-	+	-	+	-	-	+	-	-	+	-	-	+	+	+	-	+	-	+	-	+	+	+	+	
7	+	-	+	+	-	-	-	-	+	+	+	-	-	-	-	+	-	+	+	+	+	+	-	-	+	+	+	+	-	-	+	+	
32	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
14	+	+	-	+	+	-	-	+	+	-	-	-	+	+	-	-	-	+	+	-	-	-	+	+	-	-	+	+	-	+	+	+	
8	+	+	+	+	-	-	+	+	-	-	+	-	-	-	-	+	+	-	-	-	-	+	-	-	+	+	-	-	+	+	+	+	
20	+	+	+	-	-	+	+	-	-	+	-	-	+	+	-	-	-	+	+	-	-	-	+	-	-	+	+	-	-	+	+	+	
26	+	+	-	-	+	+	-	-	+	+	+	-	-	-	-	+	+	-	-	-	-	+	+	+	-	-	+	+	-	+	+	+	
1	+	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	-	
31	+	-	+	+	+	+	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	+	-	-	-	-	+	-	
11	+	-	+	-	+	-	-	+	-	+	-	+	-	-	+	-	+	-	+	+	-	+	-	+	-	+	+	-	+	-	+	-	
21	+	-	-	+	-	+	+	-	+	-	-	+	-	-	+	-	+	-	+	+	-	+	+	-	+	-	-	+	-	+	+	-	
9	+	-	-	-	+	-	+	+	-	+	+	-	+	-	+	-	-	+	-	+	-	+	+	-	+	+	-	+	-	-	-	+	
3	+	-	+	-	-	-	-	+	+	+	-	-	-	+	+	+	+	+	+	-	-	-	+	+	+	-	-	-	-	+	-	+	
23	+	-	+	+	-	+	-	-	+	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+	-	-	+	-	+	+	-	+	
29	+	-	-	+	+	+	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-	-	+	+	+	-	-	+		
6	+	+	-	+	-	-	-	+	-	-	-	+	-	-	+	-	+	+	-	-	+	+	+	-	+	+	+	-	+	-	-	-	
28	+	+	+	-	+	+	+	-	+	+	-	+	+	-	-	+	-	+	+	-	-	+	-	-	+	-	-	-	+	-	-	-	
16	+	+	+	+	+	-	+	+	+	-	+	+	-	+	-	-	+	+	-	+	-	-	+	-	-	-	+	-	-	-	-	-	
18	+	+	-	-	-	+	-	-	-	+	+	+	-	+	-	-	+	+	-	+	-	-	-	+	+	+	-	+	+	+	-	-	
17	+	-	-	-	-	+	+	+	+	-	+	+	-	+	-	-	-	+	-	+	+	+	-	+	+	+	+	-	-	-	-	+	
27	+	-	+	-	+	+	-	+	-	-	-	+	+	-	-	+	+	-	-	+	+	-	-	-	+	-	+	+	-	+	-	+	
15	+	-	+	+	+	-	-	-	-	+	+	+	-	+	-	-	-	+	-	+	+	+	+	-	-	-	-	+	+	+	-	+	
5	+	-	-	+	-	-	+	-	+	+	-	+	+	-	-	+	+	-	-	+	+	-	+	+	-	+	-	-	+	-	-	+	
4	+	+	+	-	-	-	+	-	-	-	-	-	-	+	+	+	-	-	-	+	+	+	+	+	+	-	+	+	+	-	-	-	
30	+	+	-	+	+	+	-	+	+	+	-	-	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-	-	-	+	-	-	
10	+	+	-	-	+	-	-	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	-	+	+	-	
24	+	+	+	+	-	+	+	+	-	+	+	-	+	-	+	-	+	-	+	-	+	-	-	+	-	-	-	+	-	-	-	-	

FIGURE 4.2: Plan factoriel fractionnaire 2^{5-2} à partir un plan factoriel complet 2^3

4.3.1 Résolution

Une propriété importante d'un plan fractionnaire est sa **résolution** ou sa capacité à séparer les effets principaux et les interactions d'ordre faible les uns des autres.

Résolution	Aptitude	Exemple
I	Inutile : une expérience comportant exactement une analyse ne teste qu'un niveau de facteur et ne permet donc même pas de distinguer les niveaux haut et bas de ce facteur.	2^{1-1} avec relation de définition $I = A$
II	Inutile : les effets principaux sont confondus avec d'autres effets principaux	2^{2-1} avec relation de définition $I = AB$
III	Estimer les effets principaux, mais ceux-ci peuvent être confondus avec des interactions à deux facteurs	2^{3-1} avec relation de définition $I = ABC$
IV	Estimer les effets principaux non confondus avec des interactions à deux facteurs	2^{4-1} avec relation de définition $I = ABCD$
	Estimer l'effets des interactions à deux facteurs, mais ceux-ci peuvent être confondus avec d'autres interactions à deux facteurs	
V	Estimer les effets principaux non confondus avec les interactions à trois facteurs (ou moins)	2^{5-1} avec relation de définition $I = ABCDE$
	Estimer l'effet des interactions à deux facteurs non confondus avec les interactions à deux facteurs	
	Estimer l'effet des interactions à trois facteurs, mais ils peuvent être confondus, avec d'autres interactions à deux facteurs	
VI	Estimer les effets principaux non confondus avec des interactions à quatre facteurs (ou moins)	2^{6-1} avec relation de définition $I = ABCDEF$
	Estimer l'effet des interactions à deux facteurs non confondus avec des interactions à trois facteurs (ou moins)	
	Estimer l'effet des interactions à trois facteurs, mais ils peuvent être confondus avec d'autres interactions à trois facteurs	

Formellement, la résolution d'un plan correspond à la longueur minimale des mots dans la relation de définition excluant (I). Les plans fractionnaires les plus importants sont ceux des résolu-

tions *III*, *IV* et *V* : les résolutions inférieures à *III* ne sont pas utiles et les résolutions supérieures à *V* sont inutiles, dans la mesure où l'expérimentation étendue ne présente aucun avantage pratique dans la plupart des cas. La structure 2^{5-2} de l'exemple ci-dessus correspond à la résolution *III* puisque sa relation de définition est $I = ABD = ACE = BCDE$.

