Chapitre I : Introduction aux plans d'expériences

I.1. Historique

La Méthodologie des Plans d'Expériences (MPE) est une méthode qui a été initiée dans les années 20 par **Sir Ronald Aylmer Fisher** (statisticien anglais - 1925). Les premiers utilisateurs de cette méthode furent les agronomes qui ont vite compris l'intérêt des plans d'expériences. Vers les années soixante, grâce aux travaux de Taguchi, les plans d'expériences sont utilisés au Japon dans l'industrie pour améliorer la variabilité des procédés. Après le Japon les plans d'expériences sont utilisés aux Etats Unis dans les années 80 et en Europe dans les années 90.

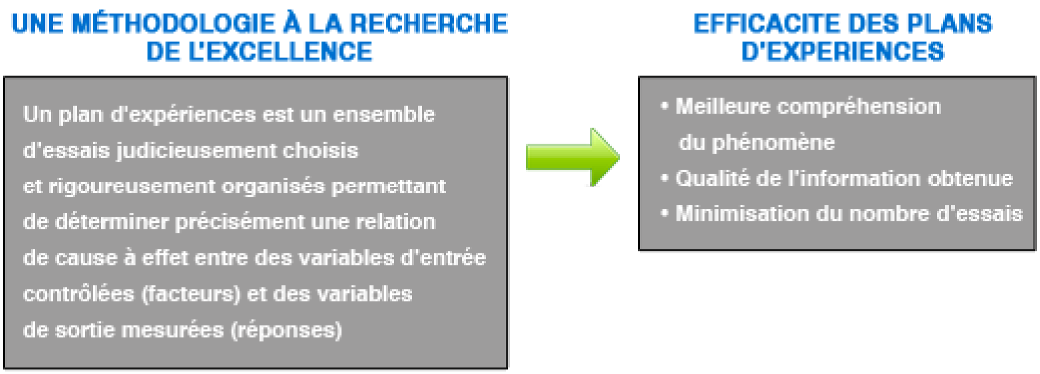
I.2 Introduction

Les plans d'expériences permettent d'organiser au mieux les essais qui accompagnent une recherche scientifique ou des études industrielles. Ils sont applicables à de nombreuses disciplines et à toutes les industries à partir du moment où l'on recherche le lien qui existe entre une grandeur d'intérêt, y, et des variables, xi. Il faut penser aux plans d'expériences si l'on s'intéresse à une fonction du type :

Réponses=f (Paramètre 1, Paramètre 2, Paramètre 3,…………..)

Avec les plans d'expériences on obtient le maximum de renseignements avec le minimum d'expériences. Pour cela, il faut suivre des règles mathématiques et adopter une démarche rigoureuse. Il existe de nombreux plans d'expériences adaptés à tous les cas rencontrés par un expérimentateur. Les principes fondamentaux de cette science seront indiqués et les principaux plans seront passés en revue.

La compréhension de la méthode des plans d'expériences s'appuie sur deux notions essentielles, celle d'espace expérimental et celle de modélisation mathématique des grandeurs étudiées.



Cette méthodologie a pour objectif d'optimiser l'efficacité de la recherche expérimentale. Pour ce faire, elle va aider l'expérimentateur à bien décrire son problème et lui proposer un choix de stratégies expérimentales optimales (enchaînements de plans d'expériences dans le temps), en fonction des **objectifs** qu'il s'est fixé et **des moyens** dont il dispose.

Les objectifs peuvent être :

- Explorer un domaine expérimental inconnu (recherche exploratoire dans le domaine de variation d'un ensemble de facteurs).
- Isoler les facteurs influents (criblage).
- Elaborer des modèles descriptifs ou prévisionnels des phénomènes étudiés (étude quantitative des facteurs, étude quantitative des réponses).
- Effectuer l'optimisation d'un produit ou d'un procédé.
- Concevoir un procédé robuste.
- Mettre au point des formulations.
- Améliorer la qualité de produits.

Comment peut-on choisir ces expériences pour :

– arriver rapidement aux meilleurs résultats possibles,
– éviter de réaliser des expériences inutiles,
– obtenir la meilleure précision possible sur les résultats,
– permettre d'avancer à coup sûr,
– établir la modélisation du phénomène étudié,
– découvrir la solution optimale.

