

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MOULOU MAMMERI, TIZI-OUZOU



FACULTE DE SCIENCES
DEPARTEMENT DE CHIMIE

Fascicule de Travaux Pratiques de Modélisations et Optimisation par la Méthode des Plans d'expérience

Dr Hamida IBOUKHOULEF Epse BEKDA

2023-2024

Préface

Ce polycopié de TP est destiné aux étudiants en deuxième année Master Chimie et à l'ensemble des doctorants et chercheurs intéressés par la modélisation avec la méthode des plans d'expérience. L'enchaînement des TP est en accord avec le cours assuré en Master 2 chimie.

Les plans d'expérience est une suite d'expérience qui permettent d'organiser au mieux les essais qui accompagnent la recherche scientifique ou des études industrielles. Ils sont appliqués dans diverses disciplines où on cherche un lien entre une grandeur d'intérêt (Y) et des variables (X_i). Ces plans d'expérience appartiennent à une branche des statistiques appliquées, ils sont basés sur l'utilisation de méthodes scientifiques pour la planification, la réalisation, l'analyse et l'interprétation des données. Ces méthodes permettent de modifier simultanément les variables d'entrée (appelées facteurs) sur les variables de sortie (appelées réponses). En combinant plusieurs variables dans une étude au lieu de créer une étude pour chacune, le nombre d'essai requis sera considérablement réduit. Ce qui permet d'obtenir un maximum de renseignement avec un minimum d'expérience et par conséquent un coût réduit.

Pour faciliter les calculs et la lecture des résultats, la méthodologie des plans d'expérience nécessitent des logiciels scientifiques (JMP, Nemrodw, Modde6...). Ces derniers sont obtenus sous licence, pour la recherche scientifique et un seul poste de travail ! Pour contourner cette situation, nous avons proposé de travailler sur Excel. C'est pour cela, que ce polycopié débute par un TP d'introduction sur Excel. L'ensemble des commandes et des fonctions statistiques utiles pour la conception d'un plan d'expérience a été introduit.

Deux parties seront faites dans le deuxième TP. La première est consacrée à l'écriture d'une matrice de criblage, par la méthode d'HADAMARD, et au calcul des coefficients d'un modèle élaboré tandis que la deuxième concerne l'estimation des facteurs influents, par le test statistique de STUDENT, en passant par le calcul de la variance expérimentale.

L'objectif du troisième TP est d'identifier les facteurs influents en utilisant les outils d'aide à l'interprétation. Cette méthode est utilisée lorsque la variance expérimentale est

inconnue et le degré de liberté (ddl) est élevé. Les méthodes Pareto, Lenth et le Normal/Half plot seront développées.

Dans le quatrième TP, nous présentons, le calcul de la variance expérimentale en doublant le plan d'expérience élaboré avec une matrice d'HADAMARD. Cette méthode est utilisée lorsque le nombre d'expérience à réaliser (N) est égale au nombre des constante du modèle élaboré (P). En effet, l'estimateur ne peut être calculé ($N-P=0$). Le cinquième TP est consacré aux plans factoriels complets.

Toutes les notions développées dans ces TP sont enseignées dans le cours. Néanmoins, chaque TP commence par une introduction, suivie de toutes les étapes nécessaires au traitement du problème posé. Un exemple d'application est traité durant la séance de TP et des exemples sont proposés aux étudiants et à remettre pour correction. Ces exemples ont été tirés des travaux de recherche publiés et cités dans les références.

Liste des abréviations

\bar{x} : La moyenne,
 σ : L'écart type,
 σ^2 : La variance,
IC : L'intervalle de confiance,
Abs : La valeur absolue,
(-1) : Le niveau bas du facteur,
(+1) : Le niveau haut du facteur,
 X_i : La variable codée,
 X^T : La matrice transposée,
N : Le nombre d'expérience,
 I_d : La matrice identité,
N : Le nombre d'expériences à réaliser,
p : Le nombre de coefficients du modèle,
 e_i : Les résidus,
 t_{exp} : Le test de Student,
ddl : Degré de liberté,
 t_{crit} : t critique donnée par la table de Student,
 S^2 : L'estimateur,
 b_j : Le coefficient du modèle,
IP : L'indice de Pareto,
IPc : L'indice de Pareto cumulé,
PSE : Le pseudo écart-type,
P : La probabilité
J : Le rang

TP N°1 : Notions de base sur Excel

I.	L'essentiel sur EXCEL	4
I.1.	Introduction.....	4
I.2.	Entrer des valeurs et des formules	5
I.3.	Format d'une cellule	5
I.4.	Utiliser des fonctions prédéfinies	5
I.5.	Application 1 : Tracer une courbe	6
II.	Statistique d'une population	7
II.1.	Moyenne.....	7
II.2.	Variance.....	7
II.3.	Ecart- type	7
II.4.	Intervalle de confiance	7
II.5.	Médiane	8
III.	Application 2 : calcul statistique.....	8
	Corrigé des applications.....	9

TP N° 2 : Criblage de facteurs par la méthode d'Hadamard

I.	Criblage de facteurs	12
II.	Estimation des facteurs influents.....	13
III.	Applications	14
	Corrigé des applications.....	15

TP 3: Identification des facteurs influents en utilisant les outils d'aide à l'interprétation

I.	Approche de Pareto	19
II.	Approche de Lenth	19
III.	Normal Plot (NP) ou Half Normal Plot (HNP).....	20
IV.	Application.....	21
	Corrigé des applications.....	22

TP N° 4 : Identification des facteurs influents -doubler un plan d'expérience-

I.	Criblage de facteurs-doubler le plan d'expérience	26
II.	Application	27
	Corrigé des applications.....	29

TP N° 5 : Plan factoriel complet 2^k

I.	Plan factoriel complet.....	31
II.	Construction de la matrice.....	31
III.	Application.....	31
	Corrigé des applications.....	32
	Comptes rendus.....	33

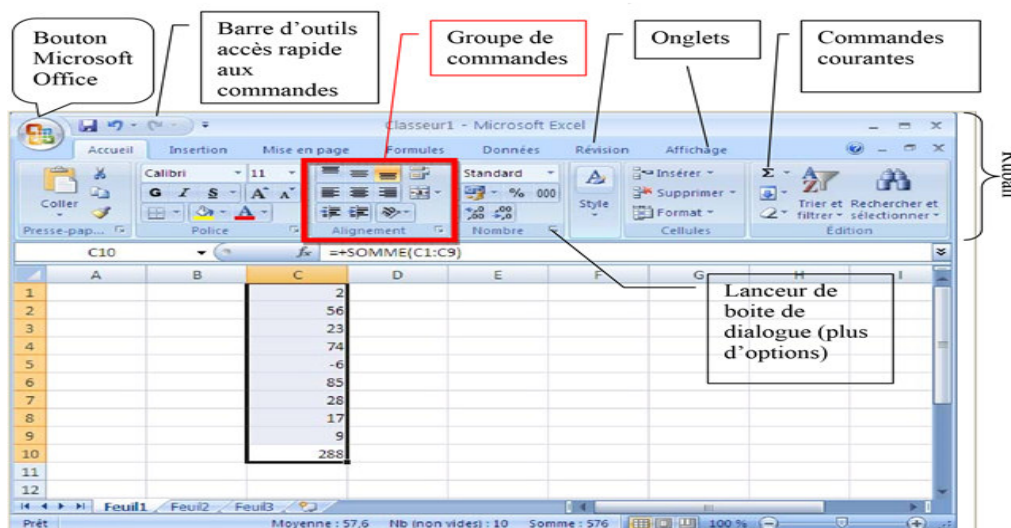
TP N°1 : NOTIONS DE BASE SUR EXCEL

- NOTIONS STATISTIQUES -

I. L'essentiel sur EXCEL

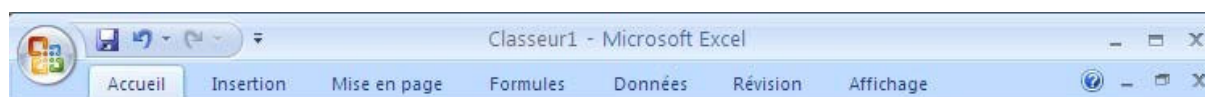
I.1. Introduction

Chaque document Excel est un classeur contenant plusieurs feuilles de calcul, indépendantes les unes des autres. Une **feuille de calcul** est constituée de lignes (numérotées à l'aide des chiffres) et des colonnes (étiquetées avec des lettres). On appelle **cellule** l'intersection d'une ligne et d'une colonne. Une cellule se désigne par sa référence, par exemple C2 = 56 (voir schéma). On appelle **cellule active** la cellule sélectionnée par le curseur, et dans laquelle on entre à la fois une formule (ou directement une valeur) et/ou éventuellement un commentaire.



Sur la copie d'écran, les cellules C1 à C10 sont sélectionnées. Parmi ces cellules, la cellule active C10 apparaît encadrée, les cellules sélectionnées sont sur fond bleuté.

Le ruban regroupe l'ensemble des fonctions spécifiques du logiciel Excel.



Avec le bouton Microsoft Office, on accède aux commandes permettant de créer un **Nouveau** document, **Ouvrir** un document déjà existant, **Fermer** un document, **Quitter** Excel, pour sauvegarder (**Enregistrer** et **Enregistrer sous...**), **Imprimer**...etc.

