

Modélisations et Optimisation par les Plans d'Expériences

1. Introduction



Lors de l'étude d'un phénomène, plusieurs questions se posent, auxquelles répondent différents types de plans. On peut distinguer 3 grandes étapes dans l'acquisition des connaissances

1. *Recherche des facteurs influents.*
2. *Modélisation : quand les facteurs influents ont été identifiés et leur importance quantifiée, on recherche ensuite l'équation permettant de décrire les variations de la réponse étudiée en fonction de celles des facteurs influents.*
3. *Optimisation : déterminer quelles conditions expérimentales (les valeurs prises par les facteurs influents) permettent d'obtenir le meilleur résultat pour la réponse.*

1.1. Recherche des facteurs influents

Parmi tous les facteurs susceptibles d'influer sur le phénomène (c'est à dire sur la ou les réponses mesurées du phénomène),

- lesquels ont une influence significative ?
- que vaut quantitativement cette influence ?
- existe-t'il des interactions entre facteurs ?

Les plans qui permettent de rechercher les facteurs influents sont:

- les plans factoriels complets,
- les plans factoriels fractionnaires.

1.2. Modélisation

Quand les facteurs influents ont été identifiés et leur importance quantifiée, on recherche ensuite l'équation permettant de décrire les variations de la réponse étudiée en fonction de celles des facteurs influents ; cette seconde étape constitue la **modélisation**.

Les plans factoriels suffisent parfois. Il se peut dans d'autres cas que l'on soit obligé de faire appel à des plans plus complexes tels que **les plans composites centrés**.

1.3. Optimisation



Pour finir, on cherche en général à déterminer quelles conditions expérimentales (les valeurs prises par les facteurs influents) permettent d'obtenir le meilleur résultat pour la réponse. Cette étape porte le nom d'**optimisation**.

Il existe plusieurs méthodes; citons:

- l'étude des courbes iso réponses,
- la méthode du simplex.

2. Généralités



- Pour un système (industriel, de laboratoire...) dont l'état dépend de *variables opératoires* le terme **Optimisation** désigne une *action* qui vise à trouver l'ensemble des *valeurs de ces variables* qui entraîne un *état souhaité* pour le système étudié.
- Les problèmes d'optimisation sont très divers et variés d'un point de vue des applications.

2. Généralités



Dans tout problème d'optimisation, il intervient 2 catégories de variables:

- Les facteurs (ou variables opératoires x). Ces facteurs doivent être de nature **quantitative** (ex : pH, température...) et il doit être possible de les faire varier à volonté (**facteurs contrôlés**) dans les limites permises par l'expérimentation.
- La réponse (ou fonction de réponse y) doit être aussi une variable **quantitative**; c'est elle qui sera dite «optimisée». La dénomination «fonction de réponse» indique que y doit dépendre des valeurs x prises par les variables opératoires.

2. Généralités



Bien que le **qualificatif** « **état souhaité pour le système** » n'implique pas nécessairement la détermination d'un **extremum** (on peut par exemple être satisfait dans certaines études par une réponse y supérieure à une valeur seuil prédéterminée ...), nous privilégierons dans la suite du chapitre l'**optimisation** comme la **recherche d'un maximum** (ex : rendement réaction) ou **d'un minimum** (ex : coût d'une opération) de la fonction de réponse y en faisant varier convenablement les variables opératoires x (les facteurs) pour faciliter cette localisation.

2. Généralités

Il s'agit donc le plus souvent de trouver l'optimum vrai de la réponse, soit d'un point de vue mathématique de rechercher l'extremum (minimum ou maximum) d'une fonction y de plusieurs variables :

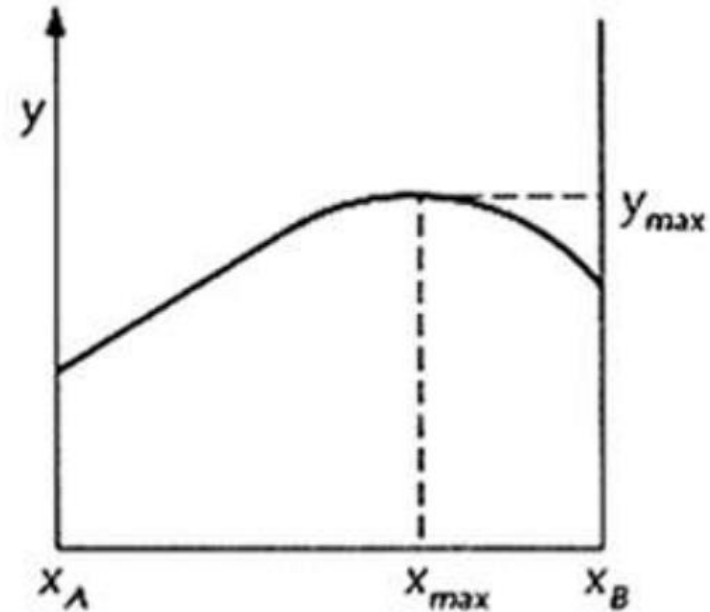
$$y = f(x_1, x_2, \dots, x_n)$$

- Cas le plus simple : une seule variable x ;
- Cas le plus fréquent : plusieurs variables opératoires, 2, 3, voire 4 ou 5.

2. Généralités

Optimisation à une variable opératoire

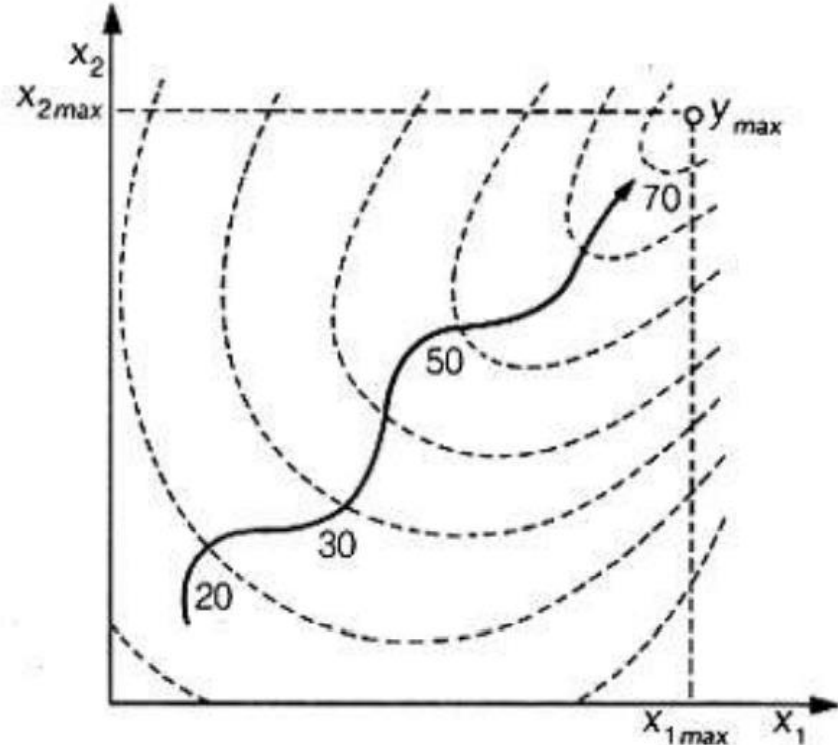
- La réponse y est fonction d'une seule variable x et on recherche la valeur x_{max} de x , comprise entre les limites expérimentales x_A et x_B , qui rend la valeur de y optimale, c'est à dire y_{max} .



2. Généralités

Optimisation à deux variables opératoires

- La fonction objectif y dépend de 2 variables x_1 et x_2 ; elle est représentée sous forme de courbes de même niveau de réponse, dites courbes d'isoréponses (en pointillés). On recherche donc les valeurs x_{1max} et x_{2max} qui correspondent à la valeur y_{max} **optimale**.



2. Généralités

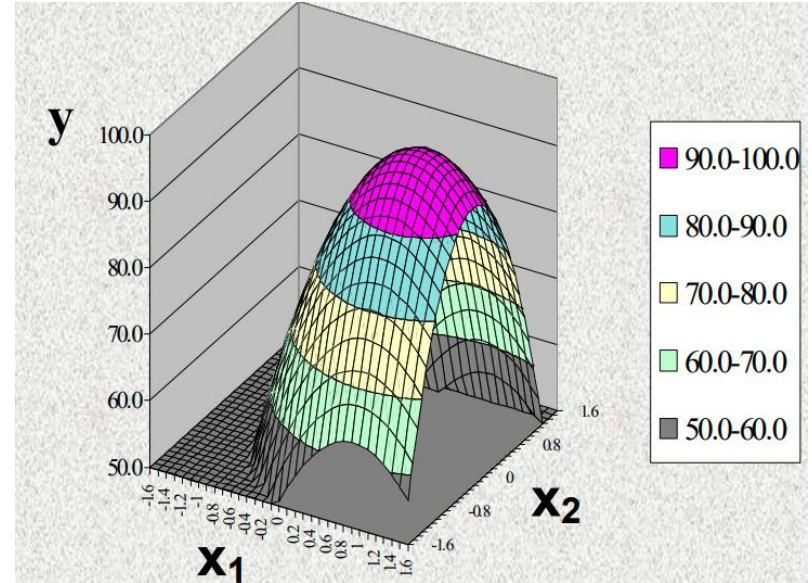
Stratégie expérimentale d'optimisation: Quand plusieurs variables opératoires x agissent sur une réponse y à optimiser:

- La **méthode intuitive** qui consiste à faire varier « *un seul facteur à la fois en maintenant tous les autres constants* » n'est pas efficace : elle nécessite d'une part d'avantage d'essais et d'autre part on risque de ne pas atteindre l'optimum s'il existe des interactions entre les facteurs.
- Il est recommandé de **faire varier simultanément tous les facteurs de façon organisée**, à l'aide d'une **méthodologie adéquate**.

2. Généralités

La notion de surface de réponse

Les courbes d'isoréponses sont les projections sur le plan $[Ox_1x_2]$ de la surface de réponse représentant la variation de y en fonction de x_1 et x_2 dans l'espace à 3 dimensions. Notons que cette notion de surface de réponse se généralise au cas de plus de 2 facteurs.



Le principe de toutes les méthodes d'optimisation consiste à explorer cette surface de façon à localiser un éventuel extremum dans un domaine expérimental donné.

3. Classification des techniques d'optimisation

Il existe 2 grandes catégories de méthodes :

- **La méthode des plans d'expériences**

méthode indirecte : modélisation puis optimisation

Déjà étudiée à propos des plans factoriels, cette méthode vise à estimer l'équation de la surface de réponse avec le modèle le plus simple possible, ***nécessairement ici au moins du 2ème degré pour restituer un maximum ou un minimum***. La position de cet optimum se déduit ensuite par dérivation de l'équation et donne les valeurs des variables opératoires.