4.3.2 Génération d'alias

Les générateurs d'alias servent à calculer la structure des alias qui décrit les confusions des plans factoriels fractionnaires. Toute lettre multipliée par elle-même désigne l'identité *I*. *I* multiplié par toute lettre désigne cette même lettre. Les lettres sont commutatives et associatives. Pour obtenir tous les alias pour chaque terme nous multiplions les termes de la relation de définition par les facteurs initiaux constituant la matrice extraite.

- **Commutativité** : $AB=BA$
- **Associativité** : $A(BC)=(AB)C=ABC$
- **Elément Neutre I** (colonne de (+)) : $IA=AI=A$
- **AA=I quel que soit A**

$$I = ABD = ACE = BCDE$$

$$A = ABD * A = BD, A = ACE * A = CE, A = BCDE * A = ABCDE$$

$$B = ABD * B = AD, B = ACE * B = ABCE, B = BCDE * B = CDE$$

$$C = ABD * C = ABCD, C = ACE * C = AE, C = BCDE * C = BDE$$

$$D = ABD * D = AB, D = ACE * D = ACDE, D = BCDE * D = BCE$$

$$E = ABD * E = ABDE, E = ACE * E = AC, E = BCDE * E = BCD$$

$$BE = ABD * BE = ADE, BE = ACE * BE = ABC, BE = BCDE * BE = CD$$

$$BC = ABD * BC = ACD, BC = ACE * BC = ABE, BC = BCDE * BC = DE$$

Termes initiaux

Exp	I	A	B	C	AB	AC	BC	ABC	Y _i
	ABD	BD	AD	AE	D	E	DE	CD	
	ACE	CE	CDE	BDE	BCE	BCD	ABE	BE	
	BCDE	ABCDE	ABCE	ABCD	ACDE	ABDE	ACD	ADE	
1	+	-	-	-	+	+	+	-	Y ₁
2	+	+	-	-	-	-	+	+	Y ₂
3	+	-	+	-	-	+	-	+	Y ₃
4	+	+	+	-	+	-	-	-	Y ₄
5	+	-	-	+	+	-	-	+	Y ₅
6	+	+	-	+	-	+	-	-	Y ₆
7	+	-	+	+	-	-	+	-	Y ₇
8	+	+	+	+	+	+	+	+	Y ₈
h _i	h ₀	h ₁	h ₂	h ₃	h ₄	h ₅	h ₆	h ₇	

FIGURE 4.3: Matrices d'expériences avec contrastes pour un plan fractionnaire 2^{5-2}

4.3.3 Contrastes h_i

Nous reconnaissons bien le modèle d'un plan factoriel complet 2^3 dont nous pourrions utiliser la matrice des effets. Les nouveaux coefficients h_i calculés avec la matrice des effets par la méthode habituelle seront nommés contrastes. Un **contraste** h_i est une somme d'effets de facteurs principaux et d'interactions.

$$h_0 = a_0 + a_{ABD} + a_{ACE} + a_{BCDE}$$

$$h_1 = a_A + a_{BD} + a_{CE} + a_{ABCDE}$$

$$h_2 = a_B + a_{AD} + a_{CDE} + a_{ABCE}$$

$$h_3 = a_C + a_{AE} + a_{BDE} + a_{ABCD}$$

$$h_4 = a_D + a_{AB} + a_{BCE} + a_{ACDE}$$

$$h_5 = a_E + a_{AC} + a_{BCD} + a_{ABDE}$$

$$h_6 = a_{BE} + a_{CD} + a_{ABC} + a_{ADE}$$

$$h_7 = a_{BC} + a_{DE} + a_{ABE} + a_{ACD}$$

Le modèle mathématique qui lie les contrastes aux différentes colonnes peut être décrit sous la même forme que celles des plans complets.

4.4 Estimation des effets et des interactions a_i

Le calcul des contrastes h_i peut être effectué en utilisant la méthode habituelle du calcul des coefficients. Quant au calcul des effets des facteurs principaux et des interactions pour obtenir un modèle en fonction des différents termes, un ensemble d'hypothèses peuvent être appliquées pour déduire les différentes valeurs des a_i .

- **Hypothèse 1** : Les interactions d'ordre 3 et plus sont supposées négligeables.
- **Hypothèse 2** : S'il s'avère qu'un contraste est négligeable, tous les termes aliasés avec ce contraste sont aussi négligeables.
- **Hypothèse 3** : Si les effets de deux facteurs sont négligeables, l'effet de leur interaction est aussi considéré comme négligeable.
- **Hypothèse 4** : Si l'effet d'un des facteurs qui est composant d'une interaction est négligeable, l'interaction est généralement aussi négligeable.