Les onglets regroupent plusieurs commandes. Chaque groupe présente les commandes courantes pour chacune des catégories (Groupe Police, Groupe Style...). En-dessous de chaque groupe se

trouve le lanceur de boîte de dialogue qui présentent l'ensemble des commandes et options, y compris celles qui sont moins courantes.

En bas de la fenêtre, on trouve les onglets de sélection d'une feuille de calcul. Avec un clic droit sur l'onglet correspondant, il est possible de changer les paramètres d'une feuille (en particulier son nom).



I.2. Entrer des valeurs et des formules

Dans la cellule active, il est possible d'entrer **soit directement une valeur** (nombre ou texte), **soit une formule** calculant une valeur.

Une formule commence toujours par le signe =, Elle comprend :

- Des valeurs
- Des opérateurs : somme (+), soustraction (-), multiplication (*), division (/) et puissance (^)...etc.
- Des références à d'autres cellules (lors d'un calcul).

Il est possible de recopier une formule dans une ou plusieurs autres cellules (soit avec les commandes copier/coller du menu édition, soit en sélectionnant la cellule à recopier pour la faire « glisser » vers les cellules cibles – le glissement s'obtient en se positionnant dans le coin inférieur droit de la cellule).

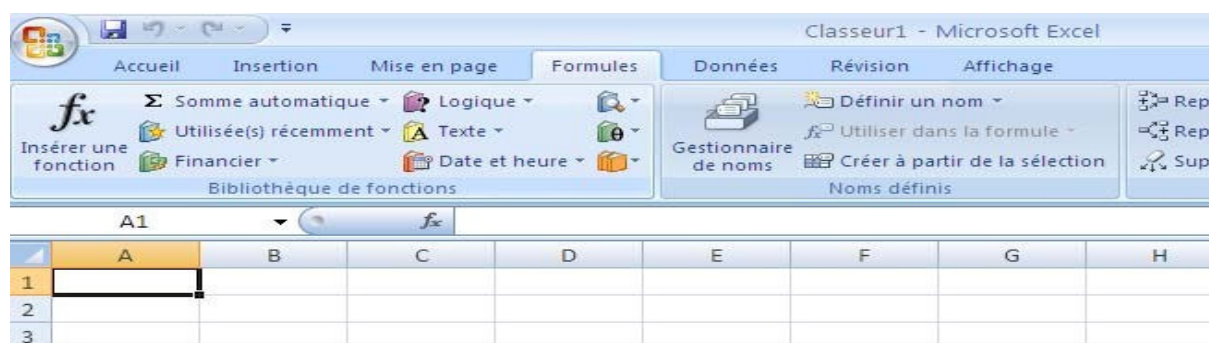
- Pour imposer l'ordre de calcul, il faut toujours utiliser les parenthèses.

I.3. Format d'une cellule

Dans une cellule, on peut placer soit une valeur ou une formule calculant une valeur. Un format précise la manière dont la valeur sera vue sur l'écran, mais n'affecte jamais la valeur. Également dans les groupes « Police », « Alignement » et « Cellules » se trouvent des commandes pour améliorer la présentation de la plage de cellules sélectionnées (nombre, alignement, police, bordure...etc.).

I.4. Utiliser des fonctions prédéfinies

Excel propose des fonctions de calcul prédéfinies accessibles à partir de l'onglet de fonctions :



Pour faire la somme de deux cellules, B1 et C1, on saisit dans la cellule D1 : « = B1+C1 ». Lors d'un calcul, la référence à une cellule est remplacée par la valeur qu'elle contient.

Les mêmes étapes sont appliquées pour la soustraction, la multiplication et la division.

Exemple 1 : Saisir le relevé de notes du semestre 1 en L1

- Ouvrir un classeur Excel, et sauvegarder ce fichier sous le : nom.xls. Nommer une feuille du classeur en lui attribuant le nom « Mes notes en L1 »,
- Saisir dans la colonne A l'ensemble des matières du semestre 1, en mettant un module par cellule,
- Dans les colonnes B et C, entrer les notes obtenues dans chaque matière (CC et EMD). Ne pas oublier de sauvegarder les éléments saisis,
- Le calcul de la moyenne (colonne D) se fait par la relation : $Moyenne\ 1 = examen * 0.67 + CC * 0.33$
- Saisir dans la colonne E le coefficient de chaque matière,
- Afin de calculer la moyenne 2, multiplier la moyenne de la matière mathématique 1 (moyenne 1) par son coefficient.
- Pour dupliquer la même commande, sélectionner la case de moyenne 2 de la matière mathématique 1 et faire glisser le curseur jusqu'à la dernière case (histoire des sciences),
- Faire la somme des résultats obtenus « = somme F2 :F9 » que vous diviserez par la somme des coefficients « =somme E2 :E9 », ceci vous donnera la moyenne du semestre.

Tableau I : Calcul de la moyenne du semestre

A	B	C	D	E	F
Modules	CC	EMD	moyenne	Coefficients	moyenne
Mathématique 1				4	
Physique 1				4	
Chimie 1				4	
Environnement				2	
Informatique				2	
Anglais				1	
Histoire des sciences				1	
Total				18	
Moyenne du semestre				18	

I.5. Application 1 : Tracer une courbe

Afin de doser les composés phénoliques contenus dans les margines des huileries d'olive sous UV, une courbe d'étalonnage est nécessaire. L'acide gallique a été choisi pour tracer cette courbe. Les concentrations (C) utilisées sont : 0.4- 0.2- 0.1- 0.05 et 0.025 g.L⁻¹. Les absorbances (Abs) correspondantes à 751nm valent respectivement 1.4- 0.84322- 0.415- 0.2001 et 0.115.

- Ouvrir un fichier Excel puis enregistré le sous votre nom,
- Tracer un tableau, faire apparaître toutes les bordures du tableau,

- Saisir les données,
- Arrondir les chiffres à 3 décimales,
- En utilisant l'onglet insertion, nuage de points, tracer la courbe $Abs = f(C)$. Que remarquez-vous ?
- Ecrire l'équation de la droite.

II. Statistique d'une population

On désire étudier une population P, on tire un échantillon E de taille n issu de la population P, On analyse les caractéristiques de E puis on généralise à P.

II.1. Moyenne : la moyenne d'une série statistique est donnée par la relation :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

II.2. Variance : la variance de la même série statistique est donnée par la relation :

$$var(x) = \sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

II.3. Ecart-type : il est donné par

$$\sigma = s = \sqrt{var}$$

II.4. Intervalle de confiance : Si l'on suppose une distribution normale d'une série de résultats, l'intervalle de confiance sur une moyenne calculée est donnée par :

$$X = \bar{x} \pm \frac{t \cdot \sigma}{\sqrt{N}}$$

t est le coefficient de Fisher-Student dépendant du nombre de mesures N qui a servi au calcul de σ . Ce coefficient se trouve dans des tables pour différents niveaux de probabilité.

II.5. Médiane

La médiane est la valeur centrale d'une série statistique ; la moitié des observations lui sont inférieures ou égales et la moitié des observations lui sont supérieures ou égales. Pour calculer la médiane, on classe les valeurs de la série statistique dans l'ordre croissant :

- Si le nombre de valeurs est impair, la médiane est la valeur du milieu.

- S'il est pair, la médiane est la demi-somme des deux valeurs du milieu.

Exemple :

1- Pour déterminer la médiane de la série : 1, 4, 2, 5, 0 : On classe les valeurs de la série statistique dans l'ordre croissant : 0, 1, 2, 4, 5.

Il y a un nombre impair de valeurs, la médiane est donc la valeur du milieu.

0, 1, **2**, 4, 5 : **la médiane est 2.**

2- Pour déterminer la médiane de la série : 10, 40, 20, 50. On classe les valeurs de la série statistique dans l'ordre croissant : 10, 20, 40, 50.

Il y a un nombre pair de valeurs, on a donc 2 valeurs centrales. La médiane est alors la moyenne de ces deux valeurs.

$$10, \{20\}, \{40\}, 50: \text{médiane} = \frac{20+40}{2} = \frac{60}{2} = 30$$

Sur Excel, on peut avoir directement la valeur de la médiane d'une série statistique. Après avoir écrit la série, on saisit la formule « = médiane (nombre1 ; nombre2 ; ; nombre n) ».

III. Application 2 : calcul statistique

Deux méthodes de dosage de l'azote ont été répétées, à partir d'un même échantillon, 25 fois avec la méthode A et 30 fois avec la méthode B. Les résultats sont rassemblés dans les tableaux ci-dessous.

1- Reprendre les tableaux dans un classeur Excel que vous allez enregistrer.

2- Calculer la moyenne de chaque population.

3- Calculer la variance, l'écart type de chaque population. Déduire l'intervalle de confiance à 95%.

