

# GdR FIBMAT 2025

Introduction aux plans d'expériences :  
de la collecte à l'exploitations des  
données expérimentales



Geoffrey GINOUX  
[geoffrey.ginoux@univ-reims.fr](mailto:geoffrey.ginoux@univ-reims.fr)

05/11/2025

# Mise en bouche aux plans d'expériences

## Recette des Bánh bao – Ingrédients (pour 6 brioches) :

- 1) Filet de poulet : 200, 300 ou 400 g
- 2) Oignon : 2, 4 ou 6
- 3) Ail : 1, 2, 3 gousses
- 4) Champignon shiitake : 3, 4 ou 5
- 5) Lap cheong : 2, 3 ou 4
- 6) Œuf dur : 1.5, 3 ou 6
- 7) Nuoc mâm : 1 càc, 1 ou 2 càs
- 8) Sucre : 1, 2, 3 càs
- 9) Huile de sésame : 5, 10 ou 15 mL
- 10) Poivre des oiseaux : 2, 5 ou 8 tours à moulin



Crédit : Jonathan Valencia

**Combien de recettes possibles ?**

**$3^{10} = 59049$  essais !** (env. 81 ans de déjeuners et dîners)

# Plan du cours

- Généralités sur les plans d'expériences
  - Intérêts et notions
  - Plans factoriels complets
- Plans de criblage
  - Plans factoriels fractionnaires
  - Méthode Taguchi
- Méthodologie des surfaces de réponse
  - Réseaux de Doehlert
- Statistiques
  - Analyse de variance

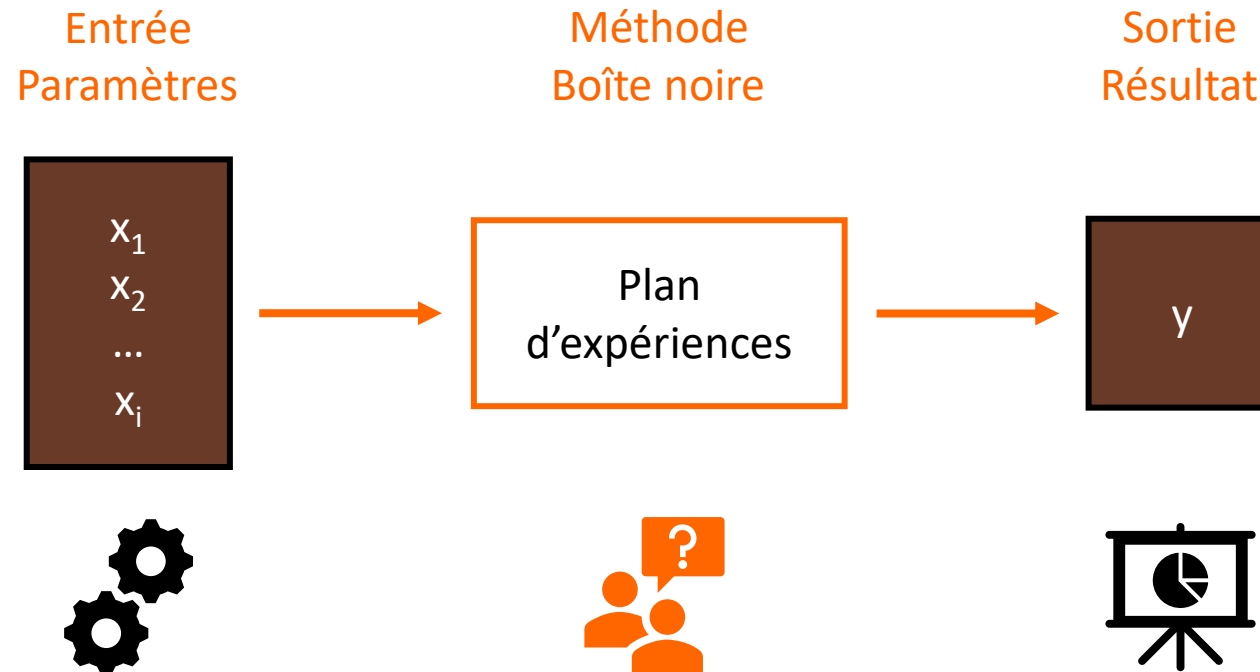


# Généralités sur les plans d'expériences

## Définition :

Un **plan d'expériences** est une méthode qui permet d'**ordonner des essais**, avec pour but d'**économiser les moyens** en ne faisant non pas varier un paramètre à la fois, mais un **ensemble de paramètres simultanément**.

## Principe :





# Intérêts des plans d'expériences

**Méthode des plans d'expériences** : ordonner les essais en faisant varier tous les paramètres en même temps.



**Méthode « intuitive »** : fixer tous les paramètres sauf un qui varie, puis passer à un autre paramètre variable, et ainsi de suite.

## Avantages des PE

Compromis optimisé nombre d'expériences/qualité-précision

Organisation des essais

Compréhension d'un phénomène/procédé (expliciter lien  $y = f(x_i)$  par un modèle)



Prédiction d'un résultat

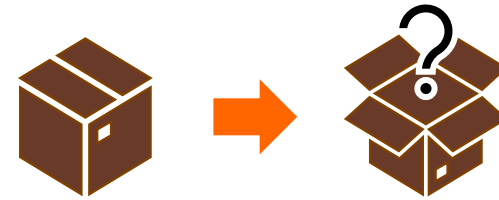
Économie  
temps/coûts

## Inconvénients des PE

Perte de données

Basés sur des hypothèses parfois semi-subjectives

Boîte noire



## Quand utiliser un plan d'expériences ?

- ✓ Questions dont on ne trouve pas de réponses toutes faites
- ✓ Beaucoup de paramètres à faire varier (explosion combinatoire)
- ✓ Essais longs et/ou coûteux
- ✓ Éviter les gâchis



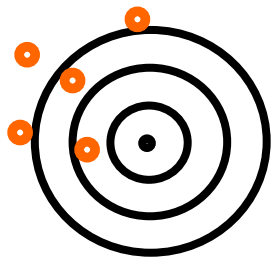
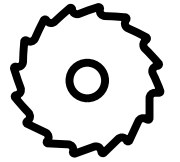
# Intérêts des plans d'expériences

## Exemples :

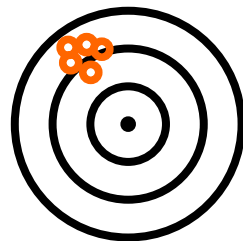
1) **Optimum** : Pour garantir la **durabilité** d'un composite à fibres végétales, un chercheur peut **maximiser le caractère hydrophobe** par l'optimisation d'un traitement chimique (concentration, type de greffons, etc.).

2) **Cible** : Pour fabriquer une pièce en composite aux **dimensions comprises dans une tolérance**, afin de respecter un cahier des charges et qu'il n'y ait **pas de rebut**, un ingénieur peut faire un **réglage des paramètres d'usinage** (type de découpe, vitesse d'avance, etc.) pour obtenir les **cotes cibles**.

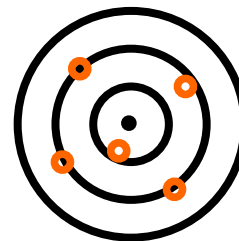
3) **Rendement** : Pour optimiser un champ de lin, un agriculteur peut maximiser son gain (quantité, taille et qualité des fibres, etc.) avec le moindre coût (type d'engrais et quantité, fréquence d'irrigation, etc.) : c'est un **rapport production/coût**. Il peut également étudier l'**influence** d'un certain climat (pluie, ensoleillement, etc.) sur la **performance** d'un engrais (**mesurable mais non contrôlable**).



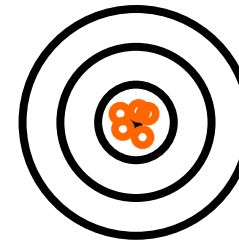
Ni juste  
Ni précis



Précis



Juste



Juste et précis  
= Exact

# Notions de plans d'expériences

## Terminologie :

Paramètre = **Facteur** (« **continu** » si valeurs continues, « **discret** » si valeurs qualitatives)

Valeur d'un paramètre = **Niveau** (aussi dit « **modalité** » si valeurs qualitatives)

Plage de valeurs que peuvent prendre les paramètres = **Espace expérimental**

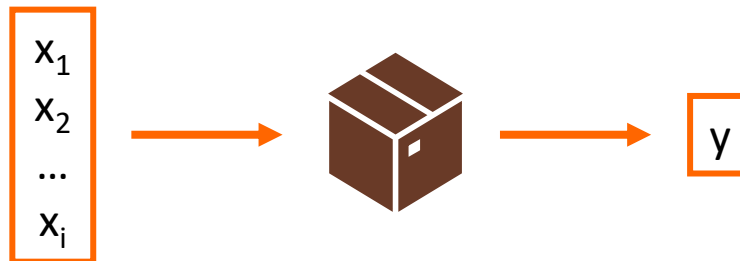
Plage de valeurs des paramètres = **Domaine d'étude** (« **domaine de facteur** » quand un seul paramètre)