4.5 Exemple applicatif

Nous nous intéressons à étudier l'effet des 4 facteurs A , B , C et D sur une réponse y . La matrice du plan fractionnaire 2^{4-1} avec réponses est donnée par la figures 4.4

Exp	I	A	B	C	D	y
1	1	-1	-1	-1	-1	55
2	1	1	-1	-1	1	65
3	1	-1	1	-1	1	42
4	1	1	1	-1	-1	44
5	1	-1	-1	1	1	58
6	1	1	-1	1	-1	74
7	1	-1	1	1	-1	52
8	1	1	1	1	1	54

FIGURE 4.4: Matrice d'expériences avec réponses pour le plan fractionnaire 2^{5-2}

La matrice du plan fractionnaire 2^{4-1} avec interactions peut être obtenue à partir d'une matrice d'un plan complet 2^3 . Le générateur d'alias pour ce plan est : $I = ABCD$. À partir de ce générateur nous pouvons définir tous les alias qui correspondent aux différents termes (A , B , C , AB , AC , BC et ABC) de la matrice en multipliant le générateur d'alias par les termes de la matrice du plan 2^3):

$$A = A * I = A * ABCD = A * A * BCD = I * BCD = BCD$$

De la même façon les autres alias pour les autres termes peuvent être déterminer pour avoir :

$$B = ACD, C = ABD, AB = CD, AC = BD, BC = AD \text{ et } ABC = D$$

Les contrastes sont calculés de la même manière que les effets dans la section 2.3. Les confusions des différents coefficients sont : $h_0 = a_0 + a_{ABCD}$, $h_1 = a_A + a_{BCD}$, $h_2 = a_B + a_{ACD}$, $h_3 = a_C + a_{ABD}$, $h_4 = a_{AB} + a_{CD}$, $h_5 = a_{AC} + a_{BD}$, $h_6 = a_{BC} + a_{AD}$, $h_7 = a_D + a_{ABC}$.

Exp	I	A	B	C	AB	AC	BC	ABC	y
	ABCD	BCD	ACD	ABD	CD	BD	AD	D	
1	1	-1	-1	-1	1	1	1	-1	55
2	1	1	-1	-1	-1	-1	1	1	65
3	1	-1	1	-1	-1	1	-1	1	42
4	1	1	1	-1	1	-1	-1	-1	44
5	1	-1	-1	1	1	-1	-1	1	58
6	1	1	-1	1	-1	1	-1	-1	74
7	1	-1	1	1	-1	-1	1	-1	52
8	1	1	1	1	1	1	1	1	54
Contraste	$h_0=55,5$	$h_1=3,75$	$h_2=-7,5$	$h_3=4$	$h_4=-2,75$	$h_5=0,75$	$h_6=1$	$h_7=-0,75$	

FIGURE 4.5: Matrice d'expériences avec contrastes pour le plan fractionnaire 2^{5-2}

Dans cette étude, les contrastes dont la valeur absolue est inférieure à 1.6 sont supposés négligeables. Un contraste présente une confusion de plusieurs effets. Pour estimer les effets purs

des facteurs ou leurs interactions, nous appliquons les hypothèses sus-présentées :

$$h_0 = a_0 + a_{ABCD}, \text{ en appliquant l'hypothèse 1 : } a_{ABCD} = 0 \Rightarrow a_0 = h_0 = 55.5.$$

$$h_1 = a_A + a_{BCD}, \text{ en appliquant l'hypothèse 1 : } a_{BCD} = 0 \Rightarrow a_A = h_1 = 3.75.$$

$$h_2 = a_B + a_{ACD}, \text{ en appliquant l'hypothèse 1 : } a_{ACD} = 0 \Rightarrow a_B = h_2 = -7.5.$$

$$h_3 = a_C + a_{ABD}, \text{ en appliquant l'hypothèse 1 : } a_{ABD} = 0 \Rightarrow a_C = h_3 = 4.$$

$$h_7 = a_D + a_{ABC}, \text{ en appliquant l'hypothèse 1 : } a_{ABC} = 0 \Rightarrow a_D = h_7 = -0.75.$$

$$|h_5| < 1.6, \text{ en appliquant l'hypothèse 2 } \Rightarrow a_{AC} = a_{BD} = 0.$$

$$|h_6| < 1.6, \text{ en appliquant l'hypothèse 2 } \Rightarrow a_{BC} = a_{AD} = 0.$$

$$|h_7| < 1.6, \text{ en appliquant l'hypothèse 2 } \Rightarrow a_D = 0.$$

$$a_D = 0, \text{ en appliquant l'hypothèse 4 } \Rightarrow a_{CD} = 0.$$

$$h_4 = a_{AB} + a_{CD} \text{ et } a_{CD} = 0 \Rightarrow a_{AB} = h_4 = -2.75.$$

Le modèle mathématique peut donc être écrit comme suit :

$$y = 55.5 + 3.75A - 7.5B + 4C - 2,75AB \quad (4.2)$$

Chapitre 5

Plans pour surface de réponse

5.1 Introduction

La méthodologie pour surfaces de réponse, ou *RSM*, est un ensemble de techniques mathématiques et statistiques dans laquelle une réponse d'intérêt est influencée par plusieurs variables. Les applications les plus étendues de *RSM* se situent dans les situations particulières où plusieurs variables d'entrée influencent potentiellement une mesure de performance ou une caractéristique de qualité du processus. Ainsi, la mesure de performance ou la caractéristique de qualité s'appelle la réponse. Les variables d'entrée sont parfois appelées variables indépendantes et sont soumises au contrôle du scientifique ou de l'ingénieur.

La méthodologie pour surfaces de réponse comprend : la stratégie expérimentale d'exploration de l'espace du processus, la modélisation empirique permettant de développer une relation d'approximation appropriée entre le rendement et les variables du processus, ainsi que des méthodes d'optimisation permettant de rechercher les valeurs des variables du processus, qui produisent des valeurs souhaitables de la réponse.

Le modèle polynomial du second ordre peut s'écrire par l'équation 5.1.

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

Où Y est la réponse prévue, β_0 est le coefficient constant du modèle, β_i est le i^{eme} coefficient linéaire du paramètre d'entrée x_i , β_{ii} est le i^{eme} coefficient quadratique du paramètre d'entrée x_i , β_{ij} est le coefficient d'interaction entre les paramètres d'entrée x_i et x_j , et ϵ est l'erreur du modèle. Nous notons par $\hat{\beta}$ le vecteur des coefficients estimés du modèle.