- **La méthode du simplex et ses variantes**

méthode directe : optimisation

L'optimum est localisé grâce à un algorithme itératif: la progression vers celui-ci sur la surface de réponse se fait pas à pas, ***par étapes successives***, sans jamais avoir à déterminer l'équation de la surface.

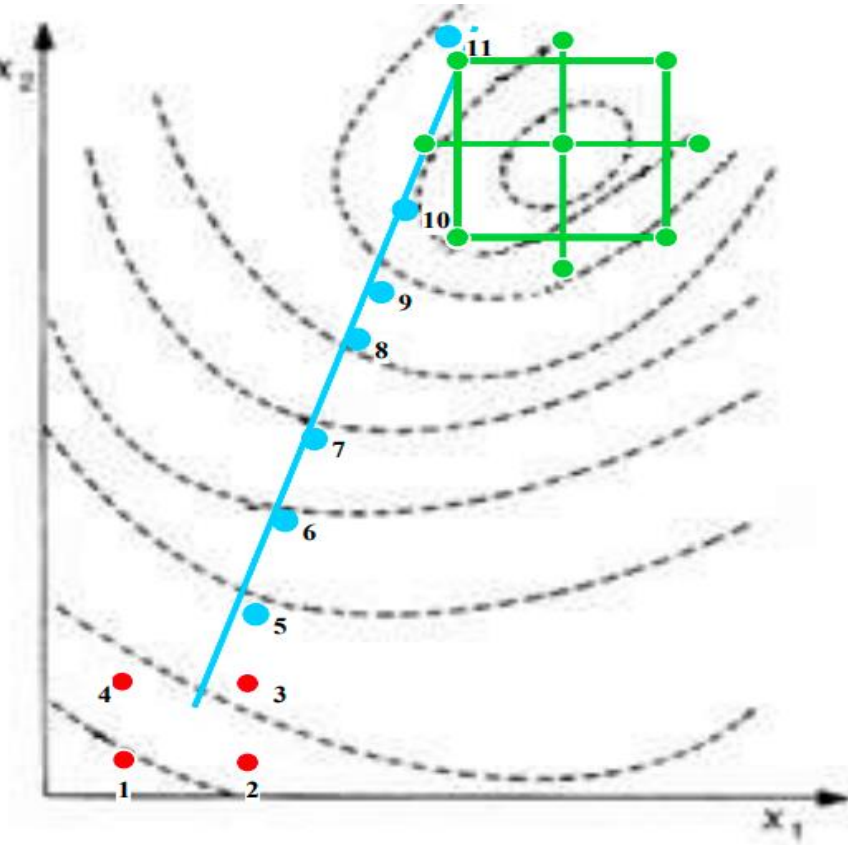
3.1 La méthode des plans d 'expériences

Dans ce cas, il faut *postuler a priori un modèle* pour la surface de réponse puis *estimer son équation* par la technique statistique de *régression linéaire multiple*.

- La valeur des coefficients de l'équation ainsi que leur précision dépendent fortement de l 'organisation des essais, d'où *la nécessité de planifier les expériences*.
- Parmi les divers plans mis au point, les **plans composites centrés (PCC)**, les plans de **BOX-Behnken**, les plans de **DOEHLERT...**, nous ne détaillerons que ces derniers parce que ce sont les plus utilisés.

Cheminement vers l'optimum

(cas avec 2 facteurs x_1 et x_2 courbes d'isoréponses en pointillé)



Localisation MAXIMUM
Plan Composite Centré

meilleure réponse



Expérimentation ligne de plus grande pente

direction optimum



équation 1^{er} degré



facteurs/interactions avec effet



Plan factoriel complet (ou fractionnaire)

3.1 La méthode des plans d 'expériences

Précisions

- **Ligne de plus grande pente.**
 - indique le chemin le plus rapide pour progresser vers l'optimum,
 - équation: obtenue à partir de l'équation du 1^{er} degré du plan factoriel initial,
 - C'est une droite en l'absence d'interaction, sinon une courbe
- Dans la suite nous supposerons avoir déjà une bonne idée de la région où se situe l'optimum et réaliserons directement le **Plan Composite Centré** qui la couvre.

3.1 Les modèles mathématiques

Les modèles utilisés appartiennent tous au *modèle linéaire*, ce qui signifie que les coefficients a_i et a_{ij} de l'équation interviennent au 1er degré. Ce modèle peut être du 1er ou du 2ème degré en x .

- **Le modèle du 1er degré** (en x) : des rappels.

Dans le cas le plus simple de 2 variables opératoires x_1 et x_2 (plan factoriel 2^2),

la réponse estimée \hat{y} s'écrit : $\hat{y} = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2$

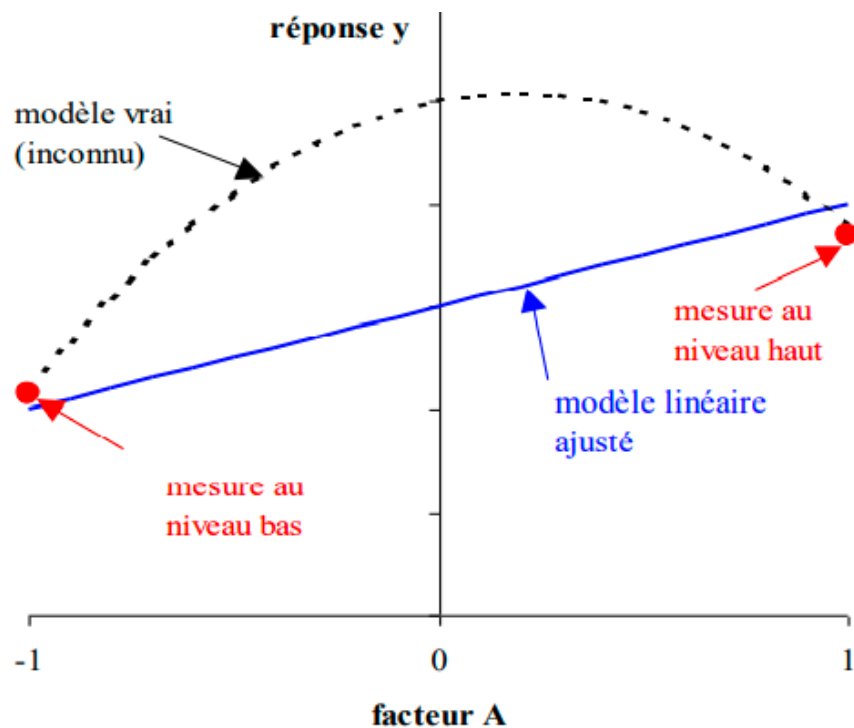
a_1 et a_2 sont les **effets principaux**, a_{12} l'**effet d'interaction** et a_0 la **moyenne** des 4 réponses, estimation de la réponse au centre du domaine expérimental.

3.1 Les modèles mathématiques

Validité du modèle du 1er degré

- Dans les plans factoriels 2^n chaque facteur n'est expérimenté qu'en 2 points, aux extrémités -1 et $+1$ du domaine. Cela explique le choix empirique d'une équation du 1er degré en fonction des facteurs centrés réduits X : par 2 points, il passe une droite et une seule ; mais il faut aussi ajouter qu'il peut y passer une infinité de courbes d'équations diverses ...
- Il est donc important de s'assurer de la linéarité de l'équation dans tout le domaine en réalisant des expériences complémentaires avec un 3ème point à l'intérieur du domaine. On choisit généralement le centre ($X = 0$) situé à égale distance des extrémités expérimentées.

3.1 Les modèles mathématiques



3.1 Les modèles mathématiques

Validité du modèle du 1er degré

En faisant des répétitions au centre du domaine expérimental, on peut donc juger la linéarité de l'équation de prédiction en comparant statistiquement

- *la moyenne y_0 de ces répétitions.*
- *la réponse prédite au centre par l'équation linéaire, égale à la moyenne \bar{y} des n réponses du plan factoriel 2^n .*

3.1 Les modèles mathématiques

Ces 2 moyennes doivent être en théorie égales et en pratique peu différentes lorsque le modèle linéaire est valide. Et, Lorsqu'on obtient des valeurs très différentes, s'écartant de plusieurs fois l'écart type, cela signifie que le modèle linéaire n'est pas valable et qu'il faut envisager un modèle empirique plus complexe, où les **facteurs centrés réduits interviennent au 2ème degré** par exemple. Notons qu'il existe des tests statistiques d'écart à la linéarité du modèle

3.1 Les modèles mathématiques

Le modèle du 2ème degré

Le modèle du 1er degré ne peut pas convenir pour caractériser un optimum : une telle fonction ne comporte ni minimum, ni maximum.

Le modèle du 2ème degré en diffère par l'existence de termes supplémentaires du 2ème degré en x :

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

Les coefficients a_{11} et a_{22} attachés aux carrés des variables x_1 et x_2 sont primordiaux pour le calcul de l'extremum.

3.1 Les modèles mathématiques



Ce modèle contient **6 paramètres** (en incluant a_0) et **un minimum de 12 essais** est donc souhaitable pour appliquer convenablement la régression linéaire multiple.

Cette condition étant remplie, il est possible **de tester la validité du modèle**, choisi empiriquement, et d'identifier les coefficients qui diffèrent statistiquement de 0 ; *seuls les termes significatifs figureront dans l'équation de la surface de réponse.*

