Les Plans d'Expériences

Sommaire:

- Introduction
- Rappels Statistiques
- Les plans complets
- Les plans factoriels 2k
- Les plans fractionnaires 2k-p
- Les plans à surface de réponse
- Les plans de mélanges

1.Présentation des plans d'expériences

1.1.Aperçu historique

Années 20 : FISHER : Application des DOE en agronomie (réduction considérable de nombre d'expériences à effectuer sans diminuer la pertinence de la campagne d'essais) ;

Années 60 : TAGUSHI : reprise de la méthode en combinant les techniques de l'ingénierie et des statistiques, Tagushi a réussi à améliorer rapidement les coûts et la qualité des produits ;

Années 70 : Expansion de l'utilisation des PE dans l'industrie au Japon

Années 80 : Introduction des PE aux Etats-Unis, au Canada et en Grande Bretagne, et enfin en France et en Espagne;

Développement de logiciels spécialisés en DOE et application en recherche et développement.

1.2. Principe des P.E

A la différence de la méthode classique on fait varier les niveaux de tous les facteurs à la fois à chaque expérience,
Avantages :

- Diminution du nombre des essais,
- Nombre de facteurs étudiés très grand,
- Détection des interactions entre facteurs,
- Détection des optimaux,
- Détermination des résultats avec une bonne précision,
- Optimisation des résultats,
- Modélisation aisée des résultats.

Les PE permettent:

- l'étude de nombreux facteurs tout en maintenant le nombre des essais faible,
- et de minimiser les erreurs.

1.2. Principe des P.E

Qu'est-ce qu'un plan d'expériences?

Un plan d'expériences est une suite d'essais organisée de manière à déterminer avec un minimum d'essais et avec un maximum de précision l'influence de plusieurs paramètres (facteurs) sur une ou plusieurs réponses

Intérêt de la méthode

- Efficacité supérieure à toute autres suite d'essais de même taille non planifiée,
- Pas de « un facteur à la fois »,
- Quantifie et hiérarchise l'influence des facteurs et de leur interaction
- Estimation à priori de la durée et des coûts des essais,
- Modélisation et optimisation des résultats,
- Conclusions fiables et reproductibles,
- Interprétation aisée des résultats.

1.3. Objectifs des plans d'expériences :

Qu'est-ce qu'un plan d'expériences?

Un plan d'expériences est une suite d'essais organisée de manière à déterminer avec un minimum d'essais et avec un maximum de précision l'influence de plusieurs paramètres (facteurs) sur une ou plusieurs réponses