Les coefficients estimés $\hat{\beta}$ sont obtenus par la régression linéaire multiple déjà présentée dans la section 3.2.1 selon la formule suivante : 5.2 :

$$\hat{\beta} = (X'X)^{-1}X'y \quad (5.2)$$

Où X' est la transposée de la matrice X , X est la matrice du modèle qui dépend des points expérimentaux choisis pour exécuter le plan du modèle considéré et y est le vecteur des réponses. L'examen d'un modèle tel que 5.1 a trois objectifs :

- **Premièrement** : établir une relation, même approximative, entre y et x_1, x_2, \dots, x_k qui peut être utilisée pour prédire les valeurs de réponse pour des valeurs données des variables.
- **Deuxièmement** : déterminer à l'aide des tests d'hypothèses la signification des facteurs.
- **Troisièmement** : déterminer les paramètres optimaux de x_1, x_2, \dots, x_k entraînant la réponse maximale (ou minimale) sur une certaine région d'intérêt.

5.2 Les plans composites centrés

Un plan composite centré de Box-Wilson, couramment appelé Plan Composite Centré, contient un plan factoriel ou fractionnaire incorporé avec des points centraux augmentés d'un groupe de points en étoile permettant d'estimer la courbure. Si la distance entre le centre de l'espace du plan et un point factoriel est de ± 1 unité pour chaque facteur, la distance entre le centre de l'espace de calcul et un point en étoile est $|\alpha| > 1$. La valeur précise de α et le nombre de points centraux que le plan doit contenir dépendent de certaines propriétés souhaitées pour le plan et du nombre de facteurs en cause.

La Figure 5.1 représente un plan composite centré pour trois facteurs. Les points 1, 2, 3, 4, 5, 6, 7, 8 sont les points expérimentaux d'un plan 2^3 . Le point 15 est le point central. Ce point peut être répliqué plusieurs fois. Les points 9, 10, 11, 12, 13 et 14 sont les points axiaux. Ces six derniers points forment le plan en étoile.

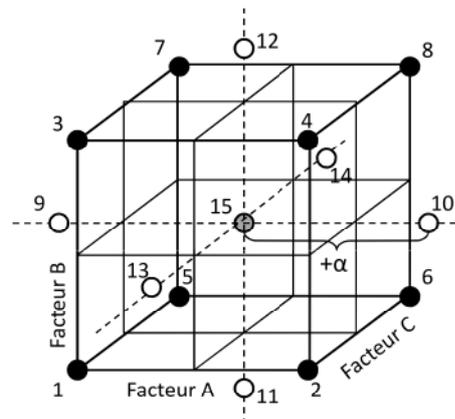


FIGURE 5.1: Plan composite centré pour 3 facteurs

Un plan composite centré contient toujours deux fois plus de points en étoiles que le nombre de facteurs du plan. Le Plan Composite Centré Circonscrit (CCC) est la forme originale du plan composite centré. Les points en étoiles sont à une certaine distance α du centre en fonction des propriétés souhaitées pour le plan et du nombre de facteurs de la conception. Ces modèles ont une symétrie circulaire, sphérique ou hyper-sphérique et nécessitent 5 niveaux pour chaque facteur. L'augmentation d'un plan factoriel ou fractionnaire de résolution V existant avec des points en étoile peut produire ce plan.

Pour maintenir la rotabilité, la valeur de α dépend du nombre d'essais expérimentaux dans la partie factorielle (complète ou fractionnaire) du plan composite centré :

$$\alpha = (2^{k-p})^{1/4} \quad (5.3)$$

Où k est le nombre de facteurs à étudier et p la taille de la fraction ($0 \Rightarrow$ pas de fraction, $1 \Rightarrow 1/2$, $2 \Rightarrow 1/4$, $3 \Rightarrow 1/8$, etc.)

La figure 5.2 montre une matrice d'expériences pour un plan composite centré (factoriel complet augmenté) pour 3 facteurs A , B et C avec trois répétitions du point central. Nous remarquons que $\alpha = (2^{k-p})^{1/4} = (2^{3-0})^{1/4} = 1.6818$.

Exp	A	B	C
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
9	-1,6818	0	0
10	1,6818	0	0
11	0	-1,6818	0
12	0	1,6818	0
13	0	0	-1,6818
14	0	0	1,6818
15	0	0	0
16	0	0	0
17	0	0	0

FIGURE 5.2: Matrice d'expériences pour un plan composite centré pour 3 facteurs

Il existe une autre forme du plan composite centré appelée Plan Composite Face Centrée, dont les points en étoiles sont au centre de chaque face de l'espace factoriel, donc $\alpha = \pm 1$. Cette variété nécessite 3 niveaux de chaque facteur. L'augmentation d'un plan factoriel ou d'un plan de résolution V existant avec des points en étoiles appropriés peut également produire ce plan.

5.3 Les plans de *Box-Behnken*

Le plan de *Box-Behnken* est une méthodologie pour surfaces de réponse (*RSM*) qui ne nécessite que trois niveaux pour exécuter une expérience (Figure 5.3). Il s'agit d'un plan spécial à 3 niveaux car il ne contient aucun point aux sommets de la région expérimentale. Cela peut être avantageux lorsque les points situés aux angles du cube représentent des combinaisons de niveaux d'un coût prohibitif ou impossibles à tester en raison des contraintes du processus.

