

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique

Université Akli Mohand Oulhadj - Bouira -

ⵍⵎⵓⵎⵉⵏⵜ ⵏ ⵉⵏⵙⵉⵎⵏⵏⵉⵎ ⵏ ⵉⵏⵙⵉⵎⵏⵏⵉⵎ
ⵏ ⵉⵏⵙⵉⵎⵏⵏⵉⵎ ⵏ ⵉⵏⵙⵉⵎⵏⵏⵉⵎ



وزارة التعليم العالي والبحث العلمي
جامعة أكلي محمد أوحاج
- البويرة -

Faculté des Sciences et des Sciences Appliquées

كلية العلوم والعلوم التطبيقية

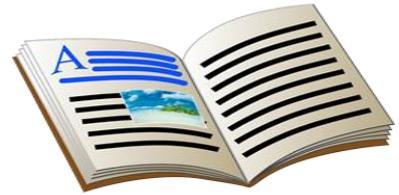
Département de Génie des Procédés

Polycopié de cours

En : Génie des Procédés

Spécialité : Génie Chimique

Niveau : Master II



Les plans d'expériences

Par : Dr ZIANI Salima

Année : 2018/2019

Table de matière

Introduction	1
--------------------	---

Chapitre I: Introduction aux plans d'expériences et plans factoriels

I. 1. Introduction.....	2
I.2. Méthodologies utilisées pour la stratégie expérimentale.....	2
I.2.1. Méthode classique (Gauss-Seidel).....	2
I.2.2. Méthode moderne-Plans d'expériences (Box-Wilson).....	3
I.3. Qu'est ce qu'un plan d'expériences ?	3
I.4. Principe des plans d'expériences	4
I.5. Etapes chronologiques des plans d'expériences.....	4
I.6. Terminologies	4
I.6.1. Facteurs ou variables naturelles (réelles)	5
I.6.2. Réponse	5
I.6.3. Domaine d'étude.....	5
I.6.4. Surfaces de réponse.....	5
I.6.5. Variables codées (variables centrées réduites).....	6
I.6.6. Matrice des expériences.....	7
I.6.7. Matrice des effets.....	7
I.6.8. Effet d'un facteur	7
I.6.9. Notion d'interaction.....	8
I.7. Notion de modèle et de régression linéaire multiple	8
I.8. Plans factoriels complets à deux niveaux 2^k	10
I.8.1. Modèles mathématiques associés aux plans factoriels complets	11
I.8.2. Calcul des coefficients du modèle	12
I.8.3. Analyse de régression sous forme matricielle	12
I.9. Optimalité des plans d'expériences.....	15
I.9.1. Critère de la matrice unité (matrice d'Hadamard)	15
I.9.2. Critère d'orthogonalité	15
I.9.3. Critère de la trace minimale.....	15
I.10. Exemple d'application.....	16
I.10.1. Problème à résoudre	16

I.10.2. Formulation du problème	16
I.10.3. Planification et expérimentation	17
I.10.4. Analyse des résultats	17

Chapitre II: Tests de signification et validation du modèle:

Analyse statistique

II.1. Introduction	19
II.2. Vérification de la signification des coefficients (effets).....	19
II.2. 1. Cas où chaque essai est répété m fois.....	19
II.2.2. Cas où l'essai au centre est répété n_0 fois.....	20
II.3. Validation du modèle	21
II.3.1. Recherche de biais	21
II.3.2. Validation de l'équation de régression	22
II.3.3. Coefficient de détermination	22
II.3.4. Analyse des résidus.....	23
II.4. Exemple d'application.....	23
II.4.1. Analyse statistique de l'équation de régression.....	23
II.4.1.1. Signification des coefficients de l'équation de régression.....	24
II.4.1.2. Validation du modèle	25
II.4.1.3. Evaluation de la qualité du modèle.....	25

Chapitre III: Les plans fractionnaires

III.1. Introduction	29
III.2. Plans factoriels fractionnaires à deux niveaux 2^{k-q}	29
III.3. Conception d'un plan fractionnaire.....	30
III.3.1. Notion de contrastes et d'aliasés	30
III.3.2. Notion de générateur d'aliasés	31
III.3.3. Hypothèses d'interprétation	32
III.4. Exemple d'application.....	33

Chapitre IV: Les plans pour surfaces de réponse

IV.1. Introduction.....	38
IV.2. Plans pour surfaces de réponse	38
IV.2.1. Plan composite centrés	38
IV.2.1.1. Propriétés des plans composites	39

IV.2.2. Plan Box- Behnken.....	41
IV.3. Notions de surfaces de réponse et courbes isoréponses.....	43
IV.3.1. Surfaces de réponse	44
IV.3.2. Courbes isoréponses.....	44
IV.4. Exemple d'application.....	45
IV.4.1. Problème à résoudre.....	45
IV.4.2. Formulation du problème.....	45
IV.4.3. Planification et expérimentation	46
IV.4.4. Analyse des résultats.....	47
IV.4.5. Optimisation	50

Chapitre V: Les plans de mélange

V.1. Introduction	53
V.2. Contrainte fondamentale des mélanges	53
V.3. Plans de mélanges sans contraintes	53
V.3.1. Représentation géométriques	53
V.3. 2. Emplacement des points expérimentaux	55
V.4. Plans de mélanges avec contraintes	56
V.4.1. Basses teneurs interdites	56
V.4.2. Hautes teneurs interdites	56
V.4.3. Hautes et basses teneurs interdites	56
V.4.4. Contraintes relationnelles.....	57
V.4.5. Emplacement des points expérimentaux.	57
V.5. Modèles mathématiques associés aux plans de mélanges	58
V.5.1. Modèle de premier degré.....	58
V.5.2. Modèle de second degré	58
V.5.3. Modèle de troisième degré.....	59

Références bibliographiques

Introduction

Si l'expérimentation est aussi ancienne que la science, la planification expérimentale, destinée à optimiser le processus d'obtention des données, ne date cependant que du XX^e siècle. Cette nouvelle méthodologie développée au cours de ces dernières années constitue un outil sûr, de portée universelle, pratique, indispensable et permettant de trouver les conditions d'optimum lors du traitement des données. Au fait, pour mener une recherche expérimentale planifiée, la méthodologie des plans d'expériences est utilisée. Cette démarche aide l'expérimentateur à structurer sa recherche de manière différente, à valider ses propres hypothèses, à mieux comprendre les phénomènes étudiés et à solutionner les problèmes.

La méthode des plans d'expériences vise à organiser de façon économique et rationnelle une expérience dépendant de plusieurs paramètres. Cette méthode permet d'assurer la qualité des résultats de l'expérience tout en apportant le maximum de renseignements pour un minimum d'expériences à réaliser. Son succès est en partie lié aux besoins de compétitivité des entreprises mais aussi, à une envie de changer la manière de réaliser des expérimentations. Cependant, il faut comprendre que les plans d'expériences ne sont pas un outil destiné a priori à la recherche fondamentale car ils ne permettront jamais une explication du phénomène physico-chimique étudié.

Ce document de cours destiné, principalement aux étudiants de Master et de Licence de toutes spécialités, et plus spécifiquement aux étudiants de Mater II Génie des Procédés, Génie Alimentaire et Ingénierie et Gestion de l'Eau, traite les plans d'expériences de manière pointue. Il peut également servir de support pour les doctorants de toutes spécialités. En revanche, il est fortement recommandé et conseillé pour les étudiants de maîtriser les outils mathématiques et informatiques afin de pouvoir maîtriser la planification statistique d'expériences qui apporte une contribution essentielle et absolument nécessaire pour celui qui entreprend un programme d'essais en laboratoire ou sur des procédés industriels. Ce document couvre, en cinq chapitres, les bases et les principaux plans utilisés par les expérimentateurs des nombreuses disciplines pour développer une procédure adéquate qui leur permet de tirer le maximum d'informations avec un minimum d'essais expérimentaux, ce qui répond parfaitement au défi qu'impose le contexte économique actuel.

Chapitre I

Introduction aux plans d'expériences et plans factoriels

Chapitre I : Introduction aux plans d'expériences et plans factoriels

I. 1. Introduction

L'expérimentation est utilisée dans de nombreuses disciplines pour aider le chercheur à améliorer ses connaissances et répondre aux questions qu'il se pose concernant un phénomène donné. Ainsi, pour étudier et surtout pour suivre l'évolution d'un processus quelconque (en Génie Chimique, Génie des Polymères, Génie Civil, Génie Mécanique, Génie Electrique, Physique, Génie Biologique,...), l'expérimentateur peut adapter une approche méthodologique rigoureuse qui l'emmène à effectuer un très grand nombre d'essais, (Exemple : étude de l'effet des paramètres opératoires sur un phénomène). Dans ce contexte, le nombre d'essais requis est soumis à deux tendances contradictoires :

1. d'un coté, il est nécessaire d'inclure tous les facteurs qui influent sur le process et de les étudier à différents niveaux afin d'obtenir un modèle représentatif ;
2. d'un autre côté, il est nécessaire de minimiser le nombre d'essais pour réduire la durée et le coût de l'expérimentation.

Dans ces conditions, l'approche méthodologique suivie ne peut être qu'une stratégie expérimentale qui va de la conception à la réalisation des expériences.

I.2. Méthodologies utilisées pour la stratégie expérimentale optimale

La stratégie optimale est définie comme une stratégie qui, pour un nombre minimum d'essais, fournit un modèle expérimental indiquant la direction à suivre pour atteindre l'optimum. Pour réaliser une modélisation expérimentale optimale, il existe deux méthodes :

I.2.1. Méthode classique (Gauss-Seidel)

La stratégie classique est définie par le principe « *un facteur à un moment donné* », qui consiste à fixer le niveau de tous les facteurs, sauf un, et de mesurer la fonction objective en fonction de l'évolution du facteur non fixé [1-3]. Ses inconvénients majeurs sont [1-3] :

- étude de plusieurs facteurs avec plusieurs points expérimentaux (travail gigantesque qui dépasse parfois la limite du raisonnable) ;
- utilisation du même algorithme pour l'étude de chaque paramètre à part.

I.2.2. Méthode moderne-Plans d'expériences (Box-Wilson)

La méthode moderne est définie par le principe « *tous les facteurs à chaque moment* ». Elle consiste à faire varier les niveaux de tous les facteurs à la fois, mais d'une manière programmée et raisonnée [1-3]. Les avantages de cette méthode sont [1-3]:

- minimisation du nombre d'essais ;
- nombre de facteurs étudiés très grand ;
- mise en évidence de possibles interactions entre les facteurs ;
- optimisation des paramètres opératoires ;
- meilleure précision sur les résultats qu'un modèle classique ;
- modélisation des résultats (conception d'un modèle mathématique).

I.3. Qu'est ce qu'un plan d'expériences ?

Un plan d'expériences n'est pas une série d'essais au hasard, ni sélectionnée par la seule intuition mais une stratégie optimale permettant de prédire, avec le maximum de précision une réponse à partir d'un nombre minimal d'essais liés au modèle postulé. De manière générale, la méthode des plans d'expériences cherche à déterminer et établir un lien qui existe entre une grandeur d'intérêt y et des variables x_i qui peuvent modifier la valeur de y via la fonction [1, 3-5] :

$$y_i = f(x_j)$$

Avec :

y_i : réponses du système (réponses mesurées) ;

x_j : facteurs ou variables d'entrées du système qui peuvent être continues^(*), discrètes^(**).

La construction d'un plan d'expériences consiste à extraire du domaine expérimental, un nombre suffisant N de combinaisons particulières afin d'estimer, avec une incertitude à la fois minimale mais aussi homogène, les M inconnues du modèle tout en respectant au mieux les contraintes techniques et économiques de l'étude.

(*) : peuvent adopter toutes les valeurs numériques réelles dans l'intervalle étudié.

(**) : ne peuvent prendre que des valeurs particulières : un nom, une lettre, une propriété....

I.4. Principe des plans d'expériences

L'utilisation des plans d'expériences dans des systèmes simples ou complexes permet d'éclaircir et de comprendre le fonctionnement des ces derniers en les simulant à une « *boite noire* » (Figure I.1). Le principe de la boite noire consiste à imposer aux facteurs d'entrées des variations particulières, à mesurer les variations induites des réponses et à en déduire les relations entre facteurs et réponses sans avoir une connaissance maximale du phénomène étudié [3, 6].

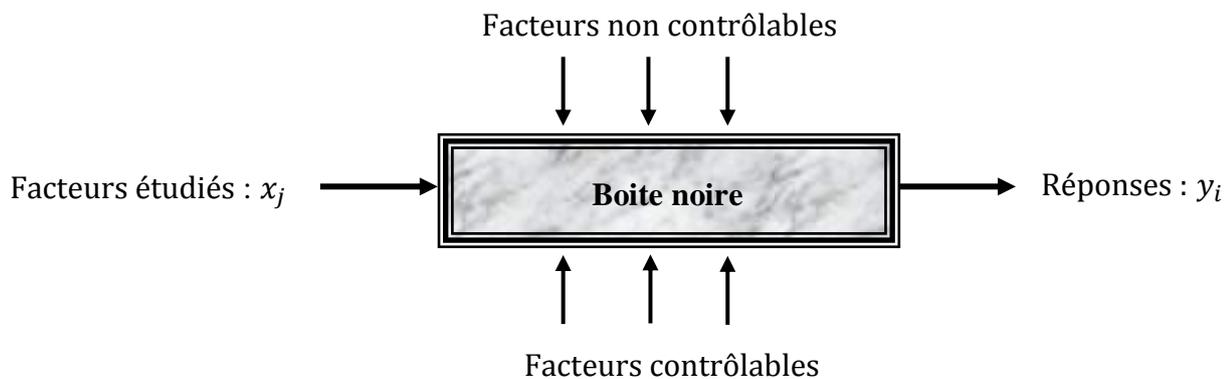


Figure I.1 : Schéma représentatif du contexte de la boite noire et du principe des plans d'expériences

I.5. Etapes chronologiques des plans d'expériences

L'objectif et le principe des plans d'expériences mettent l'accent sur le fait de programmer une méthodologie bien précise, articulée sur une succession d'étapes pour accomplir l'étude entamée. Cette succession d'étapes repose sur [1, 7, 8] :

1. la formalisation du problème (définir les objectifs, facteurs, réponses, domaine expérimental) ;
2. la planification (construction du plan d'expériences) ;
3. l'expérimentation (effectuer les essais) ;
4. l'analyse des résultats (extraire le maximum d'informations des résultats de l'expérience) ;
5. l'optimisation de la réponse (trouver des conditions expérimentales à un optimum de la réponse étudiée).

I.6. Terminologies

Les utilisateurs des plans d'expériences pratiquent un vocabulaire propre à ce domaine. Dans ce qui suit, nous illustrons les terminologies les plus importantes [2, 5-10]:

I.6.1. Facteurs ou variables naturelles (réelles)

Grandeurs modifiables par l'expérimentateur z_j , censées influencer sur les variations de la réponse y . Lorsque la variable z_j prend deux niveaux dans une expérimentation, nous désignons par :

$z_{j(\min)}$: niveau (borne) inférieur ;

$z_{j(\max)}$: niveau (borne) supérieur.

I.6.2. Réponse

Grandeur de sortie ou grandeur étudiée y . Elle représente la manifestation mesurable, observée lorsqu'on fait varier les facteurs étudiés. Elle doit être représentative, quantifiable et la moins dispersée possible pour des variables d'entrées maîtrisées et constantes.

I.6.3. Domaine d'étude

C'est la partie de l'espace expérimental retenu par l'expérimentateur pour faire ses essais. Il est délimité par les niveaux inférieur et supérieur des facteurs étudiés. La réunion des domaines des différents facteurs définit le domaine d'étude (Figure I.2).