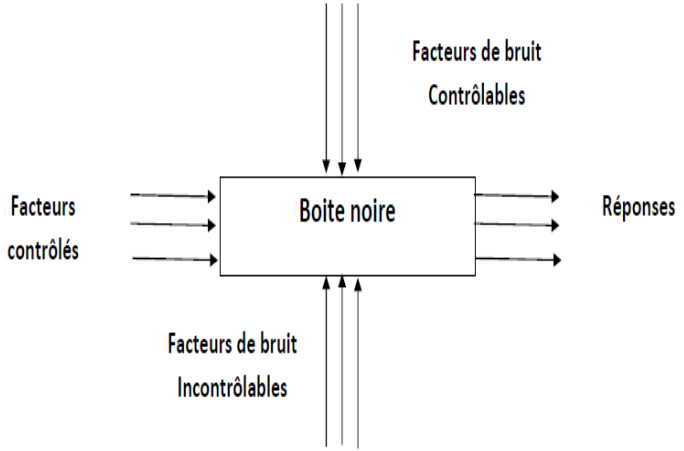


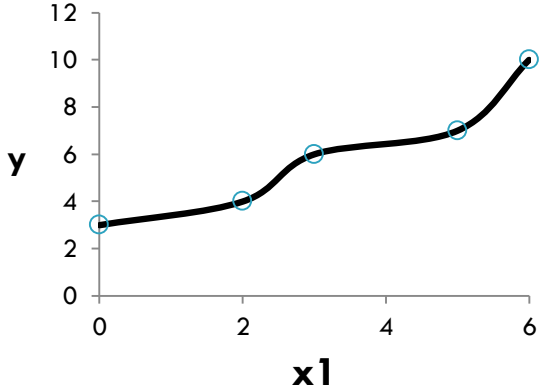
Fig.1 : système de boite noire

- Comprendre comment réagit un système en fonction des facteurs susceptibles de le modifier.
- Visualiser cette évolution en mesurant une réponse ,
- Établir des relations de cause à effet entre les réponses et les facteurs.

1.3. Objectifs des plans d'expériences :

Stratégie classique: étude d'un facteur à la fois en fixant les valeurs des autres facteurs.

Réponse $y = f(x_1)$: les autres variables fixes.



Exple:

- 7 facteurs et 5 niveaux; Nbre d'essais=5^7= 78125
- 7 facteurs et 3 niveaux; Nbre d'essais=3^7= 2187
- 7 facteurs et 2 niveaux; Nbre d'essais=2^7= 128

1.4. Terminologie :
1.4.1 Facteur:

• Facteur: c'est la variable que l'on désire étudier.

• Domaine de variation: ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut,

• Niveau (modalité): nombre de valeurs que peut prendre un facteur.

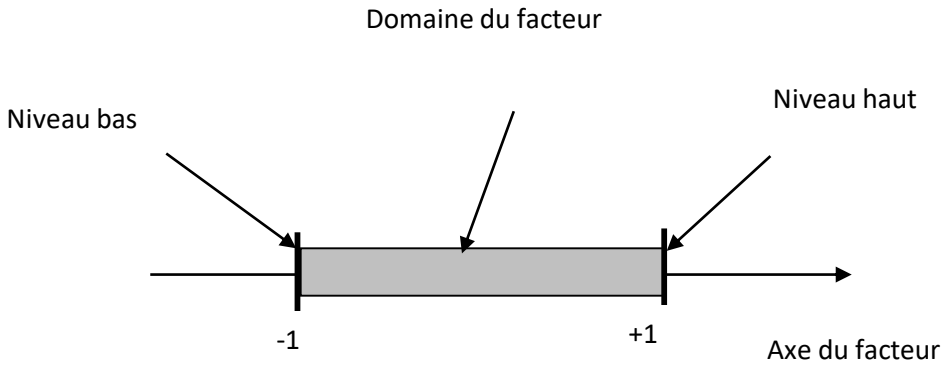


Figure 1 - Domaine de variation du facteur

1.4.1 Facteur:

1. **Facteur continu:**
   exple: longueur, concentration, température
2. **Facteur discret:**
   exple: longueur,
3. **Facteur ordonnable:**
   exple: petit, moyen grand,
4. **Facteur booléen:**
   exple: haut-bas, ouvert-fermé, Admis-ajourné

1.4.2 Niveau (modalité):

Niveau (modalité): c'est le nombre de valeurs que peut prendre un facteur.

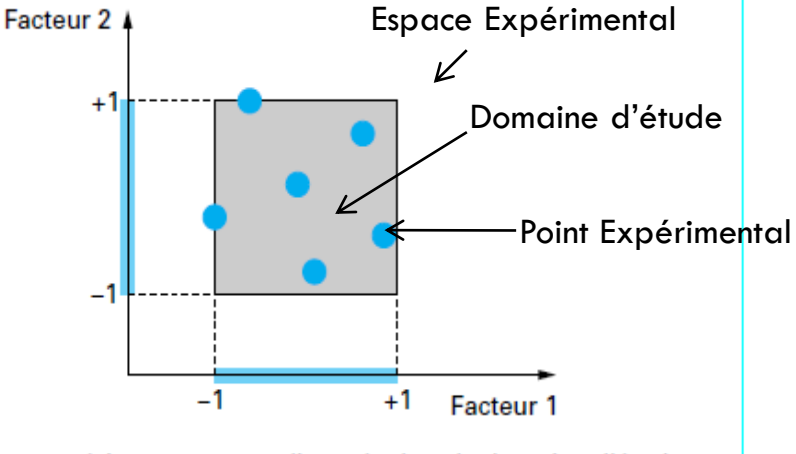
Choix des modalités des facteurs d'entrée:

- 2 niveaux: raison économique,
- 3 niveaux si courbure,
- 5 niveaux au maximum.