Expérience = **Configuration**  Peut contenir une ou des **zones exclues**

Résultat d'une expérience = **Réponse**

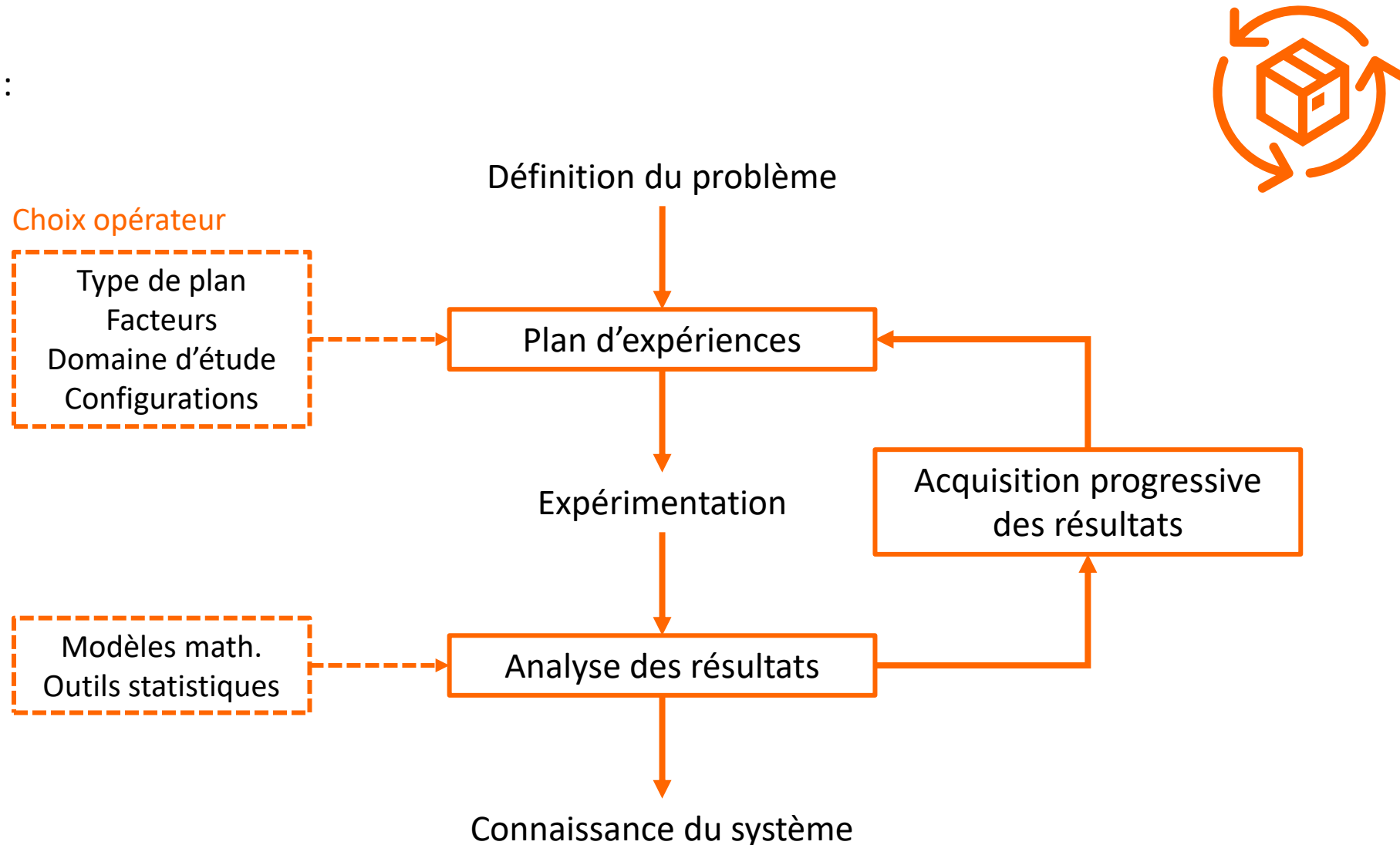
Impact d'un paramètre sur le résultat = **Effet de facteur**

Impact croisé de plusieurs paramètres sur le résultat = **Effet d'interaction**



# Notions de plans d'expériences

Méthodologie :





# Cas simple du plan complet à 2 facteurs à 2 niveaux

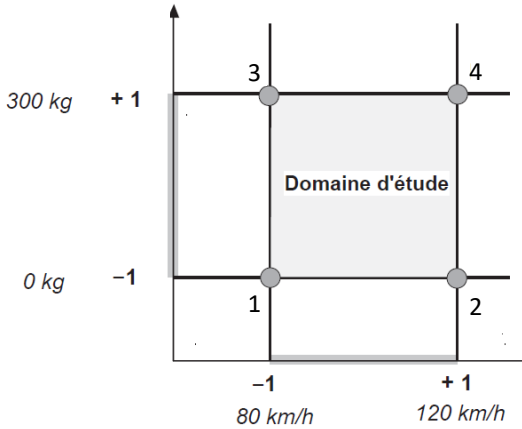
Exemple **consommation d'essence** d'un véhicule en **fonction de la vitesse et de la surcharge** :

Facteur A : Vitesse → Niveau bas : 80 km.h<sup>-1</sup> / Niveau haut : 120 km.h<sup>-1</sup>  
 Facteur B : Surcharge → Niveau bas : 0 kg / Niveau haut : 300 kg

**Notation de Yates** : centrer et réduire les variables (en unité codée : bas = -1 et haut = +1)

Vitesse :  $x_1(A) = \frac{2A - (A_{max} - A_{min})}{A_{max} - A_{min}}$        $y = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2$

Surcharge :  $x_2(B) = \frac{2B - (B_{max} - B_{min})}{B_{max} - B_{min}}$



Matrice d'expérimentation		
Config. N°	Vitesse Facteur A	Surcharge Facteur B
1	80 km.h <sup>-1</sup>	0 kg
2	120 km.h <sup>-1</sup>	0 kg
3	80 km.h <sup>-1</sup>	300 kg
4	120 km.h <sup>-1</sup>	300 kg

Matrice d'expériences		
Config. N°	Vitesse Facteur 1	Surcharge Facteur 2
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1
Niveau -1	80 km.h <sup>-1</sup>	0 kg
Niveau +1	120 km.h <sup>-1</sup>	300 kg

# Cas simple du plan complet à 2 facteurs à 2 niveaux

Matrice d'expériences et résultats

Config. N°	Vitesse Facteur 1	Surcharge Facteur 2	Consommation (L/100 km)
1	-1	-1	8,3
2	+1	-1	10,7
3	-1	+1	9,7
4	+1	+1	12,3
Niveau -1	80 km.h <sup>-1</sup>	0 kg	
Niveau +1	120 km.h <sup>-1</sup>	300 kg	

$$y = a_0 + \sum_{i=1}^f a_i x_i + \sum_{i=1}^{f-1} \sum_{j=i+1}^f a_{ij} x_i x_j$$

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

Moyenne des réponses :

$$a_0 = \frac{\sum_{i=1}^n y_i}{n} = \frac{8,3 + 10,7 + 9,7 + 12,3}{4} = 10,25$$

Calcul des effets des facteurs :

$a_i$  = le demi-différence du facteur  $i$  (différence entre le niveau haut et la moyenne). L'effet est obtenu en calculant la moyenne des réponses pondérée du niveau du facteur  $i$ .

$$a_1 = \frac{\sum y_i (A = +1) - \sum y_i (A = -1)}{n} = \frac{10,7 + 12,3 - 8,3 - 9,7}{4} = 1,25$$

$$a_2 = \frac{\sum y_i (B = +1) - \sum y_i (B = -1)}{n} = \frac{9,7 + 12,3 - 8,3 - 10,7}{4} = 0,75$$

# Cas simple du plan complet à 2 facteurs à 2 niveaux

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

$$a_0 = 10,25$$

$$a_1 = 1,25$$

$$a_2 = 0,75$$

Calcul de l'effet de l'interaction :

$a_{ij}$  = effet synergétique (quand positif) ou antagonique (quand négatif) entre deux facteurs qui évoluent dans le même sens vers le niveau haut depuis la moyenne. L'interaction se calcule en moyennant les réponses pondérées des niveaux des deux facteurs considérés.

$$a_{12} = \frac{\sum y_i(A = +1, B = +1) - \sum y_i(A = -1, B = +1) - \sum y_i(A = +1, B = -1) + \sum y_i(A = -1, B = -1)}{n}$$

$$a_{12} = \frac{12,3 - 9,7 - 10,7 + 8,3}{4} = 0,05$$

Traduction de l'interaction : Charger et rouler vite augmente davantage la consommation d'essence du véhicule.

De ce plan d'expériences, on dira que charger, rouler vite et charger en roulant vite augmentent la consommation d'essence.

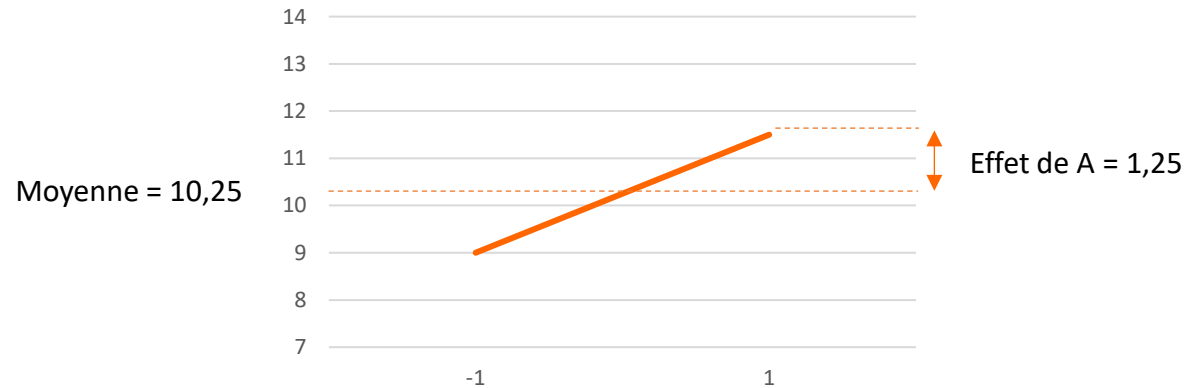
Matrice d'expériences et résultats

Config. N°	Vitesse Facteur 1	Surcharge Facteur 2	Consommation (L/100 km)
1	-1	-1	8,3
2	+1	-1	10,7
3	-1	+1	9,7
4	+1	+1	12,3