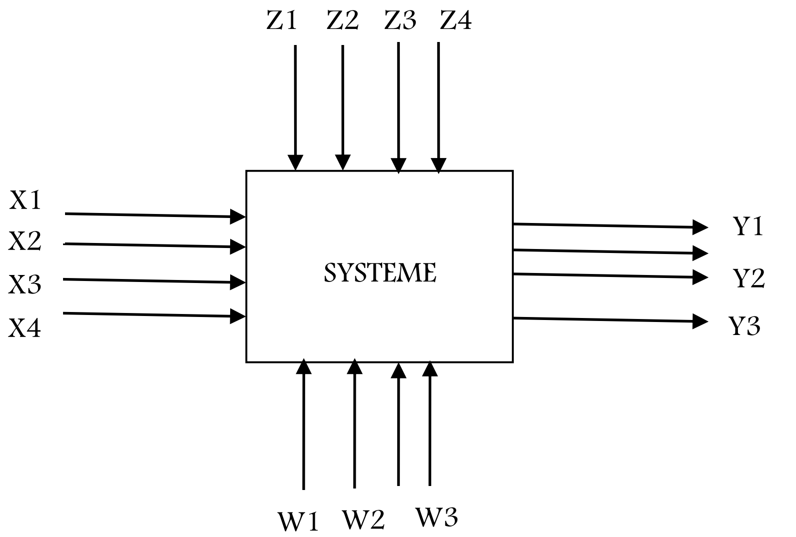


Figure I.1 : Schéma d'un système pour la modélisation

I.3. Notion d'espace expérimental

Un expérimentateur qui lance une étude s'intéresse à une grandeur qu'il mesure à chaque essai. Cette grandeur s'appelle la réponse, c'est la grandeur d'intérêt. La valeur de cette grandeur dépend de plusieurs variables. Au lieu du terme «variable» on utilisera le mot facteur. La réponse dépend donc de un ou de plusieurs facteurs. Le premier facteur peut être représenté par un axe gradué et orienté. La valeur donnée à un facteur pour réaliser un essai est appelée niveau. Lorsqu'on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes. La borne inférieure est le niveau bas. La borne supérieure est le niveau haut.

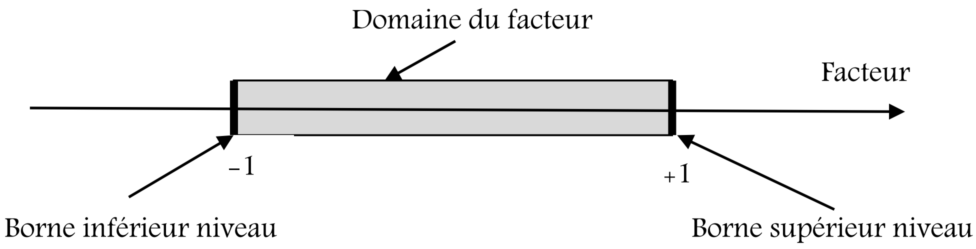


Figure I.2 : Le niveau bas du facteur est noté par - 1 et le niveau haut par +1. Le domaine de variation du facteur est constitué de toutes les valeurs comprises entre le niveau bas et le niveau haut.

L'ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut, s'appelle le domaine de variation du facteur ou plus simplement le domaine du facteur. On a l'habitude de noter le niveau bas par –1 et le niveau haut par +1.

S'il y a un second facteur, il est représenté, lui aussi, par un axe gradué et orienté. On définit, comme pour le premier facteur, son niveau haut, son niveau bas et son domaine de variation. Ce second axe est disposé orthogonalement au premier. On obtient ainsi un repère cartésien qui définit un espace euclidien à deux dimensions. Cet espace est appelé l'espace expérimental.