Méthode A : (25 expériences)

ni	1	2	2	4	7	4	2	2	1
Xi(g)	37	39	40	41	42	43	44	46	47

Méthode B : (30 expériences)

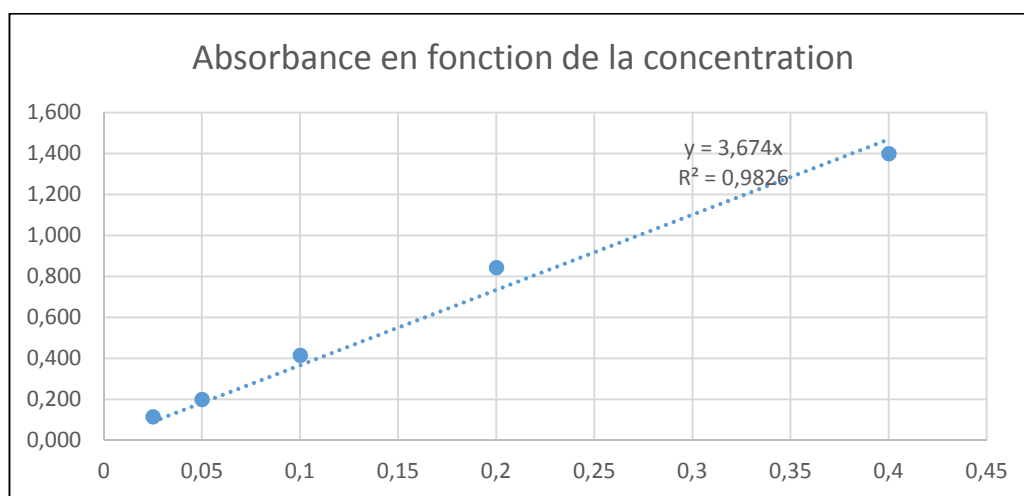
ni	2	1	6	9	8	3	1
Xi (g)	39	40	41	42	43	44	45

Corrigé des applications

1. Tracé de la courbe d'étalonnage :

C	Abs
0,4	1,400
0,2	0,843
0,1	0,415
0,05	0,200
0,025	0,115

Après avoir tracé la courbe des absorbances en fonction de la concentration, on obtient un nuage de point. L'allure de ces points est proche de celui d'une droite qui passe par l'origine. Pour déterminer l'équation de cette droite, on met le curseur sur l'un des points de la courbe, en cliquant avec la droite, on ajoute une courbe de tendance linéaire. Il faut cocher : « définir interception =0 », « afficher équation sur le graphique » et « afficher le coefficient de détermination R^2 ».



2. Deux méthodes de dosage de l'azote ont été répétées, à partir d'un même échantillon, 25 fois avec la méthode A et 30 fois avec la méthode B : L'ensemble des résultats trouvés, la moyenne, la variance et l'écart type sont regroupés dans le tableau II.

Sur Excel, on peut calculer la variance sans passer par le calcul de la moyenne. Pour cela, on utilisera la fonction « =VAR cellule 1 : cellule 25) ». Pour déterminer l'écart-type, il faut saisir « =ECARTTYPE cellule1 : cellule 25 ».

Tableau II

Méthode A				Méthode B			
N°	x	$x - \bar{x}$	$(x - \bar{x})^2$	N°	x	$x - \bar{x}$	$(x - \bar{x})^2$
1	37	-5,08	25,8064	1	39	-3,1	9,61
2	39	-3,08	9,4864	2	39	-3,1	9,61
3	39	-3,08	9,4864	3	40	-2,1	4,41
4	40	-2,08	4,3264	4	41	-1,1	1,21
5	40	-2,08	4,3264	5	41	-1,1	1,21
6	41	-1,08	1,1664	6	41	-1,1	1,21
7	41	-1,08	1,1664	7	41	-1,1	1,21
8	41	-1,08	1,1664	8	41	-1,1	1,21
9	41	-1,08	1,1664	9	41	-1,1	1,21
10	42	-0,08	0,0064	10	42	-0,1	0,01
11	42	-0,08	0,0064	11	42	-0,1	0,01
12	42	-0,08	0,0064	12	42	-0,1	0,01
13	42	-0,08	0,0064	13	42	-0,1	0,01
14	42	-0,08	0,0064	14	42	-0,1	0,01
15	42	-0,08	0,0064	15	42	-0,1	0,01
16	42	-0,08	0,0064	16	42	-0,1	0,01
17	43	0,92	0,8464	17	42	-0,1	0,01
18	43	0,92	0,8464	18	42	-0,1	0,01
19	43	0,92	0,8464	19	43	0,9	0,81
20	43	0,92	0,8464	20	43	0,9	0,81
21	44	1,92	3,6864	21	43	0,9	0,81
22	44	1,92	3,6864	22	43	0,9	0,81
23	46	3,92	15,3664	23	43	0,9	0,81
24	46	3,92	15,3664	24	43	0,9	0,81
25	47	4,92	24,2064	25	43	0,9	0,81
				26	43	0,9	0,81
\bar{x}	42,08	somme	123,84	27	44	1,9	3,61
		σ^2	5,16	28	44	1,9	3,61
		σ	2,27	29	44	1,9	3,61
		$t_{(0.05 ; 24)}$	2.064	30	45	2,9	8,41
		IC	± 0.938	\bar{x}	42,1	Somme	56,7
						σ^2	1,96
						σ	1,40
						$t_{(0.05 ; 29)}$	2.043
						IC	± 0.522

TP N° 2 : Criblage de facteurs par la méthode d'Hadamard

I. Criblage de facteurs

Un criblage de facteurs est une méthode qui nous permet de sélectionner sur l'ensemble des facteurs jugés influents, ceux qui le sont réellement. Les matrices d'Hadamard sont les matrices les plus utilisées. Dans ces matrices, les facteurs ne prennent généralement que deux niveaux distincts notés (-1) et (+1) en variables codées.

Pour k facteurs à deux niveaux, le nombre d'expériences N à réaliser doit satisfaire la condition :

$$\begin{cases} N \geq k + 1 \\ N \text{ est un multiple de } 4 \end{cases}$$

Les premières lignes des matrices d'Hadamard sont données par le tableau I.

Tableau I. Première ligne de la matrice d'Hadamard

Nombre de facteurs	Nombre d'expériences	Ligne de départ
$k \leq 3$	4	++-
$4 \leq k \leq 7$	8	+++--+
$8 \leq k \leq 11$	12	++-+++--+
$12 \leq k \leq 15$	16	++++-+--+---
$16 \leq k \leq 19$	20	++-++++-+-+----++-
$20 \leq k \leq 23$	24	+++++--+--+--+--+---

L'algorithme de construction d'une matrice de Hadamard à k facteurs et N expériences est le suivant :

- 1- Ecrire la première ligne en utilisant le tableau I,
- 2- Les autres lignes ou colonnes sont générées par permutation circulaire à droite ou à gauche ou vers le haut ou vers le bas (utiliser permutation circulaire à droite dans vos calculs),
- 3- Arrêter les permutations à la $(N-1)$ i^{ème} ligne,
- 4- La dernière ligne ne comporte que des (-1),
- 5- Supprimer les colonnes superflues,
- 6- Calculer les coefficients du modèle postulé.

Une matrice d'Hadamard est orthogonale, une matrice d'expérience X est dite orthogonale si le produit $(X^T X) = NI_d$,

(X^T étant la matrice transposée, N est le nombre d'expérience et I_d est la matrice identité).

II. Estimation des facteurs influents

Les calculs statistiques qui permettent de savoir si les effets sont significatifs, de calculer les intervalles de confiance ou de valider un modèle font intervenir d'une part les résidus e_i , c'est-à-dire la différence entre la valeur expérimentale et la valeur prédite par le modèle et, d'autre part un estimateur sans biais de la variance commune des résidus. Cet estimateur est donné par la relation :

$$S^2 = \frac{1}{N - p} \sum e_i^2$$

Où

N est le nombre d'expériences réalisées,

p est le nombre de coefficients du modèle incluant la constante b_0 ,

e_i : l'écart entre la valeur expérimentale et la valeur prédite par le modèle.

Dans ces conditions, on peut montrer que tous les effets ont la même variance σ_i^2 donnée par :

$$\sigma_i^2 = \frac{S^2}{N}$$

Pour identifier les facteurs impactant la réponse observée, il est indispensable de calculer l'écart type (σ) et le t de Student (t_{exp}) pour chaque coefficient.

Le **t_{exp} de Student** est donné par la relation :

$$t_{exp} = \frac{|b_j|}{\sigma}$$

Connaissant le nombre de degré de liberté (ddl), on calcule le $t_{critique}$ en utilisant la loi de Student en choisissant un risque α le plus souvent de 5%.

Nous considérons que si :

$t_{exp} > t_{critique(\alpha;ddl)}$: Le facteur est influent sur la réponse,

$t_{exp} < t_{critique(\alpha;ddl)}$: Le facteur est non influent sur la réponse.

III. Application

La chélation est un processus physico-chimique au cours duquel est formé un complexe, le chélate, entre un ligand, dit chélateur, et un cation métallique.

Lors de la production de l'acide phosphorique, l'éthylène diamine tétra acétique (EDTA), agent chélateur, est utilisé pour le nettoyage des installations afin d'éviter la formation d'un précipité.

Les responsables de l'unité décident de réaliser dans un premier temps un criblage. Les facteurs étudiés et leur domaine de variation sont donnés dans le tableau II.

Tableau II : Niveaux des facteurs de l'étude

	Facteurs	Niveau (-1)	Niveau (+1)
X1	pH	2	13
X2	Température de chauffage (°C)	25	50
X3	Temps d'agitation (mn)	30	120
X4	Masse de KH_2PO_4 (g)	0,05	1
X5	Temps de dépôt du métal (h)	5	24

La réponse étudiée est le rendement d'EDTA récupéré (en %).

1- Quel est le nombre d'expérience à réaliser, Construire la matrice d'Hadamard en variables codées (permutation à droite),

2- Ecrire le plan d'expérience correspondant à la matrice,

3- Après expérimentation, on trouve les résultats expérimentaux suivants :

$$y = \{36; 55; 39; 17; 49; 10; 26; 32\}.$$

4- Calculer les coefficients b_j ,

5- Retrouver les effets influents ($t_{(\alpha; dnl)} = 4,3$), déduire le modèle mathématique puis tracer qualitativement le diagramme des effets.