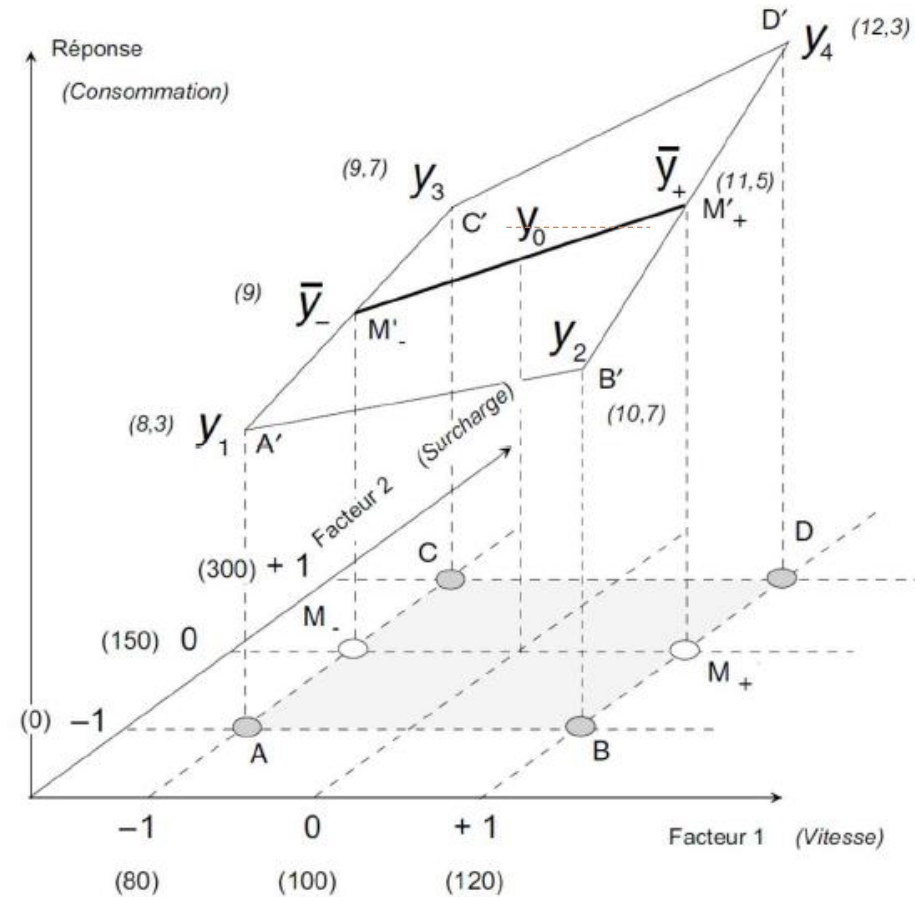
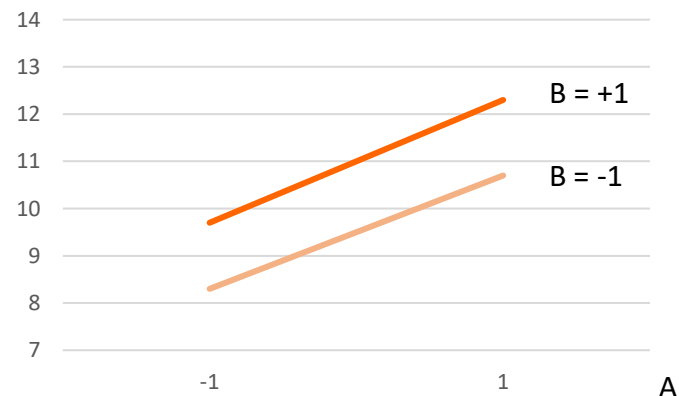
# Cas simple du plan complet à 2 facteurs à 2 niveaux

Représentation graphique des effets des facteurs et des interactions

Effet de A



Interaction entre A et B



→ Effet d'interaction : **parallèle si nul**

# Méthode de calcul par les matrices

Matrice d'expériences

Matrice du modèle = X

Vecteur des coefficients = A

Vecteur des réponses = Y

$$\begin{bmatrix} -1 & -1 \\ +1 & -1 \\ -1 & +1 \\ +1 & +1 \end{bmatrix}$$

→

$$\begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix}$$

+

$$\begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \end{Bmatrix}$$

→

$$\begin{Bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{Bmatrix}$$

Matrice d'expériences

Config. N°	Vitesse Facteur 1	Surcharge Facteur 2
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1

$$y = a_0 x_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2$$

Valeurs des coefficients à déterminer

Résultats obtenus expérimentalement à chaque configuration

Niveaux des facteurs et des interactions du modèle à chaque configuration

$$Y = X \cdot A$$

$$\hookrightarrow A = X^{-1} \cdot Y$$

# Méthode de calcul par les matrices

Dans le cas des plans où le nombre d'expériences est différent du nombre de coefficients ( $e \neq a$ ), la matrice du modèle  $X$  n'est pas inversible car non carrée ( $X^{-1}$  impossible). Il faut passer par un **estimateur** de la matrice des coefficients  $A$ , noté  $\hat{A}$  tel que :

Matrice d'information

$$\hat{A} = \underbrace{(X^T \cdot X)^{-1}}_{\text{Matrice de dispersion}} \cdot X^T \cdot Y$$

Matrice de dispersion

Cette formule plus générale reste vraie pour les plans carrés puisque  $\hat{A} = A$  (les matrices d'information et donc de dispersion sont diagonales, la variance est alors nulle :  $V = \sigma^2 = 0$ ) :

$$X = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \quad \text{et} \quad X^T = \begin{bmatrix} +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 \end{bmatrix}$$

↪ **Matrice de Hadamard**



# Méthode de calcul par les matrices

Les matrices de **Hadamard**  $H$  sont des matrices carrées d'ordre  $N$  ne possédant que des coefficients -1 ou +1 telles que  $H^T \cdot H = N \cdot I_N$  avec  $H^T$  la transposée de  $H$  et  $I_N$  la matrice identité d'ordre  $N$ .

$$X^T \cdot X = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} = 4 \times I_4 \quad \text{d'où} \quad (X^T \cdot X)^{-1} = \frac{1}{4} \times I_4 = \begin{bmatrix} 0,25 & 0 & 0 & 0 \\ 0 & 0,25 & 0 & 0 \\ 0 & 0 & 0,25 & 0 \\ 0 & 0 & 0 & 0,25 \end{bmatrix}$$

$$(X^T \cdot X)^{-1} \cdot X^T \cdot Y = \frac{1}{4} \begin{bmatrix} +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 \end{bmatrix} \times \begin{pmatrix} 8,3 \\ 10,7 \\ 9,7 \\ 12,3 \end{pmatrix} = \begin{pmatrix} 10,25 \\ 1,25 \\ 0,75 \\ 0,05 \end{pmatrix} = A \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \end{pmatrix}$$

$$X \cdot A = \begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix} \times \begin{pmatrix} 10,25 \\ 1,25 \\ 0,75 \\ 0,05 \end{pmatrix} = \begin{pmatrix} 8,3 \\ 10,7 \\ 9,7 \\ 12,3 \end{pmatrix}$$

# Méthode de calcul par les matrices

Dans le cas où la matrice n'est pas carrée, par exemple si on décide de ne pas considérer l'interaction et qu'on retire la dernière colonne de la matrice du modèle  $X$ , l'estimateur des coefficients  $\hat{A}$  se calcule classiquement :

$$X = \begin{bmatrix} +1 & -1 & -1 \\ +1 & +1 & -1 \\ +1 & -1 & +1 \\ +1 & +1 & +1 \end{bmatrix} \quad \text{d'où} \quad X^T \cdot X = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{et donc} \quad (X^T \cdot X)^{-1} = \frac{1}{4} \times I_3$$

$$(X^T \cdot X)^{-1} \cdot X^T \cdot Y = \frac{1}{4} \begin{bmatrix} +1 & +1 & +1 & +1 \\ -1 & +1 & -1 & +1 \\ -1 & -1 & +1 & +1 \end{bmatrix} \times \begin{pmatrix} 8,3 \\ 10,7 \\ 9,7 \\ 12,3 \end{pmatrix} = \begin{pmatrix} 10,25 \\ 1,25 \\ 0,75 \end{pmatrix} = \hat{A} \quad \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

$$\hat{Y} = X \cdot \hat{A} = \begin{bmatrix} +1 & -1 & -1 \\ +1 & +1 & -1 \\ +1 & -1 & +1 \\ +1 & +1 & +1 \end{bmatrix} \times \begin{pmatrix} 10,25 \\ 1,25 \\ 0,75 \end{pmatrix} = \begin{pmatrix} 8,25 \\ 10,75 \\ 9,75 \\ 12,25 \end{pmatrix} \neq Y$$

# Plans de criblage

Les plans d'expériences peuvent servir de **plans de criblage** afin d'**écarter les facteurs non ou peu influents** de ceux qui le sont.



---

## Dans quel but ?

- Savoir si un paramètre influe vraiment sur un procédé
- **Limiter le nombre de facteurs à étudier** dans l'optique d'une optimisation plus poussée



---

## Dans quel cas ?

- Quand beaucoup de paramètres à étudier
- Peu d'interactions possibles/envisagées
- Faible nombre de niveaux possibles



# Notions de plans fractionnaires

Quand toutes les combinaisons de niveaux sont réalisées dans un plan d'expériences, il s'agit d'un **plan factoriel complet** tel que :

$$e = n^k$$

avec  $e$  le nombre d'essais,  $n$  le nombre de niveaux par facteur, et  $k$  le nombre de facteurs.

Cependant,  $e$  peut devenir conséquent quand  $k$  devient grand. On introduit alors des plans réduits, appelés **plans factoriels fractionnaires**, tel que :

$$e = n^{k-p}$$

avec  $p$  le facteur de réduction.

Un plan fractionnaire est une fraction d'un plan complet mais qui **conserve les propriétés d'orthogonalité**. On introduit alors la notion d'**alias**. Les effets sont dits aliasés à un **contraste** noté  $l$  qui **confond chaque effet calculé en une somme d'effets simples** :

$$l_i = \sum a_{ijk}$$

**On perd ainsi de l'information mais le nombre d'essais est réduit (= réduction des coûts).**

# Notions de plans fractionnaires

Il existe deux critères à satisfaire pour utiliser les plans factoriels fractionnaires. Si au moins l'un des deux critères n'est pas respecté, les plans fractionnaires ne peuvent pas s'appliquer.