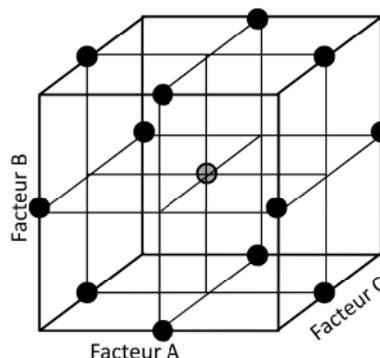


FIGURE 5.3: Plan de *Box-Behnken* pour 3 facteurs

La figure 5.4 montre une matrice d'expériences pour un plan Box-Behnken pour 3 facteurs A , B et C avec trois répétitions du point central.

Exp	A	B	C
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	0	0	0
14	0	0	0
15	0	0	0

FIGURE 5.4: Matrice d'expériences pour un plan de *Box-Behnken* pour 3 facteurs

5.4 Les plans optimaux

Un plan optimal est une forme de plan fournie par un algorithme informatique. Ce type de plan assisté par ordinateur est particulièrement utile lorsque les plans classiques ne s'appliquent pas. Contrairement aux plans standards classiques tels que les factoriels et les factoriels fractionnaires, la matrice d'un plan optimal n'est généralement pas orthogonale et les estimations des effets sont corrélées.

Ce type de plan est toujours une option, quel que soit le type de modèle que l'expérimentateur souhaite adapter (par exemple, premier ordre, premier ordre plus quelques interactions, quadratique complet, cubique, etc.) ou l'objectif spécifié pour l'expérience (par exemple, dépistage, surface de réponse, etc.). Le plan optimal est une optimisation directe basée sur un critère d'optimalité choisi et le modèle à adapter. Le critère d'optimalité utilisé pour générer un plan optimal consiste à maximiser $|X'X|$, le déterminant de la matrice d'informations $X'X$.

Ce critère d'optimalité permet de minimiser la variance généralisée des estimations des paramètres pour un modèle prédéfini. En conséquence, l'optimalité d'un plan optimal donné dépend du modèle. En d'autres termes, l'expérimentateur doit spécifier un modèle pour le plan avant qu'un ordinateur ne puisse générer les combinaisons de traitement spécifiques. Étant donné le nombre total de traitements traités pour une expérience et un modèle spécifié, l'algorithme informatique choisit l'ensemble optimal d'essais d'un plan parmi un ensemble d'essais de traitement des plans possibles candidats. Cet ensemble candidat de traitements comprend généralement toutes les combinaisons possibles des différents niveaux de facteurs que l'on souhaite utiliser dans l'expérience.

En d'autres termes, l'ensemble candidat est un ensemble de combinaisons de traitement parmi lesquelles l'algorithme optimal choisit les combinaisons de traitement à inclure dans le plan (Rien ne garantit que le plan généré par l'ordinateur est réellement optimal).

Les raisons d'utiliser des plan optimaux au lieu des plans classiques standards tombent généralement dans deux catégories :

- Les plans factoriels ou factoriels fractionnaires standards nécessitent trop d'essais pour la quantité de ressources ou le temps alloué à l'expérience.
- L'espace du plan est limité (l'espace de processus contient des paramètres de facteurs non réalisables ou impossibles à exécuter).

Supposons que : un processus industriel comporte trois variables du plan ($k = 3$), en raison des ressources limitées, seuls $n = 12$ expériences peuvent être réalisées et qu'un jugement technique spécifie le modèle suivant comme représentation appropriée du processus.

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{11}x_1^2 + \epsilon \quad (5.4)$$

Compte tenu des spécifications expérimentales, la première étape de la conception consiste à créer un ensemble de points candidats. L'ensemble candidat est une table de données avec une ligne pour chaque point (exécution) à prendre en compte pour le plan, souvent un factoriel complet. Les plans optimaux maximisent la D-efficacité, critère de volume de la variance généralisée des estimations des paramètres. Les valeurs de la D-efficacité sont en fonction du nombre de points dans le plan, du nombre de variables indépendantes dans le modèle et de l'erreur type maximale pour la prédiction sur les points du plan. Le meilleur plan est celui avec la plus grande D-efficacité.

Exp	A	B	C
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	-1	1	1
5	0	-1	-1
6	0	-1	1
7	0	1	-1
8	0	1	1
9	1	-1	-1
10	1	-1	1
11	1	1	-1
12	1	1	1

FIGURE 5.5: Matrice du plan optimal obtenu pour l'exemple étudié

Les packages logiciels peuvent avoir différentes procédures pour générer des plans optimaux, de sorte que le plan final peut être différent en fonction du logiciel utilisé.

Chapitre 6

Plans de mélange

6.1 Introduction

Dans une expérience de mélange, les facteurs indépendants sont des proportions des différents composants d'un mélange (voir figure 6.1). Par exemple, si vous souhaitez optimiser la résistance à la traction de l'acier inoxydable, les facteurs d'intérêt peuvent être les proportions de fer, de cuivre, de nickel et du chrome dans l'alliage. Le fait que les proportions des différents facteurs doivent totaliser 100% rend impossible l'utilisation des plans d'expériences ordinaires déjà vus.