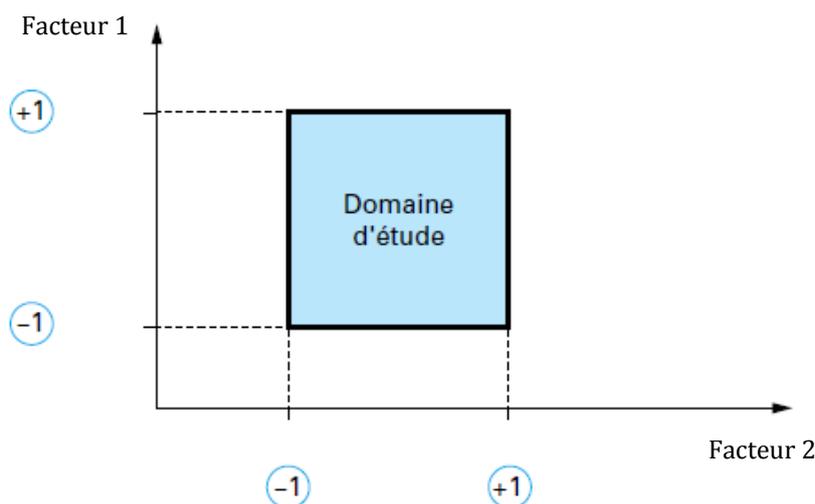


Figure I.2 : Schéma représentatif du domaine d'étude pour le cas de deux facteurs (plan 2^2)

[10].

I.6.4. Surface de réponses

Basées essentiellement sur le domaine d'étude délimité par les niveaux de chaque facteur, à chaque point du domaine d'étude correspond une réponse. A l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se

localisent sur une surface et définissent ainsi la notion de surface de réponses (Figure I.3). Ainsi, le tracé de ces surfaces aide à trouver par approximations successives les meilleures conditions opératoires pour maximiser ou minimiser une caractéristique y (rendement, pureté d'un produit, coût...etc.).

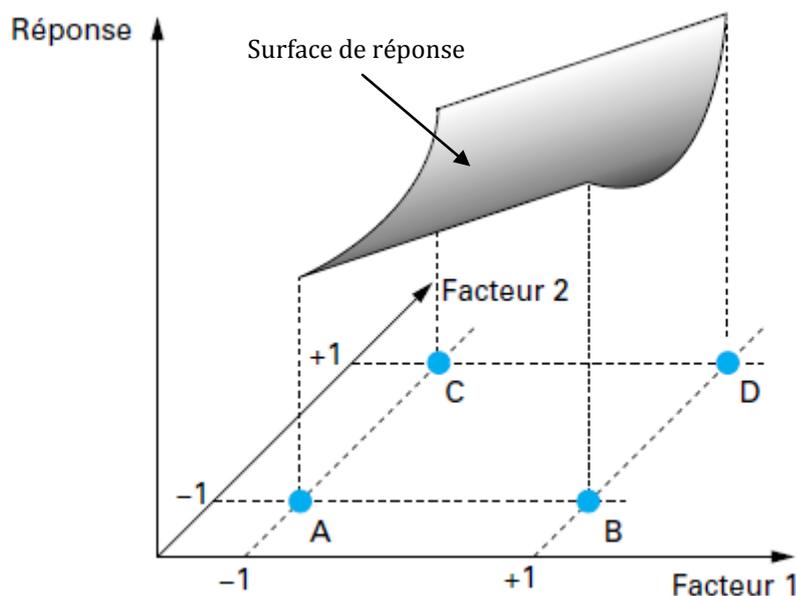


Figure I.3 : Schéma représentatif de la surface de réponse pour un domaine d'étude délimité par deux facteurs (plan 2^2) [5].

I.6.5. Variables codées (variables centrées réduites)

Les variables codées sont des variables sans dimension. Elles sont utilisées dans les modèles postulés afin de faciliter le traitement des données. A chaque variable réelle z_j , on lui associe une variable codée x_j .

$$Z_{j(\min)} \longrightarrow x_j = -1$$

$$Z_{j(\max)} \longrightarrow x_j = +1$$

Le passage des variables réelles (z_1, z_2, \dots, z_k) aux variables codées (x_1, x_2, \dots, x_k), et inversement, est donné par la formule de codage suivante :

$$x_j = \left(\frac{z_j - z_j^0}{\Delta z_j} \right), \quad j = 1, 2, \dots, k$$

Avec :

$$z_j^0 = \frac{z_{j\max} + z_{j\min}}{2} \quad \text{et} \quad \Delta z_j = \frac{z_{j\max} - z_{j\min}}{2}$$

z_j^0 : variable réelle correspondante au centre du domaine d'étude ;

Δz_j : intervalle de variation du paramètre z_j ;

z_{jmax} et z_{jmin} : valeurs maximale et minimale de la variable réelle z_j respectivement.

I.6.6. Matrice des expériences ou matrice des essais

La matrice des expériences est un tableau qui représente, sous forme codé ou réel, l'ensemble des expériences à réaliser. Elle est très recommandée lorsque le nombre de facteurs étudiés est supérieur à trois, du fait qu'au-delà de trois, la représentation géométrique devient pratiquement impossible (espaces multidimensionnels). La matrice des expériences est notée : $T(N, k)$.

Avec :

N : nombre de lignes du tableau (nombre d'expériences à réaliser) ;

k : nombre de colonnes du tableau (nombre de facteurs étudiés).

I.6.7. Matrice des effets

La matrice des effets X est un outil de calcul qui permet de déterminer, via un calcul matriciel, tous les coefficients de l'équation du modèle postulé. Elle est obtenue en ajoutant, à gauche de la matrice des expériences, une colonne ne contenant que des (+1) correspondant à la variable fictive x_0 et à droite, les différentes interactions possibles des facteurs. La matrice des effets est notée : $X(N, M)$.

Où :

M : nombre d'inconnues du modèle (nombre de coefficients) ;

I.6.8. Effet d'un facteur

L'effet d'un facteur x_1 sur la réponse y s'obtient en comparant les valeurs prises par y quand x_1 passe du niveau inférieur (-1 : y_1) au niveau supérieur (+1 : y_2) (Figure I.4).

Cet effet peut être moyen : $(\frac{1}{2}(y_2 - y_1))$ ou global : $(y_2 - y_1)$.

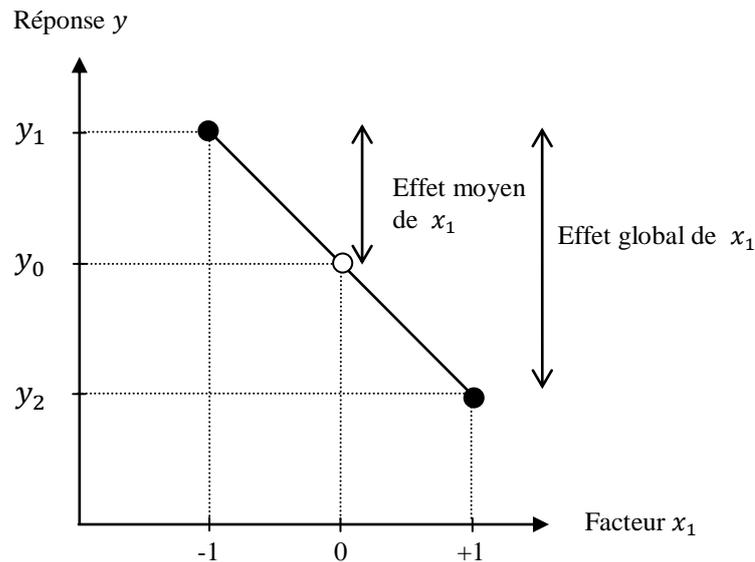


Figure I.4 : Illustration de l'effet global et de l'effet moyen d'un facteur [1, 2].

I.6.9. Notion d'interaction

Nous parlons d'interaction entre deux facteurs x_1 et x_2 si l'effet a_1 de x_1 sur la réponse y dépend du niveau de x_2 ou inversement. Elle est notée x_{12} . Ainsi, l'effet a_{12} mesure donc la variation de l'effet du facteur x_1 quand le facteur x_2 passe du niveau bas au niveau haut et inversement. La Figure I.5 illustre trois cas d'interactions : droites parallèles : pas d'interaction et droites croisées : faible et forte interactions.

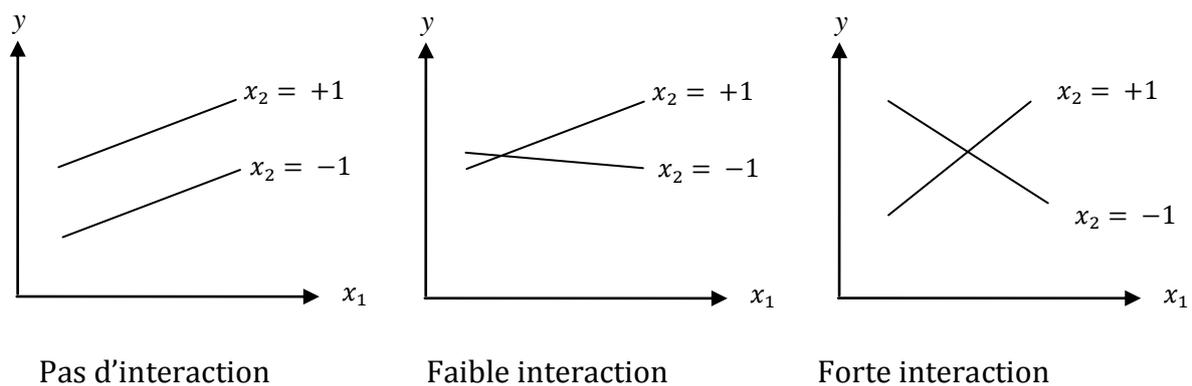


Figure I.5 : Courbes d'interaction entre x_1 et x_2 [2].

I.7. Notion de modèle et de régression linéaire multiple

En l'absence de toute information sur la fonction qui lie la réponse aux facteurs, les expérimentateurs se donnent, a priori, une loi d'évolution dont la formulation la plus générale est la suivante [1-3, 7-10] :

$$y = f(x_1, x_2, x_3, \dots, x_n) + \varepsilon$$

Avec :

y : réponse expérimentale ou réponse mesurée ;

ε : résidu ou erreur expérimentale ;

$f(x_1, x_2, x_3, \dots, x_n)$: fonction inconnue qui explique le mieux les variations de la réponse selon les différentes valeurs données aux x_i .

Cette fonction inconnue est trop générale et ne peut être qu'estimée par une fonction de lissage de forme polynomiale qui n'interprète qu'au voisinage d'un point ; toute fonction peut être développée en série de Taylor et donc remplacée par un polynôme qui s'écrit :

$$\hat{y} = a_0 + \sum_{j=1}^k a_j x_j + \sum_{\substack{u,j=1 \\ j \neq u}}^k a_{ju} x_j x_u + \sum_{\substack{l,u,j=1 \\ j \neq u \neq l}}^k a_{jul\dots k} x_j x_u x_l \dots x_k + \sum_{j=1}^k a_{jj} x_j^2 \quad (\text{Eq. 1-1})$$

Avec :

\hat{y} : réponse calculée à partir du modèle mathématique adopté ;

x_j : variables codées (variables centrées réduites) ;

$a_0, a_j, a_{ju}, a_{jul\dots k}, a_{jj}$: coefficients du modèle mathématique adopté :

a_0 : terme constant ou moyenne ;

a_j : effets linéaires ;

$a_{ju}, a_{jul\dots k}$: effets d'interactions ;

a_{jj} : effets quadratiques ;

Les plans d'expériences exigent l'utilisation de la technique de régression multilinéaire pour déterminer les coefficients du modèle polynomial postulé. Son principe repose sur la méthode des moindres carrés qui consiste à minimiser la somme des écarts quadratiques entre la variable réelle obtenue expérimentalement et celle obtenue à partir du modèle mathématique :

$$\varphi = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \min$$

L'intérêt de modéliser la réponse par un polynôme est de pouvoir utiliser tous les résultats de l'algèbre matricielle. Il est possible d'utiliser d'autres fonctions mathématiques ; toutefois, l'usage montre que les polynômes permettent de résoudre la plupart des problèmes.

I.8. Plans factoriels complets à deux niveaux 2^k

Les plans factoriels complets 2^k (PFC) sont les plans les plus utilisés car ils sont les plus simples et les plus rapides à mettre en œuvre. Ils permettent d'étudier k facteurs à raison de deux niveaux par facteur (niveau supérieur (+1) et niveau inférieur (-1)). Ils sont formés de $N = 2^k$ expériences déterminées par les différentes combinaisons possibles des niveaux minimal et maximal des k facteurs. Pour ces plans, il faut donc calculer M inconnues du modèle (coefficients) correspondants aux [1-3, 7-10] :

- k effets principaux (linéaires) ;
- $(2^k) - k - 1$ interactions ;
- une constante représentant la moyenne de la réponse.

Pour un plan factoriel complet à 3 facteurs, dont il faut réaliser un nombre d'expériences : $N = 2^3 = 8$, les points expérimentaux seront aux sommets d'un cube (Figure I.6.) et la matrice représentative des expériences sera un tableau traduisant l'emplacement de ces points en indiquant leurs trois coordonnées en variables centrées réduites (Tableau I.1).

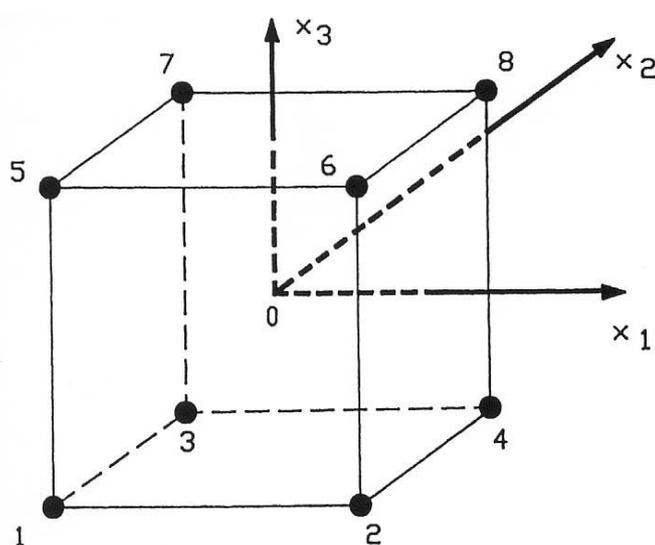


Figure I.6 : Disposition des points expérimentaux d'un plan 2^3 [1, 2].

Tableau I.1 : Matrice des expériences (N, k).

Nombre d'essai N	Valeurs des variables réelles			Valeurs des variables codées			Réponse
	Z_1	Z_2	Z_3	x_1	x_2	x_3	y_i
1	$Z_{1\min}$	$Z_{2\min}$	$Z_{3\min}$	-1	-1	-1	y_1
2	$Z_{1\max}$	$Z_{2\min}$	$Z_{3\min}$	+1	-1	-1	y_2
3	$Z_{1\min}$	$Z_{2\max}$	$Z_{3\min}$	-1	+1	-1	y_3
4	$Z_{1\max}$	$Z_{2\max}$	$Z_{3\min}$	+1	+1	-1	y_4
5	$Z_{1\min}$	$Z_{2\min}$	$Z_{3\max}$	-1	-1	+1	y_5
6	$Z_{1\max}$	$Z_{2\min}$	$Z_{3\max}$	+1	-1	+1	y_6
7	$Z_{1\min}$	$Z_{2\max}$	$Z_{3\max}$	-1	+1	+1	y_7
8	$Z_{1\max}$	$Z_{2\max}$	$Z_{3\max}$	+1	+1	+1	y_8

I.8.1. Modèles mathématiques associés aux plans factoriels complets

Les plans factoriels complets à deux niveaux peuvent être utilisés indistinctement pour les variables continues et pour les variables discrètes. Les modèles mathématiques qui leurs sont associés sont généralement des modèles polynomiaux synergiques, de premier degré qui regroupent tous les facteurs étudiés et s'écrivent sous la forme suivante [1-3, 7-10] :

$$\hat{y} = a_0 + \sum_{j=1}^k a_j x_j + \sum_{\substack{u,j=1 \\ j \neq u}}^k a_{ju} x_j x_u + \sum_{\substack{l,u,j=1 \\ j \neq u \neq l}}^k a_{jul\dots k} x_j x_u x_l \dots x_k \quad (\text{Eq. I - 2})$$

Ou :

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + \dots + a_k x_k + a_{12} x_1 x_2 + a_{13} x_1 x_3 + \dots + a_{1\dots k} x_1 \dots x_k \quad (\text{Eq. I - 3})$$

Selon le nombre de facteurs k étudiés, le polynôme (Eq. I-3) peut s'écrire de manière, plus au moins compliquée. Il doit contenir essentiellement 2^k coefficients qui sont les inconnues (M) et refléter plus clairement l'effet de chaque facteur en absolu (seul) et en interactions doubles, triples et en interactions plus élevées si elles existent. L'ensemble des résultats d'un plan factoriel complet à deux niveaux 2^k conduit donc à un système de $N = 2^k$ équations à $M = 2^k$ inconnues, avec $N=M$.