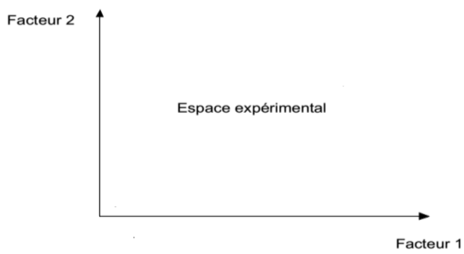


Figure I.3 : Chaque facteur est représenté par un axe gradué et orienté. Les axes des facteurs sont orthogonaux entre eux. L'espace ainsi défini est l'espace expérimental.

Le niveau x1 du facteur 1 et le niveau x2 du facteur 2 peuvent être considérés comme les coordonnées d'un point de l'espace expérimental. Une expérience donnée est alors représentée par un point dans ce système d'axes. Un plan d'expériences est représenté par un ensemble de points expérimentaux.

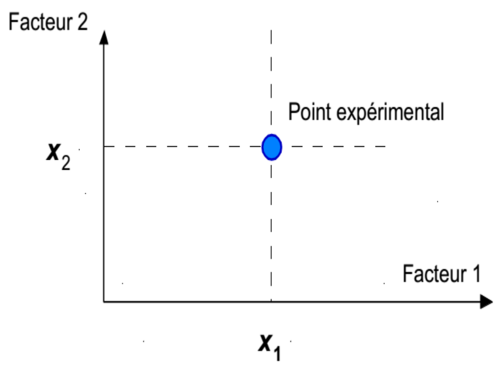


Figure I.4 : Dans l'espace expérimental, les niveaux des facteurs définissent des points expérimentaux.

Le regroupement des domaines des facteurs définit le «domaine d'étude». Ce domaine d'étude est la zone de l'espace expérimental choisie par l'expérimentateur pour faire ses essais. Une étude, c'est-à-dire plusieurs expériences bien définies, est représentée par des points répartis dans le domaine d'étude. Cette façon de représenter une expérimentation par des points dans un espace cartésien est une représentation géométrique de l'étude.

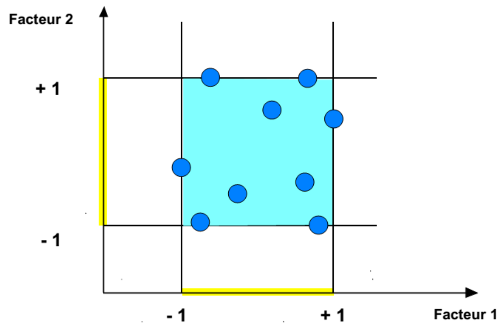


Figure I.5 : Les points expérimentaux sont disposés dans le domaine d'étude défini par l'expérimentateur.

S'il y a un troisième facteur, on le représente aussi par un axe orienté et gradué, et on le positionne perpendiculairement aux deux premiers. À partir de quatre facteurs, on opère de même, mais il n'y a plus de représentation géométrique possible et l'on doit adopter une représentation purement mathématique de l'espace expérimental qui est un hypervolume à quatre dimensions.

I.4. Étude d'un phénomène

L'étude d'un phénomène revient souvent à s'intéresser à une grandeur particulière comme la consommation d'essence d'une voiture ou comme le prix de revient d'un produit chimique. Cette grandeur, consommation, prix ou rendement, dépend d'un grand nombre de variables. La consommation de la voiture est fonction de la vitesse du véhicule, de la puissance du moteur, de la manière de conduire, de la direction et de la force du vent, etc. Le prix du produit chimique dépend de la qualité des matières premières, des rendements des unités de production, des spécifications imposées, des conditions de fabrication, etc.

Sous une forme mathématique, on peut écrire que la grandeur d'intérêt, y, que nous appellerons également réponse par la suite, est une fonction de plusieurs variables xi (variables que nous appellerons aussi facteurs par la suite).

On a : y = f(x1, x2, x3,…, xk). L'étude du phénomène se ramène à déterminer la fonction f(.....) qui lie la réponse y aux différents facteurs x1 , x2 , …, xk. Pour approfondir cette approche il faut introduire quelques notions particulières et une terminologie spécifique aux plans d'expériences.

I.5. Terminologie

La grandeur d'intérêt, qui est généralement notée y, porte le nom de réponse. Les variables qui peuvent modifier la réponse sont appelées facteurs. On parle donc des facteurs qui influent sur une réponse. Les termes facteur et réponse sont universellement employés dans le domaine des plans d'expériences.