L'intervalle de variation: ni trop rapproché ni trop large.

1.4.3 Espace expérimental:

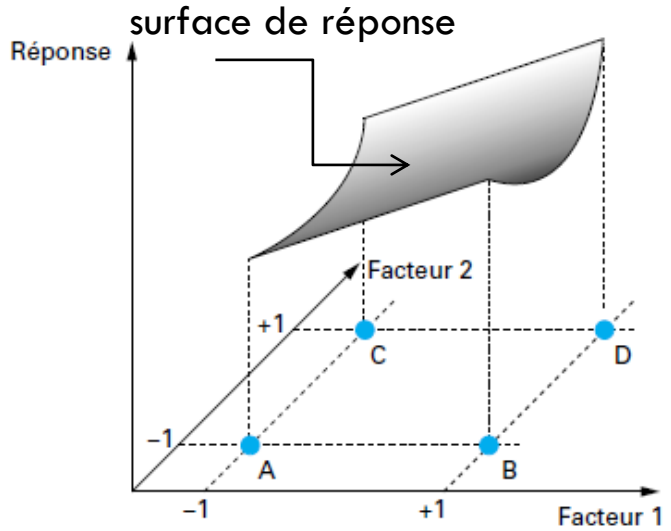
- Espace expérimental,
- Point expérimental,
- Domaine d'étude.



Les points expérimentaux sont disposés dans le domaine d'étude.

1.4.4 Surface de réponse:

Ensemble des réponses de tous les points du domaine d'étude qui se localisent sur une surface.



En général, on ne connaît que quelques réponses, celles qui correspondent aux points expérimentaux A, B, C, D retenus par l'expérimentateur.
On interpole les réponses connues pour obtenir la surface de réponse.

1.4.5 Variables centrées et réduites :

$$X_j = \frac{Z_j - Z_j^0}{\Delta Z_j}$$

$$Z_j^0 = \frac{Z_{max} + Z_{min}}{2}$$

$$\Delta Z_j = \frac{Z_{max} - Z_{min}}{2}$$

$X_1, X_2, \ldots\ldots, X_k$ Variables centrées réduites ou variables codées ;

$Z_1, Z_2, \ldots\ldots, Z_k$ Variables réelles ;

$Z_1^0, Z_2^0, \ldots\ldots, Z_k^0$ Variables réelles au centre du plan

ΔZ_j Intervalle de variation selon l'axe des Zj

Z_{min} Valeur minimale de la variable réelle

Z_{max} Valeur maximale de la variable réelle

1.4.6 Effet d'un facteur :

L'effet d'un facteur sur la réponse s'obtient en comparant les valeurs prises par **la réponse** quand ***le facteur*** passe du niveau (**-1**) au niveau (**+1**).

Soient y_1 et y_2 ces valeurs (Fig):

Nous distinguons :