I.8.2. Calcul des coefficients du modèle

Le calcul des coefficients des effets linéaires et d'interactions du modèle polynomial élaboré (Eq. I-3) s'effectue, par la résolution mathématique du système de 2^k équations à 2^k inconnus, via la formule [1, 2] :

$$A = \frac{1}{N} X^T Y$$

Avec :

A : vecteur des coefficients du modèle ;

X^T : transposé de la matrice X des effets ;

X : matrice des effets (N, M) ;

Y : vecteur des observations ou réponses.

Cependant, une résolution de la régression sous forme matricielle est plus recommandée.

I.8.3. Analyse de régression sous forme matricielle

L'analyse de régression sous forme matricielle est d'un emploi commode pour la résolution des problèmes sur ordinateur. Il s'agit de trouver par la méthode des moindres carrés les coefficients de l'équation de régression (Eq. I-3). A partir de cette équation, on peut définir le matériel statistique suivant [11, 12] :

1) Matrice des effets X (N, M):

$$X = \begin{bmatrix} x_{01} & x_{11} & x_{21} & \dots & \dots & x_{M1} \\ x_{02} & x_{12} & x_{22} & \dots & \dots & x_{M2} \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \dots & \vdots \\ x_{0N} & x_{1N} & x_{2N} & \dots & \dots & x_{MN} \end{bmatrix}$$

2) Vecteur des coefficients A (M, 1) :

$$A = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ \vdots \\ a_{1..k} \end{bmatrix}$$

Avec :

$$X^T = \begin{bmatrix} x_{01} & x_{02} & \dots & \dots & x_{0N} \\ x_{11} & x_{12} & \dots & \dots & x_{1N} \\ x_{21} & x_{22} & \dots & \dots & x_{2N} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & \vdots \\ x_{M1} & x_{M2} & \dots & \dots & x_{MN} \end{bmatrix}$$

$[X^T X]$: matrice des variances.

$$[X^T X] = \begin{bmatrix} \sum x_{0i}^2 & \sum x_{0i}x_{1i} & \dots & \dots & \sum x_{0i}x_{Mi} \\ \sum x_{1i}x_{0i} & \sum x_{1i}^2 & \dots & \dots & \sum x_{1i}x_{Mi} \\ \vdots & \vdots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & \vdots \\ \sum x_{Mi}x_{0i} & \sum x_{Mi}x_{1i} & \dots & \dots & \sum x_{Mi}^2 \end{bmatrix}$$

$[X^T Y]$: matrice colonne

$$[X^T Y] = \begin{bmatrix} \sum x_{0i}y_i \\ \sum x_{1i}y_i \\ \vdots \\ \vdots \\ \sum x_{Mi}y_i \end{bmatrix}$$

Ainsi, le vecteur des coefficients A, s'obtient à partir du produit :

$$A = [X^T X]^{-1} X^T Y \quad (\text{Eq. 1 - 5})$$

Où : $[X^T X]^{-1}$ est la matrice inverse(*) de la matrice $[X^T X]$. Dans le cas du plan factoriel complet, la matrice des variances-covariances $[X^T X]^{-1}$ est diagonale, elle se représente comme suit :

$$[X^T X]^{-1} = \begin{bmatrix} 1/N & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/N \end{bmatrix}$$

(*) : une matrice carrée A est dite inversible s'il existe une matrice carrée A^{-1} (appelée matrice inverse) telle que : $A \times A^{-1} = A^{-1} \times A = I$ (matrice unité). Seule les matrices carrées peuvent avoir des inverses.

I.9. Optimalité des plans d'expériences

Un plan d'expériences est dit optimal s'il permet de calculer des coefficients des effets des facteurs (a_j) avec la meilleure précision possible (erreur minimale sur la réponse prédite par le modèle). Pour ce faire, certains de ces critères doivent être respectés et les plus utilisés sont présentés ci-dessous [2, 3] :

I.9.1. Critère de la matrice unité (matrice d'Hadamard)

Pour atteindre en N expériences la variance minimale, la matrice carrée des effets X , dont les éléments sont $+1$ et -1 , associée aux plans factoriels doit vérifier la relation :

$$[X^T X] = N[I], \quad I : \text{matrice unité.}$$

Si cette égalité est vérifiée, la matrice X représente une matrice d'Hadamard.

I.9.2. Critère d'orthogonalité

La matrice des effets X est orthogonale si ses vecteurs colonnes sont orthogonaux deux à deux, soit le produit scalaire de deux colonnes est égale à zéro.

Soit deux vecteurs colonnes X_1 et X_2 de la matrice des effets X :

$$X_1 = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1i} \\ \vdots \\ x_{1N} \end{bmatrix} \quad \text{et} \quad X_2 = \begin{bmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2i} \\ \vdots \\ x_{2N} \end{bmatrix}$$

L'orthogonalité de ces deux colonnes se traduit par l'égalité :

$$\sum_{i=1}^N x_{1i} \times x_{2i} = 0$$

Ce critère auquel doivent répondre les plans d'expériences permet le calcul de l'effet d'un facteur sans que l'effet de l'autre facteur intervienne.

I.9.3. Critère de la trace minimale

Dans ce cas, la trace de la matrice $[X^T X]^{-1}$, qui représente la somme des éléments de sa diagonale principale, doit être la plus faible possible afin de minimiser l'erreur. Ce qui se traduit par :

$$tr = [X^T X]^{-1} = \text{minimale possible}$$

I.10. Exemple d'application

I.10.1. Problème à résoudre

Un industriel désire récupérer via le procédé d'adsorption, les molécules d'amoxicilline (AMX) présentes dans les effluents de son industrie par l'utilisation d'un adsorbant de nature minérale (hydroxyapatite : HAP). Il désire étudier l'influence de certains paramètres à savoir : le pH de la solution à traiter ($\text{pH} \in [2-10]$), la concentration en AMX ($[\text{AMX}] \in [50-300]$ mg/L) et la concentration de la suspension en HAP ($[\text{HAP}] \in [0,125-1,25]$ g/250mL) sur le taux d'adsorption de l'AMX par HAP. Pour cela, il décide d'adapter un plan factoriel complet à deux niveaux.

I.10.2. Formulation du problème

- 1) objectif de l'étude : étude de l'effet de chaque paramètre opératoire sur le taux d'adsorption de l'AMX
- 2) définir les facteurs et la réponse :
 - Réponse y : taux d'adsorption calculé comme suit :

$$y (\%) = \frac{[\text{AMX}]_0 - [\text{AMX}]_t}{[\text{AMX}]_0} \times 100$$

Avec :

$[\text{AMX}]_0$: concentration initiale en amoxicilline ;

$[\text{AMX}]_t$: concentration de la solution en AMX à l'instant t.

- Facteurs et domaines de leur étude:

Facteur 1 $\Rightarrow z_1$: pH de la solution à traiter ; $\text{pH} \in [2-10]$;

Facteur 2 $\Rightarrow z_2$: concentration en AMX ; $[\text{AMX}] \in [50-300]$ mg/L ;

Facteur 3 $\Rightarrow z_3$: concentration de la suspension en HAP ;

$[\text{HAP}] \in [0,125-1,25]$ g/250mL .

Pour mener une étude basée sur un plan factoriel complet à deux niveaux et à trois facteurs, il nous faut $N = 2^3 = 8$ expériences à réaliser. Le modèle mathématique associé à ce plan est un modèle à $M = 8$ coefficients inconnus, qui s'écrit sous la forme suivante :

$$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{123}x_1x_2x_3$$

I.10.3. Planification et expérimentation

- construction du plan d'expériences : construire la matrice d'expériences en variables réelles et en variables codées (Tableau I.2) afin de l'utiliser comme outil de démarrage pour pouvoir lancer et faciliter la mise en œuvre de l'expérimentation.

Le passage de la variable réelle z_j à la variable codée x_j est donné par l'équation (Eq. I-1). Une fois l'expérimentation effectuée, les réponses correspondant à chaque essai sont reportées sur la même matrice d'expériences.

Tableau I.2 : Matrice des expériences du plan 2^3

N° d'essai	Variables réelles			Variables codées			Réponses mesurées
	z_1	z_2	z_3	x_1	x_2	x_3	$y(\%)$
1	2	50	0,125	-1	-1	-1	55,89
2	10	50	0,125	+1	-1	-1	56,93
3	2	300	0,125	-1	+1	-1	64,70
4	10	300	0,125	+1	+1	-1	61,23
5	2	50	1,25	-1	-1	+1	85,95
6	10	50	1,25	+1	-1	+1	89,95
7	2	300	1,25	-1	+1	+1	88,50
8	10	300	1,25	+1	+1	+1	96,55

I.10.4. Analyse des résultats

Pour pouvoir extraire le maximum d'informations des résultats de l'expérience, il est recommandé de calculer les coefficients du modèle postulé. Pour cela, il faut d'abord déterminer la matrice des effets X constituée de N lignes (expériences) et de M colonnes (nombre de coefficients du modèle postulé) (Tableau I.3). A partir de cette matrice, le calcul se fait via la formule : $A = [X^T X]^{-1} X^T Y$. Les résultats trouvés sont répertoriés dans le Tableau I.4.

Tableau I.3 : Matrice des effets X (N, M)

N°d'essai	x_0	x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}	$y(\%)$
1	+1	-1	-1	-1	+1	+1	+1	-1	55,89
2	+1	+1	-1	-1	-1	-1	+1	+1	56,93
3	+1	-1	+1	-1	-1	+1	-1	+1	64,70
4	+1	+1	+1	-1	+1	-1	-1	-1	61,23
5	+1	-1	-1	+1	+1	-1	-1	+1	85,95
6	+1	+1	-1	+1	-1	+1	-1	-1	89,95
7	+1	-1	+1	+1	-1	-1	+1	-1	88,50
8	+1	+1	+1	+1	+1	+1	+1	+1	96,55

Tableau I.4 : Valeurs des coefficients du modèle de l'adsorption de l'AMX

Terme constant	Effets linéaires			Effets d'interactions			
	a_1	a_2	a_3	a_{12}	a_{13}	a_{23}	a_{123}
a_0							
74,96	1,20	2,78	15,28	-0,06	1,81	-0,49	1,07

Le modèle obtenu après calcul s'écrit alors :

$$\hat{y} (\%) = 74,96 + 1,20x_1 + 2,78x_2 + 15,28x_3 - 0,06x_1x_2 + 1,81x_1x_3 - 0,49x_2x_3 + 1,07x_1x_2x_3$$

Ce modèle postulé permet de distinguer l'influence des trois paramètres opératoires étudiés sur le taux d'adsorption de l'AMX par HAP. Ainsi, on peut déduire l'influence de chaque facteur dans l'absolu (seule), c'est-à-dire, en faisant abstraction de l'influence des autres facteurs et cerner l'influence combinée de deux facteurs à la fois, en faisant abstraction du troisième.

L'étape finale de la chronologie des plans d'expériences qui est l'optimisation de la réponse, ne peut être effectuée qu'après avoir validé le modèle postulé par les tests statiques.

Chapitre II
Tests de signification et validation du
modèle :
Analyse statistique

Chapitre II : Tests de signification et validation du modèle : Analyse statistique

II.1. Introduction

L'emploi des principes d'analyse de régression et de corrélation permet de trouver lors du traitement des données expérimentales, la dépendance entre les variables et les conditions optimales. Une fois l'équation de régression obtenue, on procède à l'analyse statistique des résultats. Pour conduire cette analyse, il faut remplir les conditions suivantes :

- Le paramètre x est mesuré avec une erreur négligeable. L'apparition d'une erreur dans la détermination de y (réponse) s'explique par la présence, dans le processus, de variables non exploitées qui ne figurent pas dans l'équation de régression ;
- Les résultats des observations sur les variables de sortie $y_1, y_2 \dots$ sont des grandeurs indépendantes à distribution normale.

II.2. Vérification de la signification des coefficients (effets)

II.2.1. Cas où chaque essai est répété m fois

Nous déterminons la moyenne arithmétique des résultats des essais parallèles \bar{y}_i [2, 11, 12]:

$$\bar{y}_i = \frac{\sum_{u=1}^m y_{iu}}{m} \quad (\text{Eq. II - 1})$$

Puis la variance de sondage S_i^2 pour l'essai i :

$$S_i^2 = \frac{\sum_{u=1}^m (y_{iu} - \bar{y}_i)^2}{m - 1}, \quad i = 1, 2, \dots, N \quad (\text{Eq. II - 2})$$

Avec :

m : nombre de répétitions pour chaque essai ;

N : nombre d'expériences ;

y_i : réponses expérimentales.

Si les variances de sondage sont homogènes, nous calculons alors la variance de reproductibilité :

$$S_{rep}^2 = \frac{\sum_{i=1}^N S_i^2}{N} \quad (\text{Eq. II - 3})$$

II.2.2. Cas où l'essai au centre est répété n_0 fois

Dans ce cas, la variance des mesures (ou de reproductibilité) est estimée par celle calculée au centre du domaine expérimental :

$$S_{rep}^2 = \frac{\sum_{i=1}^{n_0} (y_i - \bar{y}_0)^2}{n_0 - 1} \quad (\text{Eq. II - 4})$$

Avec :

$n_0 - 1$: nombre de degrés de liberté ;

n_0 : nombre de répétitions au centre du domaine d'étude;

\bar{y}_0 : moyenne arithmétique des mesures effectuées au centre du domaine, elle est calculée par :

$$\bar{y}_0 = \frac{\sum_{i=1}^{n_0} y_i}{n_0} \quad (\text{Eq. II - 5})$$

Dans les deux cas, la variance de reproductibilité est indispensable pour estimer la signification des coefficients de l'équation de régression par le test de Student. Un coefficient est dit significatif, s'il est, pour un risque donné, significativement différent de zéro. Nous testerons donc l'hypothèse :

$$H_0 = \ll b_j = 0 \gg$$

contre l'hypothèse : $H_1 = \ll b_j \neq 0 \gg$.

Pour ce faire, nous calculons le rapport :

$$t_j = \frac{|b_j|}{S_{bj}} \quad (\text{Eq. II - 6})$$

Où :

t_j : suit une loi normale ;

b_j : j^{ème} coefficient de l'équation de régression ;

S_{bj} : écart quadratique moyen qui est défini dans le cas d'un modèle du premier degré par :

$$S_{bj} = \frac{S_{rep}}{\sqrt{N}} \quad (\text{Eq. II - 7})$$

Nous utilisons la table de Student pour déterminer la valeur de $t_p(f)$ pour le niveau de signification choisi p (limite d'erreur tolérée) et le nombre de degrés de liberté f . La règle de ce test est alors la suivante :

- Si $t_j > t_p(f)$, nous rejetons H_0 au risque accepté ;
- Si $t_j < t_p(f)$, nous acceptons H_0 au risque accepté.

Si l'hypothèse H_0 est acceptée, cela veut dire que le coefficient en question n'est pas significativement différent de zéro au risque p et donc que la variable qui lui est associée n'a pas d'influence sur la réponse.

II.3. Validation du modèle

Les tests suivants peuvent être appliqués aux équations des modèles construits pour les deux cas rencontrés ci-dessus, c'est-à-dire, que les répétitions soient conduites seulement au centre du domaine ou que tous les essais soient répétés m fois. Pour adapter les formules suivantes au deuxième cas, il suffit de remplacer y_i par \bar{y}_i qui est la moyenne pour chaque expérience.

II.3.1. Recherche de biais

L'objectif de cette analyse est de vérifier si les effets retenus comme influents sont réellement associés aux facteurs et aux interactions considérées, ou s'ils ne sont que les manifestations de la variabilité naturelle du phénomène étudié. Pour cela, la variance résiduelle $S_{rés}^2$ estimée par :

$$S_{rés}^2 = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N - \ell} \quad (\text{Eq. II - 8})$$

Avec :

$N - \ell$: degrés de liberté ;

ℓ : nombre de coefficients significatifs ;

\hat{y}_i : réponses calculées à partir du modèle.

est comparée à la variance de reproductibilité S_{rep}^2 . Si l'inégalité suivante est vérifiée :

$$F = \frac{S_{rés}^2}{S_{rep}^2} < F_{0,95}(N - \ell, n_0 - 1) \quad (\text{Eq. II - 9})$$

Avec :

$F_{0,95}(N - \ell, n_0 - 1)$: valeur tabulée du test de Fisher-Snedecor (table de Fisher) pour le niveau de signification choisi $p=0,95$ et les nombres de degrés de liberté $f_1 = N - \ell$ et $f_2 = n_0 - 1$.

alors, le modèle est considéré sans biais, ce qui revient à dire que la part des variations de y non expliquée par le modèle est aléatoire.