3.2 Le plan composite centré à 2 facteurs

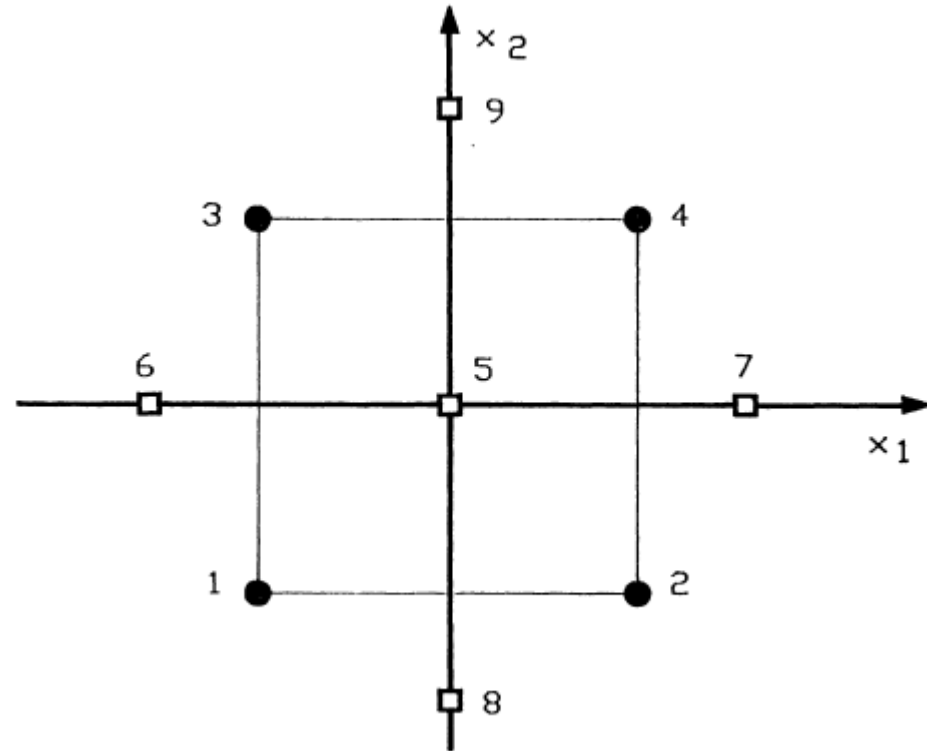
- Les plans composites centrés ont été conçus pour assurer une précision à peu près uniforme des estimations de réponse \hat{y}_e dans la totalité du domaine expérimenté, avec le plus petit nombre d'essais possible ...
- Ces plans comprennent 3 catégories d'essais établis de telle sorte qu'à chaque variable x correspondent 3 niveaux :
 - **les essais du plan factoriel** (pour 2^n et $n = 2$, il y en a 4)
 - **des essais « en étoile »** (au total $2n$, soit 4 pour 2 facteurs) et dont la distance à l'origine des coordonnées dépend de n (1,414 pour $n = 2$).
 - **des essais au centre du domaine**, 3ème niveau de chaque facteur, le nombre de répétitions augmentant avec le nombre de facteurs étudiés (pour $n = 2$, il est habituel d'en prendre 5)

3.1 Les modèles mathématiques

Le plan composite centré à 2 facteurs

Le plan composite centré à 2 facteurs cc
essais:

- 4 essais factoriels,
- 4 essais en étoiles (a),
- 5 essais au centre du domaine.



3.1 Les modèles mathématiques

- 4 **essais factoriels**, (phase 1)
- 4 **essais en étoiles** (α), (phase 3)
- 5 **essais au centre du domaine**
(phase 2)

<i>N° essai</i>	<i>X1</i>	<i>X2</i>
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5	-1,414	0
6	1,414	0
7	0	-1,414
8	0	1,414
9	0	0
10	0	0
11	0	0
12	0	0
13	0	0

3.1 Les modèles mathématiques

Analyse des résultats

- **Estimation de l'équation du modèle** par la **régression linéaire multiple**; on ne tient compte que des coefficients significativement différents de 0 ($\alpha = 0,05$ ou plus souvent $0,10$).
- **Localisation de l'extremum** : par calcul, en dérivant y par rapport à x_1 (à x_2 constant) puis par rapport à x_2 (à x_1 constant) et en recherchant la valeur du couple $[x_1, x_2]$ qui annule les 2 dérivées. Lorsque l'extremum calculé est en dehors du domaine expérimental (x_1 et/ou $x_2 > 1,414$), il faut refaire une autre expérimentation : il ne faut jamais extrapoler hors du domaine.
- **Validation** : la prédiction de l'optimum est mathématique, à l'aide d'un modèle empirique... Un essai de contrôle est indispensable à cet optimum supposé : la réponse mesurée doit correspondre à celle qui est prédite.

3.1 Les modèles mathématiques

Visualisation des résultats

Diagramme de surface de réponse 3D: la réponse passe par un maximum (ce type de graphe est difficile à exploiter)

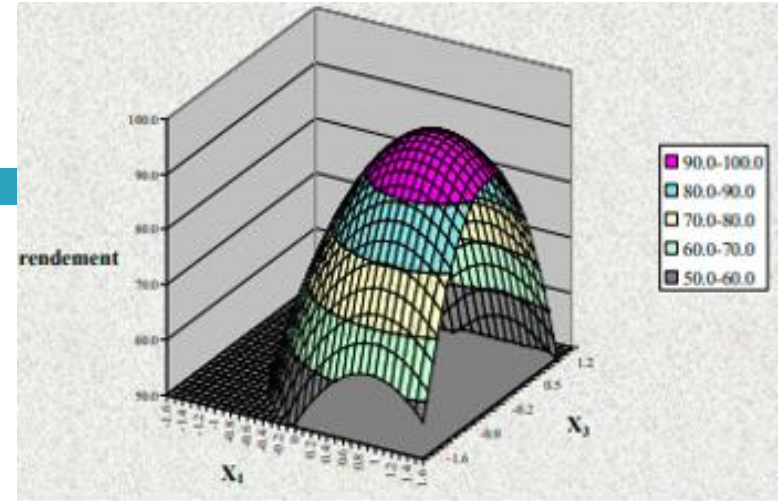
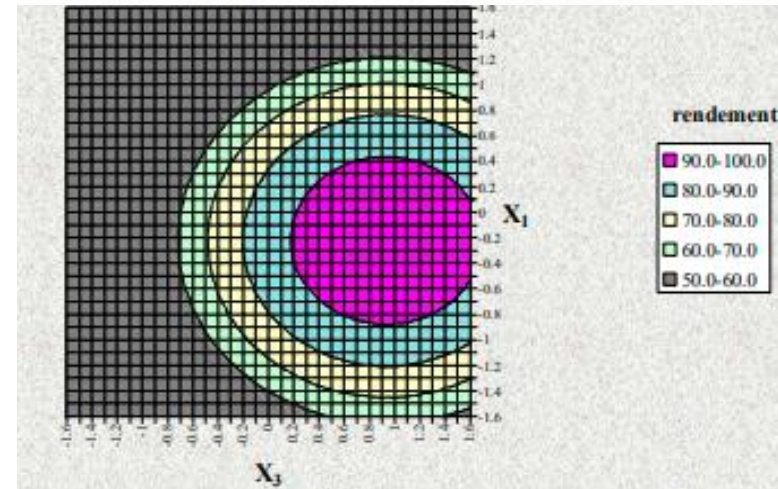


Diagramme des courbes d'isoréponses: Projection du diagramme 3D plus facile à visualiser la région du maximum avec les courbes de niveau



3.1 Les modèles mathématiques

Le plan composite centré à 3 facteurs

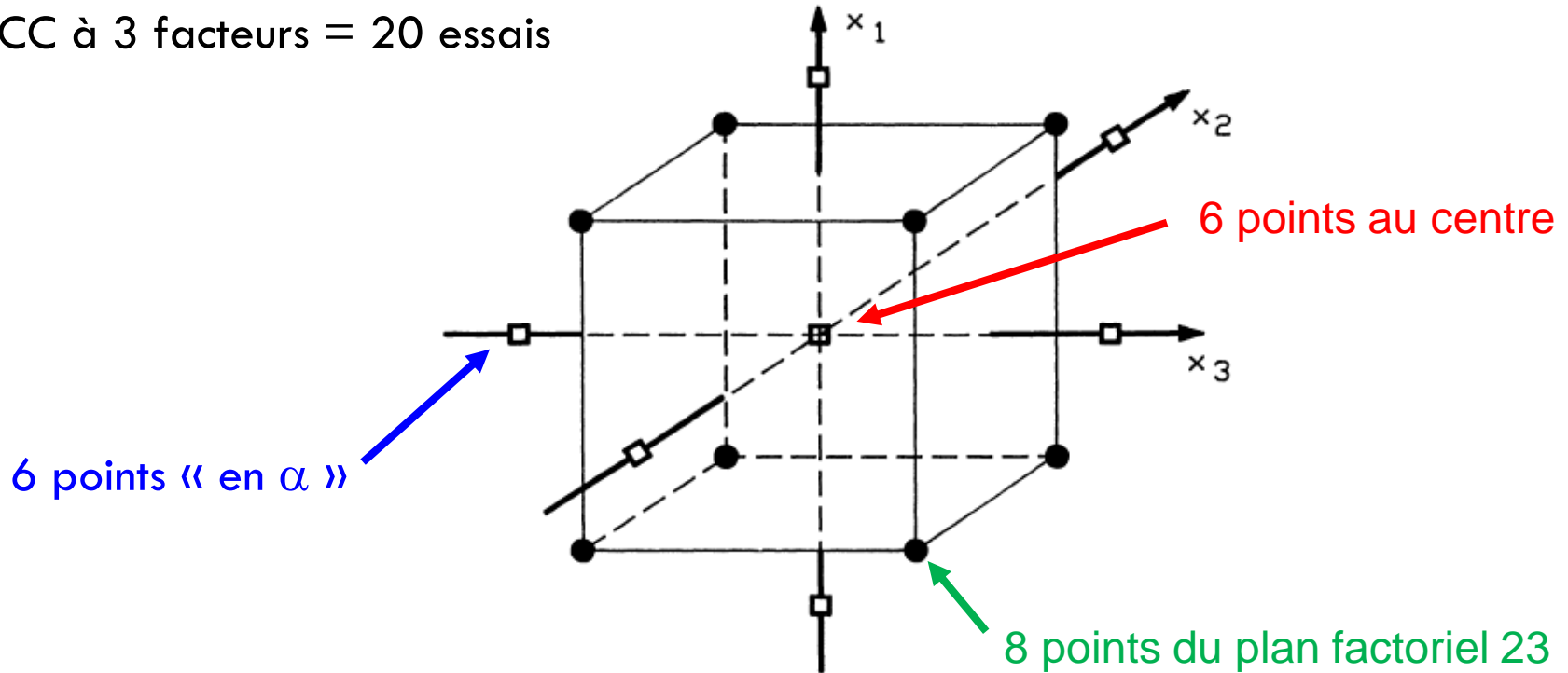
- Il comporte 3 catégories d'essais :
 - les essais du plan factoriel : 8 aux sommets du cube de côté 1.
 - Les essais « en étoile » : 6 points situés sur les 3 axes, 1 par face du cube, situés tous à la même distance de l'origine, égale à $\alpha = 1,682$ en variable réduite.
 - Les essais au centre : il faut prévoir 6 répétitions pour que la précision des estimations soit à peu près constante dans le domaine expérimental.

Au total 20 essais.

3.1 Les modèles mathématiques

Le plan composite centré à 3 facteurs

PCC à 3 facteurs = 20 essais



3.1 Les modèles mathématiques

Le plan composite centré à 3 facteurs

L'équation du modèle

- S'écrit :

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

- On tient compte des 3 interactions d'ordre 2 (a_{12} , a_{13} et a_{23}) mais pas de l'interaction d'ordre 3. Au total, ce modèle comporte 10 paramètres à estimer. Comme l'expérimentation comportera 20 réponses, la précision des estimations sera convenable ($n \geq 2p$).
- Dans l'équation du modèle, on ne tient compte que des termes dont le coefficient est significativement différent de 0.