Corrigé de l'application

1. Matrice des effets

Pour 5 facteurs à deux niveaux, le nombre d'expériences N à réaliser doit satisfaire la condition :

$$\begin{cases} N \geq 5 + 1 \\ N \text{ est un multiple de } 4 \end{cases} \rightarrow N = 8$$

- Sur un classeur Excel, on ouvre une feuille de calcul qu'il faut enregistrer,
- Définir les bordures du tableau (A3 :H3 ; A11 : H11),
- Sur la première colonne (A3 : A11), mettre le nombre d'expérience N dans l'ordre (de 1 à 8).
- Sur la première ligne (B3 :H3), écrire les variables codées (X1...X7),
- Sur la ligne (B4 : H4), écrire la première ligne de la matrice d'Hadamard donnée par le tableau V,
- Nous avons choisi la permutation de droite à gauche, commencer à permuter (voir tableau) jusqu'à la 7eme ligne. La dernière ligne ne contient que des (-1),
- Supprimer les colonnes superflues (G et H) qui ne correspondent à aucune variable.

Tableau III : Matrice d'HADAMARD

A	B	C	D	E	F	G	H
N	X1	X2	X3	X4	X5	Colonnes superflues	
1	1	1	1	-1	1	-1	-1
2	-1	1	1	1	-1	1	-1
3	-1	-1	1	1	1	-1	1
4	1	-1	-1	1	1	1	-1
5	-1	1	-1	-1	1	1	1
6	1	-1	1	-1	-1	1	1
7	1	1	-1	1	-1	-1	1
8	-1	-1	-1	-1	-1	-1	-1

- Ecrire la matrice des effets, en insérant la colonne X0 (à gauche de la matrice) et une colonne (à droite) qui va correspondre au vecteur réponse Y (rendement en EDTA).

2. Calcul des coefficients du modèle

Le calcul des coefficients du modèle de criblage sur Excel se fait avec la matrice des effets, en multipliant la première colonne (X0) par la dernière (Y), ligne par ligne, puis diviser la somme sur le nombre d'expérience (8).

Exemple : Le calcul du premier coefficient se fait :

$$b_0 = \frac{36 * B1 + 55 * B2 + 39 * B3 + 17 * B4 + 49 * B5 + 10 * B6 + 26 * B7 + 32 * B8}{8} = 33.00$$

Un seul calcul sera fait, le reste sera déduit en cliquant sur le symbole (+) en bas de la cellule B9 et le glisser jusqu'à la dernière cellule (G 9).

Tableau VI : matrice des effets

N°	X0	X1	X2	X3	X4	X5	Y _{exp}
1	1	1	1	1	-1	1	36.00
2	1	-1	1	1	1	-1	55.00
3	1	-1	-1	1	1	1	39.00
4	1	1	-1	-1	1	1	17.00
5	1	-1	1	-1	-1	1	49.00
6	1	1	-1	1	-1	-1	10.00
7	1	1	1	-1	1	-1	26.00
8	1	-1	-1	-1	-1	-1	32.00
bj	33.00	-10.75	8.50	2.00	1.25	2.25	

Le modèle théorique élaboré par le criblage est donné par :

$$y_{the} = 33.00 - 10.75 X1 + 8.50 X2 + 2X3 + 1.25X4 + 2.25X5$$

3. Recherche des facteurs influents

Pour connaître les facteurs influents sur le rendement en EDTA, il faut calculer l'écart type. Ce dernier se déduit de la variance expérimentale calculée à partir d'un estimateur S^2 :

- Pour calculer cet estimateur, il faut passer par le calcul du rendement théorique en EDTA (Y_{the}), calculé par le modèle élaboré pour chaque expérience,
- Sur la même feuille de calcul Excel, on insère une nouvelle colonne à droite (Y_{the}),

- Le calcul des valeurs de (Y_{the}) se fait en multipliant les coefficients du modèle élaboré par les lignes de la matrice d'Hadamard dans l'ordre,
- La première valeur du rendement théorique est calculé par :

$$Y_{the1} == 33 * B1 - 10.75 * C1 + 8.5 * D1 + 2 * E1 + 1.25 * F1 + 2.25 * G1$$

- Les autres valeurs seront déduites en tirant sur le symbole (+) en bas de la cellule (I1) jusqu'à la dernière cellule (I8),
- Dans la colonne J, mettre les valeurs des résidus (e_i^2) qui correspondent à l'écart entre les valeurs expérimentales et les valeurs théoriques calculées par le modèle élaboré :

$$e_i^2 = (Y_{exp} - Y_{the})^2$$

- Calculer l'estimateur S^2 par :

$$S^2 = \frac{1}{N - p} \sum (Y_{exp} - Y_{the})^2 = \frac{1}{N - p} \sum (e_i)^2$$

e_i : l'écart entre la valeur expérimentale et la valeur prédite par le modèle.

Tableau V : Détermination des facteurs influents

A	B	C	D	E	F	G	H	I	J
Exp	X0	X1	X2	X3	X4	X5	Y exp	y _{thé}	(ei) ²
1	1	1	1	1	-1	1	36	33,750	5,063
2	1	-1	1	1	1	-1	55	53,250	3,063
3	1	-1	-1	1	1	1	39	40,750	3,063
4	1	1	-1	-1	1	1	17	15,250	3,063
5	1	-1	1	-1	-1	1	49	51,250	5,063
6	1	1	-1	1	-1	-1	10	12,250	5,063
7	1	1	1	-1	1	-1	26	27,750	3,063
8	1	-1	-1	-1	-1	-1	32	29,750	5,063
Bj	33,00	-10,75	8,50	2,00	1,25	2,25			
								somme ei ²	32,50
								S ²	16,250
								var (bj)	2,031
								Ecart- type	1,425

Pour identifier les facteurs impactant le rendement de l'EDTA, il faut calculer le t de Student (t_{exp}) pour chaque coefficient, il est donné par la relation :

$$t_{exp} = \frac{|b_j|}{\sigma}$$

Le tableau VI résume l'ensemble des résultats ainsi que la conclusion tirée.

Tableau VI : Test de STUDENT

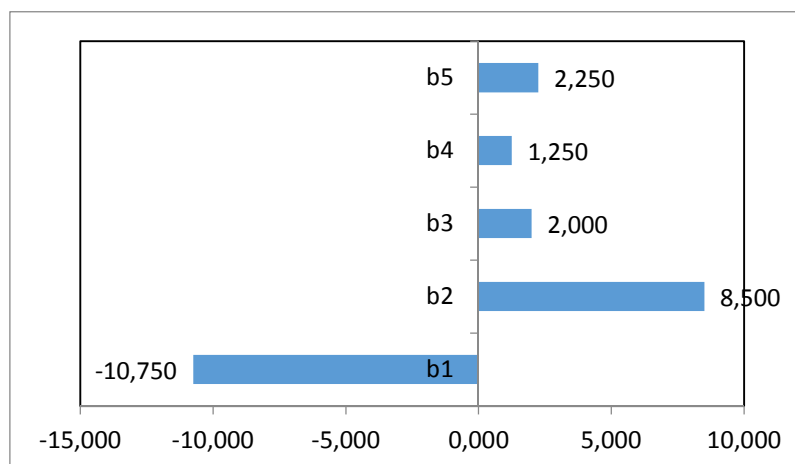
Facteur		Abs (bj)	σ	t_{exp}	t(0.05 ; 2)	conclusion
pH	b1	10.750	1.430	7.517	4.30	influent
Température de chauffage	b2	8.500	1.430	5.944	4.30	influent
Temps d'agitation	b3	2.000	1.430	1.399	4.30	non influent
Masse de KH ₂ PO ₄	b4	1.250	1.430	0.874	4.30	non influent
Temps de dépôt du métal	b5	2.250	1.430	1.573	4.30	non influent

4. Conclusion

Les facteurs influents sur le rendement en EDTA sont le pH et la température. Après élimination des facteurs non influents, le modèle mathématique du criblage est donné par :

$$y_{the} = 33.00 - 10.75 X_1 + 8.50 X_2$$

Diagramme des effets : tracé qualitatif du diagramme des effets est le suivant :



TP 3: Identification des facteurs influents en utilisant les outils d'aide à l'interprétation

Pour identifier les effets principaux impactant la réponse lorsque le nombre de degré de liberté est élevé et **la variance expérimentale n'est pas connue**, nous pouvons utiliser les outils d'aide à l'interprétation. Dans ces outils, on considère que l'erreur expérimentale est une variable aléatoire distribuée selon une loi normale.

I. Approche de Pareto

Pareto a montré que dans la majorité des situations, un petit nombre de facteurs a une influence majeure sur les résultats. C'est la loi dite de Pareto des 80-20, où, 20% des facteurs expliquent 80% des résultats (20 % des étudiants demandent 80% de l'effort d'un enseignant, 20% de la population mondiale a accès à 80% de l'eau potable, 20% de la population paie 80% des impôts, 20% des problèmes représentent 80% des préoccupations, 20% des pays émettent 80% des gaz à effet de serre...etc).