II.3.2. Validation de l'équation de régression

Lorsque le modèle est sans biais, nous pouvons donc vérifier la signification de la régression via l'équation suivante :

$$F = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2 / (\ell - 1)}{\sum_{i=1}^N (y_i - \hat{y}_i)^2 / (N - \ell)} \quad (\text{Eq. II - 10})$$

Avec :

$$\bar{y} = \frac{\sum_{i=1}^N y_i}{N} \quad (\text{Eq. II - 11})$$

Si le rapport de l'équation (Eq.II-10) est supérieur à la valeur tabulée $F_p(f_1, f_2)$ pour le niveau de signification choisi $p=0,95$ et les nombres de degrés de liberté $f_1 = \ell - 1$ et $f_2 = N - \ell$, l'équation est adéquate. Ce nouveau test indique de combien est réduite la variance relative à l'équation obtenue en comparaison avec la variance relative à la moyenne. Plus la valeur de F excède celle de $F_p(f_1, f_2)$, plus l'équation de régression est efficiente.

II.3.3. Coefficient de détermination

Le coefficient de corrélation R^2 interprète le rapport de la variance expliquée par la régression sur la variance totale. Il se calcule comme suit :

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (\text{Eq. II - 12})$$

Plus la valeur de ce coefficient s'approche de 1, plus le modèle est représentatif. L'estimation de la force de la liaison par le coefficient de corrélation multiple est d'autant plus exagérée que le nombre de degrés de liberté de l'échantillon $f_2 = N - \ell$ est

plus petit. Lorsque l'échantillon est d'étendue assez faible, il est nécessaire d'adopter une correction pour l'erreur systématique. La formule de correction du coefficient de régression multiple est la suivante :

$$\bar{R}^2 = R^2 - (1 - R^2) \frac{\ell - 1}{N - \ell} \quad (\text{Eq. II - 13})$$

II.3.4. Analyse des résidus

L'évaluation de la qualité de prévision d'un modèle peut se faire par une analyse des résidus qui consiste à porter ces derniers en fonction des réponses prédites ou en fonction d'autres grandeurs. On peut ainsi détecter s'il existe une relation entre les résidus et la grandeur portée en abscisse. On doit arriver à des diagrammes ne laissant apparaître aucune relation ou tendance entre les grandeurs étudiées.

II.4. Exemple d'application

Nous revenons à l'exemple déjà entamé dans le chapitre I (paragraphe I.10.1 : Adsorption de l'AMX sur HAP) pour vérifier la validité du modèle retenu par l'analyse statistique présentées dans ce chapitre. Le modèle postulé obtenu après calcul est :

$$\hat{y} \text{ (\%)} = 74,96 + 1,20x_1 + 2,78x_2 + 15,28x_3 - 0,06x_1x_2 + 1,81x_1x_3 - 0,49x_2x_3 + 1,07x_1x_2x_3$$

II.4.1. Analyse statistique de l'équation de régression

Pour estimer la variance de reproductibilité, l'essai effectué au centre du domaine d'étude est répété quatre fois. Les valeurs du taux d'adsorption de l'AMX pour chacun de ces essais sont rassemblées dans le Tableau II.1.

Tableau II.1 : Valeurs du taux d'adsorption de l'AMX au centre du domaine.

n ₀	y ₀ (%)
1	97,78
2	97,75
3	97,74
4	97,99

La moyenne des 4 essais au centre est : $\bar{y}_{0(\text{AMX})}(\%) = 97,81$

II.4.1.1. Signification des coefficients de l'équation de régression

Le test de Student permet de déterminer si parmi les coefficients du modèle il existe des coefficients non significatifs, lesquels seront éliminés de l'équation car leur influence sur l'indice taux d'adsorption de l'AMX est négligeable. Pour cela, nous déterminons la variance de reproductibilité et l'écart type moyen des essais :

- $S_{rep}^2 = 1,43 \cdot 10^{-2}$;
- $S_{bj}^2 = 4,23 \cdot 10^{-2}$.

puis nous calculons les t_j correspondants à chaque effet. Les résultats obtenus sont répertoriés dans le Tableau II.2.

Tableau II.2 : Valeurs des t_j

Terme constant	Effets linéaires			Effets d'interactions			
t_0	t_1	t_2	t_3	t_{12}	t_{13}	t_{23}	t_{123}
1771,77	28,39	65,76	361,04	1,37	42,78	11,69	25,29

Pour le niveau de signification $p=0,95$ (correspondant à 0,05 dans la table de Student) et le nombre de degré de liberté $f = n_0 - 1 = 3$, la valeur tabulée du test de Student relative à un test bilatéral est égale à : $t_p(f) = t_{0,05}(3) = 3,18$. La valeur de t_{12} calculée étant inférieure à la valeur tabulée, le coefficient correspondant a_{12} n'est donc pas significatif, ce qui revient à l'éliminer de l'équation du modèle.

II.4.1.2. Validation du modèle

- **Recherche de biais**

Avant de valider le modèle postulé, nous devons vérifier l'absence du biais. Pour cela, nous avons calculé la variance résiduelle représentative de notre système ($S_{res}^2 = 2,70 \cdot 10^{-2}$), que nous avons comparée à la variance de reproductibilité (S_{rep}^2) à l'aide du test de Fisher-Snedecor : $F = \frac{S_{rés}^2}{S_{rep}^2}$.

Comme la valeur calculée du test de Fisher- Snedecor : $F = 1,89$ est inférieure à la valeur tabulée pour le niveau de signification $p=0,95$, soit : $F_{0,95}(N - \ell, n_0 - 1) = F_{0,95}(1, 3) = 10,1$ (Table de Fisher), nous pouvons confirmer que le modèle obtenu pour l'adsorption de l'AMX sur HAP n'est pas biaisé.

- **Test de signification de la régression**

Le modèle étant sans biais, nous pouvons donc vérifier la signification de la régression

$$\frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{(\ell - 1)}$$

par le test de Fisher : $F = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{(N - \ell)}$

La valeur calculée du test de Fisher est égale à $F = 12201,82$ tandis que la valeur tabulée $F_{0,95}(\ell - 1, N - \ell) = F_{0,95}(6, 1)$ est égale à 234. Cette dernière étant inférieure à celle calculée, nous pouvons donc conclure que l'équation de régression adoptée peut être considérée comme adéquate et valide à 95%. L'équation de ce modèle s'écrit alors :

$$(\%) = 74,96 + 1,20x_1 + 2,78x_2 + 15,28x_3 + 1,81x_1x_3 - 0,49x_2x_3 + 1,07x_1x_2x_3$$

II. 4. 1. 3. Evaluation de la qualité du modèle

- **Calcul de coefficient de détermination**

Pour évaluer les réponses données par le modèle postulé par rapport aux réponses mesurées, nous avons calculé le coefficient de détermination : $R^2 = 0,999$. Sa valeur est pratiquement proche de 1, ce qui signifie que toutes les variations de la réponse y , dans le domaine d'étude choisi, sont expliquées par l'équation de régression postulée.

- **Analyse des résidus**

L'analyse des résidus du modèle prédit est illustrée sur la Figure II-1. Les résidus (ε) sont calculés à partir de la relation suivante :

$$\varepsilon_i = y_i - \hat{y}_i$$

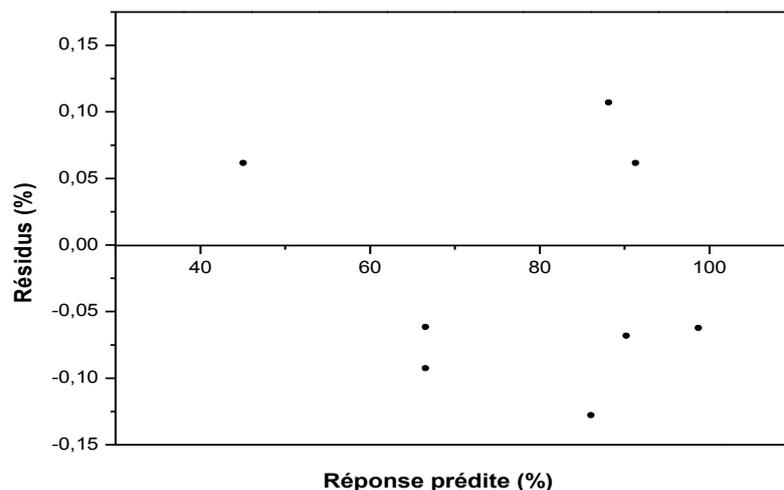


Figure II.1 : Diagramme des résidus.

Le diagramme de la Figure II.1 ne laisse apparaître aucune relation entre les valeurs prédites \hat{y}_i et les résidus (ε), car les points semblent disposés aléatoirement. Ce résultat confirme l'absence d'informations dans les résidus ce qui nous permet d'affirmer que toutes les variations de la réponse sont expliquées par le modèle de régression retenu.

Tableau 1. Table de loi de Student

Unilatéral Bilatéral	0,01 0,20	0,05 0,10	0,025 0,05	0,01 0,02	0,005 0,01
f_1					
1	3,08	6,31	12,7	31,8	63,7
2	1,89	2,92	4,30	6,97	9,92
3	1,64	2,35	3,18	4,54	5,84
4	1,53	2,13	2,78	3,75	4,60
5	1,48	2,02	2,57	3,37	4,03
6	1,44	1,94	2,45	3,14	3,71
7	1,42	1,90	2,37	3,00	3,50
8	1,40	1,86	2,31	2,90	3,36
9	1,38	1,83	2,26	2,82	3,25
10	1,37	1,81	2,23	2,76	3,17
11	1,36	1,80	2,20	2,72	3,10
12	1,36	1,78	2,18	2,68	3,06
13	1,35	1,77	2,16	2,65	3,01
14	1,35	1,76	2,15	2,62	2,98
15	1,34	1,75	2,13	2,60	2,95
16	1,34	1,75	2,12	2,58	2,92
17	1,33	1,74	2,11	2,57	2,90
18	1,33	1,73	2,10	2,55	2,88
19	1,33	1,73	2,09	2,54	2,86
20	1,33	1,73	2,09	2,53	2,85
21	1,32	1,72	2,08	2,52	2,83
22	1,32	1,72	2,07	2,51	2,82
23	1,32	1,71	2,07	2,50	2,81
24	1,32	1,71	2,06	2,49	2,80
25	1,32	1,71	2,06	2,49	2,79
26	1,32	1,71	2,06	2,48	2,78
27	1,31	1,70	2,05	2,47	2,77
28	1,31	1,70	2,05	2,47	2,76
29	1,31	1,70	2,05	2,46	2,76
30	1,31	1,70	2,04	2,46	2,75
40	1,30	1,68	2,02	2,42	2,70
∞	1,28	1,65	1,96	2,33	2,58

Tableau 2. Table de la loi de Fisher-Snedecor pour $p = 0,95$ f_1 : degrés de liberté au numérateur ; f_2 : degrés de liberté au dénominateur.

$f_1 \rightarrow$ $f_2 \downarrow$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19,00	19,2	19,20	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
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	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
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	4,53	4,50	4,46	4,43	4,40	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
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	3,41	3,38	3,34	3,3	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
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	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,8	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,7	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
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	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,1	2,06	2,01	1,96
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	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
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	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
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	1,95	1,9	1,85	1,8	1,75	1,69
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	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
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	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Chapitre III

Les plans fractionnaires

donc évaluer les risques avant de démarrer l'expérimentation et les minimiser en construisant le plan fractionnaire adéquat [3, 5, 9].

III.3. Conception d'un plan fractionnaire

Comme dans les plans factoriels complets, les plans fractionnaires sont présentés par des modèles mathématiques de forme polynômiale contenant 2^k coefficients : une constante, des effets principaux et des interactions. Si nous réalisons $N=2^{k-q}$ expériences, nous obtiendrons un système de N équations à M inconnus, avec $M > N$ [1-3, 5, 9].

Soit le système matriciel : $Y = X.A$ (Eq. III – 1)

$$\text{Avec : } \begin{cases} Y (N, 1) \\ X (N, M) \\ A (M, 1) \end{cases} \quad \text{et} \quad M > N$$

Mathématiquement, ce système d'équations est impossible à résoudre et à fin de pouvoir trouver sa solution, il est recommandé d'adapter un autre modèle ne contenant que N inconnues. **Alors, comment peut-on faire ça ?**

III.3.1. Notion de contrastes et d'aliasés

Pour réduire le nombre d'inconnues M et le ramener à N , il est conseillé de regrouper les coefficients du modèle initial dans de nouvelles inconnues appelées les contrastes (théorie des contrastes). Par définition, un contraste représente un effet apparent L_i , du modèle mathématique réduit qui, regroupe et associe deux effets réels (un effet principal a_i et un autre d'interaction a_{ij}). Ces derniers sont dits confondus ou aliasés dans le contraste. L'utilisation de la théorie des contrastes nous mène à un nouveau système matriciel qui s'écrit sous cette forme [1-3, 5, 9] :

$$Y = X_{sup} \cdot L \quad \text{(Eq. III – 2)}$$

$$\text{Avec : } \begin{cases} Y (N, 1): \text{ vecteur des réponses;} \\ X_{sup} (N, N): \text{ matrice supérieur orthogonale d'Hadamard;} \\ L (N, 1): \text{ vecteur des contrastes.} \end{cases}$$

Le passage du système (Eq. III-1) au système (Eq. III-2) génère un calcul matriciel qui permet de déterminer les contrastes L_i du modèle réduit en fonction des coefficients du

modèle initial (modèle de départ). Pour cela, la matrice des effets X (N, M) est décomposée en deux sous matrices supérieure : $X_{sup}(N, N)$ et inférieure : $X_{inf}(N, M-N)$.

$$\text{On a : } X = [X_{sup} : X_{inf}] \quad \text{et} \quad A = \begin{bmatrix} A_{sup} \\ A_{inf} \end{bmatrix}$$

L'équation (Eq. III-1) s'écrit alors :

$$Y = X.A = [X_{sup} : X_{inf}] \times \begin{bmatrix} A_{sup} \\ A_{inf} \end{bmatrix} = X_{sup}.A_{sup} + X_{inf}.A_{inf} \quad (\text{Eq. III - 3})$$

Combinons (Eq. III-2) et (Eq. III-3), nous obtiendrons :

$$Y = X_{sup}.L = X_{sup}.A_{sup} + X_{inf}.A_{inf} \quad (\text{Eq. III - 4})$$

Ainsi, le vecteur des contrastes L s'obtient à partir de l'équation suivante :

$$L = A_{sup} + [X_{sup}^T X_{sup}]^{-1} X_{sup}^T X_{inf}.A_{inf} \quad (\text{Eq. III - 5})$$

Ou : $[X_{sup}^T X_{sup}]^{-1} X_{sup}^T X_{inf}$ est la matrice des aliases.

Comme il a déjà été signalé ci-dessus, un plan fractionnaire réalise, certes, des économies d'expériences, mais en revanche, il présente quelques inconvénients. Des ambiguïtés peuvent survenir lors de l'interprétation des résultats, chose due essentiellement au jumelage des coefficients du modèle initial qui, lui ne permet ni d'obtenir les informations exactes ni de soutirer les renseignements recherchés sur le phénomène étudié. **Alors, comment peut-on lever ces ambiguïtés et arriver à interpréter convenablement le modèle mathématique postulé dans ce type de plan ?**

III.3.2. Notion de générateur d'aliases

Pour palier ce problème d'interprétation, il est recommandé de faire appel au calcul de Box qui, via un générateur d'alias et une relation de définition, permet de prévoir rapidement quels seront les couples d'aliases des coefficients du modèle initial et la manière dont ils sont aliasés dans les contrastes sans utiliser la matrice des aliases .