- **Les différents types de facteurs**

La construction des plans et l'interprétation des résultats dépendent en grande partie des types de facteurs rencontrés dans l'étude. On distingue plusieurs types de facteurs. Nous retiendrons les types de facteurs suivants : les facteurs continus, les facteurs discrets, les facteurs ordonnables, les facteurs booléens.

Facteurs continus

La pression est un exemple de facteur continu. Dans un intervalle de pression donné, on peut choisir toutes les valeurs possibles. Il en est de même d'une longueur, d'une concentration ou d'une température. Les valeurs prises par les facteurs continus sont donc représentées par des nombres continus.

Facteurs discrets

Au contraire, les facteurs discrets ne peuvent prendre que des valeurs particulières. Ces valeurs ne sont pas forcément numériques : on peut représenter un facteur discret par un nom, une lettre, une propriété ou même par un nombre qui n'a alors en soi aucune valeur numérique mais qu'une signification de repère. Par exemple, on peut s'intéresser aux couleurs d'un produit : bleu, rouge et jaune sont des facteurs discrets.

Facteurs ordonnables

Il s'agit de facteurs discrets que l'on peut mettre dans un ordre logique. Par exemple, grand, moyen, petit, ou encore premier, deuxième, troisième et quatrième.

Facteurs booléens

Les facteurs booléens sont des facteurs discrets qui ne peuvent prendre que deux valeurs : haut ou bas, ouvert ou fermé, blanc ou noir, etc.

I.6. Modélisation mathématique

En l'absence de toute information sur la fonction qui lie la réponse aux facteurs, on se donne a priori une loi d'évolution dont la formulation la plus générale est la suivante :

$$y = f(x1, x2, x3, \ldots, xn)$$

Cette fonction est trop générale et il est d'usage d'en prendre un développement limité de Taylor-Mac Laurin, c'est-à-dire une approximation. Si les dérivées peuvent être considérées comme des constantes, le développement précédent prend la forme d'un polynôme de degré plus ou moins élevé :

$$y = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \cdots + \sum a_{ii} x_i^2 + a_{ij\ldots z} x_i x_j \ldots x_z$$

Où :

– y est la grandeur à laquelle s'intéresse l'expérimentateur ; c'est la réponse ou la grandeur d'intérêt,

– x_i : représente un niveau du facteur i,

– x_j : représente un niveau du facteur j,

– a_0, a_i, a_{ij}, a_{ii} sont les coefficients du polynôme.

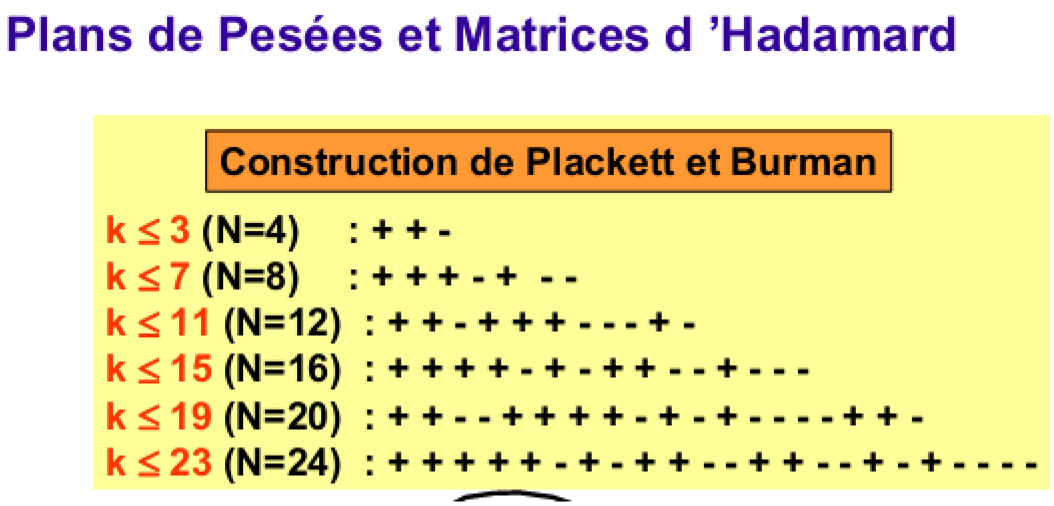
Ce modèle est appelé le modèle a priori ou le modèle postulé.