- l'effet global: $(y_2 - y_1)$
- l'effet moyen: $\frac{(y_2 - y_1)}{2}$

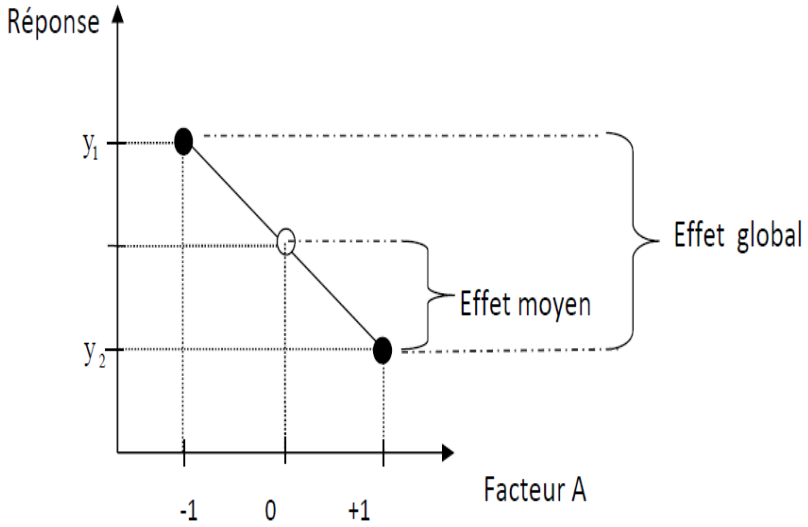


Fig.7 : effet d'un facteur

1.4.7 Effet d'interaction :

Il y a interaction entre deux facteurs **A** et **B** si l'effet de **A** sur la réponse dépend du niveau de **B** ou inversement (Figure 8)

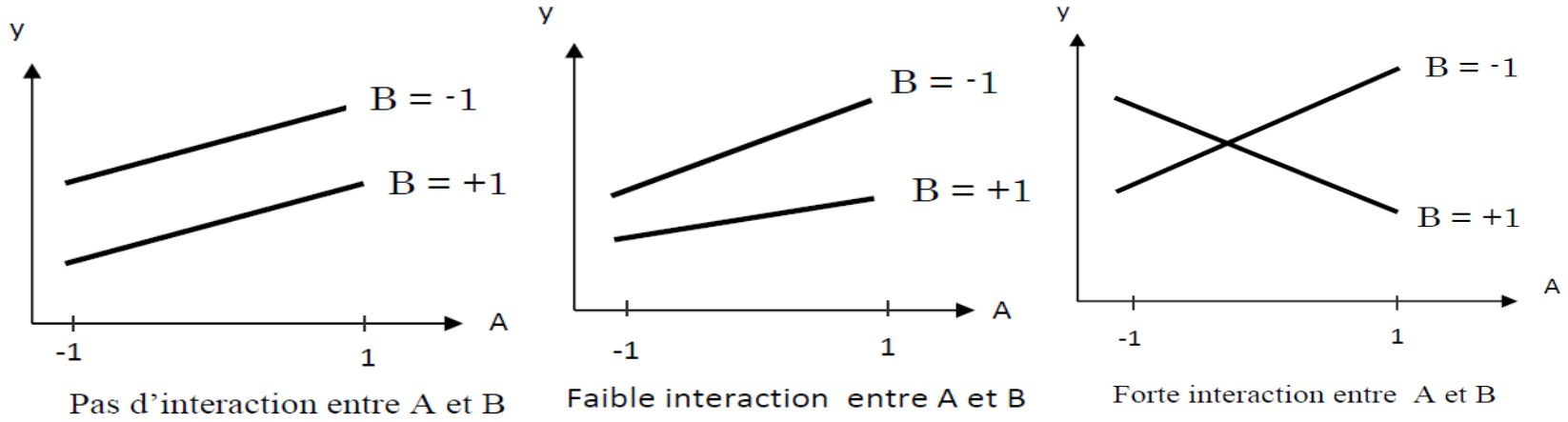
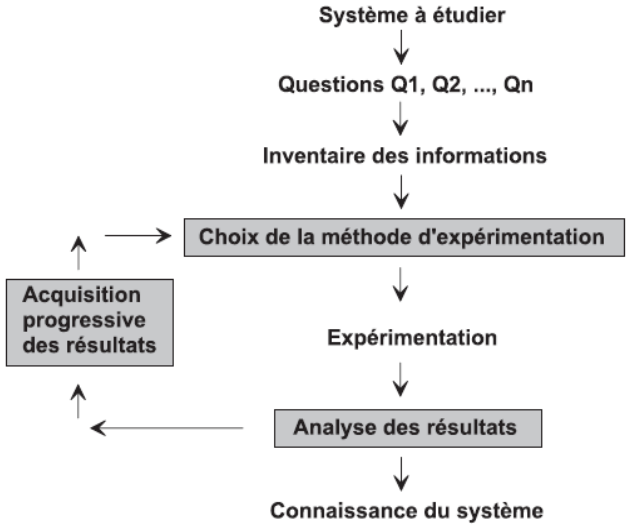