Pour identifier les facteurs influents, nous devons :

- Ecrire la matrice d'Hadamard en adoptant la permutation de droite à gauche,
- Calculer les coefficients b_j et leurs carrés b_j^2 ,
- Classer les valeurs b_j^2 par ordre décroissant (de Z à A) puis calculer leurs sommes $\sum b_j^2$,
- Calculer l'indice de Pareto (IP) donné par la relation:

$$IP = 100 * \frac{b_j^2}{\sum b_j^2}$$

- Calculer l'indice de Pareto cumulé (IPc).
- Tracer les histogrammes des b_j en fonction de IPc.

II. Approche de Lenth

Lenth a mis au point une méthode qui consiste à estimer le pseudo écart-type des estimations des coefficients calculés. Sous cette hypothèse, Lenth démontre qu'il y a un lien entre la médiane des valeurs absolues des b_j et l'écart-type donné par :

$$S_0 = 1.5 * \text{médiane} |b_j|$$

S_0 est appelé « pseudo écart-type » noté PSE,

$|b_j|$ représente la valeur absolue des b_j .

Pour identifier les facteurs ayant un poids, par l'approche de Lenth, nous devons :

- Ecrire la matrice d'Hadamard en adoptant la permutation à droite,
- Calculer les coefficients b_j ,
- Calculer $Abs|b_j|$ puis classer les valeurs par ordre croissant (A à Z),
- Calculer la médiane,
- Calculer l'écart type S_0 ,
- Calculer $2.5*S_0$,
- Eliminer les b_j supérieurs à $(2.5*S_0)$ puis reprendre les calculs.
- Le PSE correspond à la valeur de S_0 trouvée dans la dernière itération,

Les limites de signification sont calculées à l'aide de la relation :

$$limites = \pm t_{\alpha,d} * PSE$$

α = seuil de signification qui est généralement de 5%,

d = ddl = nombre d'effets restants/3.

III. Normal Plot (NP) ou Half Normal Plot (HNP)

Comme dans l'approche de Lenth, on considère que tous les b_j sont distribués selon une loi normale autour de zéro. Les graphes Normal Plot et Half Normal Plot permettent de tester la répartition des valeurs de ces effets. La construction du Normal Plot se fait de la manière suivante :

- Ecrire la matrice d'Hadamard en adoptant la permutation à droite,
- Calculer les coefficients b_j ,
- Classer les coefficients par ordre croissant (A à Z),
- Affecter à chaque coefficient un rang j ($j=1,2,\dots,k$). La valeur des effets constitue l'abscisse du graphe,
- Calculer la probabilité correspondant au rang j à l'aide de la relation :

$$P = \frac{j - 0.5}{k}$$

- Tracer la probabilité en fonction de b_j .

Les points qui s'écartent de la droite correspondent aux effets influents. Pour la construction du Half Normal Plot, on adopte la même démarche mais le classement se fait avec la valeur absolue des valeurs des effets.

IV. Application

L'étude de la robustesse d'un procédé de préparation de sulfate amide des acides gras de l'huile de grignons d'olive a été étudiée. Ce procédé est réalisé en plusieurs étapes : la saponification, l'hydrolyse des savons obtenus, l'estérification des acides gras, l'amidation et la sulfatation des diéthanolamines des acides gras et enfin une neutralisation avec une solution de KOH.

Nous avons appliqué un plan de criblage pour distinguer parmi onze facteurs jugés potentiellement influents ceux qui le sont réellement. Ces facteurs ainsi que leurs niveaux bas et haut sont donnés par le tableau I.

Tableau I : Domaine expérimental

	Facteurs	Niveau (-1)	Niveau (+1)
X1	Durée d'hydrolyse (h)	1	1,5
X2	Température d'estérification (°C)	115	125
X3	Durée d'estérification (h)	3	5
X4	Température d'amidation (°C)	120	140
X5	Durée d'amidation (h)	10	12
X6	Méthanoate de sodium (%)	0,25	0,5
X7	Amine /Ester	1,2	1,6
X8	Réactif sulfatation	oléum	acide chlorosulfonique
X9	SO ₃ /Ester	1,5	2
X10	Température de sulfatation (°C)	5	15
X11	Durée de sulfatation (h)	2	3

Les rendements de la réaction globale de préparation des sulfate-amides sont :

$$y = \{76, 80, 106, 113, 140, 86, 92, 134, 96, 88, 91, 73\}.$$

- 1- Quel est le nombre d'expérience à réaliser,
- 2- Construire la matrice d'Hadamard en variables codées et réelles,
- 3- Calculer les coefficients b_j ,
- 4- Quels sont les facteurs influents en utilisant les outils d'aide à l'interprétation en utilisant : Approche de Pareto, l'approche de Lenth, Normal Plot (NP) et Half normal Plot (HNP).
- 5- Conclure.

Corrigé de l'application

1- Construction de la matrice et calcul des coefficients du modèle

Le nombre d'expériences N à réaliser doit satisfaire la condition :

$$\begin{cases} N \geq k + 1 \\ N \text{ est un multiple de } 4 \end{cases}$$

$$k = 11 \rightarrow N = 12$$

Le nombre d'expérience à réaliser est de 12, la construction de la matrice d'Hadamard et le calcul des coefficients du modèle se fait comme précédemment (voir TP 2).

La matrice d'Hadamard en variables codées et les valeurs des coefficients b_j sont regroupés dans le tableau suivant :

Tableau II : matrice d'Hadamard

	X0	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	Y
1	1	1	1	-1	1	1	1	-1	-1	-1	1	-1	76
2	1	-1	1	1	-1	1	1	1	-1	-1	-1	1	80
3	1	1	-1	1	1	-1	1	1	1	-1	-1	-1	106
4	1	-1	1	-1	1	1	-1	1	1	1	-1	-1	113
5	1	-1	-1	1	-1	1	1	-1	1	1	1	-1	140
6	1	-1	-1	-1	1	-1	1	1	-1	1	1	1	86
7	1	1	-1	-1	-1	1	-1	1	1	-1	1	1	92
8	1	1	1	-1	-1	-1	1	-1	1	1	-1	1	134
9	1	1	1	1	-1	-1	-1	1	-1	1	1	-1	96
10	1	-1	1	1	1	-1	-1	-1	1	-1	1	1	88
11	1	1	-1	1	1	1	-1	-1	-1	1	-1	1	91
12	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	73
b_j	97,92	1,25	-0,08	2,25	-4,58	0,75	5,75	-2,42	14,25	12,08	-1,58	-2,75	

2. Détermination des facteurs influents en utilisant les outils d'aide à l'interprétation

➤ Approche de PARETO

En respectant la démarche donnée précédemment, Les résultats trouvés sont regroupés dans le tableau III :

Tableau III : Méthode de Pareto

Nom	b_j	Nom	b_j (tri A à Z)	b_j^2	IP	IP _{Cumulé}
b1	1.25	b8	14,25	203,0625	47,652	47,652
b2	-0.08	b9	12,08	145,9264	34,244	81,896
b3	2.25	b6	5,75	33,0625	7,759	89,655
b4	-4.58	b4	-4,58	20,9764	4,922	94,577
b5	0.75	b11	-2,75	7,5625	1,775	96,352
b6	5.75	b7	-2,42	5,8564	1,374	97,726
b7	-2.42	b3	2,25	5,0625	1,188	98,914
b8	14.25	b10	-1,58	2,4964	0,586	99,500
b9	12.08	b1	1,25	1,5625	0,367	99,867
b10	-1.58	b5	0,75	0,5625	0,132	99,999
b11	-2.75	b2	-0,08	0,0064	0,002	100,00
			Somme b_j^2	426,137		

Selon Pareto, seul le facteur b8 est considéré influent car son indice de Pareto est inférieur à 80. Pour mieux voir les résultats, on trace l'histogramme de Pareto.

Le réactif sulfatation est un facteur influent sur la réponse, sa valeur est positive. La réponse augmente en passant de l'oléum (niveau bas) à l'acide chlorosulfonique (niveau haut).

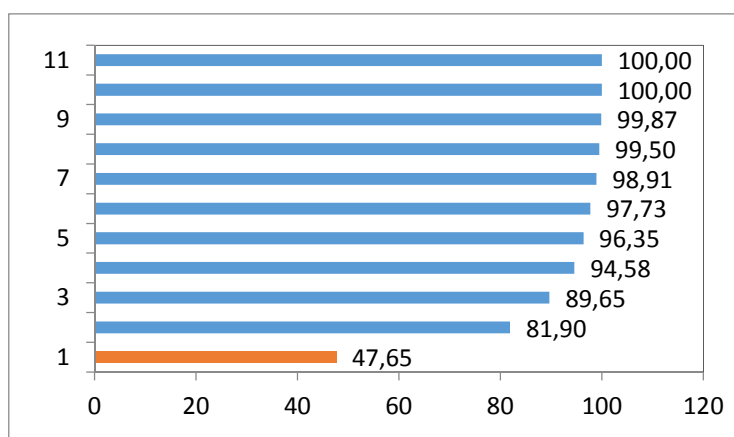


Fig.1.Histogramme de Pareto

➤ Approche de LENTH

Tableau IV : Approche de Lenth

Nom	b_j	Nom	b_j	$ b_j $ (tri de Z à A)	
b1	1.25	b2	-0,08	0,08	0,08
b2	-0.08	b5	0,75	0,75	0,75
b3	2.25	b1	1,25	1,25	1,25
b4	-4.58	b10	-1,58	1,58	1,58
b5	0.75	b3	2,25	2,25	2,25
b6	5.75	b7	-2,42	2,42	2,42
b7	-2.42	b11	-2,75	2,75	2,75
b8	14.25	b4	-4,58	4,58	4,58
b9	12.08	b6	5,75	5,75	5,75
b10	-1.58	b9	12,08	12,08	
b11	-2.75	b8	14,25	14,25	
			Mediane	2,42	2,25
			S0	3,63	3,375
			2.5*S0	9,075	8,4375

- Sur Excel, le calcul de la médiane se fait par « = mediane (Cellule 1 : cellule11) »,
- L'écart type S_0 est déterminé par la relation : $S_0 = 1.5 * médiane|b_j|$
- Le pseudo écart type (PSE) correspond à la valeur de S_0 trouvée à la dernière itération,

Deux itérations sont faites, on conclut que :

$PSE = 3.375$
$ddl = \frac{11 - 2}{3} = 3$
$t_{(0.05; 3)} = 3.18$
$Limites = \pm PSE * t_{0.05; 3} = \pm 10.76$

Les facteurs Réactif sulfatation (b8) et SO₃/Ester (b9) sont considérés influents car les valeurs de leurs coefficients (12.08 et 14.25) sont en dehors des limites (± 10.75) données par l'approche de LENTH.