Par définition, un générateur d'aliases représente l'interaction d'ordre le plus élevé du modèle initial $x_{i...k}$ qui, après un tri sélectif et successif de (+1) et (-1) dans la matrice

des effets du PFC, permet de déterminer l'ensemble des aliases grâce à la relation de définition qui l'associe à la colonne x_0 (notée I : élément signifiant +1), utile au calcul de la moyenne des réponses a_0 . Ainsi, la relation de définition s'écrit sous la forme : $I = x_{i...k}$ [1-3, 5, 9].

D'une manière générale, le calcul de Box nous mène à la relation suivante :

$$L_i = a_i \pm \sum_{i=1}^k a_{i...k}$$

Remarque :

Le calcul de Box permet de retrouver les colonnes de signes identiques, il est uniquement valable pour les plans factoriels fractionnaires à deux niveaux dont les points expérimentaux sont exactement situés aux sommets du domaine d'étude.

III.3.3. Hypothèses d'interprétation

Comme nous n'effectuons pas toutes les expériences du plan complet, nous ne pouvons pas obtenir la valeur de toutes les interactions. Il faut créer soi même des informations supplémentaires de remplacement. Ces informations doivent être réalistes et compatibles avec l'étude menée. Elles sont introduites sous forme d'hypothèses et elles demandent à être vérifiées avant la conclusion de l'étude. Ces hypothèses sont [1-3, 5, 9] :

1. les interactions d'ordre 3 ou d'ordre plus élevé sont considérées comme négligeables ;
2. si un contraste est nul, cela peut signifier que :
 - les effets et les interactions aliasés sont tous nuls ;
 - les effets et les interactions aliasés se compensent ;
3. si deux contrastes sont faibles, on supposera que leur interaction l'est aussi ;
4. si un contraste est faible et l'autre est fort, on supposera que leur interaction est faible ;
5. si deux contrastes sont forts, on se méfiera de leur interaction qui peut l'être également ;
6. si deux interactions d'ordre deux sont aliasées dans un contraste de valeur significative (on ne sait pas alors quelle est l'interaction responsable de l'effet

observé ou si les deux interactions ont chacune un effet), il sera indispensable d'effectuer des essais complémentaires.

III.4. Exemple d'application

Soit la matrice de l'exemple pris pour l'étude du plan factoriel complet 2^3 (adsorption de l'amoxicilline par une hydroxyapatite).

Tableau III.1 : Matrice des effets X (N, M) du PFC

N°d'essai	x_0	x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}	y(%)
1	+1	-1	-1	-1	+1	+1	+1	-1	55,89
2	+1	+1	-1	-1	-1	-1	+1	+1	56,93
3	+1	-1	+1	-1	-1	+1	-1	+1	64,70
4	+1	+1	+1	-1	+1	-1	-1	-1	61,23
5	+1	-1	-1	+1	+1	-1	-1	+1	85,95
6	+1	+1	-1	+1	-1	+1	-1	-1	89,95
7	+1	-1	+1	+1	-1	-1	+1	-1	88,50
8	+1	+1	+1	+1	+1	+1	+1	+1	96,55

Tableau III.2 : Valeurs des coefficients du modèle de l'adsorption de l'AMX

Terme constant	Effets linéaires			Effets d'interactions			
	a_1	a_2	a_3	a_{12}	a_{13}	a_{23}	a_{123}
74,964	1,201	2,782	15,276	0	1,809	-0,494	1,069

Pour fractionner ce plan factoriel complet en deux parties afin d'obtenir un plan fractionnaire 2^{3-1} , il est nécessaire de :

1. trier les lignes de ce plan d'après les valeurs de la colonne d'interaction la plus élevée (calcul de Box). Dans ce cas, le générateur d'alias est l'interaction triple 123 ;

Le trie se fait par rapport aux (+1) et (-1) afin d'obtenir une partie de demi plan supérieur ne contenant que (+1) au nombre de 4 lignes dans la colonne 123 et une partie de demi plan inférieur ne contenant que (-1) au nombre de 4 dans la même colonne (Tableau III.3).

Tableau III.3 : Matrice des effets X (N, M) triée selon le PFF

N°d'essai	x_0	x_1	x_2	x_3	x_{12}	x_{13}	x_{23}	x_{123}	y(%)
2	+1	+1	-1	-1	-1	-1	+1	+1	56,93
3	+1	-1	+1	-1	-1	+1	-1	+1	64,70
5	+1	-1	-1	+1	+1	-1	-1	+1	85,95
8	+1	+1	+1	+1	+1	+1	+1	+1	96,55
1	+1	-1	-1	-1	+1	+1	+1	-1	55,89
4	+1	+1	+1	-1	+1	-1	-1	-1	61,23
6	+1	+1	-1	+1	-1	+1	-1	-1	89,95
7	+1	-1	+1	+1	-1	-1	+1	-1	88,50

Demi plan sup
Demi plan inf

- Choisir un demi-plan supérieur ou inférieur puis faire l'étude par rapport à cette partie (réaliser les expériences correspondants à la partie choisie) ;
- Déterminer les colonnes aliasées de la nouvelle matrice des effets qui présentent une séquence de signe (+1) et (-1) identique 2 à 2. Dans le tableau ci-dessus : la colonne x_0 et la colonne x_{123} sont identiques, ce qui revient à écrire $0=123$. De même les colonnes x_1 et x_{23} , les colonnes x_2 et x_{13} et les colonnes x_3 et x_{12} sont identiques, ce qui revient à écrire : $1=23$, $2=13$ et $3=12$;
- Déterminer les contrastes (effets apparents qui résultent du fait que les effets sont confondus 2 par 2).

- **Demi-plan supérieur :**

Générateur d'aliasés c'est : colonne 123.

La relation de définition: $0=I=123$

D'après Box, son calcul est basé sur les relations suivantes :

- la multiplication d'une colonne par elle-même donne une colonne de signes (+) ;
 $I : 1 \times 1 = 1^2 = I$;
- la multiplication d'une colonne par une colonne de signes (+) ; I, redonne la colonne initiale : $1 \times I = I \times 1 = 1$.

Multiplions les 2 membres de la relation de définition $I = 123$ par 1 ; nous obtiendrons
 $1 \times I = 1 \times 123 = 1 \times 1 \times 23 = 1^2 \times 23$ et en simplifiant en fonction des règles ci-dessus

(algèbre de Box) : $1 = 23$; ce qui signifie que l'effet principal 1 est donc aliasé avec l'interaction 23.

Faisons de même avec tous les facteurs, on obtient :

$2 = 13$; l'effet principal 2 est aliasé avec l'interaction 13 ;

$3 = 12$; l'effet principal 3 est aliasé avec l'interaction 12.

Une fois les aliasés retrouvés, nous déterminons les contrastes L_i :

$$\begin{cases} L_1 = L_{23} = a_1 + a_{23} \\ L_2 = L_{13} = a_2 + a_{13} \\ L_3 = L_{12} = a_3 + a_{12} \\ L_0 = L_{123} = a_0 + a_{123} \end{cases}$$

Le système obtenu après avoir réduit le nombre d'inconnues M , se présente comme suit :

Tableau III.4 : Matrice des effets $X_{sup}(N, N)$ du PFF

N° d'essai	x_0	x_1	x_2	x_3	y(%)
2	+1	-1	-1	-1	56,93
3	+1	+1	+1	-1	64,70
5	+1	+1	-1	+1	85,95
8	+1	-1	+1	+1	96,55

A partir de cette matrice réduite X_{sup} , le calcul des contrastes se fait via la formule :

$A = [X^T X]^{-1} X^T Y$. Les résultats trouvés sont répertoriés dans le Tableau III.5.

Tableau III.5 : Valeurs des contrastes du modèle de l'adsorption de l'AMX pour le demi-plan supérieur.

Effets apparents	L_0	L_1	L_2	L_3
valeurs	76,03	0,71	4,59	15,22

Le modèle obtenu après calcul s'écrit alors :

$$\hat{y} (\%) = 76,03 + 0,71x_1 + 4,59x_2 + 15,22x_3$$

- **Demi-plan inférieur :**

Dans le demi-plan inférieur, la colonne 123 ne comprend que des (-1) contrairement à la colonne 0=I qui ne contient que des (+1) ; la relation de définition s'écrit alors : $0=I=-123$.

Multiplions les 2 membres de la relation de définition $I=-123$ par 1 ; nous obtiendrons : $1 \times I = -1 \times 123 = -1 \times 1 \times 23 = -1^2 \times 23$ et en simplifiant en fonction des règles de l'algèbre de Box citées ci-dessus : $1 = -23$; ce qui signifie que l'effet principal 1 est donc aliasé avec l'interaction -23.

Faisons de même avec tous les facteurs, on obtient :

$2 = -13$; l'effet principal 2 est aliasé avec l'interaction -13 ;

$3 = -12$; l'effet principal 3 est aliasé avec l'interaction -12.

Comme pour le demi-plan supérieur, les valeurs des contrastes sont fonction des effets réels du plan complet. Nous pouvons écrire :

$$\begin{cases} L_1 = -L_{23} = a_1 - a_{23} \\ L_2 = -L_{13} = a_2 - a_{13} \\ L_3 = -L_{12} = a_3 - a_{12} \\ L_0 = -L_{123} = a_0 - a_{123} \end{cases}$$

Ainsi, le système obtenu après avoir réduit le nombre d'inconnues M , se présente comme suit :

Tableau III.6 : Matrice des effets X_{inf} (N, N) du PFF

N°d'essai	x_0	x_1	x_2	x_3	y(%)
2	+1	+1	-1	-1	55,89
3	+1	-1	+1	-1	61,23
5	+1	-1	-1	+1	89,95
8	+1	+1	+1	+1	88,50

Demi plan inf

Les contrastes calculés pour ce demi-plan sont répertoriés dans le Tableau III.7.

Tableau III.7 : Valeurs des contrastes du modèle d'adsorption de l'AMX pour un demi-plan inférieur.

Effets apparents	L_0	L_1	L_2	L_3
valeurs	73,89	1,69	0,97	15,33

Le modèle obtenu après calcul s'écrit alors :

$$\hat{y} (\%) = 73,89 + 1,69x_1 + 0,97x_2 + 15,33x_3$$

Remarque :

Dans le demi-plan inférieur, la colonne 123 ne comprend que des (-1) contrairement à la colonne I qui ne contient que des (+1) ; ce qui nous donne la relation de définition : $I = -123$. A partir de là, nous pouvons retrouver tous les aliases en posant : $1 \times I = -1 \times 123 = -1^2 \times 23$, soit $1 = -23$ par exemple.

Une fois les valeurs des contrastes déterminées, l'analyse statistique est la même que celle présentée dans le chapitre II, pour valider le modèle mathématique postulé avant qu'il soit interprété.

Chapitre IV

Les plans pour surfaces de réponse

Chapitre IV : Les plans pour surfaces de réponse

IV.1. Introduction

Les premiers plans d'expériences diffusés dans la littérature ont été destinés à l'estimation des effets des facteurs. Cependant, du point de vue industriel, tous les problèmes rencontrés ne consistent pas uniquement à estimer puis à comparer les effets des facteurs. De nombreuses études consistent à trouver, s'il en existe, un optimum dans un domaine d'étude appelé domaine expérimental. Pour ce faire, la méthodologie des plans d'expériences pour surfaces de réponse a été créée afin de déterminer d'une façon quantitative les variations de la réponse vis-à-vis des facteurs d'influence significative, identifiés si besoin lors d'une première étude de screening.

IV.2. Plans pour surfaces de réponse

L'étude des surfaces de réponse est associée à l'utilisation des polynômes de second degré qui s'écrivent sous la forme générale suivante [1-3, 5-9, 13]:

$$\hat{y} = b_0 + \sum_{j=1}^k b_j x_j + \sum_{j=1}^k \sum_{u \neq j}^k b_{ju} x_j x_u + \sum_{j=1}^k b_{jj} x_j^2 \quad (\text{Eq. IV} - 1)$$

Les plans correspondants à ce type de modèle doivent :

- permettre d'étudier chacun des facteurs sur au moins trois niveaux ;
- comporter un nombre d'essais raisonnable ;
- satisfaire aux critères d'optimalité, sachant que la matrice \mathbf{X} ne peut plus être orthogonale.

Il existe plusieurs types de plans du second degré dont les principaux sont :

IV.2.1. Plan composite centré

Le plan composite centré consiste à réaliser les essais d'un plan factoriel complété par des expériences au centre du domaine d'étude et des essais en étoiles (Figure IV.1). Le nombre total d'essais N à réaliser est la somme de :

- n_f essais d'un plan factoriel complet ;
- $n_\alpha = 2k$ essais en étoiles sur les axes à une distance α du centre du domaine ;
- n_0 essais au centre du domaine.

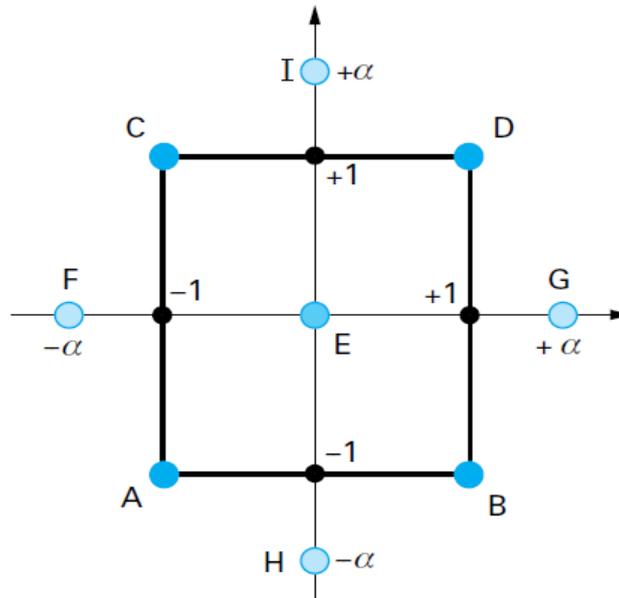


Figure IV.1 : Schéma représentatif d'un plan composite centré pour deux facteurs [5].

La figure ci-dessus illustre le domaine d'étude d'un plan composite à deux facteurs. Les points *A*, *B*, *C* et *D* sont les points expérimentaux d'un plan factoriel complet 2^2 . Le point *E* est le point central. Ce point peut être répété une ou plusieurs fois. Les points *F*, *G*, *H* et *I* sont les points axiaux situés à une distance particulière α du centre du domaine d'étude (α joue un rôle important pour la qualité du modèle et ne doit pas être choisi au hasard). Ces quatre derniers points forment ce que l'on appelle le plan en étoiles.

Par ailleurs, le modèle mathématique classique issu du modèle général (Eq. IV-1) et représentant le plan composite à deux facteurs s'écrit :

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

Basé sur le schéma représentatif du domaine d'étude du plan composite et le modèle mathématique le représentant, il ressort qu'on est ramené à réaliser 9 essais pour déterminer 6 coefficients. Il faut donc résoudre un système de 9 équations à 6 inconnues.

IV.2.1.1. Propriétés des plans composites

- **Orthogonalité de la matrice de calcul**

Pour un plan composite à deux facteurs, 9 expériences sont nécessaires pour déterminer les valeurs des 6 coefficients de l'équation du modèle de second degré postulé. Ainsi, la matrice des effets est la suivante :

$$X = \begin{vmatrix} + & - & - & + & + & + \\ + & + & - & - & + & + \\ + & - & + & + & + & + \\ + & + & + & - & + & + \\ + & 0 & 0 & 0 & 0 & 0 \\ + & -\alpha & 0 & 0 & +\alpha^2 & 0 \\ + & +\alpha & 0 & 0 & +\alpha^2 & 0 \\ + & 0 & -\alpha & 0 & 0 & +\alpha^2 \\ + & 0 & +\alpha & 0 & 0 & +\alpha^2 \end{vmatrix}$$

Comme la matrice X n'est pas orthogonale, la matrice $[X^T X]^{-1}$ n'est plus diagonale. Les variances des coefficients du modèle sont obtenues en multipliant les éléments C_{jj} de la diagonale de cette matrice de dispersion par la variance de reproductibilité :

$$S_{bj}^2 = C_{jj} \cdot S_{rep}^2 \quad (\text{Eq. IV - 2})$$

Le vecteur de la solution se calcule de la même manière que pour le plan factoriel à l'aide de l'équation suivante :

$$A = [X^T X]^{-1} X^T Y$$

- **Critères d'optimalité**

- ✓ **Critère d'isovariance par rotation** : pour que le plan composite satisfasse ce critère, il faut placer les points en étoiles à une distance α égale à la racine quatrième du nombre de points du plan factoriel : $\alpha = n_f^{1/4}$;
- ✓ **Critère de précision uniforme** : obtenu lorsqu'on augmente le nombre de points centraux. Son intérêt est d'assurer une même précision de réponse prédite sur tout le domaine ;
- ✓ **Critère de presque orthogonalité** : le critère d'orthogonalité est respecté si la matrice $[X^T X]^{-1}$ est diagonale. Dans le cas des plans composites centrés, ce critère est respecté si α choisi satisfait la condition suivante :

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

Les valeurs du paramètre α et du nombre n_0 de points au centre sont fonction du nombre de facteurs k du plan factoriel de base et du critère d'optimalité rempli par le plan (Tableau IV.1).