Fig.8 : Effets d'interactions

5. Démarche méthodologique d'un plan d'expérience:

- Choix des facteurs d'entrée, des modalités et de l'intervalle de variation(ni trop rapproché ni trop large),
- Choix du plan d'expérimentation
  - Plan de criblage
  - Plan de modélisation
    - PFC (FFD)
    - PFF
    - P hybride pour surface de réponse
  - Plan de mélange: facteurs avec contraintes.
- Analyse des résultats
- Validation du modèle si le problème est résolu, sinon passer à l'acquisition progressive des connaissances.



5.1 Notion de modélisation mathématique:

y=f(x1,x2, …, xn) + erreur (modèle à priori)……. (1)

- y grandeur d'intérêt ou réponse et xi facteurs ou variables d'entrée
- Erreur=$\Delta+\varepsilon$ (où: Δ: erreur d'ajustement (lack of fit), ε: erreur aléatoire (pure error)).
- Polynôme (développement de Taylor)

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

Modèle à priori (postulé) : polynôme

$$\hat{y} = a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \cdots + e \quad (3)$$

a_0, a_i, a_{ij}, a_{ii} Coefficients du modèle à priori:

5.2 Notion de modélisation mathématique:

Calcul des coefficients aj: méthode des moindres carrés:

$$f = \sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2 = min \qquad (4)$$

N: nbre d'essais (de relations) ≥p: nbre de coefficients (inconnues).

$\hat{y}_i$: réponse calculée avec le modèle postulé

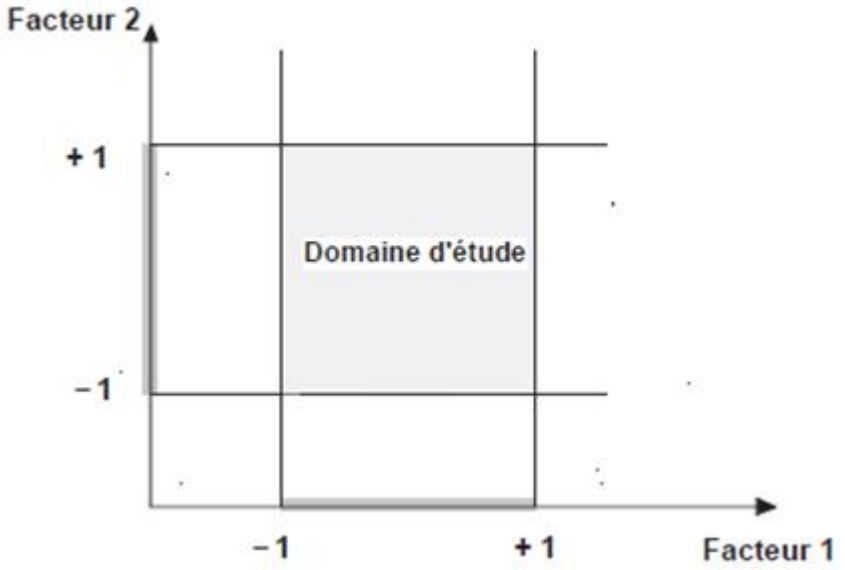
Donc: $$f = \sum_{i=1}^{i=n} \left[y_i - \left(a_0 + \sum a_i x_i + \sum a_{ij} x_i x_j + \sum a_{ii} x_i^2 + \cdots \right) \right] \quad (5)$$