➤ **NORMAL ET HALF NORMAL PLOT (NHP/NP)**

En respectant la démarche décrite précédemment, il faut :

- Ordonner les valeurs des coefficients dans un ordre croissant (A à Z) en affectant à chaque effet un rang j ($j=1,2,\dots,k$),

- Calculer la probabilité correspondant au rang j à l'aide de la relation :

$$P = \frac{j - 0.5}{k}$$

- Tracer la probabilité en fonction de b_j . Les points qui s'écartent de la droite correspondent aux facteurs influents.

Les résultats sont regroupés dans le tableau V.

Pour la construction du Half Normal Plot (HNP), il faut adopter la même démarche mais le classement se fait avec les valeurs absolues des coefficients.

Tableau V : Méthode de Normal plot

Nom	b_j	Tri	b_j (tri de Z à A)	rang	P
b1	1.25	b4	-4,58	1	0,045
b2	-0.08	b11	-2,75	2	0,136
b3	2.25	b7	-2,42	3	0,227
b4	-4.58	b10	-1,58	4	0,318
b5	0.75	b2	-0,08	5	0,409
b6	5.75	b5	0,75	6	0,500
b7	-2.42	b1	1,25	7	0,591
b8	14.25	b3	2,25	8	0,682
b9	12.08	b6	5,75	9	0,773
b10	-1.58	b9	12,08	10	0,864
b11	-2.75	b8	14,25	11	0,955

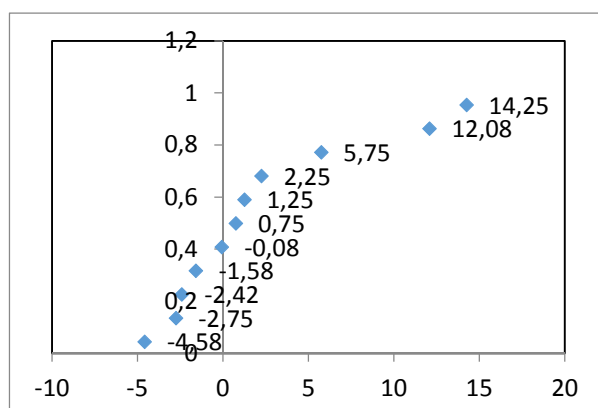


Fig.2: Normal Plot (NP)

Encore une fois, on confirme l'influence des deux facteurs réactif sulfatation (b8) et le rapport SO_3/Ester (b9) sur le rendement de la réaction globale de préparation des sulfate-amides des acides gras de l'huile de grignons d'olive.

TP N° 4 : Identification des facteurs influents -doubler un plan d'expérience-

I. Criblage de facteurs-doubler le plan d'expérience

Un criblage de facteurs est une méthode qui nous permet de sélectionner sur l'ensemble des facteurs jugés influents, ceux qui le sont réellement. Dans les TP précédents, la connaissance de l'estimateur S^2 nous a permis de calculer la variance expérimentale et déduire l'écart-type. Dans la pratique, plusieurs cas peuvent se présenter :

- **L'écart-type est connu** par des expériences antérieures de même type que celles du plan et suffisamment nombreuses pour être fiables. C'est le meilleur des cas.
- **l'écart-type n'est pas connu** : pour l'estimer, l'expérimentateur doit :
 - Soit faire des expériences au centre du domaine,
 - Soit doubler le plan d'expérience.

L'objectif de ce TP est de déterminer les facteurs influents sur une réponse donnée en doublant le plan d'expérience et respectant les méthodes expliquées précédemment dans les TP2 et TP3.

II. Application

Afin de conserver les composés phénoliques des margines des huileries d'olive, une encapsulation a été faite en utilisant la pectine extraite à partir d'un zeste de citron. Plusieurs facteurs interviennent dans le processus de l'extraction. C'est pour cela, qu'un criblage de facteurs est proposé pour déterminer les facteurs réellement influents.

Ces facteurs ainsi que leurs niveaux bas et haut sont donnés par le tableau I.

Tableau I. Domaine expérimental de chaque facteur.

Facteurs	Valeurs codées	Unités	Niveau	
			(-1)	(+1)
Temps de coagulation	X1	h	1	2
pH	X2	-	1.5	3
Rapport liquide/solide	X3	Mg.L ⁻¹	25	50
Température	X4	°C	40	80
Temps d'extraction	X5	mn	20	60
Vitesse d'agitation	X6	Tr.mn ⁻¹	200	400
Température de séchage	X7	°C	40	60

La réponse étudiée est le rendement de l'extraction de la pectine (en %).

- 1- Quel est le nombre d'expérience à réaliser, Construire la matrice d'Hadamard en variables codées (permutation à droite). Ecrire le plan d'expérimentation.
- 2- Le plan a été doublé. Après expérimentation, on trouve les résultats expérimentaux suivants :
$$y = \left\{ \begin{array}{cccccccc} 6.20; & 7.40; & 9.20; & 9.00; & 21.40; & 20.20; & 29.80; & 20.20; & 5.40; \\ & 5.80; & 12.60; & 13.80; & 6.80; & 6.60; & 11.80; & 12.00 & \end{array} \right\}$$
- 3- Calculer les coefficients du modèle,
- 4- Retrouver les facteurs influents ($t_{(\alpha;ddl)} = 2.31$). Déduire le modèle mathématique,
- 5- Conclure.

Corrigé de l'application

1. Pour sept facteurs à deux niveaux, le nombre d'expériences N à réaliser doit satisfaire la condition :

$$\begin{cases} N \geq 7 + 1 \\ N \text{ est un multiple de } 4 \end{cases} \rightarrow N = 8$$

Pour connaître les facteurs influents, Le calcul de la variance expérimentale de la réponse s'impose. Le nombre de degré de liberté (ddl) étant peu élevé, nous ne pouvons pas utiliser les outils d'aide à l'interprétation. Il faut doubler le plan et réaliser 16 expériences.

L'écriture de la matrice d'expérience se fait en se basant sur les TP précédents. La matrice est donnée dans le tableau II.

Tableau II. Matrice d'expérience

N°	X1	X2	X3	X4	X5	X6	X7
1	1	1	1	-1	1	-1	-1
2	1	1	1	-1	1	-1	-1
3	-1	1	1	1	-1	1	-1
4	-1	1	1	1	-1	1	-1
5	-1	-1	1	1	1	-1	1
6	-1	-1	1	1	1	-1	1
7	1	-1	-1	1	1	1	-1
8	1	-1	-1	1	1	1	-1
9	-1	1	-1	-1	1	1	1
10	-1	1	-1	-1	1	1	1
11	1	-1	1	-1	-1	1	1
12	1	-1	1	-1	-1	1	1
13	1	1	-1	1	-1	-1	1
14	1	1	-1	1	-1	-1	1
15	-1	-1	-1	-1	-1	-1	-1
16	-1	-1	-1	-1	-1	-1	-1

1. Le calcul des coefficients doit passer par l'écriture de la matrice des effets (tableau II), le calcul se fait en respectant la démarche faite auparavant.

Pour connaître les facteurs influents sur le rendement de cette extraction, il faut calculer l'écart type.

Ce dernier se déduit de la variance expérimentale calculée à partir de l'estimateur S^2 .