Critères d'applicabilité :

- Le nombre d'essais doit être supérieur au nombre de degrés de liberté du modèle :  $e \geq ddl_{\text{modèle}}$   
tel que le  $ddl_{\text{modèle}} = ddl_{\text{moyenne}} + \sum ddl_{\text{facteur}} + \sum ddl_{\text{interaction}}$   
avec  $ddl_{\text{moyenne}} = 1$ ,  
 $ddl_{\text{facteur}} = (\text{nombre de niveaux du facteur} - 1)$  et  
 $ddl_{\text{interaction}} = (\text{produit des ddl des facteurs en jeu dans l'interaction})$
- **Règle d'orthogonalité** : Un plan fractionnaire est et doit être orthogonal, i.e. quand, pour un facteur fixé à un même niveau, tous les autres facteurs se retrouvent le même nombre de fois à des niveaux différents.  
(Pour aller plus loin : Il s'agit alors de déterminer le plus petit commun multiple du produit du nombre de niveaux de toutes les actions disjointes prises deux à deux tel que  $e_{\min} = PPCM$ ,  $e$  étant multiple du PPCM)

# Méthode Taguchi – Introduction

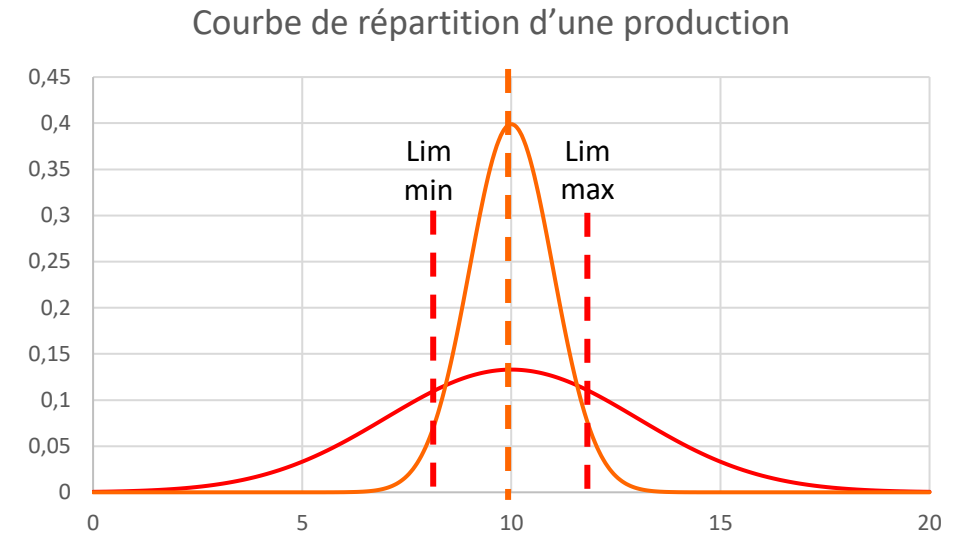
**Genichi Taguchi** (1924-2012) est un statisticien japonais qui a développé des **tables prêtes à l'emploi** dans les années 50-60.

Les **tables de Taguchi** étaient initialement prévues pour l'amélioration continue de la qualité en rendant les **procédés plus robustes aux facteurs extérieurs**, c'est-à-dire en réduisant les effets des facteurs non contrôlables. La méthode permet ainsi de réduire les perturbations (appelées « **bruit** » qui correspond à la tolérance ou à l'écart-type autour d'une moyenne) pour **minimiser les pertes d'une production sans modifier le produit** (valeur toujours aussi juste = moyenne inchangée).

↪ **Exactitude augmentée**

La méthode est facile d'utilisation et est depuis répandue dans le monde industriel en management de la qualité. Elle est d'ailleurs attractive pour le criblage de par son efficacité et sa simplicité.

La méthode est basée sur des **tables orthogonales** remplies à l'aide de **graphes linéaires** et de **tables triangulaires d'interactions**.

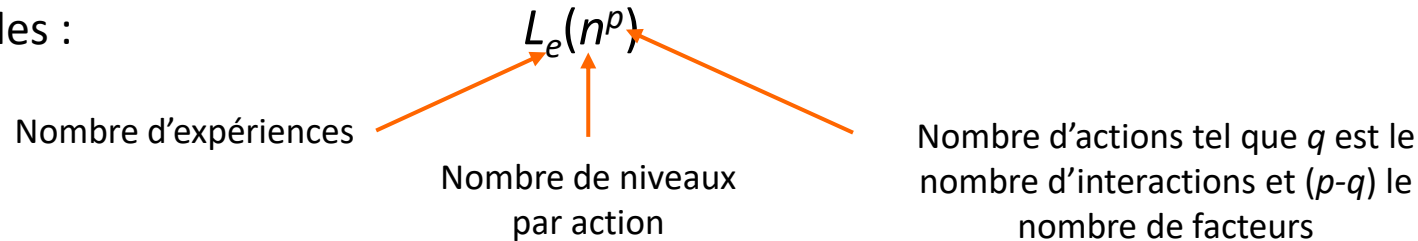


Même moyenne, mais bruit réduit  
=  
Augmente la conformité des produits  
Moins de perte en dehors des limites



# Méthode Taguchi – Tables orthogonales

Nomenclature des tables :



Les interactions sont possibles avec les tables de Taguchi, mais pas systématiquement. Quel que soit le cas, les interactions resteront limitées aux interactions d'ordre 2, les **interactions d'ordre 3+ sont négligées**. Une liste de 18 plans préconstruits existe :



Même si les tables de Taguchi sont des plans fractionnaires particuliers, les critères d'applicabilité des tables sont les mêmes que pour les plans factoriels fractionnaires :

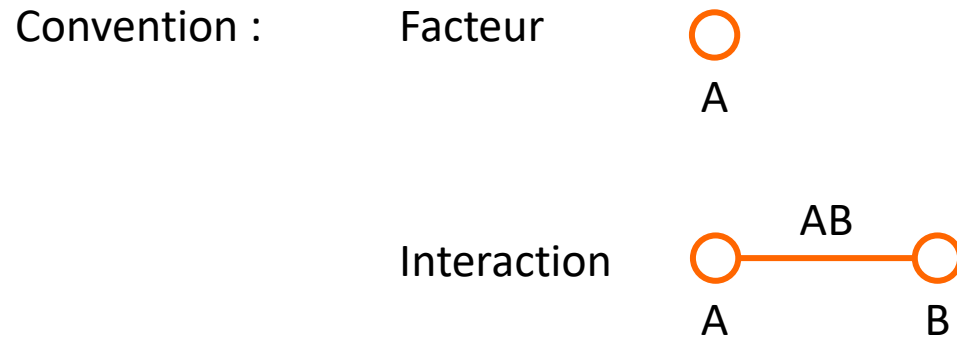
- $e \geq ddl_{modèle}$
- $e_{min} = PPCM$  (orthogonalité)

Interaction	Plans à 2 niveaux	Plans à 3 niveaux	Plans à 4+ niveaux	Plans mixtes
Impossible	$L_{12}(2^{11})$			$L_{36}(2^{11} \times 3^{12})$
Limitée				$L_{18}(2^1 \times 3^7)$ $L_{32}(2^1 \times 4^9)$ $L_{50}(2^1 \times 5^{11})$
Possible	$L_4(2^3)$ $L_8(2^7)$ $L_{16}(2^{15})$ $L_{32}(2^{31})$ $L_{64}(2^{63})$	$L_9(3^4)$ <div style="border: 2px solid orange; padding: 2px;"><math>L_{27}(3^{13})</math></div> $L_{81}(3^{40})$	$L_{16}(4^5)$ $L_{64}(4^{21})$ $L_{25}(5^6)$	$L_{36}(2^3 \times 3^{13})$ $L_{54}(2^1 \times 3^{25})$

**Cas de la recette : 27 essais suffisent au lieu de 59049 ! (2 semaines)**

# Méthode Taguchi – Graphes linéaires

Les **graphes linéaires** sont une **représentation graphique** du modèle à étudier. Les graphes aident à l'**organisation des expériences** en permettant de placer les interactions possibles et les facteurs faciles ou difficiles à modifier.



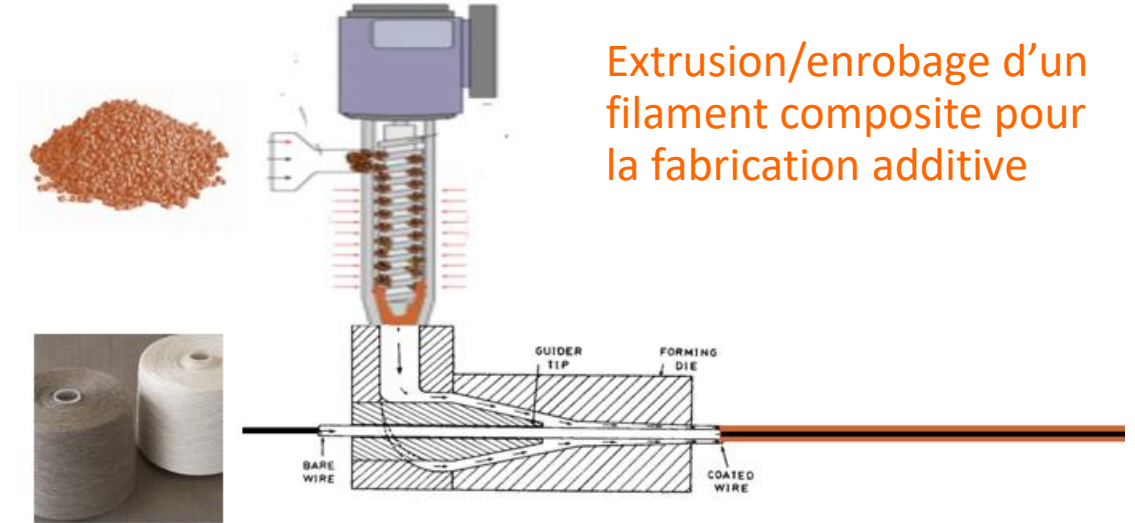
Groupes de facteurs :