3.1 Les modèles mathématiques

Le plan composite centré à 3 facteurs

Analyse et interprétation: quand les 20 réponses y sont connues, il faut :

- 1 - calculer les coefficients a_{ij} du modèle au moyen de la régression linéaire multiple. Ne retenir que les coefficients significativement différents de 0.
- 2 - Déterminer par calcul les valeurs des facteurs réduits qui rendent la réponse y optimale (par dérivation) à l 'aide de l'équation déterminée. Transformer ces valeurs réduites en valeurs réelles.
- 3 – Tracer les graphiques (courbes de niveaux en prenant les variables 2 par 2).
- 4 – Valider en faisant un ou plusieurs essais de contrôle à l'extremum.

3.1 D'autres modèles

Plan de Box-Behnken

Les plans de Box-Behnken sont des plans économiques comportant moins d'essais que le PCC correspondant au même nombre de facteurs pour localiser un optimum; par exemple pour 3 facteurs, il n'y a que 15 essais à réaliser au lieu de 20.

Il s'analyse par la régression linéaire multiple comme le PCC.

Du point de vue statistique, ces plans sont moins sûrs et couvrent un domaine expérimental moins étendu que le plan composite.

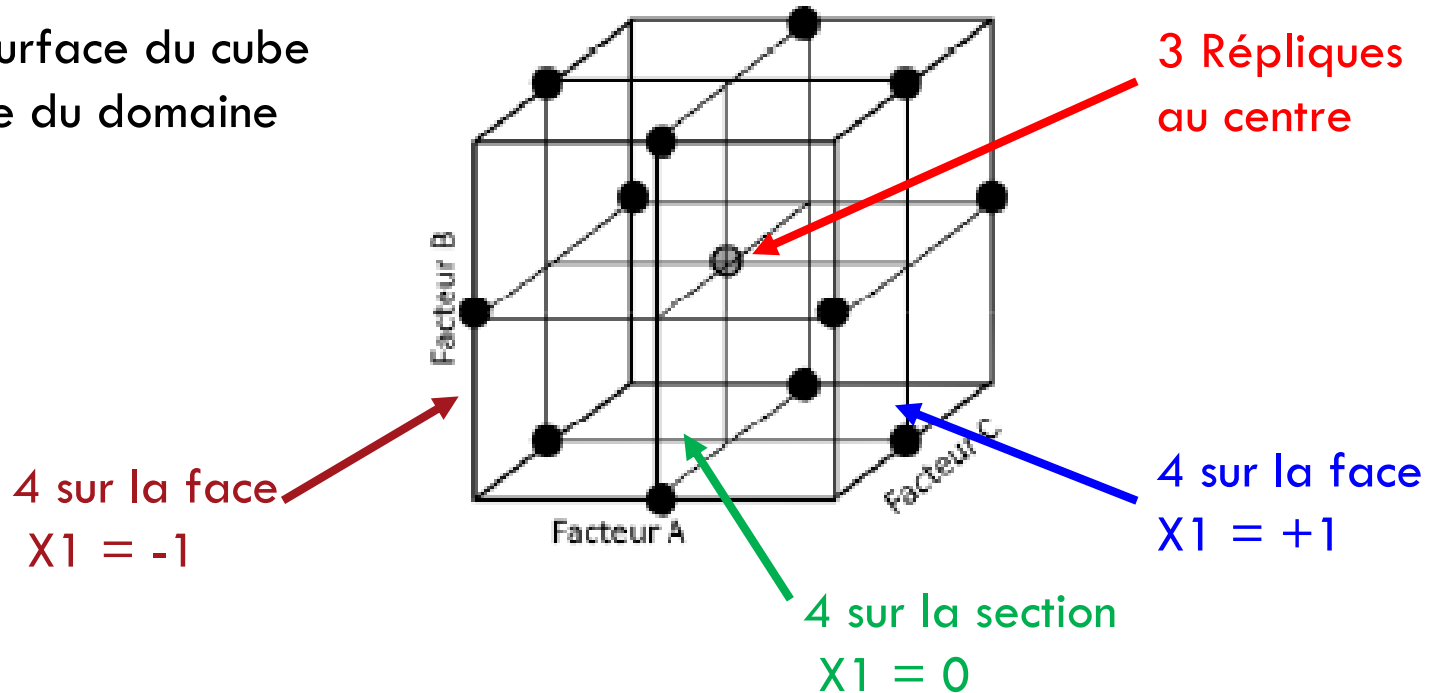
3.1 D'autres modèles

Plan de Box-Behnken

Plan de **BOX-BEHNKEN** à 3 facteurs: **15 essais**

12 points sur la surface du cube

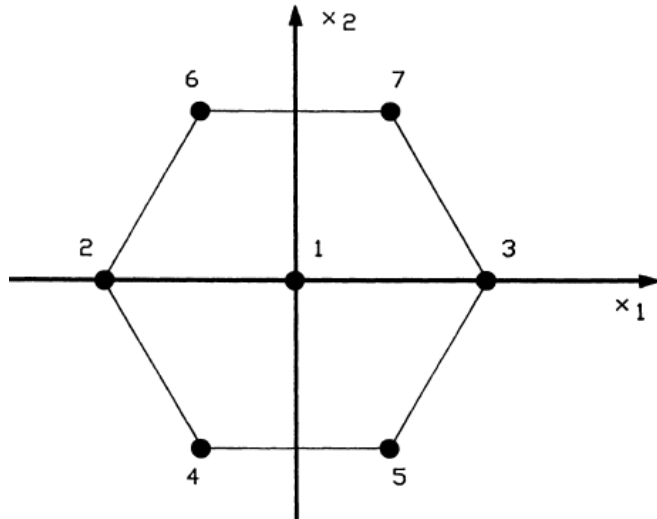
3 points au centre du domaine



3.1 D'autres modèles

Plan de Doehlert

Répartition uniforme des points expérimentaux dans l'espace expérimental. Tous les points sont à la même distance du centre du domaine et sont régulièrement disposés sur le cercle trigonométrique. Ils forment un hexagone régulier.

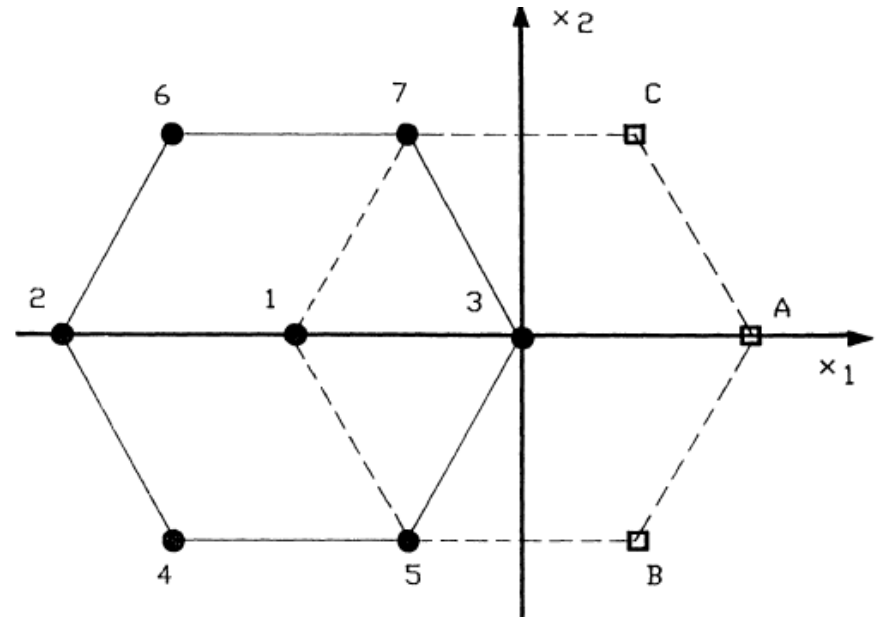


N° de l'essai	x_1	x_2
1	0	0
2	-1	0
3	+1	0
4	-1/2	$-\sqrt{3}/2$
5	+1/2	$-\sqrt{3}/2$
6	-1/2	$+\sqrt{3}/2$
7	+1/2	$+\sqrt{3}/2$

3.1 D'autres modèles

Plan de Doehlert

L'expérimentateur peut facilement ajouter des essais en conservant la disposition régulière des points expérimentaux dans le plan. Par exemple, les trois points A, B et C ajoutés aux points déjà existants 1, 3, 5, et 7 donnent à nouveau un hexagone régulier.



3.1 D'autres modèles

Plan de Doehlert

Dans un espace à 3 dimensions, les points expérimentaux se répartissent régulièrement sur une sphère de rayon unité (en variables centrées réduites).

- 6 points de l'hexagone régulier
- 6 points supplémentaires
- 1 point au centre

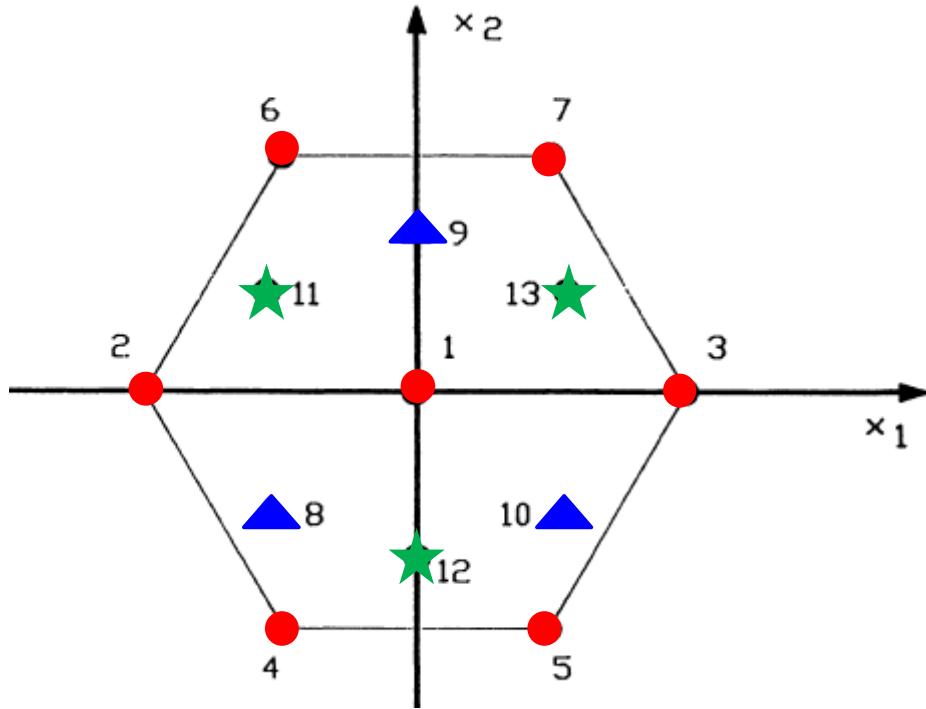
Soit 13 points au total

N° de l'essai	Facteur 1	Facteur 2	Facteur 3
1	0	0	0
2	-1	0	0
3	1	0	0
4	-1/2	$-\sqrt{3}/2$	0
5	+1/2	$-\sqrt{3}/2$	0
6	-1/2	$+\sqrt{3}/2$	0
7	+1/2	$+\sqrt{3}/2$	0
8	-1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
9	0	$+1/\sqrt{3}$	$-\sqrt{2/3}$
10	+1/2	$-1/2\sqrt{3}$	$-\sqrt{2/3}$
11	-1/2	$+1/2\sqrt{3}$	$+\sqrt{2/3}$
12	0	$-1/\sqrt{3}$	$+\sqrt{2/3}$
13	+1/2	$1/2\sqrt{3}$	$+\sqrt{2/3}$

3.1 D'autres modèles

Plan de Doehlert

- Points expérimentaux dans le plan x_1x_2
- ▲ Points expérimentaux sous le plan x_1x_2
- ★ Points expérimentaux au dessus du plan x_1x_2



4. Plan Composite Centré (PCC)

Il s'agit de déterminer les conditions expérimentales optimales de l'élimination du Pb (II) sur un matériau adsorbant à base d'oxyde de zirconium hydraté.