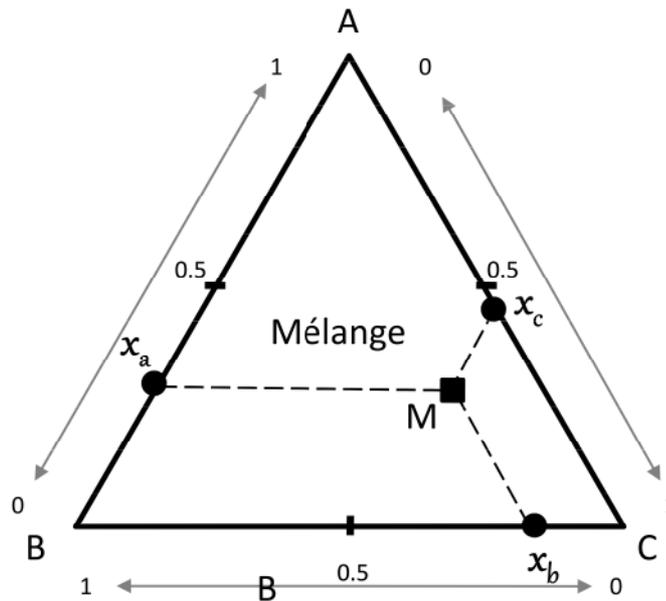


FIGURE 6.1: Valeurs des composants dans un plan de mélange

6.2 Types des plans de mélange

Lorsque les composants du mélange sont soumis à la contrainte qu'ils doivent s'additionner, il existe des plans de mélange standard pour l'ajustement, telles que les **Plans de Mélange en Réseaux** et les **Plans de Mélange Centrés**. Lorsque des composants de mélange sont soumis à des contraintes supplémentaires, telles qu'une valeur maximale et / ou minimale pour chaque composant, des plans autres que les plans de mélange standard, appelées plans de mélange contraints ou **Plans Sommets Extrêmes**, sont appropriés.

Pour évaluer le manque d'ajustement, une estimation de l'erreur pure est nécessaire. Cette estimation provient de la réplication de plusieurs points du plan (Cette réplication peut être appliquée aussi sur des points ajoutés à un plan standard tels que les points axiaux et le point central).

6.2.1 Plans de mélange en réseau

Un Plan de Mélange en Réseau (q, m) pour q composants est constitué de points définis par les paramètres de coordonnées suivants : les proportions assumées par chaque composant

prennent les $m + 1$ valeurs également espacées de 0 à 1 .

$x_i = 0, 1/m, 2/m, \dots, 1$ pour $i = 1, 2, \dots, q$ et toutes les combinaisons possibles (mélanges) des proportions de cette équation sont utilisées. À l'exception du centre, tous les points du plan se trouvent sur les limites du simplexe.

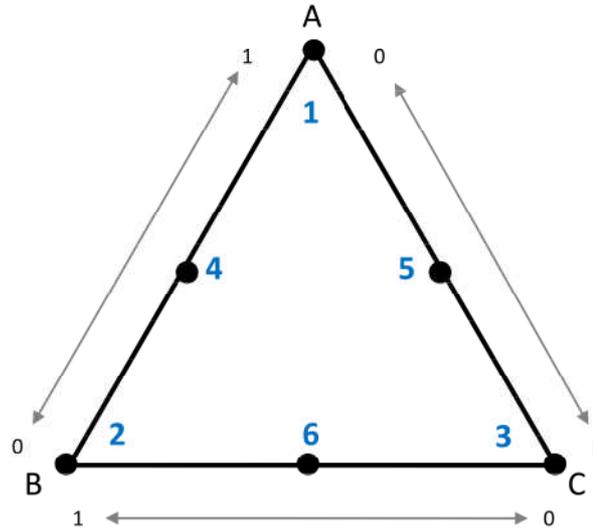


FIGURE 6.2: Plan de mélange en réseau (3, 2)

Considérons un mélange à trois composants pour lequel le nombre de niveaux équidistants pour chaque composant est de deux (c'est-à-dire, $x_i = 0, 0.5, 1$). La figure 6.2 montre la région expérimentale et la distribution des points du plan sur la région simplexe. Il existe 6 points expérimentaux pour le plan de mélange en réseau (3, 2).

Dans cet exemple, $q = 3$ et $m = 2$. Si nous utilisons tous les mélanges possibles des trois composants avec ces proportions, le simplexe réseau contient alors les 6 expériences de mélange énumérées dans la matrice d'expériences de la figure 6.3.

Exp	A	B	C
1	1	0	0
2	0	1	0
3	0	0	1
4	0,5	0,5	0
5	0,5	0	0,5
6	0	0,5	0,5

FIGURE 6.3: Matrice d'expériences pour un plan de mélange en réseau (3, 2)

Les points comprennent les composantes pures et suffisamment de points entre eux pour estimer une équation de degré m . Un plan de mélange en réseau (3,3) diffère du plan de mélange centré en ayant suffisamment de points pour estimer un modèle cubique complet.

L'augmentation du plan de mélange en réseau revient à ajouter des points intérieurs ; des points de mélange de contrôle et un point central. Les points de mélange de contrôle sont à mi-chemin entre le point central et chaque sommet de la simplexe. La figure 6.4 montre un plan de mélange en réseau (3, 2) augmenté de 4 points (un point central + trois points de mélange de contrôle).

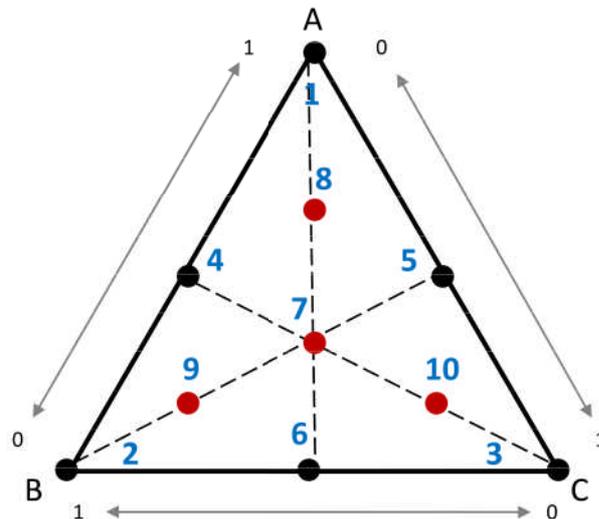


FIGURE 6.4: Plan de mélange en réseau (3, 2) augmenté

La nouvelle matrice d'expériences qui correspond au nouveau plan de mélange en réseau (3, 2) augmenté est illustré dans la figure 6.5