Tableau IV.1 : Valeurs de α et de n_0 selon les propriétés recherchées pour le plan composite **[2, 12]**.

k		2	3	4	5	$2^{(5-1)}$	6	$2^{(6-1)}$
n_f	2^k (ou 2^{k-p})	4	8	16	32	16	64	32
n_α		4	6	8	10	10	12	12
	isovariant par rotation	≥ 1	≥ 1	≥ 1				
n_0	précision uniforme	5	6	7	10	6	15	9
	ortogonalité	8	12	12	17	10	24	15
α		1,41	1,68	2	2,38	2	2,83	2,38

Remarque :

L'inconvénient majeur de la stratégie du plan composite réside dans sa nécessité de réaliser plus de points expérimentaux que de coefficients à déterminer.

IV.2.2. Plan Box-Behnken

Le plan Box-Behnken est économique et ne nécessite que 3 niveaux par facteur contrairement au plan composite qui, lui nécessite 5 niveaux. Le plan Box-Behnken est un plan de second degré qui ne repose pas sur l'étape du plan factoriel complet. Il est réalisé lorsque le chercheur est sûr de connaître tous les facteurs qui influent sur le phénomène étudié. Il commence avec trois facteurs, comme il peut utiliser quatre, cinq ou plus.

Le plan de Box-Behnken possède la propriété de séquentialité par rapport aux facteurs, c'est-à-dire que l'on peut entreprendre l'étude de k facteurs en se réservant la possibilité d'en rajouter d'autres tout en gardant les résultats des essais déjà effectués.

Pour un nombre de trois facteurs, le plan Box-Behnken se construit de la manière suivante (Figure IV.2) :

- écrire un plan 2^2 sur deux facteurs et mettre $k-2$ variables restantes au niveau 0 (dans ce cas $k-2= 3-2=1$) ;

- répéter l'opération pour toutes les combinaisons possibles des facteurs (dans ce cas, répéter deux fois) ;
- ajouter des points au centre.

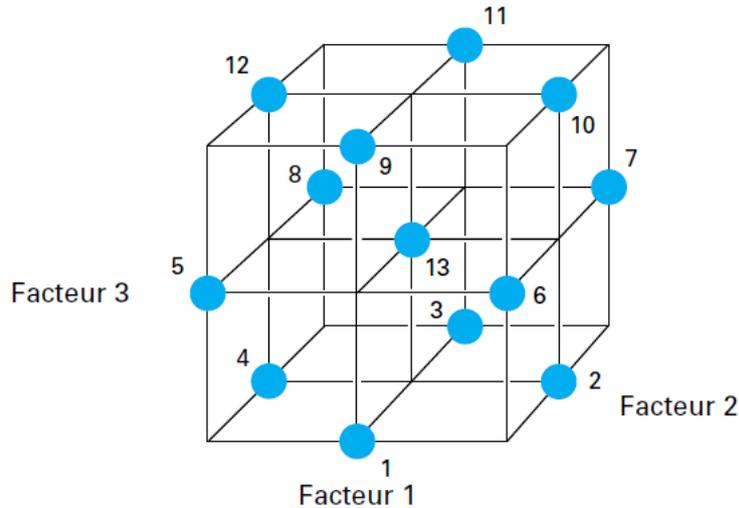


Figure IV.2 : Plan de Box-Behnken pour trois facteurs [5].

La figure IV.2 illustre le domaine d'étude d'un plan Box-Behnken à trois facteurs. Les points expérimentaux (points 1 jusqu'à 12) sont tous placés à égale distance du centre du domaine d'étude : milieu des arêtes du cube. Le point 13, en revanche, représente l'essai au centre que messieurs Box et Behnken recommandent de répéter trois fois. Ainsi, ce plan comporte douze essais auxquels nous pouvons ajouter un (ou plusieurs) point(s) central (centraux). Le Tableau IV.2 présente ces douze essais accompagnés de 3 points centraux. La construction de la matrice du Plan Box- Behnken lui confère la particularité d'être isovariant par rotation (symétrie sphérique).

Comme pour le plan composite, le modèle mathématique classique issu du modèle général (Eq. IV-1) et représentant le plan Box-Behnken à trois facteurs s'écrit :

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

En combinant le domaine d'étude schématisé sur la figure IV.2 (Tableau IV.2) et le modèle mathématique représentatif pour un nombre de facteurs égale à trois, il ressort qu'un plan de Box-Behnken nécessite 13 à 15 essais pour déterminer 10 coefficients. Il faut donc résoudre un système de 13 à 15 équations à 10 inconnues.

Tableau IV.2: Matrice de Box- Behnken pour 3 facteurs.

N°d'essai	x_1	x_2	x_3
1	-1	-1	0
2	+1	-1	0
3	-1	+1	0
4	+1	+1	0
5	-1	0	-1
6	+1	0	-1
7	-1	0	+1
8	+1	0	+1
9	0	-1	-1
10	0	+1	-1
11	0	-1	+1
12	0	+1	+1
13	0	0	0
14	0	0	0
15	0	0	0

Remarque :

Pour déterminer le nombre total d'essais N à réaliser, il est préférable d'adopter d'abord le modèle mathématique de second degré correspondant au nombre de facteurs étudiés, puis compter le nombre d'interactions que comporte ce modèle. A partir de là, $N = (\text{nombre d'interaction} \times 2^2) + n_0$.

Comme la matrice X du plan Box-Behnken n'est pas orthogonale, la matrice $[X^T X]^{-1}$ n'est plus diagonale. Ce qui revient à faire le même calcul que le plan composite pour déterminer les coefficients du modèle postulé.

IV.3. Notions de surfaces de réponse et courbes isoréponses

Une fois le modèle mathématique de second degré établi, il est utilisé pour accéder à la combinaison optimale des paramètres opératoires conduisant à une meilleure réponse y . La recherche d'optimum peut se faire, soit par une résolution mathématique de

l'équation représentative du phénomène étudié, soit par une résolution graphique en utilisant le tracé des surfaces de réponse et des courbes isoréponses [1-3, 5-9].

IV.3.1. Surfaces de réponse

L'exploitation du modèle de second degré obtenu après expérimentation permet de tracer dans l'espace des variables x_1, x_2, \dots, x_k et de \hat{y} , la relation $\hat{y} = f(x_1, x_2, \dots, x_k)$. Cette relation est donc illustrée par une surface appelée *surfaces de réponse* qui n'est autre qu'un ensemble de graphiques représentés dans un espace à trois dimensions (Figure IV.3) ; deux dimensions pour les facteurs et une pour la réponse. Ainsi, le plan horizontal de la figure matérialise le domaine de variation de deux facteurs et l'axe vertical matérialise la variation de la réponse à partir du modèle postulé. Il est à noter que l'allure des surfaces de réponse (paraboloïdes ou hyperboloïdes) peut nous renseigner sur l'existence d'un maximum ou d'un minimum ou des deux à la fois. Cependant, leurs courbures sont dues à la présence des termes carrés dans le modèle.

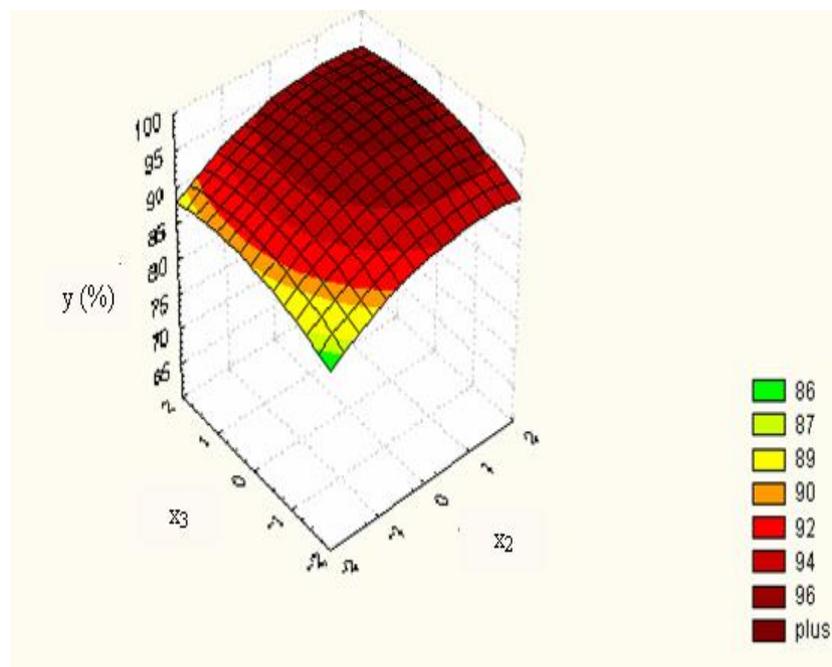


Figure IV.3 : Courbe de Surfaces de réponse [13].

IV.3.2. Courbes isoréponses

Les courbes isoréponses interprètent la projection de l'ensemble des points constituant la surface de réponse dans le plan horizontal des variables x_1, x_2, \dots, x_k (Figure IV.4). Elles s'interprètent comme des courbes de niveaux sur lesquelles est projetée la valeur de la réponse.

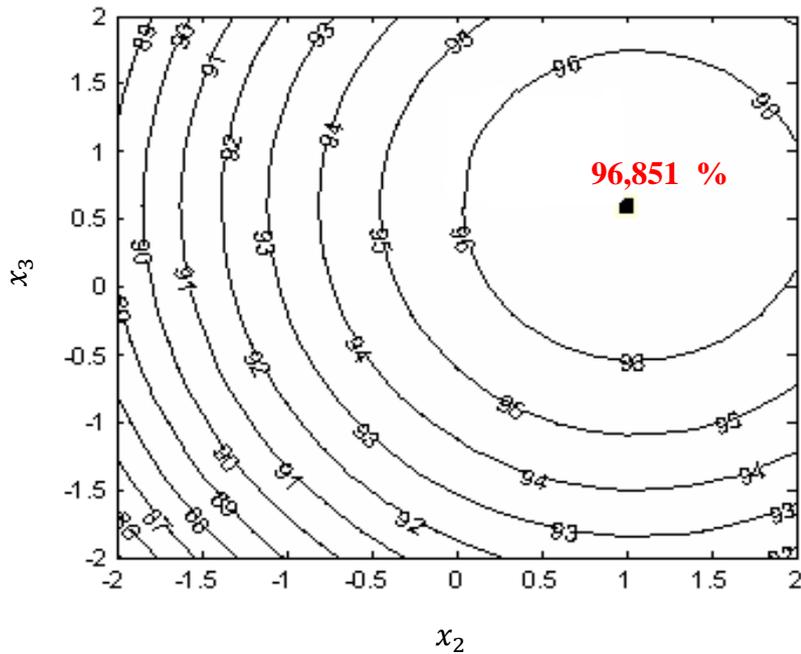


Figure IV.4 : Courbe des isoréponses [13]

IV.4. Exemple d'application

IV.4.1. Problème à résoudre [14]

Un chercheur désire exploiter le déchet des semelles de chaussures pour récupérer un polluant de nature organique : jaune bézacryl (JB) dans un réacteur agité fermé. Son intérêt majeur est d'obtenir les conditions optimales de pH, de concentration en JB et de concentration en déchet de semelles de chaussures (sous forme de suspension) qui donnent un meilleur taux d'adsorption du jaune bézacryl. Le chercheur utilise pour réaliser cette étude un plan composite centré. Les domaines de variation de chacun des paramètres étudiés sont : $[JB] \in [10-50] \text{ mg/L}$, $[\text{susp}] \in [0,1-1] \text{ g/250mL}$ et $\text{pH} \in [2-10]$.

IV.4.2. Formulation du problème

- 1) objectif de l'étude : optimisation des paramètres opératoires pour un meilleur rendement d'adsorption du jaune bézacryl par le déchet de semelles de chaussures ;
- 2) définir les facteurs et la réponse :

Réponse y : taux d'adsorption du jaune bézacryl par le procédé de l'adsorption, calculé comme suit :

$$y (\%) = \frac{[JB]_0 - [JB]_t}{[JB]_0} \times 100$$

Avec :

$[JB]_0$: concentration initiale en jaune bézacyl ;

$[JB]_t$: concentration de la solution en jaune bézacyl à l'instant t.

Facteurs et domaines de variation :

Facteur 1 $\Rightarrow z_1$: concentration en jaune bézacyl, $[JB] \in [10-50]$ mg/L ;

Facteur 2 $\Rightarrow z_2$: concentration de la suspension en semelles de chaussures ; $[susp] \in [0,1-1]$ g/250mL ;

Facteur 3 $\Rightarrow z_3$: pH de la solution ; $pH \in [2-10]$

Pour mener une étude basée sur un plan composite centré à trois facteurs et à précision uniforme, il nous faut $N = n_f + n_\alpha + n_0 = 8 + 6 + 6 = 20$ expériences à réaliser. Le modèle mathématique associé à ce plan est un modèle à $M = 10$ coefficients qui s'écrit sous la forme suivante :

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

Les niveaux des facteurs et leurs correspondances en variables codées sont répertoriés dans le Tableau IV.3

Tableau IV.3 : Valeurs des paramètres opératoires à différents niveaux.

Variables réelles	Variables codées	Niveau bas		Point central	Niveau haut	
		-1,68	-1		0	+1
$z_1 : [JB]_0$ (mg/L)	x_1	10	18,1	30	41,9	50
$z_2 : [Susp]_0$ (g/250mL)	x_2	0,1	0,29	0,55	0,81	1
$z_3 : pH$	x_3	2	3,62	6	8,38	10

Les valeurs -1,68 et +1,68 sont respectivement, le niveau bas et haut pour les points axiaux α pour un nombre de facteurs $k=3$.

IV.4.3. Planification et expérimentation

- construction du plan d'expériences : construire la matrice d'expériences en unités réelles et en codées (Tableau IV.4) afin de l'utiliser comme outil de démarrage pour pouvoir lancer et faciliter la mise en œuvre de l'expérimentation.

Tableau IV.4 : Matrice des expériences d'un plan composite centré.

N	z_1	z_2	z_3	x_0	x_1	x_2	x_3	y (%)
1	18,1	0,29	3,62	+1	-1	-1	-1	93,92
2	18,1	0,29	8,38	+1	+1	-1	-1	82,45
3	18,1	0,81	3,62	+1	-1	+1	-1	70,11
4	18,1	0,81	8,38	+1	+1	+1	-1	97,70
5	41,9	0,29	3,62	+1	-1	-1	+1	85,52
6	41,9	0,29	8,38	+1	+1	-1	+1	87,85
7	41,9	0,81	3,62	+1	-1	+1	+1	87,40
8	41,9	0,81	8,38	+1	+1	+1	+1	86,42
9	30	0,55	6	+1	0	0	0	88,26
10	30	0,55	6	+1	0	0	0	91,63
11	30	0,55	6	+1	0	0	0	89,30
12	30	0,55	6	+1	0	0	0	87,40
13	30	0,55	6	+1	0	0	0	90,50
14	30	0,55	6	+1	0	0	0	90,00
15	10	0,55	6	+1	-1,68	0	0	81,20
16	50	0,55	6	+1	+1,68	0	0	85,82
17	30	0,1	6	+1	0	-1,68	0	94,30
18	30	1	6	+1	0	+1,68	0	90,00
19	30	0,55	2	+1	0	0	-1,68	23,43
20	30	0,55	10	+1	0	0	+1,68	80,36

IV.4.4. Analyse des résultats

- **Calcul des coefficients du modèle**

Comme nous l'avons déjà signalé précédemment, l'estimation des coefficients de l'équation de régression se fait par le calcul du produit matriciel suivant :

$B = [X^T X]^{-1} X^T Y$. Les résultats de ce calcul sont rassemblés dans le Tableau IV.5.