Tableau II : Matrice des effets

N	X0	X1	X2	X3	X4	X5	X6	X7	y _{exp}	y _{the}	e _i ²
1	1	1	1	1	-1	1	-1	-1	6,20	6,80	0,36
2	1	1	1	1	-1	1	-1	-1	7,40	6,80	0,36
3	1	-1	1	1	1	-1	1	-1	9,20	9,10	0,01
4	1	-1	1	1	1	-1	1	-1	9,00	9,10	0,01
5	1	-1	-1	1	1	1	-1	1	21,40	20,80	0,36
6	1	-1	-1	1	1	1	-1	1	20,20	20,80	0,36
7	1	1	-1	-1	1	1	1	-1	29,80	25,00	23,04
8	1	1	-1	-1	1	1	1	-1	20,20	25,00	23,04
9	1	-1	1	-1	-1	1	1	1	5,40	5,60	0,04
10	1	-1	1	-1	-1	1	1	1	5,80	5,60	0,04
11	1	1	-1	1	-1	-1	1	1	12,60	13,22	0,38
12	1	1	-1	1	-1	-1	1	1	13,80	13,22	0,34
13	1	1	1	-1	1	-1	-1	1	6,80	6,70	0,01
14	1	1	1	-1	1	-1	-1	1	6,60	6,70	0,01
15	1	-1	-1	-1	-1	-1	-1	-1	11,80	11,90	0,01
16	1	-1	-1	-1	-1	-1	-1	-1	12,00	11,90	0,01
b _j	12,39	0,54	-5,34	0,09	3,01	2,16	0,84	-0,81		Somme	48,38
										S ²	6,05
										Var (y)	0,38
										σ	0,61

Le tableau III résume l'ensemble des résultats ainsi que la conclusion tirée :

Tableau III : Etude de effets influents

Facteur		Abs (b _j)	σ	t	t _(0.05 ;8)	Conclusion
Temps de coagulation	b1	0,54	0,61	0,885	2,31	Non influent
pH	b2	5,34	0,61	8,754	2,31	influent
Rapport liquide/solide	b3	0,09	0,61	0,148	2,31	Non influent
Température	b4	3,01	0,61	4,934	2,31	influent
Temps d'extraction	b5	2,16	0,61	3,541	2,31	influent
Vitesse d'agitation	b6	0,84	0,61	1,377	2,31	Non influent
Température de séchage	b7	0,81	0,61	1,328	2,31	Non influent

Conclusion

Sur l'ensemble des facteurs sélectionnés, uniquement le pH, la température et le temps de l'extraction sont jugés influents par le criblage. Le modèle est donné par :

$$y = 12.39 - 5.34 X_2 + 3.01 X_4 + 2.16 X_5$$

TP N° 5 : Plan factoriel complet 2^k

I. Plan factoriel complet

Un plan est dit factoriel complet, noté (2^k) , lorsque les niveaux de tous les facteurs sont programmés pour varier simultanément et de manière équilibrée dans l'ensemble de l'expérimentation. Ces plans sont à utiliser pour rechercher les facteurs agissant sur une réponse mesurée. Ce sont les plus simples à interpréter et ils présentent le meilleur rapport coût/efficacité.

II. Construction de la matrice

Il est commode de symboliser par (-1) le niveau bas et par (+1) le niveau haut de chaque facteur. Ce qui permet de rassembler les éléments relatifs à chaque facteur dans une matrice d'expérience que l'on présente en correspondance avec une colonne donnant les résultats expérimentaux de la réponse y .

En général, la matrice d'expérience comporte k colonnes pour les facteurs principaux et 2^k lignes correspondantes à k essais. La construction se fait avec l'algorithme de YATES.

Colonne 1 du premier facteur : alternance -1 et +1,

Colonne 2 du deuxième facteur : alternance -1 et +1 de 2 en 2,

Colonne 3 du troisième facteur : alternance -1 et +1 de 4 en 4,

Colonne 4 du quatrième facteur : alternance -1 et +1 de 8 en 8.

Au-delà de 5 facteurs, il faut passer aux plans fractionnaires.

III. Application

On se propose d'étudier le rendement de l'extraction d'une huile essentielle à partir de *Cymbopogon Citratus*. Les facteurs choisis (deux niveaux) sont consignés dans le tableau I.

Tableau I : Domaine expérimental de chaque facteur

Facteurs	Niveau (-1)	Niveau (+1)
A : % d'eau perdue par le végétal (%)	23	70
B : Débit de condensat (mL/h)	110	210
C : Longueur des feuilles coupées (cm)	1.7	6.3

La réponse étudiée correspond au volume de l'huile essentielle extraite de 250 g de matériel végétal.

Les résultats obtenus sont : $Y = \{2.2; 1.5; 2.0; 1.1; 3.4; 1.8; 3.2; 1.6\}$.

- 1- Construire la matrice des effets,
- 2- Calculer les effets des facteurs et interactions (négliger toute interaction d'ordre ≥ 2)
- 3- Retrouver les effets réellement influents puis donner le modèle mathématique élaboré,
- 4- Indiquer les conditions opératoires souhaitables pour augmenter la masse d'huile extraite.

Corrigé de l'application

1- La matrice des effets

Le nombre d'expérience à réaliser est $2^3 = 8$ expériences, la matrice des effets ainsi que le calcul des coefficients du modèle sont donnés dans le tableau suivant II.

Tableau II : matrice des effets

	X0	X1	X2	X3	X1X2	X1X3	X2X3	y
1	1	-1	-1	-1	1	1	1	2,2
2	1	1	-1	-1	-1	-1	1	1,5
3	1	-1	1	-1	-1	1	-1	2
4	1	1	1	-1	1	-1	-1	1,1
5	1	-1	-1	1	1	-1	-1	3,4
6	1	1	-1	1	-1	1	-1	1,8
7	1	-1	1	1	-1	-1	1	3,2
8	1	1	1	1	1	1	1	1,6
bj	2,100	-0,600	-0,125	0,400	-0,025	-0,200	0,025	

2- Les effets réellement influents seront connus après avoir calculé l'estimateur, la variance, l'écart type ainsi que le t de Student (tableau IV et V).

Tableau IV : calcul des résidus

N°	X0	X1	X2	X3	X1X2	X1X3	X2X3	y	y _{the}	ei ²
1	1	-1	-1	-1	1	1	1	2,2	2,225	0,000625
2	1	1	-1	-1	-1	-1	1	1,5	1,475	0,000625
3	1	-1	1	-1	-1	1	-1	2	1,975	0,000625
4	1	1	1	-1	1	-1	-1	1,1	1,125	0,000625
5	1	-1	-1	1	1	-1	-1	3,4	3,375	0,000625
6	1	1	-1	1	-1	1	-1	1,8	1,825	0,000625
7	1	-1	1	1	-1	-1	1	3,2	3,225	0,000625
8	1	1	1	1	1	1	1	1,6	1,575	0,000625
bj	2,100	-0,600	-0,125	0,400	-0,025	-0,200	0,025		somme	0,005
									S ²	0,005
									σ^2	0,000625
									σ	0,025

Tableau V : Effets influents

Effet	$ b_j $	σ	t_{exp}	$t_{(0.05; 1)}$	Conclusion
b1	0,6	0,025	24	12,71	influent
b2	0,125	0,025	5	12,71	non influent
b3	0,4	0,025	16	12,71	influent
b12	0,025	0,025	1	12,71	non influent
b13	0,2	0,025	8	12,71	non influent
b23	0,025	0,025	1	12,71	non influent

Le modèle postulé est :

$$Y = 2.1 - 0.6 X1 - 0.4 X3$$

- 3- Pour augmenter la masse de l'huile extraite, il faut prendre les facteurs A et B à leurs niveaux bas (effet < 0) et le facteur C à son niveau haut (effet > 0).

TP1

I. Soit la réaction : $CH_3I + C_2H_5ONa \rightarrow CH_3OC_2H_5 + NaI$

L'ordre total est égal à 2 et k est la constante de vitesse.

On a établi les résultats expérimentaux suivants :

T (°C)	0	6	12	18	24	30
k (mol ⁻¹ . L. s ⁻¹) 10 ⁺⁵	5,60	11,8	24,5	48,8	100	208

Déterminer graphiquement les paramètres de l'équation d'Arrhenius de la réaction (l'énergie d'activation et le facteur de fréquence).

II. La détermination du contenu en ion de sodium dans un échantillon d'urine avec une électrode sélective d'ions proportionne les résultats suivants :

N°	1	2	3	4	5	6
[Na]	102mM	97 mM	99 mM	98 mM	101 mM	106 mM

Quelles sont les limites de l'intervalle de confiance du sodium à 99% ? ($t_{99\%}=4,03$).

TP2

Nous avons utilisé la matrice d'Hadamard pour étudier le dépôt de l'alumine sur le silicium. La méthode des dépôts chimiques en phase vapeur (CVD) permet de réaliser des couches minces d'alumine sur le silicium chauffé à partir d'un précurseur gazeux et d'une réaction chimique. Pour que les réactions chimiques aient lieu, le silicium doit être porté à des températures allant de 500 à 2000°C. Pour réduire ces températures, on utilise la méthode PECVD dans laquelle l'énergie thermique (chauffage) est remplacée par l'action énergétique générée par une décharge électrique : le plasma. Le protocole de dépôt fait intervenir 8 facteurs (voir tableau) et la réponse constitue la quantité de matière déposée par unité de surface et de temps sur le substrat ($\mu\text{g}\cdot\text{cm}^{-2}\cdot\text{mn}^{-1}$).

	Facteurs	Niveau (-1)	Niveau (+1)
X1	Mode de préparation	plasma	Sans plasma
X2	Température (°C)	100	800
X3	Distance injecteur de TMA et porte substrat (cm)	4	8
X4	Puissance microonde (W)	1000	1600
X5	Vitesse de rotation de la pompe (1000 rpm)	16	42
X6	Débit de l'oxygène (O ₂) (sccm)	125	250
X7	Polarité (V)	0	500
X8	Débit de triméthyl aluminium (TMA) (sccm)	2	5

- 1- Quel est le nombre d'expérience à réaliser, Construire la matrice d'Hadamard en variables codées,
- 2- Ecrire la matrice transposée de la matrice d'Hadamard puis vérifier l'orthogonalité,
- 3- Ecrire le plan d'expérimentation correspondant la matrice d'expérience. Pourquoi il est important de toujours écrire le plan d'expérimentation ?
- 4- Après expérimentation, on trouve les résultats expérimentaux suivants :
 $Y = \{81.5 ; 29.2 ; 14.4 ; 11.8 ; 29.4 ; 30.9 ; 14.0 ; 31.0 ; 2.9 ; 2.4 ; 18.3 ; 14.5\}$.
- 5- Calculer les coefficients b_j ,
- 6- Retrouver les effets influents ($t_{\text{critique}} = 4,3$), déduire le modèle mathématique puis tracer qualitativement le diagramme des effets.
- 7- Dans quelles conditions doit-on travailler pour augmenter la réponse y .