- Trois (3) facteurs sont étudiés :
 - x_1 est le pH de la solution de Pb (II) .
 - x_2 est la masse de l'adsorbant.
 - x_3 est la concentration initiale du polluant.
- La réponse Y est le taux d'élimination du Pb (II) en solution aqueuse.

Domaines d'étude:

Le tableau suivant permet de déterminer les conditions expérimentales à réaliser :

Variables	Unit	Factor	Range and level				
			- α	-1	0	+1	+ α
pH (-)	(-)	X_1	0.745	3.200	6.800	10.400	12.854
Dosage	(g)	X_2	0.009	0.012	0.016	0.020	0.023
Initial Concentration	(mg/L)	X_3	9.546	30.000	60.000	90.000	110.454

Matrice des essais:

n	X1	X2	X3
1	-1	-1	-1
2	1	-1	-1
3	-1	1	-1
4	1	1	-1
5	-1	-1	1
6	1	-1	1
7	-1	1	1
8	1	1	1
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0
14	0	0	0
15	- α	0	0
16	+ α	0	0
17	0	- α	0
18	0	+ α	0
19	0	0	- α
20	0	0	+ α

Critère d'optimalité:

- Nombre de points expérimentaux:

$$N = 2^k + 2.k + n_0$$

pts factoriels pts en étoile points au centre

- Critère d'isovariance par rotation:

$$\alpha = \sqrt[4]{2^k}$$

- Critère de presque orthogonalité :

$$\alpha = \left(\frac{2^k (\sqrt{2^k + 2k + n_0} - \sqrt{2^k})^2}{4} \right)^{\frac{1}{4}}$$

- Critère de précision uniforme:

$$n_0 = \frac{(k + 3) + \sqrt{9k^2 + 14k - 7}}{4(k + 2)} \left(\sqrt{2^k + 2} \right)^2 - 2^k - 2k$$

Valeurs de α et de n_0

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_e		4	6	8	10	10	12	12
n_0	isovariant par rotation	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1	≥ 1
	précision uniforme	5	6	7	10	6	15	9
	orthogonalité	8	12	12	17	10	24	15
ϖ		1.41	1.68	2	2.38	2.00	2.83	2.38

Plan de second ordre:

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

Matrice des essais:

Exp. Run	Space type	Coded factors			Uncoded factors			Observed response (%)	Predicted response (%)
		X ₁	X ₂	X ₃	pH	Adsorbent dose (g)	Initial concentration (mg/L)		
1	Axial	0	0	+ α	6.8	0.016	110.5	57.36	57.52
2	Center	0	0	0	6.8	0.016	60	99.71	99.37
3	Factorial	-1	+1	-1	3.2	0.020	30	46.94	46.97
4	Factorial	-1	-1	-1	3.2	0.012	30	40.25	40.28
5	Axial	0	0	- α	6.8	0.016	9.5	65.47	65.38
6	Factorial	+1	+1	+1	10.4	0.020	90	65.26	65.19
7	Factorial	+1	+1	-1	10.4	0.020	30	81.21	81.24
8	Axial	0	- α	0	6.8	0.009	60	66.36	66.39
9	Factorial	-1	-1	+1	3.2	0.012	90	47.06	46.98
10	Center	0	0	0	6.8	0.016	60	99.71	99.37
11	Center	0	0	0	6.8	0.016	60	99.71	99.37
12	Axial	+ α	0	0	12.9	0.016	60	36.79	36.82
13	Factorial	+1	-1	+1	10.4	0.012	90	47.29	47.21
14	Axial	- α	0	0	0.75	0.016	60	7.78	7.81
15	Center	0	0	0	6.8	0.016	60	99.71	99.37
16	Center	0	0	0	6.8	0.016	60	99.71	99.37
17	Axial	0	+ α	0	6.8	0.022	60	87.1	87.13
18	Center	0	0	0	6.8	0.016	60	99.71	99.37
19	Factorial	+1	-1	-1	10.4	0.012	30	43.74	43.77
20	Factorial	-1	+1	+1	3.2	0.020	90	34.26	34.18

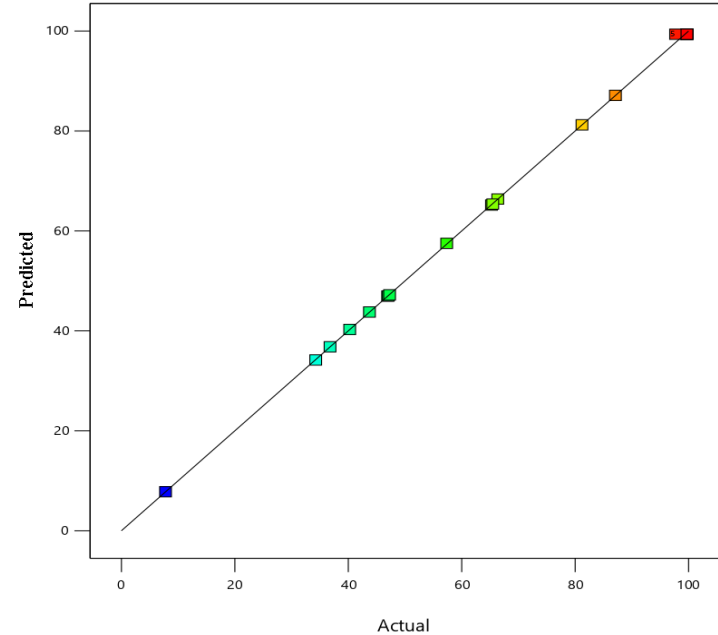
Source	Sum of square	df	Mean square	F-value	P-value	Remarks
Sequential sum of square						
Mean vs total	87533.65	1	87533.65			
Linear vs Mean	1609.55	3	536.52	0.6449	0.5973	
2FI vs Linear	668.91	3	222.97	0.2293	0.8743	
Quadratic vs 2FI	12638.02	3	4212.67	12396.85	< 0.0001	Suggested
Residual	3.34	6	0.5574			
Total	1.025E + 05	16	5122.66			
Analysis of variance (ANOVA)						
Model	14916.48	9	1657.39	4877.28	<0.0001	Significant
X ₁ -pH	1015.74	1	1015.74	2989.08	<0.0001	
X ₂ -Adsorbent dose	519.25	1	519.25	1528.04	< 0.0001	
X ₃ -Initial Concentration	74.56	1	74.56	219.40	<0.0001	
X ₁ X ₂	473.55	1	473.55	1393.54	<0.0001	
X ₁ X ₃	5.33	1	5.33	15.69	0.0027	
X ₂ X ₃	190.03	1	190.03	559.20	<0.0001	
X ₁ ²	10696.25	1	10696.25	31476.42	< 0.0001	
X ₂ ²	921.03	1	921.03	2710.35	<0.0001	
X ₃ ²	2591.19	1	2591.19	7625.22	<0.0001	
Residual	3.40	10	0.3398			
Lack of Fit	0.0648	5	0.0130	0.0195	0.9997	Not significant
Pure Error	3.33	5	0.6667			
Cor Total	14919.88	19				
Goodness of fit						
Std. Dev.	0.5829		R ²	0.9998		
Mean	66.16		Adjusted R ²	0.9996		
C.V. %	0.8812		Predicted R ²	0.9996		
PRESS	5.29		Adeq Precision	222.1264		
VIF	1.0					

Interprétation:

- La p-valeur $< 0,0001$ (valeur de F élevée) du **modèle quadratique** confirme que le modèle quadratique est hautement significatif.
- La p-valeur $> 0,9997$ du **manque d'ajustement** indique que le manque d'ajustement est insignifiant.
- Tous les termes du modèle ont des valeurs $p < 0,05$ montrent qu'ils sont significatifs.
- Les valeurs élevées de R^2 (0,9997), R^2 ajusté (0,9996) et R^2 prévu (0,9996) démontrent qu'il y a une excellente relation entre les réponses expérimentales et les réponses prédites.
- Les faibles valeurs du coefficient de variation (C.V = 0,8812%) et de l'écart type (0,5829) suggèrent que la précision est satisfaisante.

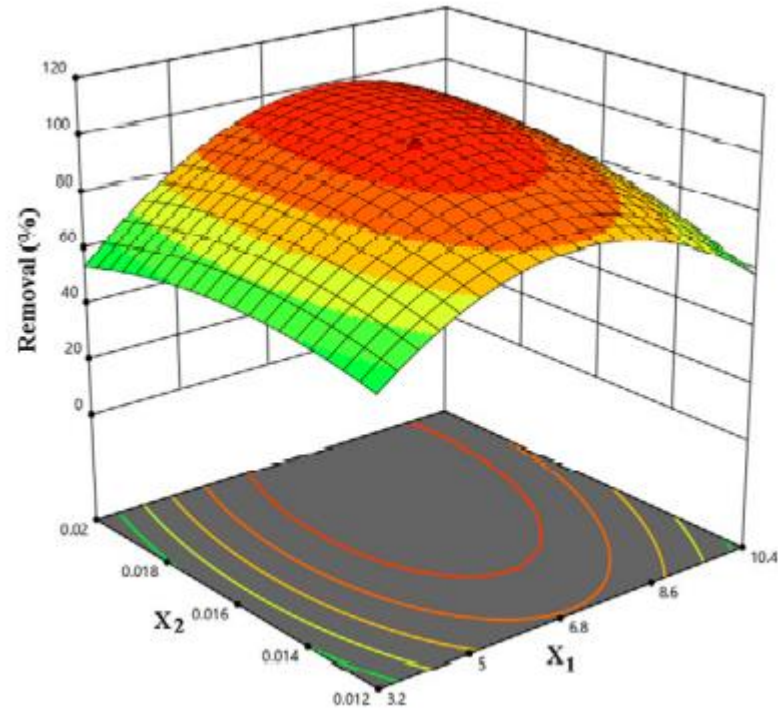
$$\text{Removal (\%)} = + 99.37 + 8.62 X_1 + 6.17 X_2 - 2.34 X_3 + 7.69 X_1 X_2 - 0.8162 X_1 X_3 - 4.87 X_2 X_3 - 27.24 X_1^2 - 7.99 X_2^2 - 13.41 X_3^2$$

La faible valeur d'erreur résiduelle prédite de somme des carrés (PRESS = 9,13) indique que le modèle quadratique correspond à chaque point du plan.