Groupe 1 : Facteur difficile à régler ○

Groupe 2 : Facteur assez difficile à régler ◎

Groupe 3 : Facteur assez facile à régler ◉

Groupe 4 : Facteur facile à régler ●



Exemple d'un procédé d'extrusion

Changement de profil de vis

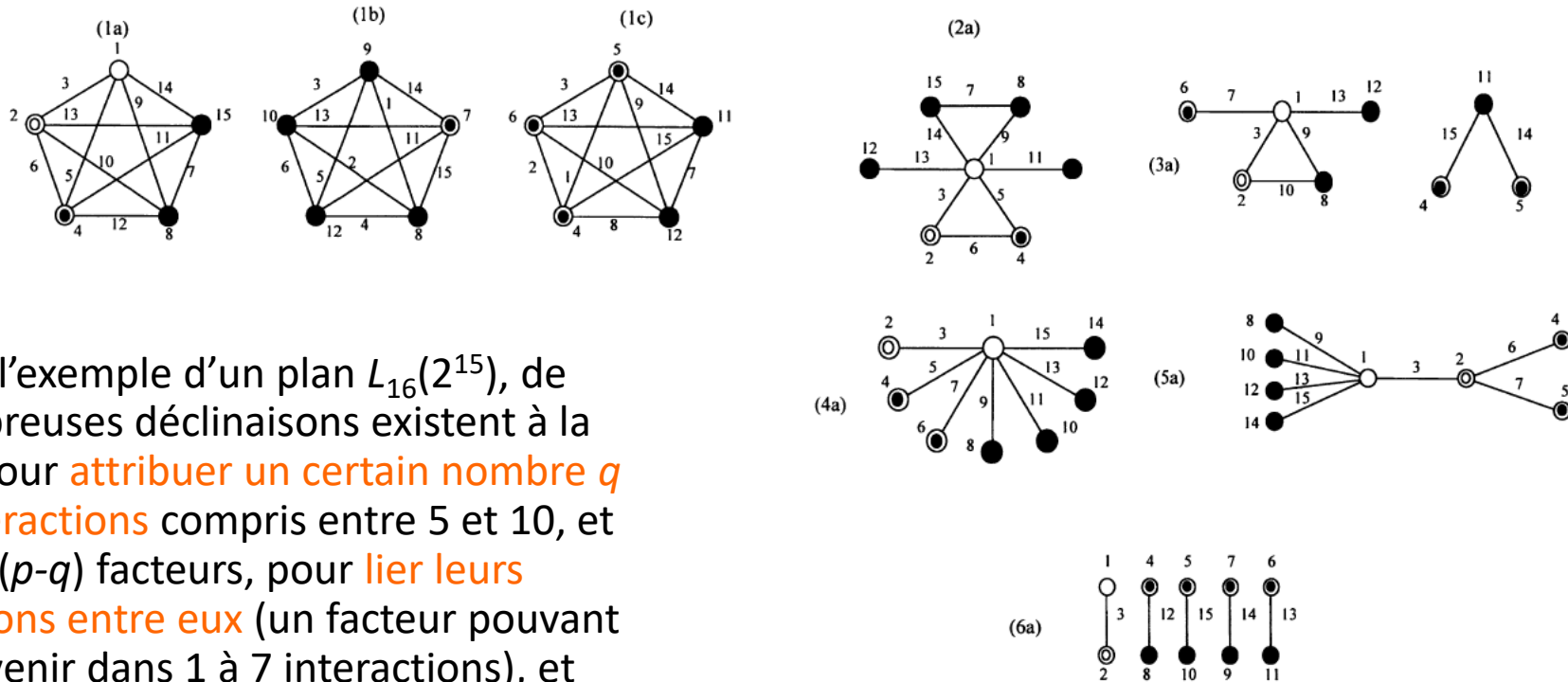
Changement de matière

Changement de température

Changement de rotation de vis

# Méthode Taguchi – Graphes linéaires

Chacune des 18 tables orthogonales de Taguchi possède des graphes linéaires, mais il peut aussi en exister plusieurs pour chaque table orthogonale. Il convient alors de **choisir le graphe linéaire le plus adapté** au plan.



Dans l'exemple d'un plan  $L_{16}(2^{15})$ , de nombreuses déclinaisons existent à la fois pour **attribuer un certain nombre  $q$  d'interactions** compris entre 5 et 10, et donc  $(p-q)$  facteurs, pour **lier leurs relations entre eux** (un facteur pouvant intervenir dans 1 à 7 interactions), et pour **assigner un groupe de difficulté**.

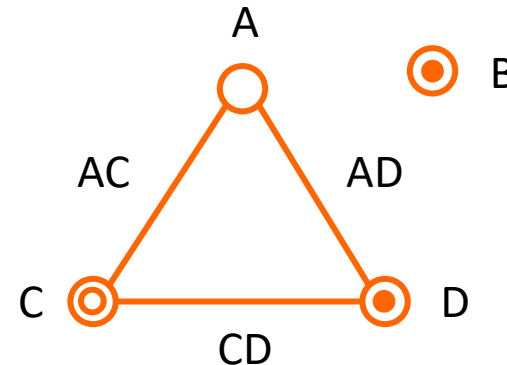
# Méthode Taguchi – Tables triangulaires des interactions

Application : Nous souhaitons toujours optimiser le procédé d'extrusion en faisant varier le profil de vis (A), la nature de la fibre (B), la température d'extrusion (C) et la vitesse de rotation de la vis (D). Nous considérons uniquement les interactions AC, AD et CD. Chaque facteur possède deux niveaux et la vérification des critères d'applicabilité nous a conduit à sélectionner une table  $L_8(2^7)$ .

$$Y = M + A + B + C + D + AC + AD + CD$$

Table orthogonale  $L_8(2^7)$

Config.	A	C	AC	D	AD	CD	B
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2



	2	3	4	5	6	7
1	3	2	5	4	7	6
2		1	6	7	4	5
3			7	6	5	4
4				1	2	3
5					3	2
6						1

(dans les faits, les graphes linéaires donnent souvent et directement le numéro de colonne le plus optimal)

Notation équivalente Taguchi ↔ Yates : 1 = +1 et 2 = -1

# Méthode Taguchi – Calcul des effets et interactions

Une fois le plan établi, les expériences sont réalisées et le taux de conformité (en %) est récupéré. Il s'agit maintenant de calculer les effets des facteurs et des interactions pour déterminer les paramètres maximisant le taux de conformité.

Le modèle s'écrit donc :  $Y = \bar{Y} + E_A + E_B + E_C + E_D + I_{AC} + I_{AD} + I_{CD}$

La table étant orthogonale, la matrice du modèle est carrée et il n'y a pas de résidu. Le calcul peut se faire sans estimateur comme pour un plan complet (moyennes pondérées).

Config.	A	C	AC	D	AD	CD	B	Y
1	1	1	1	1	1	1	1	82
2	1	1	1	2	2	2	2	85
3	1	2	2	1	1	2	2	96
4	1	2	2	2	2	1	1	79
5	2	1	2	1	2	1	2	88
6	2	1	2	2	1	2	1	90
7	2	2	1	1	2	2	1	52
8	2	2	1	2	1	1	2	72

1 = +1 et 2 = -1

$$\bar{Y} = \frac{82+85+96+79+88+90+52+72}{8} = 80,5$$

$$E_A = \frac{82+85+96+79-88-90-52-72}{8} = +5,00$$

$$E_B = \frac{82+79+90+52-85-96-88-72}{8} = -4,75$$

$$E_C = \frac{82+85+88+90-96-79-52-72}{8} = +5,75$$

$$E_D = \frac{82+96+88+52-85-79-90-72}{8} = -1,00$$

$$I_{AC} = \frac{82+85+52+72-96-79-88-90}{8} = -7,75$$

$$I_{AD} = \frac{82+96+90+72-85-79-88-52}{8} = +4,50$$

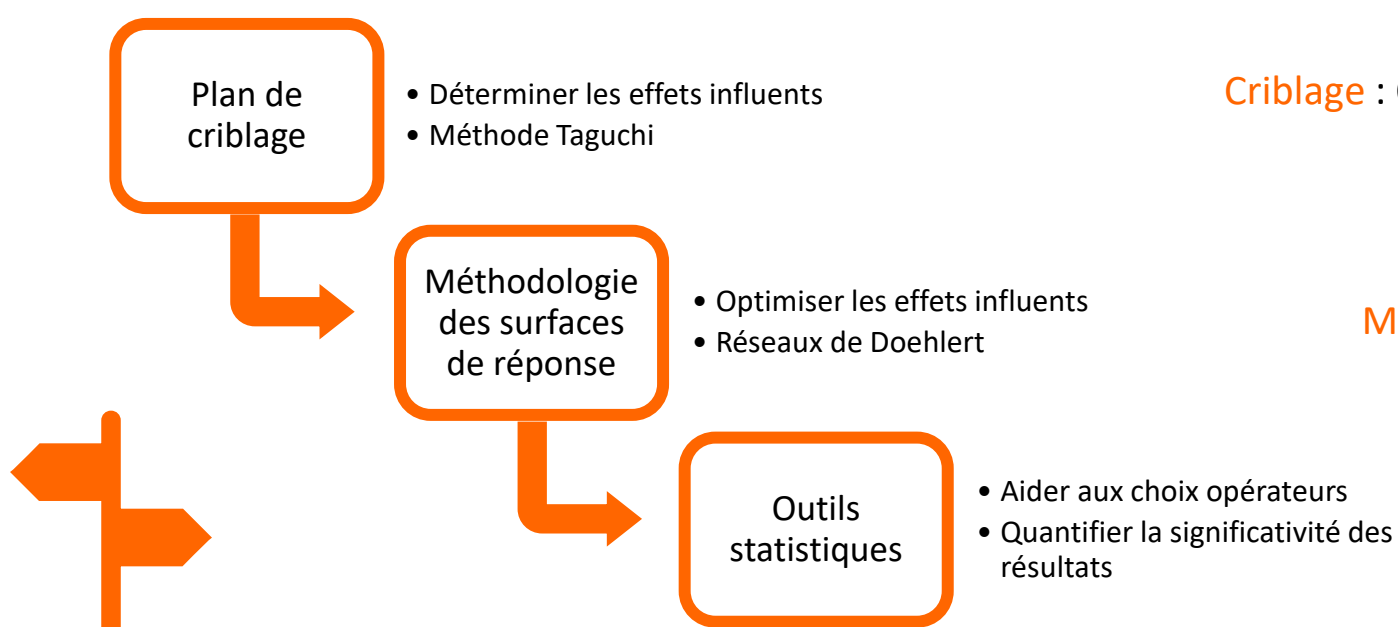
$$I_{CD} = \frac{82+79+88+72-85-96-90-52}{8} = -0,25$$

→ A2 B2 C1 D2  
Y = 99,5

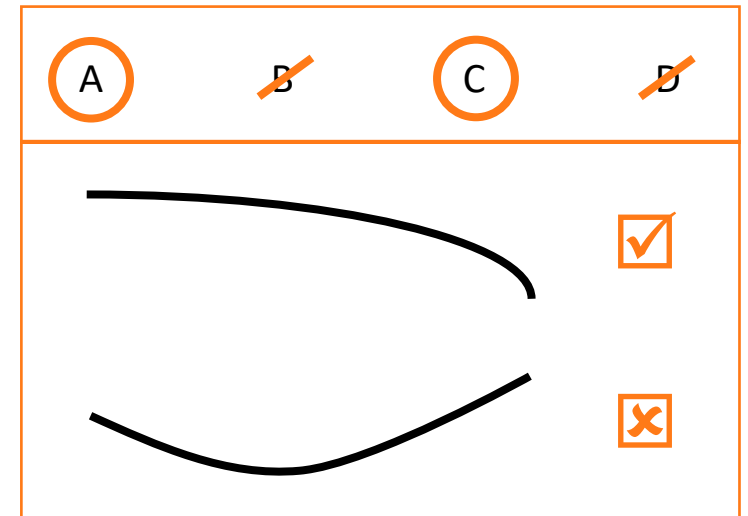
# Vers la méthodologie des surfaces de réponse

La méthode Taguchi est ainsi une méthode efficace pour le criblage, permettant de récupérer les effets les plus influents afin de les étudier plus en détails dans la méthodologie des surfaces de réponse (MSR) qui est une méthode d'optimisation. La MSR peut être lourde en expériences, c'est pour cela que des méthodes efficaces comme Taguchi pour du criblage sont intéressantes, afin de réduire le nombre d'expériences en réduisant le nombre de facteurs.