Tableau IV.5 : Valeurs des coefficients du modèle de second degré.

Terme constant		Effets linéaires			
a_0		a_1	a_2	a_3	
89,13		1,85	-1,12	7,23	
Effets d'interactions et quadratiques					
a_{12}	a_{13}	a_{23}	a_{11}	a_{22}	a_{33}
4,47	-1,85	1,13	0,30	3,35	-10,91

- **Analyse statistique de l'équation de régression**
 - **Vérification de la signification des coefficients**

La valeur de la variance de reproductibilité est $S^2_{rep} = 2,35$. Cependant, pour déterminer la variance des coefficients du modèle, il suffit de multiplier les éléments C_{jj} de la diagonale de la matrice de dispersion $[X^T X]^{-1}$ (EXCEL) par la variance de reproductibilité : $S^2_{bj} = C_{jj} S^2_{rep}$. Les valeurs de t_j obtenues pour chaque effet sont rassemblées dans le Tableau IV.6.

Tableau IV.6 : Valeurs de t_j

Terme constant		Effets linéaires			
t_0		t_1	t_2	t_3	
227,61		10,73	6,51	41,90	
Effets d'interactions					
t_{12}	t_{13}	t_{23}	t_{11}	t_{22}	t_{33}
15,18	6,27	3,83	1,80	20,46	66,56

Pour le niveau de signification $\alpha = 0,05$ et le nombre de degré de liberté $f = (n_0 - 1) = 5$, la valeur tabulée du test de Student $t_\alpha(f)$ est égale à 2,57. Comme la valeur de t_{11} est inférieure à $t_\alpha(f)$, le coefficient a_{11} est donc non significatif et par conséquent éliminé de l'équation de régression.

- **Test de validation de l'équation de régression**
 - ✓ **Recherche de biais**

La valeur calculée de la variance résiduelle pour $N = 20$ et $\ell = 9$ est égale à $S_{res}^2 = 142,26$, ce qui donne : $F = 1,59$. La valeur tabulée du test de Fisher Snedecor pour le niveau de signification $\alpha = 0,05$ et les nombres de degrés de liberté $(N - \ell) = 11$ et $(n_0 - 1) = 5$ varie

entre 4,74 et 4,68. Cet intervalle étant supérieur à la valeur calculée, le modèle est alors considéré sans biais.

✓ Test de signification de la régression

Pour le niveau de signification $\alpha = 0,05$ et les nombres de degrés de liberté $(\ell - 1) = 8$ et $(N - \ell) = 11$, la valeur du test de Fisher tabulée est de 2,95. La valeur de F calculée étant supérieure à celle tabulée, l'équation de régression est adéquate et le modèle est validé à 95%. L'équation retenue pour le modèle s'écrit donc :

$$\hat{y} = 89,13 + 1,85 x_1 - 1,12x_2 + 7,23x_3 + 4,74x_1x_2 - 1,85x_1x_3 + 1,13x_2x_3 \\ + 3,35 x_2^2 - 10,91 x_3^2$$

L'analyse du modèle indique que les paramètres les plus influents sont la concentration initiale en colorant et le pH de la solution avec des effets positifs 1,85, et 7,23 respectivement. La concentration en suspension influe négativement avec un effet négatif de - 1,12.

• Analyse des résidus

L'évaluation de la qualité du modèle de second degré peut être faite par une analyse des résidus (Figure IV.5).

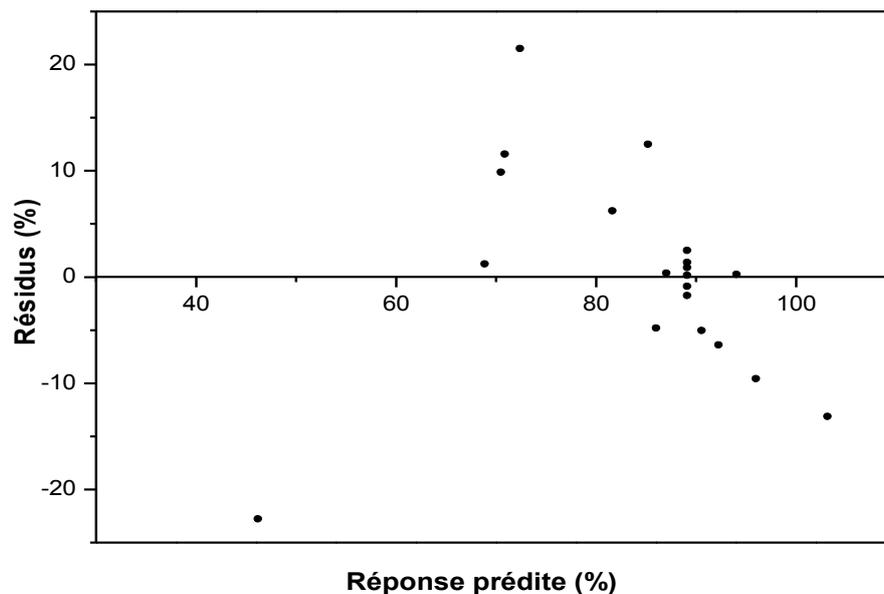


Figure IV.5 : Diagramme des résidus

Le diagramme de la Figure IV.5 ne laisse apparaître aucune relation entre les valeurs prédites et les résidus, car les points semblent être disposés aléatoirement. Ce résultat traduit l'absence d'informations dans les résidus donc toute l'information est expliquée par le modèle de régression retenu. En conclusion, le modèle de second degré obtenu simule parfaitement le procédé d'adsorption du JB sur le déchet de semelles; il sera donc exploité pour l'optimisation des paramètres opératoires.

IV.4.5. Optimisation

Après l'étape de validation qui nous a permis d'aboutir à un modèle de second degré traduisant l'effet de la variation des trois paramètres opératoires étudiés sur le taux d'adsorption du JB par le déchet des semelles de chaussures, la recherche des valeurs optimales de ces paramètres peut se faire alors soit par :

- **Résolution mathématique**

Rappelons que la modélisation du processus étudié a abouti à l'équation suivante :

$$\hat{y} = 89,13 + 1,85 x_1 - 1,12x_2 + 7,23x_3 + 4,74x_1x_2 - 1,85x_1x_3 + 1,13x_2x_3 + 3,35 x_2^2 - 10,91 x_3^2$$

Cette équation mathématique prédictive permet de calculer les valeurs optimales des paramètres opératoires conduisant à l'optimum de la réponse dans le domaine d'étude. Il suffit de dériver la fonction par rapport à chacune des variables x_1 , x_2 et x_3 . Le système d'équations obtenu après dérivation est représenté comme suit :

$$\begin{cases} \frac{\partial \hat{y}}{\partial x_1} = 1,85 + 4,74x_2 - 1,85x_3 = 0 \\ \frac{\partial \hat{y}}{\partial x_2} = -1,12 + 4,74x_1 + 1,13x_3 + 6,70 x_2 = 0 \\ \frac{\partial \hat{y}}{\partial x_3} = 7,23 - 1,85x_1 + 1,13x_2 - 21,82 x_3 = 0 \end{cases}$$

La résolution de ce système ne peut se faire analytiquement, il faut donc envisager une résolution numérique.

- **Résolution graphique**

Les courbes de surfaces de réponse tracées via le logiciel STASTICA laissent apparaître un maximum qui traduit l'existence d'un rendement maximal.

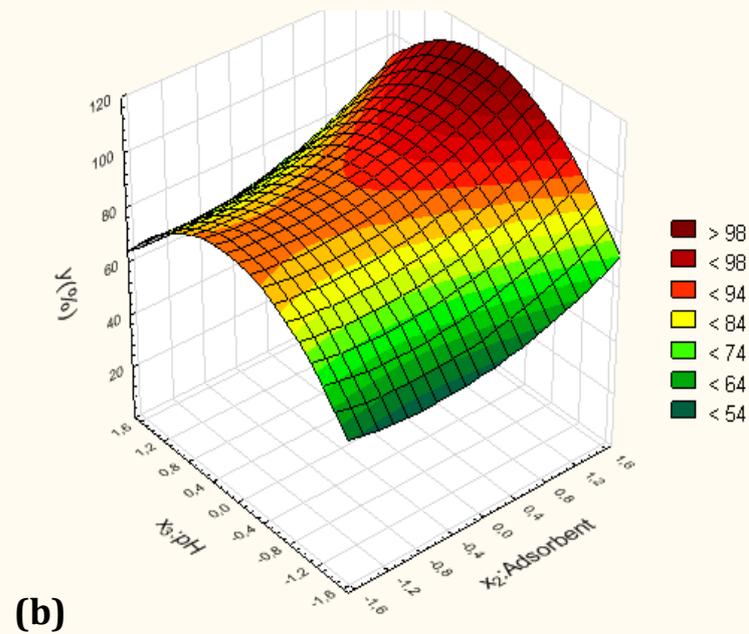
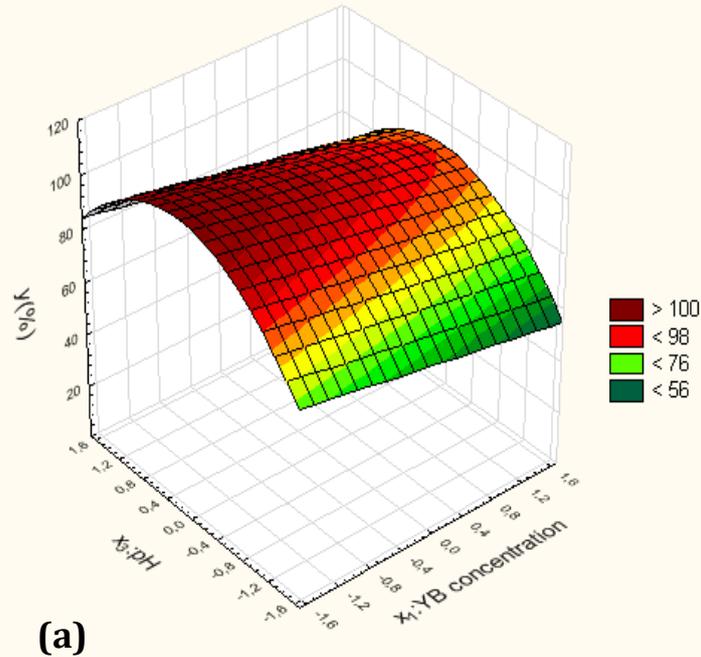


Figure IV.6 : Courbes de Surfaces de réponse illustrant l'effet de :

(a) : $[susp]_0$ - pH à $[JB]_0 = 50$ mg/L ;

(b) : $[JB]_0$ - pH à $[susp]_0 = 0,3$ g/250mL sur le taux d'adsorption du JB.

Selon ces courbes, un taux d'adsorption de JB d'environ 98 % (zone rouge) est obtenu lorsqu'on travaille avec une concentration en JB égale à 30 mg/L, une concentration en suspension de semelles de chaussures égale à 0,1 g/250 mL et un pH égale à 6. Cette combinaison des trois paramètres représente les conditions optimales recherchées.

Chapitre V

Les plans de mélanges

Chapitre V : Les plans de mélanges

V.1. Introduction

Contrairement aux plans d'expériences classiques (plans factoriels complets et fractionnaires, plans pour surfaces de réponse,...) qui se basent sur l'étude de l'influence des facteurs indépendants sur la réponse mesurée, les plans de mélanges sont des plans particuliers, adaptés aux facteurs dépendants. Ils sont surtout employés pour traiter les cas où les propriétés d'un produit quelconque sont dépendantes uniquement des proportions de ses constituants et non pas de ses quantités utilisées, chose qui leur permet d'atteindre d'autres objectifs que l'étude des effets principaux et des interactions.

V.2. Contrainte fondamentale des mélanges

Si l'on note par x_i la teneur en constituant i d'un mélange quelconque, la somme des teneurs de tous les constituants de ce dernier satisfait la relation [2, 3, 5, 9, 10, 15]:

$$\sum_{i=1}^N x_i = 100\% \quad \text{ou} \quad \sum_{i=1}^N x_i = 1 \quad (\text{Eq.V} - 1)$$

Ainsi, la teneur de chaque constituant reste comprise entre 0 et 100 %. Lorsqu'on augmente la teneur de l'un des constituants, automatiquement les autres teneurs diminuent pour que la somme de toutes les teneurs reste égale à 100 %.

Cette relation appelée *contrainte fondamentale des mélanges* a des conséquences importantes sur les représentations géométrique et matricielle de ces plans. Elle a également des répercussions sur la modélisation des résultats.

V.3. Plans de mélanges sans contraintes

Dans ce type de plans, les mélanges étudiés sont constitués de produits dont les teneurs peuvent varier de 0 jusqu'à 1 [2, 3, 5, 9, 10, 15].

V.3.1. Représentation géométrique

Comme pour les plans d'expériences, on peut utiliser des axes cartésiens pour présenter graphiquement la composition du mélange étudié. Le premier axe est attribué au premier constituant (facteur 1 : x_1). Le second axe, orthogonal au premier est attribué au second constituant (facteur 2 : x_2) et ainsi de suite. La Figure V.1 illustre la

représentation cartésienne d'un mélange à trois constituants A, B et C. La Figure V.1.(a) indique l'emplacement d'un point d'un mélange à trois constituants tandis que la Figure V.1.(b) révèle l'ensemble des mélanges à trois constituants dans l'espace cartésien des trois constituants A, B et C.

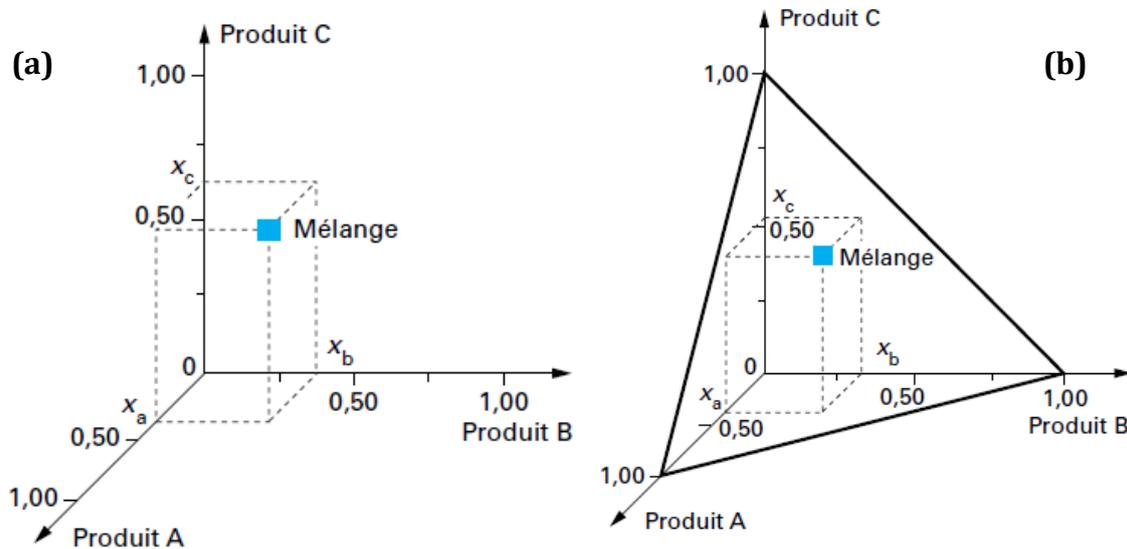


Figure V.1: Représentation de mélanges à trois constituants dans un espace cartésien[5].

Il est à noter que la représentation géométrique d'un mélange doit également tenir compte de la contrainte fondamentale des mélanges (Eq. V-1.). Celle-ci fait disparaître une dimension dans l'espace des x_i . Ainsi, pour définir un domaine d'étude d'un mélange composé de deux constituants A, B (binaire), on utilise un segment de droite (Figure V.2. (a)) et pour un mélange à trois constituants A, B et C (ternaire), on adopte la représentation du triangle équilatéral ayant pour sommets les trois points d'abscisses 1 (Figure V.2.(b)). En effet, les constituants purs sont représentés par les extrémités du segment de droite pour le binaire et les sommets du triangle équilatéral dans le cas du ternaire. Par exemple, un mélange binaire à 50 % en A est situé au milieu du segment délimité par les constituants A et B. Le point M de coordonnées x_a , x_b et x_c correspond,, en revanche, à un mélange tertiaire.