Pour que: $f(a_0, a_i, a_{ij}, a_{ii}) = min$ (6) il faut que: $\frac{\partial f}{\partial a_i} = 0$ (7)

5.1 Matrice des expériences :

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

	Facteur 1	Facteur 2
Niveau bas	Z_{1min}	Z_{2min}
Niveau haut	Z_{1max}	Z_{2max}



5.1 Matrice des expériences :

Tableau 1 : Matrice d'expériences

	Facteur A	Facteur B	Facteur C	Réponses
N° Essais	x_1	x_2	x_3	y
1	-1	-1	-1	y_1
2	1	-1	-1	y_2
3	-1	1	-1	y_3
4	1	1	-1	y_4
5	-1	-1	1	y_5
6	1	-1	1	y_6
7	-1	1	1	y_7
8	1	1	1	y_8

5.1 Matrice des expériences :

Tableau 2 : Matrice des effets.

		Facteur A	Facteur B	Facteur C	Interactions Entre A, B et C				Réponses
N° Essais	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2\ x_3$	y
1	1	-1	-1	-1	1	1	1	-1	y_1
2	1	1	-1	-1	-1	-1	1	-1	y_2
3	1	-1	1	-1	-1	1	-1	1	y_3
4	1	1	1	-1	1	-1	-1	-1	y_4
5	1	-1	-1	1	1	-1	-1	1	y_5
6	1	1	-1	1	-1	1	-1	-1	y_6
7	1	-1	1	1	-1	-1	1	-1	y_7
8	1	1	1	1	1	1	1	1	y_8

5.1 Matrice des expériences :

Tableau 2 : Matrice des effets servant au calcul (ajout de la colonne X0).

		Facteur A	Facteur B	Facteur C	Interactions Entre A, B et C				Réponses
N° Essais	x_0	x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	$x_1x_2x_3$	y
1	1	-1	-1	-1	1	1	1	-1	y_1
2	1	1	-1	-1	-1	-1	1	-1	y_2
3	1	-1	1	-1	-1	1	-1	1	y_3
4	1	1	1	-1	1	-1	-1	-1	y_4
5	1	-1	-1	1	1	-1	-1	1	y_5
6	1	1	-1	1	-1	1	-1	-1	y_6
7	1	-1	1	1	-1	-1	1	-1	y_7
8	1	1	1	1	1	1	1	1	y_8

5.3. Analyse de la régression sous forme matricielle

$$\hat{y} = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \quad (8)$$

$$Y = A \cdot X \quad (9)$$

Y : vecteur de réponses ;
X : matrice des expériences
A : vecteur des coefficients.

Donc, le modèle matricielle peut s'écrire :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} x_{01} & x_{11} & \cdots & x_{k1} \\ x_{02} & x_{12} & \cdots & x_{k2} \\ \vdots & \vdots & \cdots & \vdots \\ x_{0n} & x_{1n} & \cdots & x_{kn} \end{pmatrix} \quad (10)$$

le modèle matricielle permet de calculer les coefficients :

$$a_0, a_1, a_2, \ldots, a_n$$

5.3. Analyse de la régression sous forme matricielle

L'équation (10) peut s'écrire:

$$({}^{t}X.X)\cdot A = {}^{t}X.Y \qquad (11)$$

Matrice des variances (d'information):

$${}^{t}X\cdot X = \begin{pmatrix} \sum x_{0i}^2 & \sum x_{0i}x_{1i} & \cdots & \sum x_{0i}x_{ki} \\ \sum x_{1i}x_{0i} & \sum x_{1i}^2 & \cdots & \sum x_{1i}x_{ki} \\ \vdots & \vdots & & \vdots \\ \sum x_{ki}x_{0i} & \sum x_{ki}x_{1i} & \cdots & \sum x_{ki}^2 \end{pmatrix} \qquad (12)$$