La sélection des effets influents peut simplement se faire par la sélection des plus hauts effets (les deux à trois effets les plus importants), mais il existe des outils statistiques pour déterminer plus rigoureusement les effets influents et qui seront discutés à la fin.



Criblage : Quels facteurs ?



MSR : Comment ?



# MSR – Introduction

La **méthodologie des surfaces de réponse** (MSR) est une méthode permettant de modéliser un plan d'expériences à l'aide d'un **modèle polynomial d'ordre 2** (éventuellement plus) vis-à-vis des effets de facteurs.

$$y_{th} = a_0 + \sum_{i=1}^f a_i x_i + \sum_{i=1}^f a_{ii} x_i^2 + \sum_{i=1}^{f-1} \sum_{j=i+1}^f a_{ij} x_i x_j$$



---

## Dans quel but ?

- Savoir comment évolue une réponse avec un groupe de facteurs continus
- **Optimiser** un procédé



---

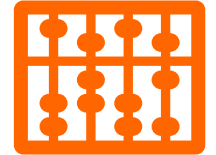
## Dans quel cas ?

- Peu de facteurs à étudier
- Plusieurs niveaux continus
- Effets complexes sur la réponse



# MSR – Calcul matriciel

La MSR nécessite *a minima 3 niveaux* pour chacun des facteurs afin d'en déterminer une courbure. Le nombre d'expériences dans le plan d'expériences devient alors supérieur à ceux vus jusque-là. Dans quasiment tous les cas en MSR, le nombre de coefficient est inférieur au nombre d'expériences ( $a < e$ ) et la matrice du modèle n'est alors pas inversible. Les coefficients du modèle sont donc estimés et *le calcul passe nécessairement par les matrices*. Les calculs matriciels fonctionnent de la même manière que vu précédemment avec les plans non orthogonaux.



Rappel du calcul de l'estimateur du vecteur des coefficients :

$$\hat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y$$

Un *résidu*  $\varepsilon$  est alors ajouté au modèle pour rendre compte de ce *manque d'approximation* :

$$y_{th} = a_0 + \sum_{i=1}^f a_i x_i + \sum_{i=1}^f a_{ii} x_i^2 + \sum_{i=1}^{f-1} \sum_{j=i+1}^f a_{ij} x_i x_j + \varepsilon$$

Le résidu est défini comme l'écart entre la réponse théorique calculée par le modèle et la réponse expérimentale mesurée, c'est-à-dire qu'il est un *indicateur de la qualité de prédiction* de la régression polynomiale du modèle sur les expériences.



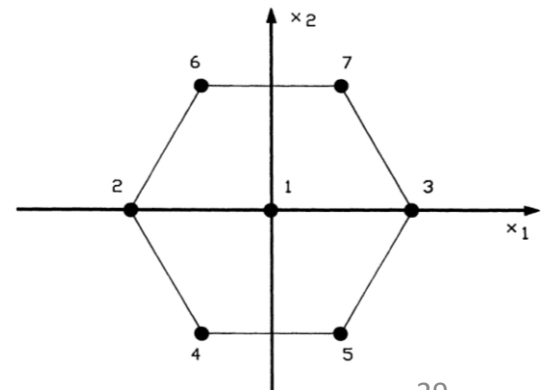
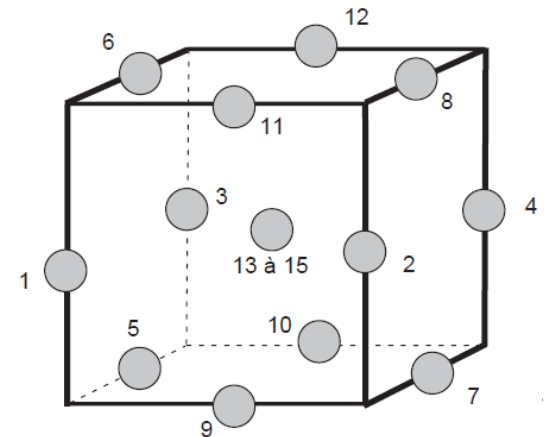
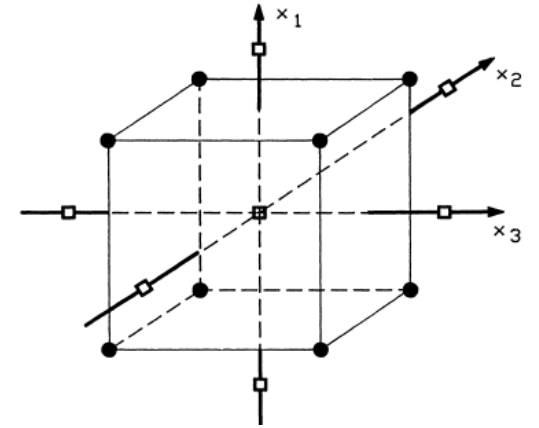
# MSR – Optimalité

Le calcul de l'estimateur est **dépendant du placement des points** dans le plan d'expériences, c'est-à-dire de la matrice du modèle  $X$ . Un placement optimal est conditionné par le **critère d'optimalité** choisi dont les plus courants peuvent permettre de réduire au maximum le résidu, et donc d'**estimer avec la meilleure précision possible**, en minimisant :

- La somme des variances des coefficients
- La variance globale de tous les coefficients
- La plus grande variance de coefficient...

Il existe **plusieurs plans pour la MSR** répondant à différents critères d'optimalité tels que :

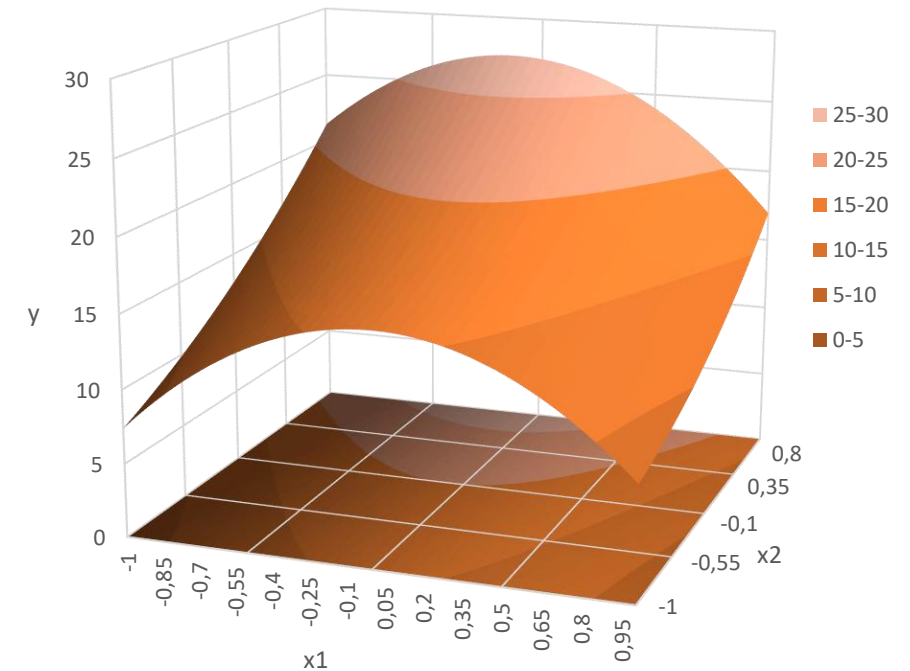
- Les **plans composites centrés** (ou plans en étoile) : basés sur des plans factoriels additionnés d'un point central et de points axiaux extérieurs disposés en étoile.
- Les **plans de Box-Behnken** : plans séquentiels permettant l'ajout de nouveaux facteurs au fur et à mesure tout en gardant les anciens points expérimentaux grâce à un placement à la surface d'une hypersphère.
- Les **réseaux de Doehlert** : permet l'exploration d'un domaine expérimental *via* une disposition homogène des points dans l'espace à l'aide de simplex.



# MSR – Surface de réponse

Après calcul matriciel, à partir du vecteur des coefficients, la surface de réponse issue du modèle mathématique théorique peut être tracée *via* une **représentation graphique de la réponse en fonction de deux variables**. Dans le cas des modèles avec plus de deux variables, les surfaces de réponse sont représentées par couple de variables. Il faut alors fixer les autres variables à une valeur définie (souvent 0 pour annuler tous les autres termes).

$$y_{th} = a_0 + \sum_{i=1}^f a_i x_i + \sum_{i=1}^f a_{ii} x_i^2 + \sum_{i=1}^{f-1} \sum_{j=i+1}^f a_{ij} x_i x_j + \varepsilon$$



Selon les coefficients du modèle, il existe plusieurs types de surface de réponse :

- **Maximum** : les effets d'ordre 2 sont prédominants et de même signe négatif.
- **Minimum** : les effets d'ordre 2 sont prédominants et de même signe positif.
- **Minimax** (surface en « selle de cheval » ou « paraboloïde hyperbolique ») : les effets d'ordre 2 sont prédominants et de signes opposés, et/ou l'interaction d'ordre 2 est prédominante.
- **Sans optimum** : les effets d'ordre 1 sont prédominants et l'optimum est donc hors domaine expérimental.