Remarque :

Les sommets du triangle équilatéral correspondent aux constituants purs ;

Les côtés du triangle correspondent aux mélanges binaires ;

L'intérieur du triangle correspond à tous les mélanges ternaires

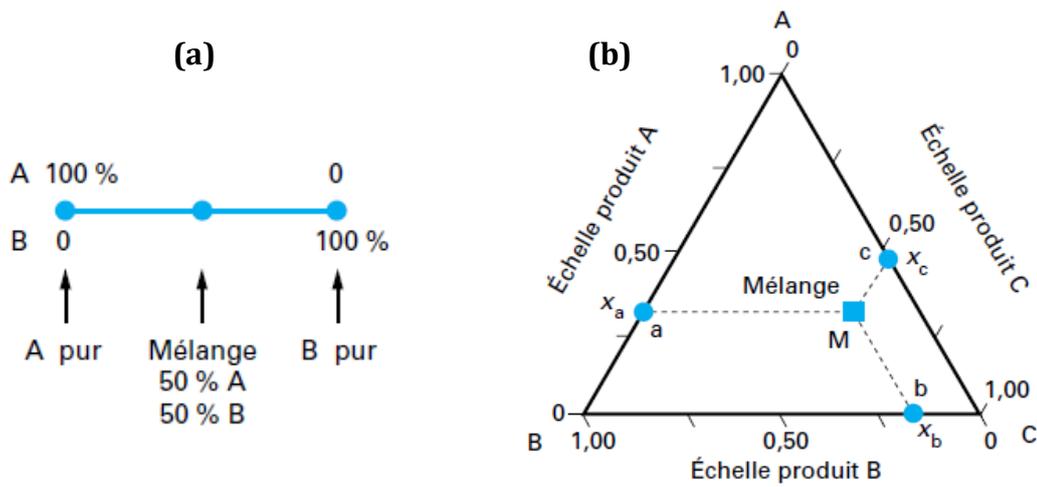


Figure V.2: Domaine d'étude de mélanges composés : (a) : deux constituants et (b) : trois constituants [5, 10].

Au delà de quatre constituants, la représentation géométrique devient difficile ce qui nécessite l'utilisation de tableau (matrice des essais) pour regrouper tous les mélanges étudiés, comme pour les plans d'expériences classiques.

V.3.2. Emplacement des points expérimentaux

Pour un mélange à trois constituants et lorsqu'il n'y a pas de contraintes, les points d'expériences sont répartis dans tout le domaine d'étude (Figure V.3).

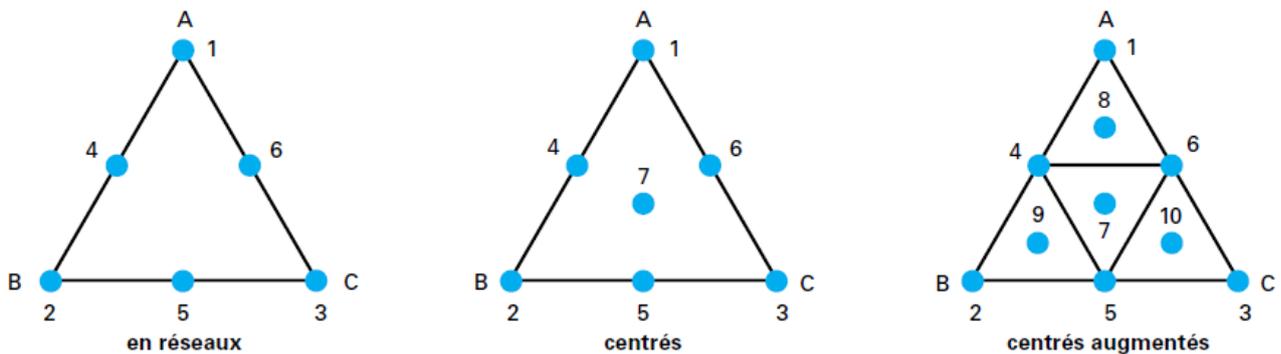


Figure V.3: Plan de mélanges en réseaux (à gauche), plan de mélanges centrés (au milieu), plan de mélanges centrés augmentés (à droite) [5].

Selon la disposition de ces points, on distingue plusieurs types de plans de mélanges :

- plans de mélanges en réseaux (Simplex lattice designs) : ils comportent les trois constituants purs et les mélanges moitié-moitié ;

- plans de mélanges centrés (Simplex-Centroid Designs) : ils se distinguent des plans en réseaux par la présence systématique d'un point central composé de $\frac{1}{3}$ A, $\frac{1}{3}$ B et $\frac{1}{3}$ C ;
- plans de mélanges centrés augmentés (Augmented Simplex-Centroid Designs) : ce sont des plans de mélanges centrés auxquels on ajoute les centres de gravité des simplexes unitaires

V.4. Plans de mélanges avec contraintes

Dans ce type de plans, les mélanges sont caractérisés par de nombreuses contraintes (contraintes basses, hautes ou relationnelles) qui peuvent peser sur le choix des proportions des constituants. En fonction de ces contraintes, la planification de l'étude est modifiée et doit être adaptée à chaque cas [2, 3, 5, 9, 10, 15].

V.4.1. Basses teneurs interdites

Ce cas traduit le fait que la composition de un ou plusieurs constituants ne peut pas être inférieure à une valeur donnée : c'est la limite inférieure notée L_i . Cette condition réduit le domaine d'étude tout en gardant sa forme de départ (pour un mélange à trois constituants, on a un triangle équilatéral : Figure V.4.(a)).

V.4.2. Hautes teneurs interdites

Dans ce cas la composition de un ou plusieurs constituants ne peut pas être supérieure à une valeur donnée : c'est la limite supérieure notée U_i . Dans ce cas, le domaine d'étude est réduit, mais la forme du domaine est complètement modifiée (pour un mélange à trois constituants, ce n'est plus un triangle mais un polygone : Figure V.4.(b)).

V.4.3. Hautes et basses teneurs interdites

Cette contrainte révèle la possibilité que toute composition de un ou de plusieurs constituants peut être soumise, à la fois, à des contraintes inférieures et supérieures. L'introduction des contraintes doubles a en général pour effet de modifier la forme du domaine expérimental. Il peut être un triangle équilatéral, un trapèze, un pentagone,(Figure V.4.(c)).

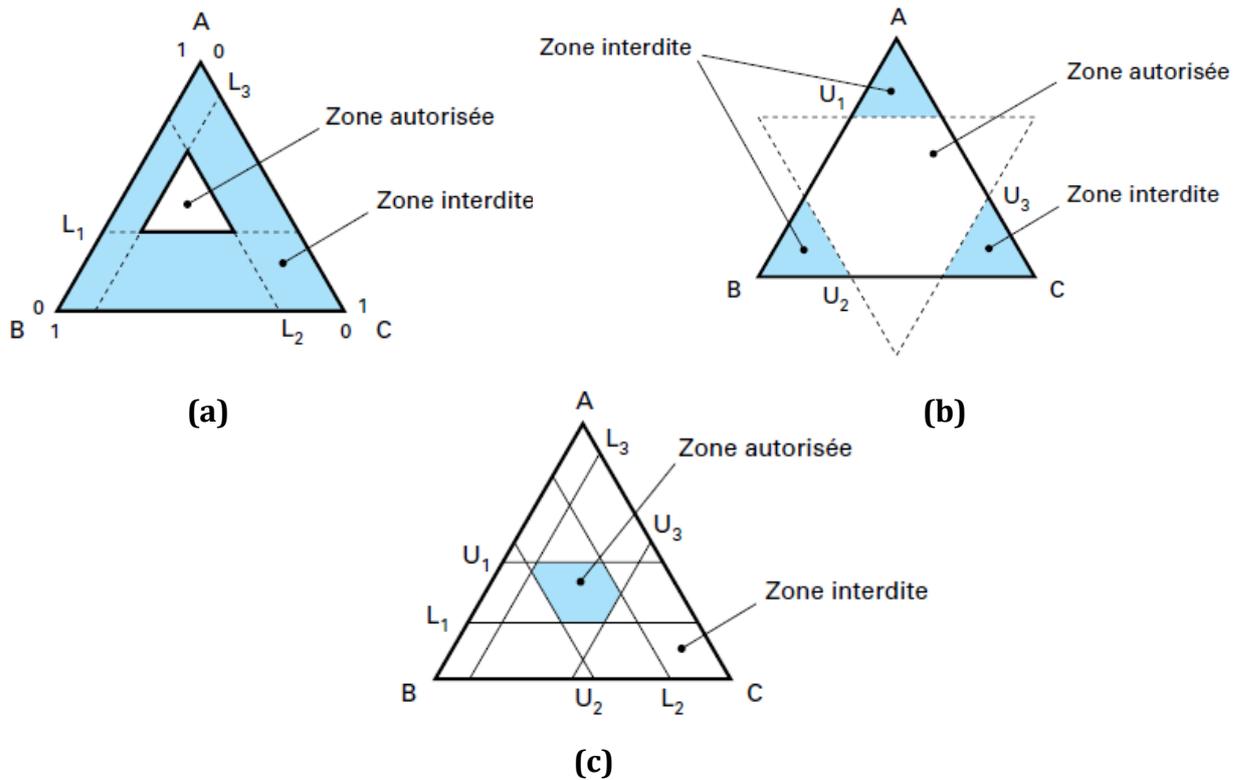


Figure V.4 : Représentation des domaines d'étude des mélanges avec contraintes [5]

V.4.4. Contraintes relationnelles

En plus aux contraintes hautes et basses teneurs interdites, les mélanges présentent d'autres contraintes relationnelles liées soit :

- à la conservation d'un rapport constant entre les proportions de deux constituants ;
- au respect d'une relation d'addition entre les proportions de deux ou de plusieurs constituants.

Ces nouvelles contraintes entraînent de nouvelles restrictions sur le domaine d'étude et modifient l'emplacement des points d'expérimentation.

V.4.5. Emplacement des points expérimentaux.

La démarche générale adoptée pour sélectionner les points expérimentaux d'un domaine de forme quelconque découle de la proposition de Scheffé pour les domaines réguliers ; les points informatifs sont situés aux sommets du domaine, aux milieux des arêtes, aux centres des faces ou des hyperfaces et au centre de gravité du volume ou de l'hypervolume.

V.5. Modèles mathématiques associés aux plans de mélanges

Comme tous les plans d'expériences classiques, pour traduire les variations de la propriété \hat{y} en fonction de la composition du mélange, les expérimentateurs choisissent des modèles polynomiaux classiques qui peuvent être écrits sous une forme simplifiée appelée forme canonique et qui résulte de l'existence de la contrainte égalité sur les fractions de constituants en mélange [2, 3, 5, 9, 10, 15].

V.5.1. Modèle de premier degré

Pour l'étude d'un mélange à trois constituants, le modèle mathématique de premier degré postulé s'obtient à partir d'un polynôme classique d'ordre un, soit :

$$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_3 \quad (\text{Eq.V} - 2)$$

auquel nous appliquons la contrainte fondamentale des mélanges :

$$x_1 + x_2 + x_3 = 1$$

Avec :

\hat{y} : réponse calculée à partir du modèle au point de composition choisie ;

La relation (Eq.V-2) peut alors s'écrire :

$$\hat{y} = a_0(x_1 + x_2 + x_3) + a_1x_1 + a_2x_2 + a_3x_3 \quad (\text{Eq.V} - 3)$$

En regroupant les coefficients :

$$\hat{y} = (a_0 + a_1)x_1 + (a_0 + a_2)x_2 + (a_0 + a_3)x_3$$

le modèle prend alors la forme suivante :

$$\hat{y} = b_1x_1 + b_2x_2 + b_3x_3 \quad (\text{Eq.V} - 4)$$

Avec :

$$\begin{cases} b_1 = a_0 + a_1 \\ b_2 = a_0 + a_2 \\ b_3 = a_0 + a_3 \end{cases} : \text{coefficients linéaires du modèle.}$$

V.5.2. Modèle de second degré

Le modèle de second degré avec interactions (synergique) qui s'écrit sous la forme :

$$\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3 + a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 \quad (\text{Eq. V} - 5)$$

devient , après l'application de la contrainte fondamentale des mélanges :

$$\hat{y} = b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \quad (\text{Eq. V} - 6)$$

Avec :

$$\begin{cases} x_1 = (1 - x_2 - x_3) \\ x_1^2 = x_1(1 - x_2 - x_3) \\ x_1^3 = x_1 - x_1x_2 - x_1x_3 \end{cases}$$

et

$$\begin{cases} b_1 = a_0 + a_1 + a_{11} \\ b_2 = a_0 + a_2 + a_{22} \\ b_3 = a_0 + a_3 + a_{33} \\ b_{12} = a_{12} - a_{11} - a_{22} \\ b_{13} = a_{13} - a_{11} - a_{33} \\ b_{23} = a_{23} - a_{22} - a_{33} \end{cases}$$

V.5.3. Modèles de troisième degré

Le modèle mathématique de troisième degré, particulier aux plans de mélanges, obtenu en appliquant la contrainte fondamentale des mélanges à un polynôme de troisième degré s'écrit sous la forme :

$$\hat{y} = b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 \quad (\text{Eq. V} - 7)$$

Ce modèle simplifié est appelé *modèle cubique restreint*. Il ne comprend que les termes linéaires et d'interactions double et triple.

Remarque :

La contrainte fondamentale des mélanges fait disparaître la constante et les termes du second degré sont réduites aux termes rectangles (interactions doubles et triples).

Les coefficients du modèle postulé sont calculés à partir de la relation de régression appliquée dans le cas des plans d'expériences classiques : $B = [X^T X]^{-1} X^T Y$.

Références bibliographiques

- [1] J. Goupy, Introduction aux plans d'expériences, Edition DUNOD, Paris, 2001.
- [2] G. Sado, M. C. Sado, Les plans d'expériences : De l'expérience à l'assurance qualité, Edition AFNOR TECHNIQUE, Paris, 1996.
- [3] Z. R. Lazic, Design of Experiments in Chemical Engineering, Edition WILEY-VCH, Weinheim, 2004.
- [4] P. Ozil, Plans d'expériences : Méthode de TAGUCHI, Ecole Nationale Supérieure d'Electrochimie et d'Electrometallurgie, INPG, 1997.
- [5] J. Goupy, Modélisation par les plans d'expériences, Techniques de l'Ingénieur, R275, 2000.
- [6] D. Benoit, Y. Tourbier, S. Germain, Plans d'expériences, construction et analyse, Technique et Documentation, Edition LAVOISIER, Paris 1994.
- [7] J. Goupy, Introduction aux plans d'expériences avec application, Edition DUNOD, Paris, 2013.
- [8] J. Goupy, Plans d'expériences pour surfaces de réponse, Edition DUNOD, Paris, 1999.
- [9] J. Goupy, Pratiquer les plans d'expériences, Edition DUNOD, Paris, 2005.
- [10] J. Goupy, Plans d'expériences, Techniques de l'Ingénieur, PE 230, 1997.
- [11] K. Hinkelmann, O. Kempthorne, Design and Analysis of Experiments: Advanced Experimental Design, Edition WILEY INTERSCIENCE, New Jersey, 2005.
- [12] V. Kafarov. Méthodes cybernétiques et technologie chimique, Edition MIR, Moscou, 1974.
- [13] I. Yahiaoui, F. Aissani-Benissad, H. Ait amar, Optimization of silver cementation yield in fixed bed reactor using factorial design and central composite design, The Canadian Journal of Chemical Engineering 6, 1099-1106, 2010.

[14] F. Boudrahem, S. Ziani, F. Aissani-Benissada, Application of Response Surface Methodology Based on Central Composite Design for Optimization of Yellow Bezacryl Sorption on Shoe Soles Waste, Environmental Progress and Sustainable Energy, DOI 10.1002/ep.

[15] J. Goupy, Les plans d'expériences : Les mélanges, Edition DUNOD, Paris, 2000.