Exp	A	B	C
1	1	0	0
2	0	1	0
3	0	0	1
4	0,5	0,5	0
5	0,5	0	0,5
6	0	0,5	0,5
7	0,3333	0,3333	0,3333
8	0,6667	0,1667	0,1667
9	0,1667	0,6667	0,1667
10	0,1667	0,1667	0,6667

FIGURE 6.5: Matrice d'expériences pour un plan de mélange en réseau (3, 2) augmenté

6.2.2 Plans de mélange centrés

Un deuxième type de plan de mélange est le plan de mélange centré. Dans le plan de mélange centré à q composants, le nombre de points distincts est $2^q - 1$ points ou mélanges, ces points sont organisés de la manière suivante :

- q composants purs : q permutations de $(1,0,0,0, \dots, 0)$.
- Mélanges binaires à proportions égales $\begin{bmatrix} q \\ 2 \end{bmatrix}$ permutations de $(1/2, 1/2, 0, 0, \dots, 0)$.
- Mélanges ternaires à proportions égales $\begin{bmatrix} q \\ 3 \end{bmatrix}$ permutations de $(1/3, 1/3, 1/3, 0, \dots, 0)$.
- ...
- Un mélanges q -aire avec des proportions égales $(1/q, 1/q, 1/q, \dots, 1/q)$.

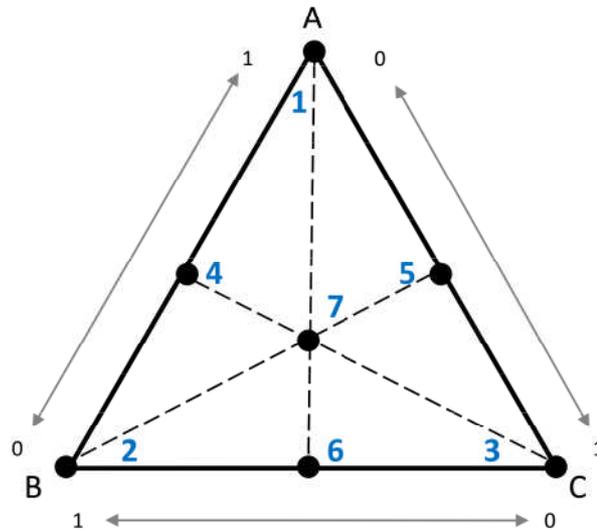


FIGURE 6.6: Plan de mélange centré à 3 composants

La figure 6.7 montre les différents points expérimentaux pour un plan de mélange centré $(3, 2)$.

Exp	A	B	C
1	1	0	0
2	0	1	0
3	0	0	1
4	0,5	0,5	0
5	0,5	0	0,5
6	0	0,5	0,5
7	0,3333	0,3333	0,3333

FIGURE 6.7: Matrice d'expériences pour un plan de mélange centré à 3 composants

Pour un nombre donné de composants, il n'existe qu'un seul plan de mélange centré, alors que différents plans de mélange en réseaux sont possibles. Les points d'un plan de mélange centré prendront en charge le polynôme suivant :

$$y = \sum_{i=1}^q a_i x_i + \sum_{i=1}^q \sum_{i < j}^q a_{ij} x_i x_j + \dots + \sum_{i=1}^q \dots \sum_{i < j < \dots < q}^q a_{i j \dots q} x_i x_j \dots x_q \tag{6.1}$$

Les plans de mélange centrés peuvent être augmentés en ajoutant des points intérieurs ; ces points sont à mi-chemin entre le point central et chaque sommet. La figure 6.8 montre un plan de mélange centré à 3 composants augmenté avec les 3 points.

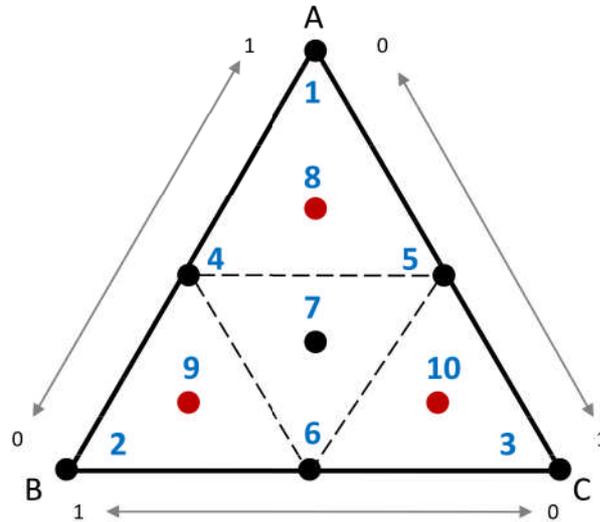


FIGURE 6.8: Plan de mélange centré à 3 composants (augmenté)

La nouvelle matrice d'expériences qui correspond au nouveau plan de mélange centré à 3 composants augmenté est illustré par la figure 6.9

Exp	A	B	C
1	1	0	0
2	0	1	0
3	0	0	1
4	0,5	0,5	0
5	0,5	0	0,5
6	0	0,5	0,5
7	0,3333	0,3333	0,3333
8	0,6667	0,1667	0,1667
9	0,1667	0,6667	0,1667
10	0,1667	0,1667	0,6667

FIGURE 6.9: Matrice d'expériences pour un plan de mélange centré à 3 composants (augmenté)