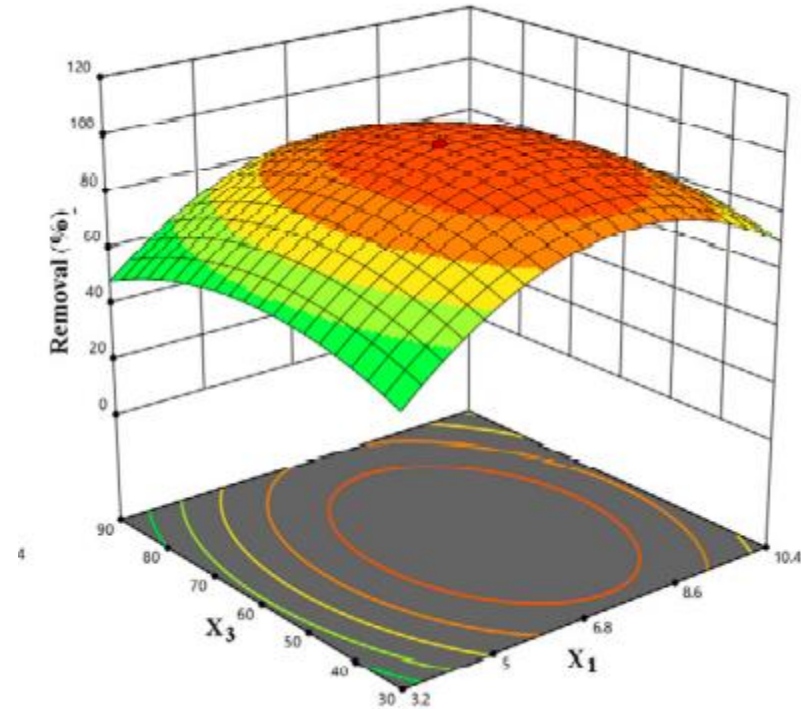


Graphique du taux d'élimination du Pb (II) prévu vs réel

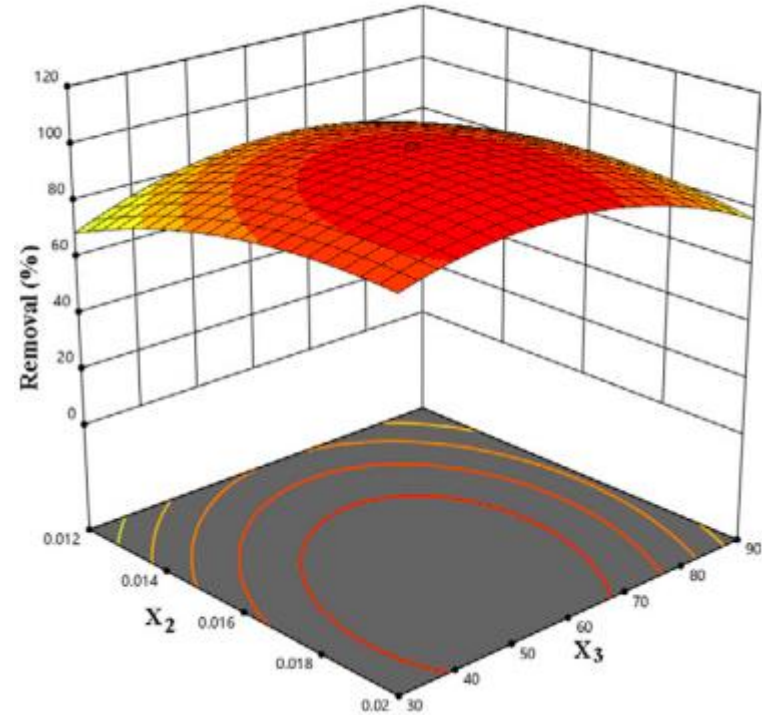
- Cette figure montre les effets interactifs du pH et de la masse d'adsorbant sur le taux d'élimination la concentration initiale étant maintenue au niveau central.
- Pour toute quantité fixe d'adsorbant, l'absorption augmente avec le pH et l'élimination la plus élevée est observée à pH 6,8.
- À pH < 5 et à pH > 7, on voit une diminution du taux d'élimination.
- L'élimination maximale (99,37 %) est réalisée à pH = 6,8.



- Les effets interactifs du pH et de la concentration initiale montrent qu'à dose constante d'adsorbant, le taux d'adsorption augmente avec l'augmentation de la concentration de Pb(II) à n'importe quelle valeur de pH donnée.
- L'adsorption maximale a été observée lorsque la concentration initiale de Pb(II) était de 60 mg/L.



- Les effets combinés de la masse et de la concentration initiale sur l'absorption de Pb(II) révèlent qu'à pH constant (6,8), le taux d'élimination du Pb(II) augmente avec l'augmentation à la fois de la masse et de la concentration initiale.
- À la concentration optimale (60 mg/L) et à 0,016 g de masse d'adsorbant, le pourcentage d'élimination est de 99,37 %.



Conclusion: Taux d'élimination max = 99,37 %),
pH = 6,8 masse d'adsorbant = 0,016 g et la concent init = 60 mg/L.

5. Plan Box-Behnken (PBB)

Dans cette étude, la biomasse morte de *Vibrio alginolyticus* PBR1 (microorganismes marins) a été utilisée pour éliminer les métaux lourds, le cadmium (Cd).

- Trois (3) facteurs sont étudiés :
 - x_1 est le pH de la solution.
 - x_2 est la masse du biosorbant.
 - x_3 est la concentration initiale du métal.

- La réponse Y est le taux d'élimination du métal en solution aqueuse.

Domaines d'étude:

Le tableau suivant permet de déterminer les conditions expérimentales à réaliser :

Variables	Unité	Facteur	niveau		
			-1	0	+1
pH (-)	(-)	X_1	4	6	8
Dosage	(g/L)	X_2	0.50	1.25	2.00
Concentration initiale	(mg/L)	X_3	50.00	125.00	200.00

Matrice des essais:

Run	Factor 1 A: pH	Factor 2	Factor 3	Metal Removal (%)	
		B:Biosorbent dosage g L ⁻¹ (w/v)	C: Metal concentration mg L ⁻¹ (w/v)	Actual Values Cd	Predicted Values Cd
1	4	1.25	200	3.21	4.8775
2	6	2	200	11.43	10.145
3	6	1.25	125	30.24	32.106
4	6	1.25	125	31.02	32.106
5	6	0.5	200	7.87	8.205
6	4	2	125	13.24	12.8575
7	8	1.25	50	50.84	49.1725
8	6	1.25	125	32.51	32.106
9	6	1.25	125	33.52	32.106
10	8	0.5	125	4.02	4.4025
11	4	0.5	125	2.30	0.2975
12	6	1.25	125	33.24	32.106
13	8	2	125	17.24	19.2425
14	4	1.25	50	34.79	35.5075
15	6	2	50	61.29	60.955
16	8	1.25	200	2.42	1.7025
17	6	0.5	50	34.21	35.495

ANOVA

Source	Sum of Squares	df	Mean Square	F Value	p-value > F	Prob
Response 1: Cd Removal (%)						
Model	4988.12	9	554.24	146.18	< 0.0001	
A-pH	55.02	1	55.02	14.51	0.0066	
B-Biosorbent dosage	375.38	1	375.38	99.01	< 0.0001	
C-Metal concentration	3049.80	1	3049.80	804.39	< 0.0001	
AB	1.30	1	1.30	0.34	0.5766	
AC	70.90	1	70.90	18.70	0.0035	
BC	138.30	1	138.30	36.48	0.0005	
A 2	872.55	1	872.55	230.14	< 0.0001	
B 2	304.96	1	304.96	80.43	< 0.0001	
C 2	109.71	1	109.71	28.94	0.0010	
Residual	26.54	7	3.79			
Lack of Fit	18.43	3	6.14	3.03	0.1561	
Pure Error	8.11	4	2.03			
Cor Total	5014.66	16				

$R^2 = 0.9947,$
 $R^2_{aj} = 0.9879,$
 $R^2_{pred} = 0.9387$

Interprétation

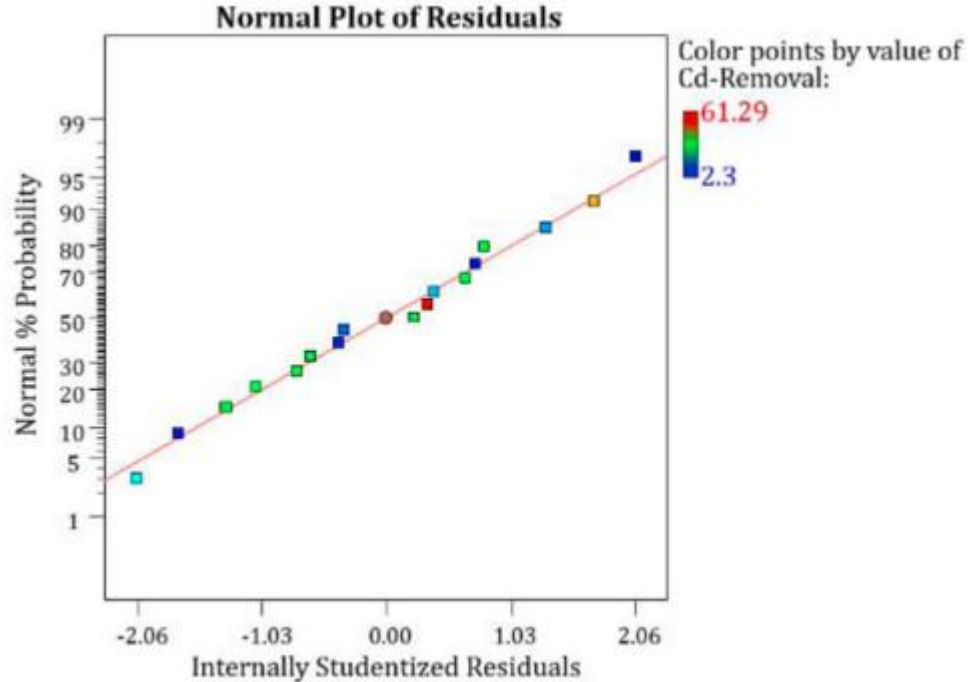
$$Y1 (\%) = +32.11 + 3,93 * A + 6,85 * B - 19,53 * C + 0,85 * A * B - 6,32 * A * C - 5,88 * B * C - 32,39 * A^2 - 8,51 * B^2 + 5,10 * C^2$$

- Le modèle s'est avéré significatif; valeur $F = 146,18$, ce qui implique que le modèle est significatif.
- Les valeurs de p-valeur $< 0,0500$ indiquent que les termes sont significatifs. En effet, les termes du modèle A , B , C , AC , BC , A^2 , B^2 et C^2 sont significatifs.
- La «valeur F du manque d'ajustement» de $0,1561$ implique que le manque d'ajustement n'est pas significatif par rapport à l'erreur pure.