I.7. *Criblage d'un grand nombre de facteurs*

I.7.1. Plan de criblage de Plackett-Burman (matrice de Hadamard)

Lorsque vous avez besoin d'examiner un grand nombre de facteurs pour identifier ceux qui sont potentiellement importants (c'est-à-dire ceux qui sont liés à la variable dépendante étudiée), vous souhaitez pouvoir utiliser un plan vous permettant de tester le plus grand nombre de facteurs d'effets principaux avec le plus petit nombre d'essais, c'est-à-dire construire un plan de résolution avec le plus petit nombre d'essais possible. Une manière de concevoir ces expériences consiste à confondre toutes les interactions avec de "nouveaux" effets principaux.

Les facteurs ajoutés sont créés en mettant en équation (voir ci-dessous), les "nouveaux" facteurs avec les interactions d'un plan factoriel complet, ces plans auront toujours 2^k essais (par exemple, 4, 8, 16, 32, et ainsi de suite). Plackett et Burman (1946) ont montré comment un plan factoriel complet pouvait être fractionné de diverses manières, pour produire des plans saturés où le nombre d'essais est un multiple de 4, et non une puissance de 2. Ces plans sont également parfois appelés plans de la matrice de Hadamard. De toute évidence, vous n'êtes pas obligé(e) d'utiliser tous les facteurs disponibles dans ces plans, et en fait, vous pouvez parfois souhaiter générer un plan saturé pour un facteur de plus que ce que vous vous cherchez à tester. Vous pouvez ainsi estimer la dispersion de l'erreur aléatoire, et tester la significativité statistique des paramètres estimés.



I.8. Etude de l'influence des facteurs

I.8.1. Plans Factoriels Complets à deux niveaux « 2^k »

Le plan d'expériences le plus utilisé est le plan factoriel complet à deux niveaux noté 2^k ; la lettre **k** représente le nombre de facteurs à étudier, le chiffre 2 signifie que chaque facteur prend deux niveaux, un niveau bas désignant la borne inférieure retenue par l'expérimentateur et un niveau haut, indiquant la borne supérieure. Cette stratégie de choix des valeurs extrêmes pour chaque paramètre conduit à une meilleure estimation des effets principaux et des interactions des variables considérées.

I.8.2. Plan à deux facteurs

Pour deux facteurs, le domaine d'étude est un carré Le modèle mathématique postulé est un modèle du premier degré par rapport à chaque facteur :

$$y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + e$$

y : la réponse

x_i : représente le niveau attribué au facteur i.

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

a_1 : L'effet (ou effet principal) du facteur 1.

a_2 : L'effet (ou effet principal) du facteur 2.

a_{12} : Interaction entre les facteurs1 et 2.

e : l'écart.

Exemple 1 : «Amélioration du rendement»

Description de l'étude

Un industriel cherche à augmenter le rendement de sa fabrication. Il prépare un médicament à partir de plantes naturelles et cherche à améliorer le rendement d'extraction du principe actif. L'extraction est effectuée en présence de chlorure de sodium dont la concentration est de 50 grammes par litre et à une température de 70°C. L'industriel décide d'étudier ces deux facteurs et de les faire varier autour des consignes normales de fonctionnement. D'où les facteurs et le domaine d'étude :

- ✓ Facteur 1 : concentration en chlorure de sodium entre 40 et 60 grammes.
- ✓ Facteur 2 : température entre 60°C et 80°C.

La réponse est la masse de produit actif fabriqué.

- ✓ **Représentation d'une étude sous forme de tableau :**

Les représentations géométriques sont commodes et très parlantes mais dès que le nombre de facteurs est supérieur à trois, elles ne peuvent plus être employées. Pour les espaces multidimensionnels, on adopte une représentation en forme de tableau. Pour montrer la correspondance entre les deux représentations, géométrique et tableau, nous allons expliquer la construction du tableau rassemblant les expériences du plan 2^2 associé au tableau II.1.