6.2.3 Plans de mélange contraints (Sommets extrêmes)

Les plans de sommets extrêmes sont des plans de mélange qui couvrent uniquement une sous-partie ou un espace plus petit dans le simplexe. Ces plans doivent être utilisés lorsque l'espace du plan choisi n'est pas lui-même un plan simplexe. La présence de contraintes de limite inférieure et supérieure sur les composants crée souvent cette condition. Les contraintes sont de la forme $I_i \leq x_i \leq S_i, (i = 1, 2, \dots, q)$ avec $I_i \leq x_i \leq S_i, i = 1, 2, \dots, q$, où I_i est la limite inférieure du

i^{eme} composant et S_i la limite supérieure du i^{eme} composant. La forme générale d'un problème de mélange contraint (Plan de sommets extrêmes) est :

$$x_1 + x_2 + \dots + x_q = 1$$

$$I_i \leq x_i \leq S_i \text{ pour } i = 1, 2, \dots, q \text{ Avec : } I_i \geq 0 \text{ et } S_i \leq 1$$

Par exemple, vous devez déterminer les proportions de la farine, du lait, de la levure chimique, d'œufs et d'huile dans un mélange à crêpes qui donneront un produit optimal à base de goût. Comme les expériences précédentes indiquent qu'un mélange ne contenant pas tous les ingrédients ou contenant trop de levure chimique ne répondra pas aux exigences de la saveur, vous décidez de limiter le plan en définissant des limites inférieures et supérieures.

L'objectif d'un plan de sommets extrêmes est de choisir des points du plan qui couvrent de manière adéquate l'espace du plan. La figure suivante montre les sommets extrêmes d'un plan à trois composants avec des contraintes supérieures et inférieures :

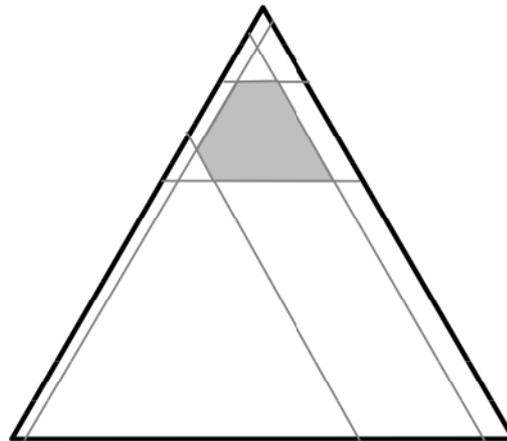


FIGURE 6.10: Limites inférieures et supérieures dans un plan de mélange contraint

Les lignes en gris clair représentent les limites inférieures et supérieures des composants. La zone en gris foncé représente l'espace du plan. Les points sont placés aux sommets extrêmes de l'espace du plan.

6.3 Combinaison des plans de mélange avec d'autres plans

Il est possible de combiner p facteurs ordinaires avec q composants de mélange dans un plan. Dans une expérience factorielle (ou RSM) avec les p facteurs, les propriétés des facteurs doivent être examinées pour chaque niveau de tous les composants q du mélange. Le nombre de variables du processus augmente exponentiellement la quantité des points expérimentaux. Par conséquent, le nombre des facteurs ordinaires doit être réduit au minimum. Le plan expérimental combiné implique la mise en place d'une expérience factorielle à chaque point d'une configuration dans les composants du mélange ou la mise en place d'une expérience de mélange à chaque point

d'une configuration dans les points factoriels

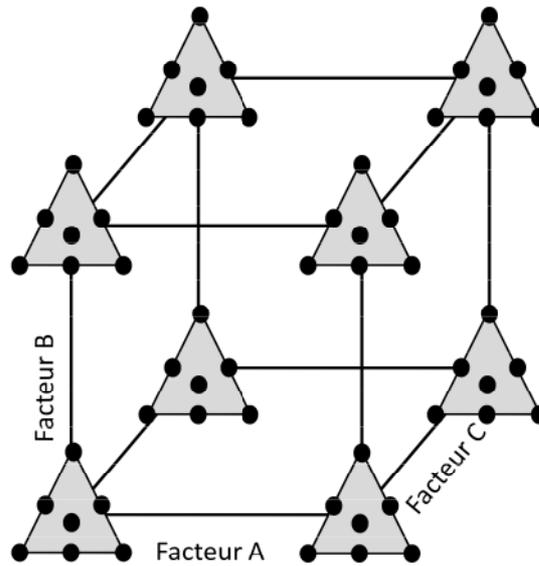


FIGURE 6.11: Un plan combiné de variables factorielles et de composants de mélange

Bibliographie

- [1] N.H. Bingham and J.M. Fry. *Regression : Linear Models in Statistics*. Springer Undergraduate Mathematics Series. Springer London, 2010.
- [2] ReliSoft Corporatuon. Experiment design analysis reference. ”[http ://www.reliawiki.org](http://www.reliawiki.org)”, 2015. (Accessed on 20/08/2015).
- [3] J. Goupy and L. Creighton. *Introduction aux plans d’expériences - 3ème édition - Livre+CD-Rom*. Technique et ingénierie. Dunod, 2006.
- [4] Minitab Inc. Minitab 18 support. ”[https ://support.minitab.com/en-us/minitab/18](https://support.minitab.com/en-us/minitab/18)”. (Accessed on 20/10/2019).
- [5] Moresteam. Design of experiments (doe). ”[https ://www.moresteam.com/toolbox/design-of-experiments.cfm](https://www.moresteam.com/toolbox/design-of-experiments.cfm)”. (Accessed on 17/12/2018).
- [6] R.H. Myers, D.C. Montgomery, and C.M. Anderson-Cook. *Response Surface Methodology : Process and Product Optimization Using Designed Experiments*. Wiley Series in Probability and Statistics. Wiley, 2009.
- [7] Stat-Ease. *Handbook for Experimenters*. Stat-Ease, Inc, T2021 East Hennepin Ave, Suite 480 Minneapolis, MN 55413, 2016.