- Par conséquent, on peut dire que le pH (A) du système, la concentration en biomasse (B) et la concentration initiale en métal (C) jouent un rôle important dans le taux d'élimination du cadmium sur le biosorbant.
- La valeur élevée du $R^2_{Adj} = 0,9879$ indique que 98,79 % de la variation totale des données d'élimination du Cd peuvent être expliquées par le modèle postulé.
- La valeur élevée du coefficient de détermination R^2 ($R^2 = 99,47$ %) et celle de R^2_{Adj} indique une forte dépendance et corrélation entre les réponses observées et les réponses prédites.

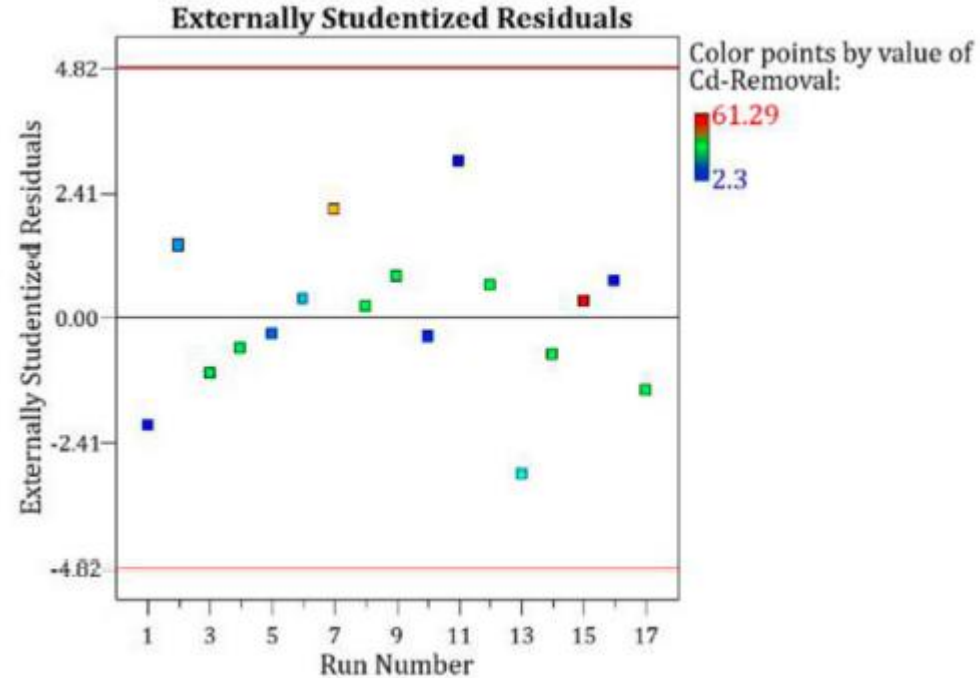
Analyse des résidus

Le graphique de probabilité normale révèle que les points de données sont alignés sur une ligne ce qui suggère que les données sont distribuées normalement, et les hypothèses de normalité ont été satisfaites.

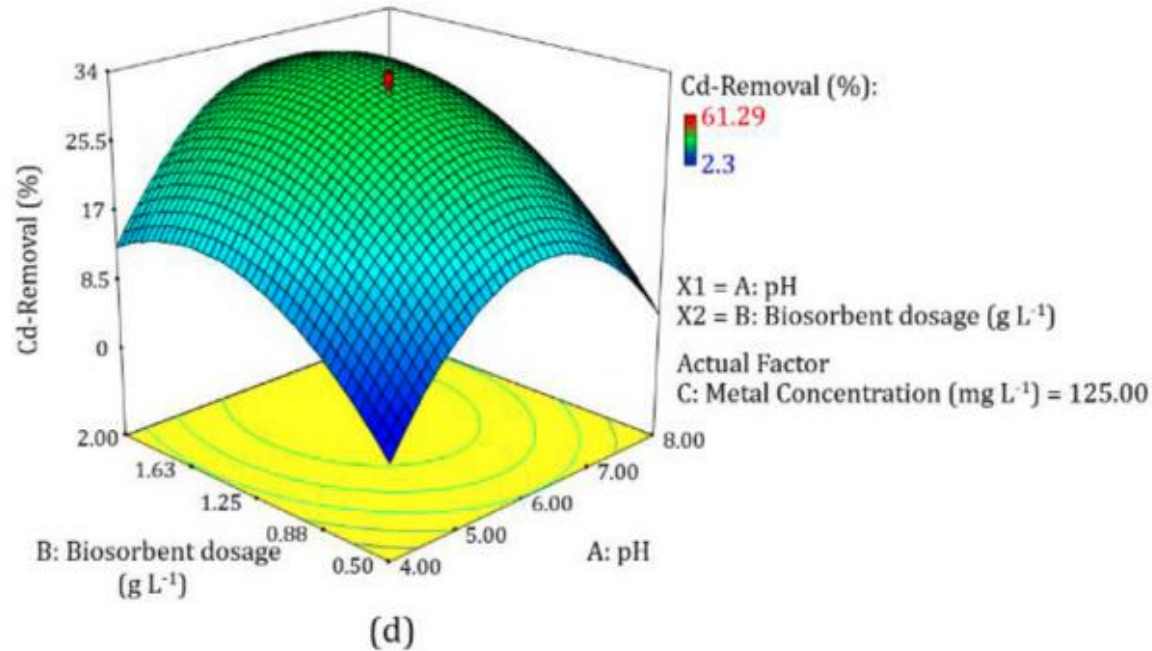


Graphique de la réponse prévue par rapport a la réponse réelle.

- Sur ce graphique, les valeurs réelles sont représentées par des points carrés et les valeurs prédites sont représentées par une ligne droite.
- Les points sont dispersés de façon aléatoire ce qui indique un bon accord entre les valeurs prédites et réelles.
- Les points expérimentaux sont dispersés dans une plage acceptable, ce qui indique la qualité de l'ajustement du modèle.

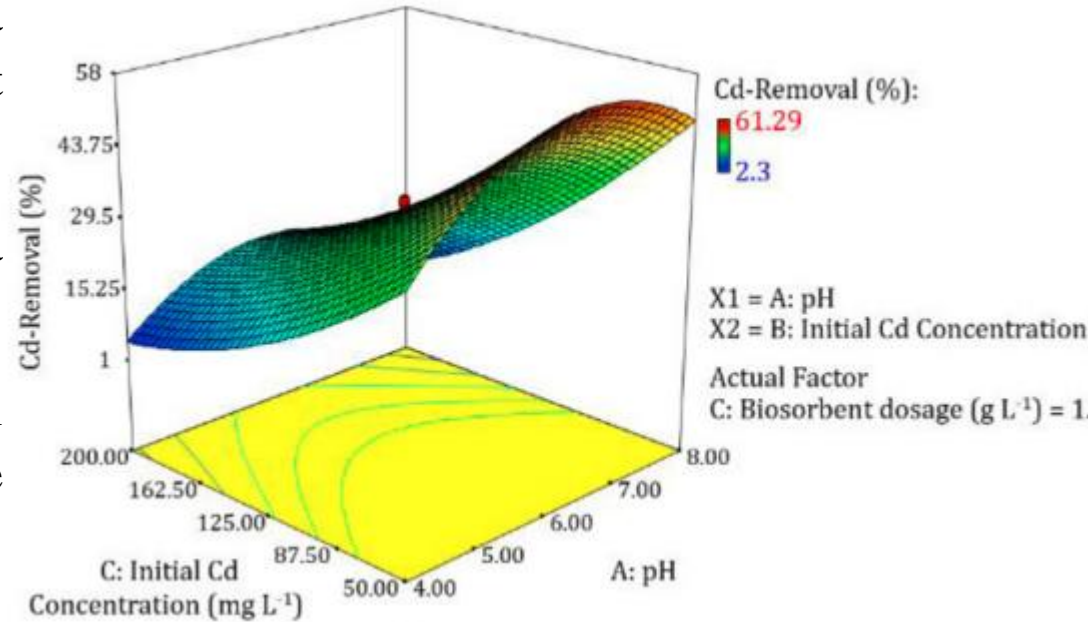


- Effet interactif du pH et du dosage du biosorbant sur la réponse, la concentration initiale du Cd est maintenue constante.
- La réponse augmente avec une augmentation du pH de 4 à 8
- La réponse augmente avec l'augmentation de la masse de biosorbant.



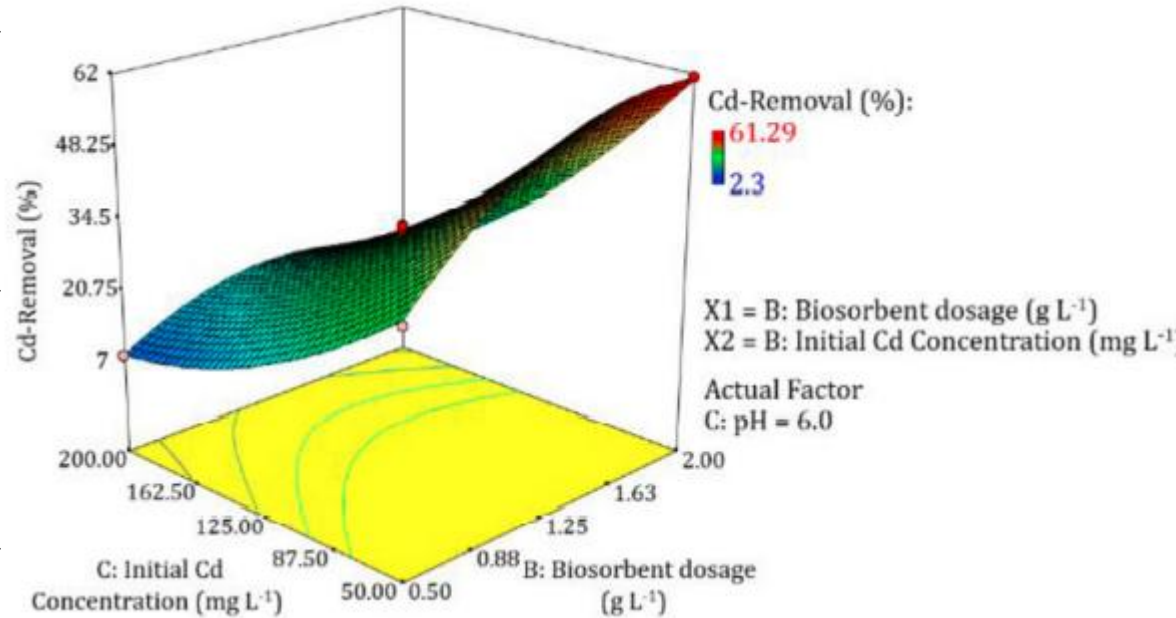
La réponse maximale est de 32,11 % obtenue à **pH 6** et à un **dosage de 1,25 g L⁻¹** de biosorbant.