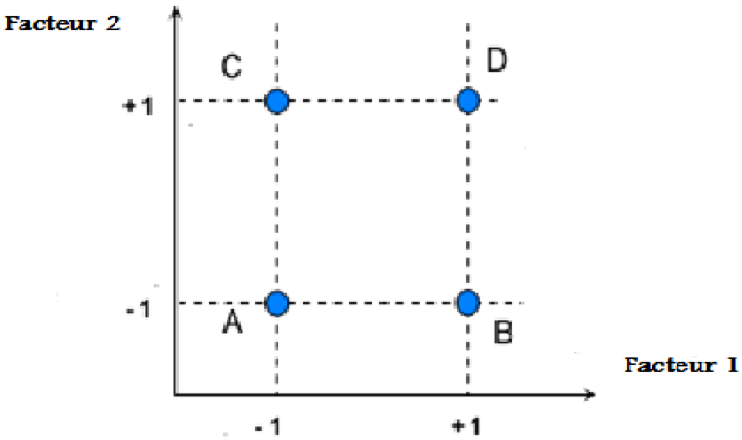


Figure I.6 : Les meilleurs emplacements des points expérimentaux sont les sommets

L'industriel exécute un plan factoriel complet 2^2. Ce plan est définit par Tableau 3.

Les résultats expérimentaux sont consignés dans la colonne "Réponse".

Tableau I.2. Tableau d'expérimentation (unités courantes).

N° Essai	Facteur (1)	Facteur (2)
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1

Réponse
y_1
y_2
y_3
y_4

Niveau -1	40 g/L	60 °C
Niveau +1	60 g/L	80 °C

N° Essai	Facteur (1)	Facteur (2)
1 (A)	-1	-1
2 (B)	+1	-1
3 (C)	-1	+1
4 (D)	+1	+1

Réponse
$y_1 = 115$
$y_2 = 185$
$y_3 = 104$
$y_4 = 156$

Ce tableau comprend trois colonnes, la première identifie les essais, la seconde et la troisième indiquent les coordonnées des points d'expériences. L'essai n°1 est celui pour lequel les deux facteurs étudiés sont aux niveaux bas, 40 g/L °C (ou - 1 en unités codées) et 60 °C (ou - 1 en en unités codées). Cet essai n°1 correspond au point A de la Figure I.6. L'essai n°2 est celui pour lequel le premier facteur est fixé au niveau haut, 60 g/L (ou +1 en unités codées) et le second facteur est fixé au niveau bas : 60 °C (ou - 1 en unités codées). Cet essai n°2 correspond au point B. Ce tableau s'appelle **« Tableau d'expérimentation »** s'il est construit avec les unités physiques habituelles (Tableau III.1) et **« Plan d'expérience »** s'il emploie les unités codées (Tableau III.2). Dans ce dernier cas, on rappelle la signification des unités codées en indiquant, pour chaque facteur, leurs valeurs en unités physiques habituelles en bas du tableau.

Les représentations géométriques et les représentations par tableaux sont équivalentes. Les tableaux (ou matrices) présentent l'avantage de pouvoir être utilisés quel que soit le nombre de facteurs, c'est-à-dire quel que soit le nombre de dimensions de l'espace expérimental.

Coordonnées codées :

Dans la théorie des plans d'expériences, les coordonnées réduites notées X_i, sont présentées de la même manière quel que soit le paramètre étudié. Elles sont définies et calculées à l'aide de relation (1) :

$$X_i = \frac{V_i - \overline{V}_i}{\Delta V_i} \tag{I.1}$$

Où :

V_i : Valeur réelle du paramètre « i ».