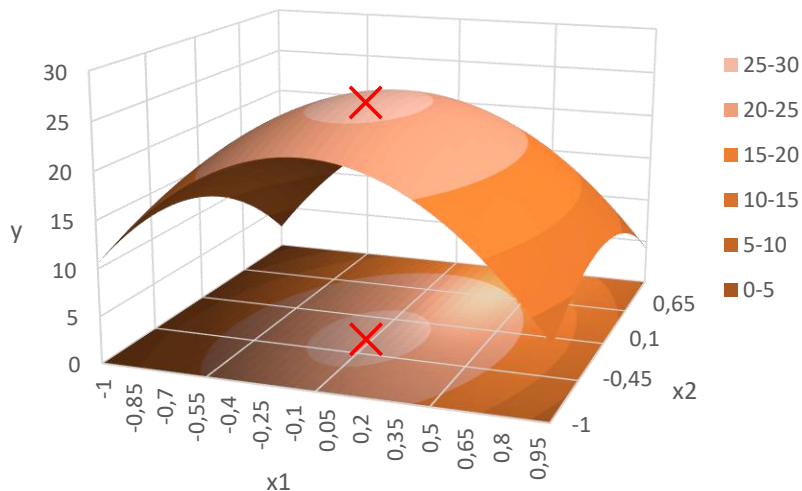
# MSR – Optimisation

Connaissant le modèle mathématique théorique, il est possible d'**estimer un optimum de réponse** à partir des dérivées partielles en  $x_i$  pour chaque facteur  $i$  (i.e.  $x_1, x_2 \dots x_f$ ) par la résolution d'un système à  $f$  équations et  $f$  inconnues :

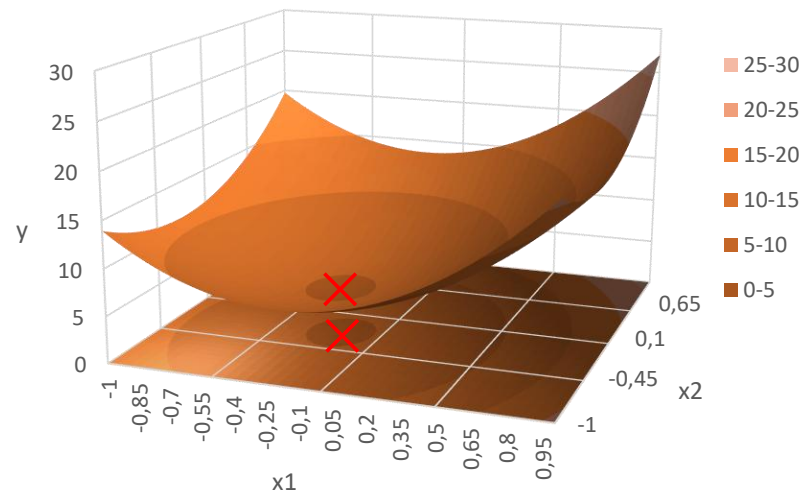
$$\begin{cases} \frac{\partial y_{th}}{\partial x_1} = a_1 + \sum_{i=2}^f a_{1i}x_i + 2a_{11}x_1 = 0 \\ \vdots \\ \frac{\partial y_{th}}{\partial x_f} = a_f + \sum_{i=1}^{f-1} a_{if}x_i + 2a_{ff}x_f = 0 \end{cases}$$

La résolution permet de prédire les niveaux des facteurs donnant la réponse optimale et la valeur optimale correspondante.

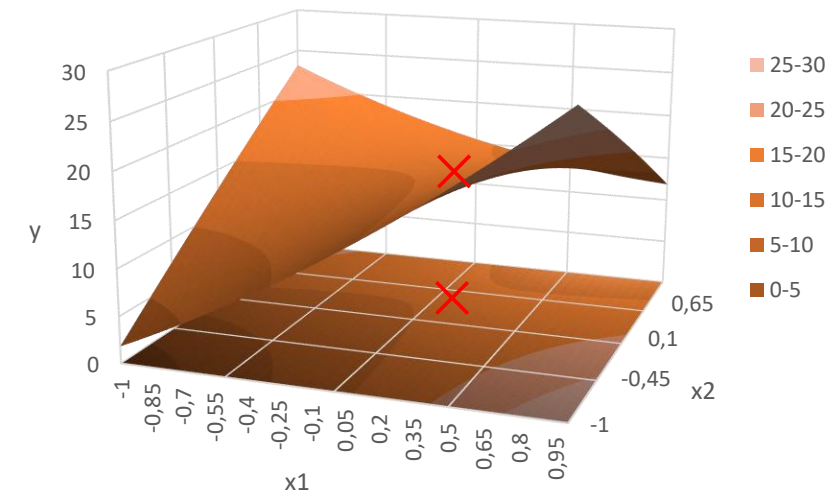
Maximum



Minimum



Minimax



# Réseaux de Doehlert – Introduction

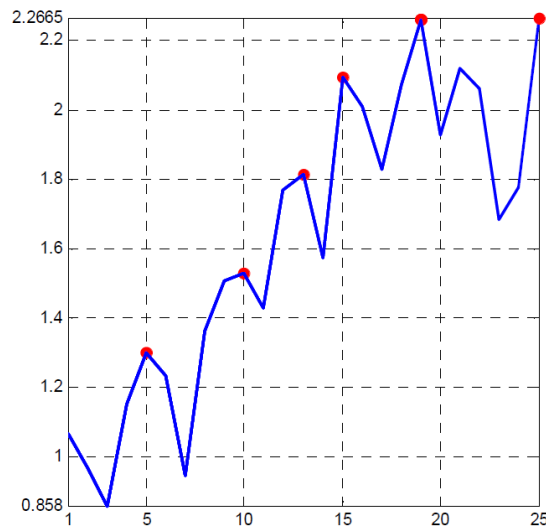
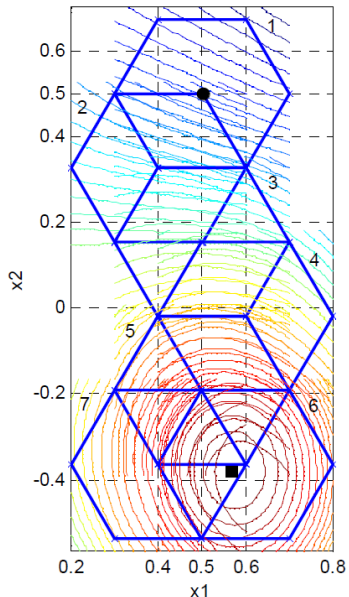
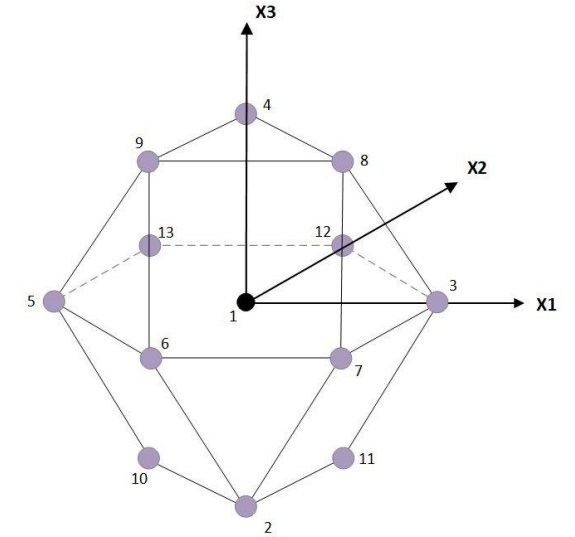
David H. **Doehlert** a proposé un modèle de plan d'expériences en 1970 utilisable pour la MSR et dont le principe est basé sur un **remplissage uniforme de l'espace via des simplexes**.

La méthode permet de compléter le plan d'expériences au fur et à mesure **vers un point optimum** : la méthode est dite « **itérative** ».



Permet d'économiser les essais en ne réalisant que les points d'intérêts.

$N = k^2 + k + 1$  essais à la base ( $k$  le nombre de facteurs) puis incrémentation vers la zone d'intérêt.



[Vivier 2002]



Formation d'un réseau à la recherche d'un optimum.

Méthode efficace pour trouver un optimum quand les **effets sont complexes** et difficiles à prédire par un modèle quadratique ou que la **localisation des optima est multiple** (optima dans plusieurs domaines expérimentaux) **et/ou non évidentes**.



# Réseaux de Doehlert – Propriétés

Les réseaux de Doehlert **ne respectent pas la règle d'orthogonalité** contrairement aux tables de Taguchi.

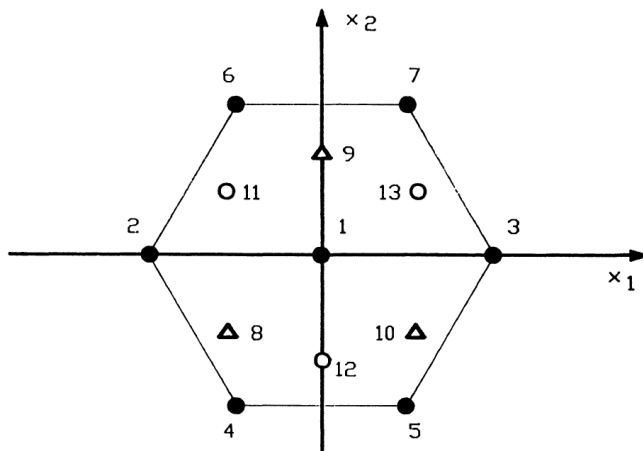
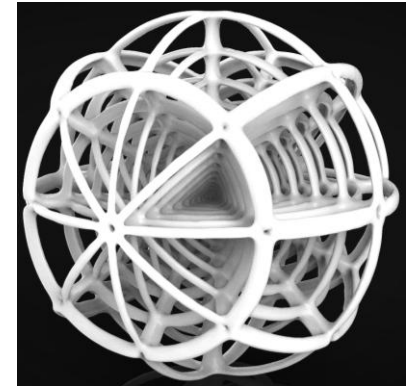
Critère d'optimalité :

**Isovariance par rotation** : La rotation du plan d'expériences permet de conserver la précision du modèle. Les erreurs de prédiction sont minimisées et égales pour chaque configuration hors point central ( $\varepsilon_1 = 0$ ).

$$\varepsilon_{i \neq 1} = cste$$

**Placement des points :**

Pour respecter le critère d'optimalité, les points sont placés à la surface d'une hypersphère. Le placement suit donc les règles de la trigonométrie.



$$y_{th} = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2 + \varepsilon$$

$$X(k=2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & -0,5 & -\sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \\ 1 & 0,5 & -\sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\ 1 & -0,5 & \sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\ 1 & 0,5 & \sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \end{bmatrix}$$