- Effet interactif du pH et de la concentration initiale en métal sur la réponse, le dosage du biosorbant est maintenu constant.
- Quand le pH passe de 4 à 8, la réponse augmente.
- Quand la concentration de métal passe de 50 à 200 mg L⁻¹ la réponse diminue.



La réponse max de **57,18 %** est obtenue à un pH de **6,5** et à une concentration initiale de **50 mg L⁻¹**.

- Effet de la masse et de la concentration initiale sur la réponse, le pH maintenu constant à 6,0.
- La réponse augmente quand la masse du biosorbant augmente de 0,50 à 2,0 g L⁻¹.
- La réponse diminue lorsque la concentration initiale augmente de 20 à 200 mg L⁻¹.



La réponse maximale de 60,58 % a été obtenue à une dose de 1,73 g L⁻¹ de biosorbant et à une concentration initiale de Cd de 50 mg L⁻¹.

6. Plan Doehlert

Dans ce travail on s'intéresse à une optimisation multivariée d'une méthode spectrophotométrique pour la détermination du cuivre dans l'alcool de canne à sucre, explorant la réaction du Cu(II) avec la phénylfluorone (PF) en présence de chlorure de cétypyridinium (CPC) et de Triton X-100. Pour déterminer les conditions optimales, on a appliqué un plan de Doehlert.

- Trois (3) facteurs sont étudiés :
 - x_1 est le pH de la solution.
 - x_2 est la concentration du phénylfluorone (PF).
 - x_3 est la concentration Triton X-100.

La concentration du chlorure de cétypyridinium (CPC) est maintenue constante

- La réponse (Y) utilisée pour analyser les effets est l'absorbance de Cu(II).

Matrice des essais:

Exp.	PF (mol L ⁻¹)	TX100 (mol L ⁻¹)	pH	Abs
1	8.0 × 10 ⁻⁶ (0.866)	1.0 × 10 ⁻² (-0.5)	8.0 (0)	0.301 ± 0.005
2	8.0 × 10 ⁻⁶ (0.866)	2.0 × 10 ⁻² (0.5)	8.0 (0)	0.277 ± 0.001
3	7.0 × 10 ⁻⁶ (0.577)	1.5 × 10 ⁻² (0)	11.0 (0.817)	0.194 ± 0.008
4	6.0 × 10 ⁻⁶ (0.289)	1.0 × 10 ⁻² (-0.5)	5.0 (-0.817)	0.240 ± 0.006
5	6.0 × 10 ⁻⁶ (0.289)	2.0 × 10 ⁻² (0.5)	5.0 (-0.817)	0.163 ± 0.001
6	5.0 × 10 ⁻⁶ (0)	5.0 × 10 ⁻³ (-1.0)	8.0 (0)	0.298 ± 0.004
7a	5.0 × 10 ⁻⁶ (0)	1.5 × 10 ⁻² (0)	8.0 (0)	0.304
7b	5.0 × 10 ⁻⁶ (0)	1.5 × 10 ⁻² (0)	8.0 (0)	0.309
7c	5.0 × 10 ⁻⁶ (0)	1.5 × 10 ⁻² (0)	8.0 (0)	0.293
8	5.0 × 10 ⁻⁶ (0)	2.5 × 10 ⁻² (1.0)	8.0 (0)	0.294 ± 0.004
9	4.0 × 10 ⁻⁶ (-0.289)	1.0 × 10 ⁻² (-0.5)	11.0 (0.817)	0.178 ± 0.010
10	4.0 × 10 ⁻⁶ (-0.289)	2.0 × 10 ⁻² (0.5)	11.0 (0.817)	0.117 ± 0.005
11	3.0 × 10 ⁻⁶ (-0.577)	1.5 × 10 ⁻² (0)	5.0 (-0.817)	0.115 ± 0.003
12	2.0 × 10 ⁻⁶ (-0.866)	1.0 × 10 ⁻² (-0.5)	8.0 (0)	0.120 ± 0.004
13	2.0 × 10 ⁻⁶ (-0.866)	2.0 × 10 ⁻² (0.5)	8.0 (0)	0.120 ± 0.003

ANOVA

Factor	SS ^a	df ^a	MS ^a	F ratio	p value ^b
pH (L)	0.000145	1	0.000145	2.164801	0.279039
pH(Q)	0.037651	1	0.037651	561.9503	0.001775
TX-100 (L)	0.001838	1	0.001838	27.4368	0.034569
TX-100 (Q)	4.69×10^{-5}	1	4.69×10^{-5}	0.699627	0.490925
PF (L)	0.03416	1	0.03416	509.8508	0.001956
PF (Q)	0.019686	1	0.019686	293.827	0.003386
pH-TX-100	1.38×10^{-5}	1	1.38×10^{-5}	0.206063	0.694373
pH-PF	0.000131	1	0.000131	1.958986	0.296565
PF-TX-100	0.00015	1	0.00015	2.239739	0.273176
Lack of Fit	0.003385	3	0.001128	16.83986	0.056574
Pure Error	0.000134	2	6.70×10^{-5}		
Total SS	0.089491	14			

^a SS = sum of squares, df = degrees of freedom and MS = mean squares.

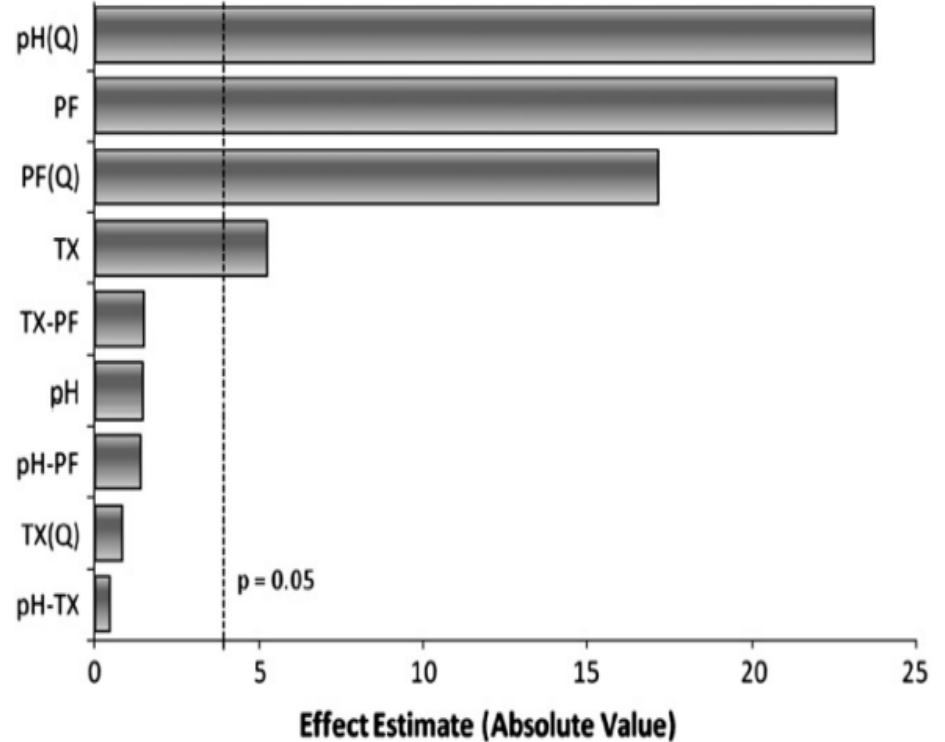
^b $\alpha = 0.05$.

$$A = -0,832 + 0,190X_1 + 0,728X_2 + 1,30 \cdot 10^5 X_3 \\ - 1,24 \cdot 10^{-2} X_1^2 - 62,50 X_2^2 - 1,07 \cdot 10^{10} X_3^2 + 1,30 \cdot X_1 X_2 \\ + 1,16 \cdot 10^3 X_1 X_3 - 4,08 \cdot 10^5 X_2 X_3$$

- Les valeurs de p-valeur $< 0,0500$ indiquent que les termes sont significatifs (termes en gras).
- La p-valeur du manque d'ajustement de $0,056574$ implique que le manque d'ajustement n'est pas significatif.
- La valeur élevée du $R^2A = 0,945$ indique que les observations expérimentales sont bien expliquées par le modèle postulé.

Analyse des résidus

Le diagramme de Pareto des effets a également été utilisé pour évaluer l'influence de chaque variable sur la réponse. Comme est montré sur la figure, les trois variables étudiées ont influencé la réponse de manière significative, bien qu'aucune interaction entre elles n'ait présenté d'effet significatif.



Optimisation

Les coordonnées du point critique ($\{pH\}_c$, $\{PF\}_c$ et $\{TX-100\}_c$) ont été calculées en résolvant le système d'équations linéaires obtenu après avoir dérivé la fonction de réponse en termes de chaque variable considérée dans le modèle, comme suit :

$$\delta A / \delta \{PF\} = 0; \delta A / \delta \{TX - 100\} = 0 \text{ and } \delta A / \delta \{pH\} = 0$$

A partir des applications de ces dérivés, un système à trois variables et trois équations linéaires a été obtenu. La résolution de ce système a été effectuée pour trouver les conditions critiques par simple application de la méthode des moindres carrés

$$Y^{-1}X = \begin{bmatrix} \{PF\}_c \\ \{TX - 100\}_c \\ \{pH\}_c \end{bmatrix}$$

Valeurs optimales :

- $X_1 = \text{pH} = 8,1$
- $X_2 = \text{concentration PF} = 6,2 \times 10^{-6} \text{ mol L}^{-1}$
- $X_3 = \text{concentration Triton X-100} = 1.03 \times 10^{-2} \text{ mol L}^{-1}$

FIN!