Matrice colonne:

$${}^{t}X\cdot Y = \begin{pmatrix} \sum x_{0i}y_i \\ \sum x_{1i}y_i \\ \vdots \\ \sum x_{ki}y_i \end{pmatrix} \qquad (11)$$

on calcule :

$$({}^{t}X\cdot X)^{-1} \qquad (12)$$

on obtient :

$$\hat{A} = ({}^{t}X\cdot X)^{-1}\cdot {}^{t}X\cdot Y \qquad (13)$$

6. Analyse statistique :

6.1. Vérification de la signification des coefficients:

Cas où chaque essai est répété m fois :

-La moyenne des résultats des essais:

$$\overline{y_i} = \frac{\sum_{u=1}^{m} y_{iu}}{m} \quad (14) \quad i = 1,2,....,N$$

- les variances de sondage :

$$s_i^2 = \frac{\sum_{u=1}^{m} \left(y_{iu} - \overline{y_i}\right)^2}{m-1} \quad (15) \quad i = 1,2,....,N$$

où: m : nbre de répétitions /essai,

N : nbre d'expériences.

Si les variances de sondage sont homogènes, on calcule :

- Les variances de reproductibilité:

$$s_{rep}^2 = \frac{\sum_{i=1}^{m} s_i^2}{N} \quad (16)$$

6. Analyse statistique :

6.1. Vérification de la signification des coefficients:

Cas où l'essai au centre est répété n_0 fois :

- dans ce cas, la variance de reproductibilité est estimée par celle calculée au centre du domaine expérimental

$$s_{rep}^2 = \frac{\sum_{i=1}^{n0} \left(y_i - \overline{y_0}\right)^2}{n_0 - 1} \qquad (17)$$

où: $\nu = (n_0 - 1)$ ddl

$$\overline{y_0} = \frac{\sum_{i=1}^{n0} y_i}{n_0} \qquad (18)$$

moyenne des mesures au centre

6. Analyse statistique :

6.1. Vérification de la signification des coefficients:

Test de Student :

- tester l'hypothèse:

$$H_0 = ''aj = 0''$$

contre:

$$H_1 = ''aj \neq 0''$$

- calcul de:

$$t_j = \frac{|a_j|}{s_{aj}}$$

où: a_j j[ième] coeff de l'équation de régression ,

S_{aj} écart quadratique moyen (modèle 1[er] degré)

$$s_{aj} = \frac{s_{rep}}{\sqrt{N}}$$

- Comparaison de $t_j > t_\alpha (\nu)$ (test bilateral)

$t_j > t_\alpha$ (f), H_0 rejeté au risque α

$t_j < t_\alpha$ (f), H_0 accepté au risque α

- Si H_0 est acceptée : aj~0 et la variable associée n'a pas d'influence sur Y

6.2. Validation du modèle :

a- Recherche de biais du modèle : (test de Fisher Snedecor)

- Calcul de s^2_{res}

$$s^2_{res} = \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N - l}$$

Où: $(N - l)$: nbre de ddl (nbre de coefficients significatifs) à s^2_{rep}

- Calcul de F: $\mathbf{F = \frac{s^2_{res}}{s^2_{rep}}}$

-Comparaison avec Ftab: si $F < F(0{,}95, N - l, n_0 - 1)$

Absence de biais (erreurs aléatoires)

6.2. Validation du modèle :

b- Test de signification de la régression (test de Fischer):

$$F = \frac{\sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 / (l-1)}{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2 / (N-l)}$$

$$\bar{y} = \frac{\sum_{i=1}^{N} y_i}{N}$$

- Comparaison si: $F_{cal} \geq F(\alpha, l-1, N-l)$

- Conclusion: les variables du modèle ont un effet significatif sur Y et l'équation est adéquate.

6.2. Validation du modèle :

c- Coefficient de détermination:

$$R^2 = \frac{\sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \qquad R^2 = \frac{\text{Variation due à la régression}}{\text{Variation totale}}$$

d- Coefficient de détermination corrigé:

$$\overline{R}^2 = R^2 - (1 - R^2)\frac{(l-1)}{(N-l)}$$