$\overline{V_i}$: La valeur réelle moyenne des deux bornes du paramètre « i » ; elle est donnée par :

$$\overline{V}_i = \frac{V_{iSup} + V_{iInf}}{2} \qquad (I.2)$$

ΔV_i : représente l'écart moyen entre les deux niveaux du paramètre «i » ; elle est définie par la relation (3) :

$$\Delta V_i = \frac{V_{iSup} - V_{iInf}}{2} \qquad (I.3)$$

- ✓ **Calcul des coefficients :**

Les quatre points d'expériences apportent quatre équations.

$$y_1 = a_0 + a_1 x_{1,-1} + a_2 x_{2,-1} + a_{12} x_{1,-1} x_{2,-1} + e_1$$
$$y_2 = a_0 + a_1 x_{1,+1} + a_2 x_{2,-1} + a_{12} x_{1,+1} x_{2,-1} + e_2$$
$$y_3 = a_0 + a_1 x_{1,-1} + a_2 x_{2,+1} + a_{12} x_{1,-1} x_{2,+1} + e_3$$
$$y_4 = a_0 + a_1 x_{1,+1} + a_2 x_{2,+1} + a_{12} x_{1,+1} x_{2,+1} + e_4$$

La résolution de ce système donne la valeur des coefficients :

$$\hat{a}_0 = \frac{1}{4}[+y_1 + y_2 + y_3 + y_4] \quad \{6\ a\}$$
$$\hat{a}_1 = \frac{1}{4}[-y_1 + y_2 - y_3 + y_4] \quad \{6\ b\}$$
$$\hat{a}_2 = \frac{1}{4}[-y_1 - y_2 + y_3 + y_4] \quad \{6\ c\}$$
$$\hat{a}_{12} = \frac{1}{4}[+y_1 - y_2 - y_3 + y_4] \quad \{6\ d\}$$

Connaissant les coefficients, on peut écrire le modèle de régression qui servira à faire des prévisions

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$$

Signification de a_0 :

Si l'on donne à x_1 et à x_2 la valeur zéro, on définit le centre du domaine d'étude. La relation (1) devient alors : $y = a_0$

Le coefficient a_0 est la valeur calculée de la réponse au centre du domaine d'étude.

- ✓ **Signification de a_1 :**

Plaçons nous maintenant au niveau moyen du facteur 2, pour cela donnons la valeur zéro à x_2. :

$$y = a_0 + a_1 x_1$$

Cette relation permet de tracer l'évolution de la réponse prédite dans un plan de coupe $x_2 = 0$ (figure I.7). L'effet du facteur 1 apparaît comme la variation de la réponse quand on passe du niveau zéro au niveau haut du facteur 1.

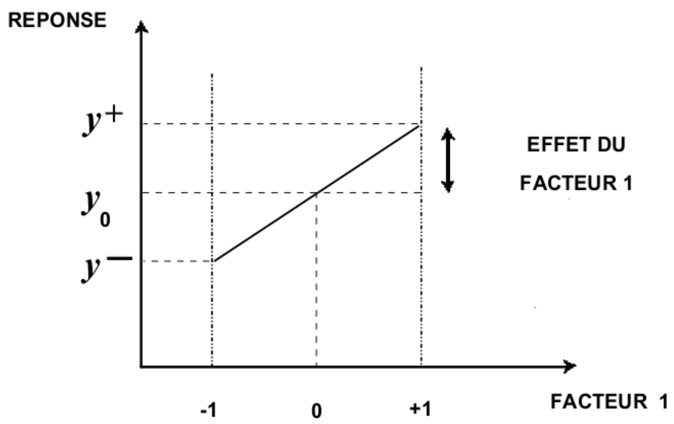


Figure I.7 : Effet du facteur 1

Signification de a_{12} :

$$\hat{a}_{12} = \frac{1}{2}\left[\frac{1}{2}(y_4 - y_3) - \frac{1}{2}(y_2 - y_1)\right] = \frac{1}{2}\left[+ef^+ - ef^-\right]$$

L'interaction apparaît comme la demi différence entre l'effet du facteur 1 au niveau haut du facteur 2 (effet noté ef^+) et l'effet du facteur 1 au niveau bas du facteur 2 (effet noté ef^-). Elle traduit une variation de l'effet d'un facteur en fonction du niveau d'un autre facteur. L'interaction entre les deux facteurs 1 et 2 est une interaction d'ordre 2.

Interprétation :

Les valeurs des coefficients sont :

Tableau I.3. Tableau des effets

Moyenne a_0	140
Effet 1 a_1	30,5
Effet 2 a_2	-10
Interaction a_{12}	-4.5