TP3

La filtration des boues, provenant de l'étape d'épuration, permet de récupérer le filtrat (eau sucrée) et les tourteaux. Des analyses quantitatives de taux d'humidité et du pourcentage de sucre s'effectuent de façon journalière sur les tourteaux. Ces analyses ont montré que les recommandations de l'entreprise en taux de sucre et d'humidité sont largement dépassées. Ceci prouve un dysfonctionnement des paramètres des filtres presse. Une optimisation des paramètres du fonctionnement des filtres s'impose en utilisant la méthode des plans d'expérience.

Une étude d'influence de dix facteurs sur le pourcentage de sucre en utilisant la matrice d'HADAMARD dont la variance expérimentale est inconnue. Le domaine de variation des facteurs sont regroupés dans le tableau 1.

Tableau 1 : Les facteurs et leur domaine expérimental

	Facteurs	Niveau (-1)	Niveau (+1)
X1	Pression de boues (bar)	4	5
X2	Quantité d'eau désucrage (m ³ /h)	2	4
X3	Débit d'eau de désucrage (m ³ /h)	20	60
X4	Temps de compactage préliminaire (s)	20	70
X5	Temps de compactage intermédiaire (s)	30	80
X6	Temps de compactage final (s)	60	110
X7	Quantité d'eau pour lavage tourteaux (m ³)	1	3
X8	Débit d'eau pour lavage tourteaux (m ³ /h)	20	60
X9	Pression d'eau pour lavage tourteaux (bar)	3	5
X10	Temps de séchage tourteaux (s)	50	150

Les pourcentages en sucres (%) sont :

$$y = \{0.5; 0.43; 1.84; 0.98; 1.23; 1.41; 1.16; 1.22; 0.22; 0.20; 3.08; 3.2\}.$$

- Quel est le nombre d'expérience à réaliser ?
- Construire la matrice d'expériences en variables codées, vérifier l'orthogonalité.
- Calculer les coefficients b_j ,
- Quels sont les facteurs influents en utilisant les outils d'aide à l'interprétation en utilisant : l'approche de Pareto et le Half normal plot.
- Conclure

TP4

Lors de l'élimination des composés phénoliques des margines issues des huileries d'olive, nous avons associé des radiations UV à un système Fenton en phase homogène. Un plan de criblage, par la matrice Hadamard, a été choisi pour sélectionner les variables significatives pour cette oxydation photo catalytique. Les facteurs sélectionnés ainsi que leurs niveaux de variation sont indiqués dans le tableau I.

Tableau I. Domaine expérimental de chaque facteur.

Facteurs	Valeurs codées	Unités	Niveau	
			(-1)	(+1)
Type de catalyseur	X1	g L ⁻¹	Fe(II)	Cu(II)
Vitesse d'agitation	X2	Tr mn ⁻¹	300	700
pH	X3		2	5
Dilution	X4	Fois	60	80
UV	X5	nm	sans	avec

La réponse étudiée est la masse en composés phénoliques restante après traitement (en eq.g d'AG).

- 1- Quel est le nombre d'expérience à réaliser, Construire la matrice d'Hadamard en variables codées (permutation à droite),
- 2- Après expérimentation, on trouve les résultats expérimentaux suivants :

$$y = \left\{ \begin{array}{l} 1.70; 1.69; 4.64; 4.66; 3.14; 3.02; 2.12; 1.66; 2.17; \\ 2.43; 1.52; 1.36; 2.00; 2.26; 2.75; 2.88 \end{array} \right\}$$
- 3- Calculer les coefficients b_j ,
- 4- Retrouver les effets influents ($t_{critique} = 2.22$), déduire le modèle mathématique,
- 5- Conclure.

TP 5

La technologie micro-ondes (MO) est une technique mature qui trouve une large application dans les domaines de la science des matériaux et dans diverses applications environnementales. Cette technique peut être couplée avec les procédés d'oxydation avancée. Le principal avantage de cette combinaison est la stimulation de la production de radicaux libres oxydants et la rapide polarisation des molécules polluantes. Cette combinaison conduit à une obtention d'une température de réaction désirée en un délai plus court par rapport à la méthode d'oxydation thermique ou catalytique traditionnel. Ce qui favorise la dégradation rapide des polluants.

Dans cette étude, nous nous proposons de déterminer les effets des facteurs et leurs interactions sur l'élimination des composés phénoliques des margines en utilisant un système Fenton-like (H_2O_2/Cu) assisté par micro-ondes. Pour réaliser cette étude quantitative, nous avons opté pour un plan factoriel 2^k où k est le nombre de facteurs. Les conditions opératoires choisies pour exécuter ce plan d'expérience sont données par le tableau I.

Tableau I : Domaine expérimental de chaque facteur

Facteurs	Symbole	Niveau	
		(-1)	(+1)
[Cu ²⁺] (gL ⁻¹)	X1	0	0.5
[H ₂ O ₂] M	X2	4	12
Temps d'irradiation (min)	X3	1	8
Puissance d'irradiation (W)	X4	340	680

La réponse mesurée est la concentration en composés phénoliques. Les résultats obtenus sont : $Y =$

$$\left\{ \begin{array}{l} 15.33 ; 14.49 ; 11.95 ; 10.27 ; 13.56 ; 10.42 ; 14.49 ; 5.68 ; 16.51 ; 13.21 ; 15.23 ; \\ 11.65 ; 16.50 ; 10.93 ; 16.50 ; 8.48 \end{array} \right\}$$

1. Construire la matrice des effets,
2. Calculer les effets des facteurs et interactions,
3. La matrice étant saturée (autant d'effet que d'expérience), pour éviter d'utiliser les outils d'aide à l'interprétation, nous préférons utiliser des points au centre du domaine pour déterminer la variance expérimentale. Donner le centre du domaine de chaque facteur,
4. Cinq expériences au centre du domaine ont été faites, les résultats obtenus sont :

$$y = \{11.2 ; 11.018 ; 9.892 ; 10.677 ; 10.10\}$$
5. Retrouver les effets réellement influents puis donner le modèle mathématique élaboré (négliger toute interaction d'ordre \geq à 2),
6. Indiquer les conditions opératoires souhaitables pour diminuer la réponse y .

Références

- H. Iboukhoulef, Traitement des margines des huileries d'olive par les procédés d'oxydation avancée basés sur le système Fenton-like (H₂O₂/Cu), Thèse de Doctorat, UMMTO, 2014 (TP1).*
- S. Dkhissi, Etude des paramètres influents sur le recyclage de l'EDTA par le plan Plackett et Burman, mémoire de fin d'études en Sciences et Techniques CAC Agiq- Chimiométrie et Analyse Chimique, Maroc. 2015 (TP2).*
- B. Wahdam, Analyse et optimisation du fonctionnement de piles à combustibles par la méthode des plans d'expérience, Thèse de Doctorat, Université de technologie de Belfort Montbeliard et de l'université de Franche-Comte, 2006 (TP2).*
- F. Rais, A. Kamoun, M. Chaabouni, M. Claeys-Bruno, R. Luu, M. Sergent, Etude par plans d'expériences de la robustesse d'un procédé de préparation de sulfate-amide des acides gras de l'huile de grignons d'olives, Journal de la Société Chimique de Tunisie, 2011, 13, 191-202 (TP3).*
- T. Khalil, Optimisation des paramètres du fonctionnement des filtres presse par les plans d'expérience, mémoire de fin d'études, Université Sidi Mohammed Ben Abdellah, Maroc, 2015 (TP3).*
- O. Sebaoui, R. Moussaoui, H. Kadi, P. Michaud, C. Delattre, Kinetic Modeling of Pectin Extraction from Wasted Citrus Lemon, Waste and Biomass Valorization, 2017, 8, 2329- 2337 (TP4).*
- H. Iboukhoulef, H. Kadi, R. Moussaoui, Homogenous photo-Fenton catalytic oxidation of olive mill wastewater using Plackett and Burman's design, materials research and applications, materials horizons: from nature to nanomaterial's, Springer, https://doi.org/10.1007/978-981-15-9223-2_15, 273-281, 2021 (TP4).*
- T. Silou, M. Malanda, L. Loubaki, Optimisation de l'extraction de l'huile essentielle de *Cymbopogon citratus* grâce à un plan factoriel complet 2³, Journal of Food Engineering, Volume 65, Issue 2, 2004, 219-223 (TP5).*
- H. Iboukhoulef, A. Amrane, H. Kadi, Optimization of phenolic compounds abatement in olive mill wastewater by Fenton's like treatment with H₂O₂/Cu under microwave using experimental design, Environmental Engineering and Management Journal, Vol.17, No. 5, 1267-1274, 2018 (TP5).*
- J. Goupy, Introduction aux plans d'expériences, Dunod- Paris, 2001,*
- J. Goupy, Plans d'expériences pour surfaces de réponse, Dunod. Paris, 1999,*
- J. Goupy, Les plans d'expériences, Revue modulad, 34, 2006, pp74-116.*