# Réseaux de Doehlert – Propriétés

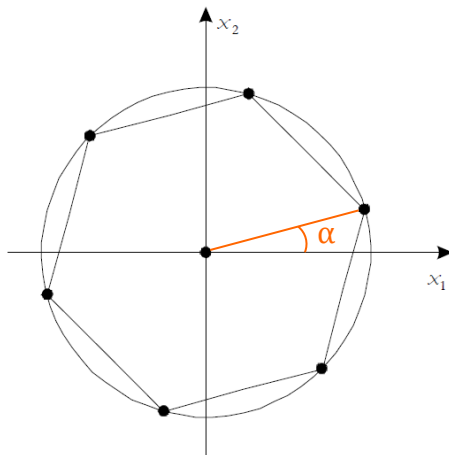
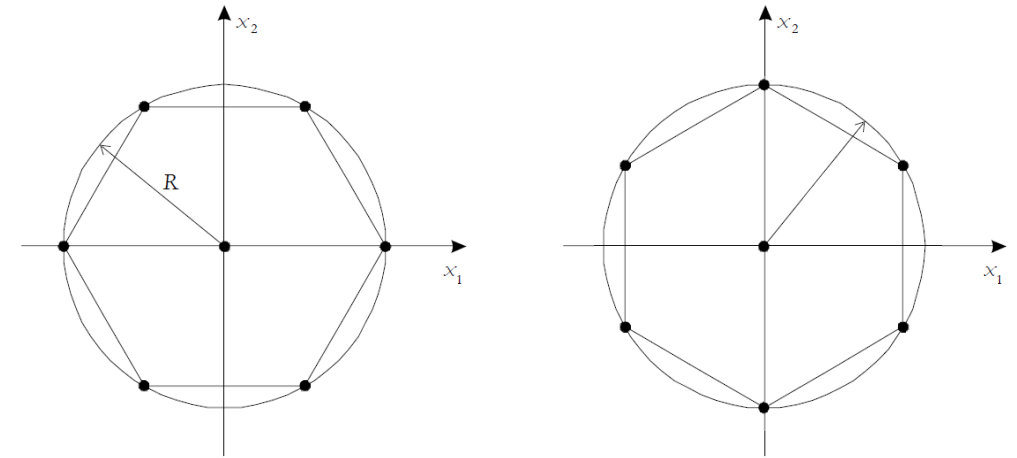
## Placement des points :

Grâce à la propriété d'isovariance par rotation, le plan de Doehlert peut être **librement orienté** pour **attribuer différents niveaux à chacun des facteurs**.

Exemple du cas à 2 facteurs :

## 3 et 5 niveaux ou 5 et 3 niveaux

= revient à mettre  $x_1$  à la place de  $x_2$  ou de faire pivoter le plan de  $90^\circ$ .



## 7 et 7 niveaux

Le plan peut être pivoté d'un angle  $\alpha$  compris entre  $0$  et  $60^\circ$

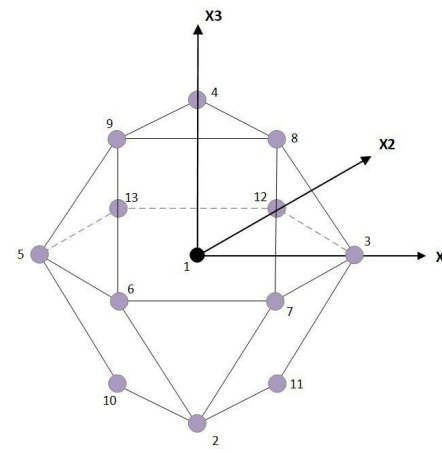
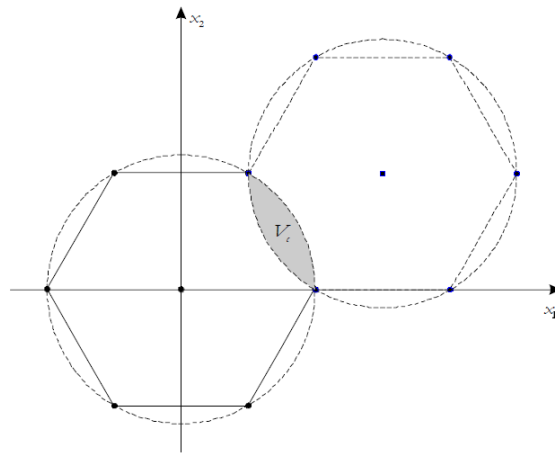
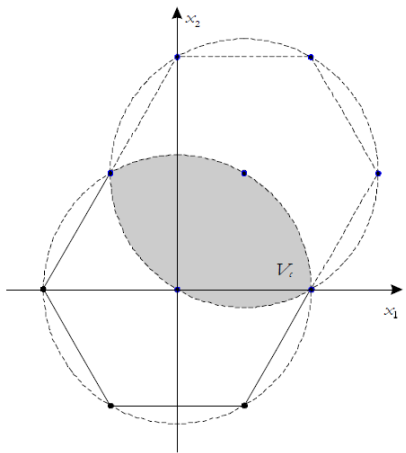
$$\alpha = ]0^\circ; 60^\circ[$$

Permet d'attribuer plus de valeurs aux facteurs, sans modifier le nombre d'expériences ou le domaine expérimental, mais peut vite devenir, sans logiciel, fastidieux en calculs.

# Réseaux de Doehlert – Itération

La méthode permet d'étendre le domaine expérimental en conservant un certain nombre de points du précédent plan et en ajoutant des nouveaux, soit :

- En ajoutant des **niveaux aux facteurs**
- En ajoutant un **nouveau facteur**

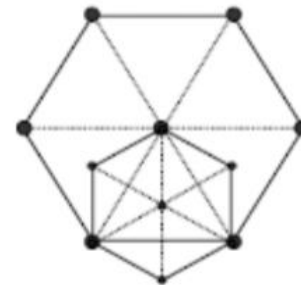


Tout point placé dans un plan de dimension  $n$  considère un niveau égal à 0 pour le facteur dans la dimension  $(n+1)$ .

Il suffit alors d'**ajouter les points dans la dimension  $(n+1)$**  pour **ajouter un nouveau facteur**.



Le **volume commun** entre deux plans consécutifs dépend du **nombre d'expériences conservées**. Le **glissement du plan** dépend de la direction souhaitée pour rechercher l'optimum.

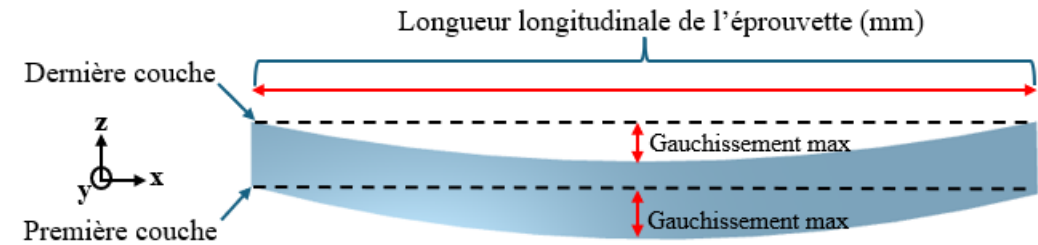
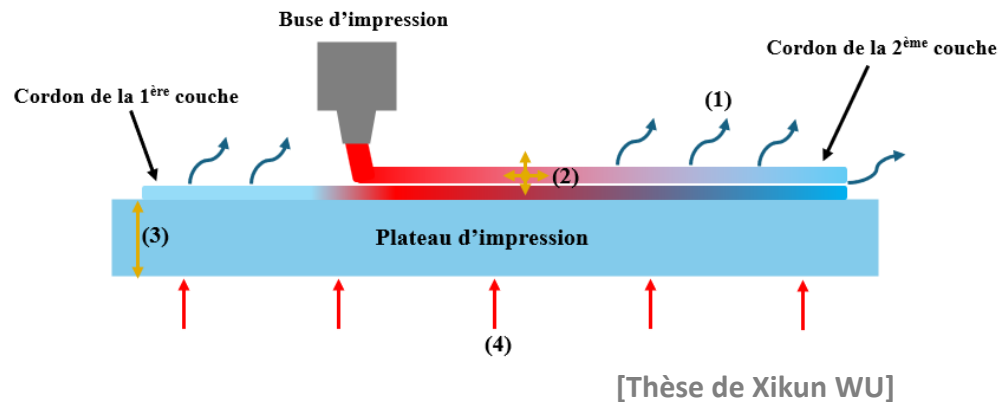


Une méthode dérivée permet de « **zoomer** » dans un domaine expérimental pour **affiner la prédiction**.

# Réseaux de Doehlert – Application

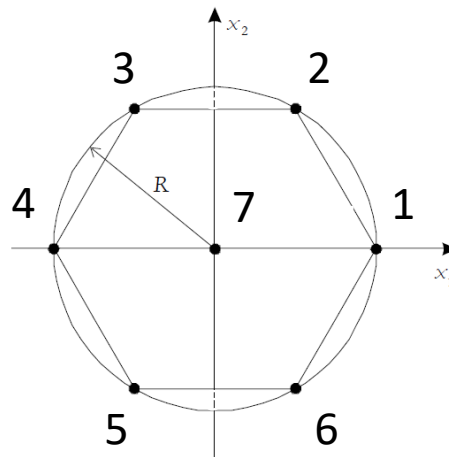
Cas d'un plan à 2 facteurs :

Étude de la hauteur de couche (facteur 1) et de la température (facteur 2) sur la déflexion d'une éprouvette imprimée.



On répartit 7 points aux coordonnées :

$x_1$	$x_2$
1	0
0,5	$\sqrt{3}/2$
-0,5	$\sqrt{3}/2$
-1	0
-0,5	$-\sqrt{3}/2$
0,5	$-\sqrt{3}/2$
0	0



avec pour réponses :  $Y = \begin{Bmatrix} 3,1 \\ 2,2 \\ 3,5 \\ 3,8 \\ 2,7 \\ 3,1 \\ 4,4 \end{Bmatrix}$  (en mm)

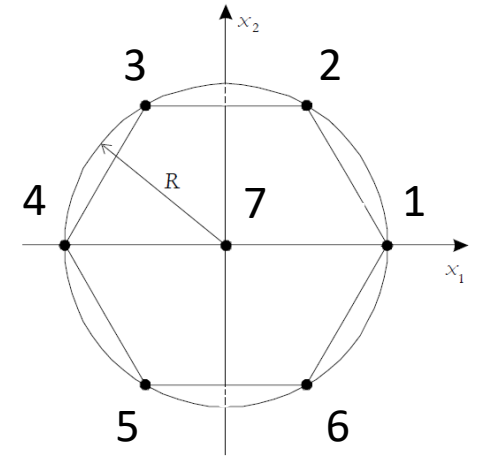
# Réseaux de Doehlert – Application

On veut déterminer les coefficients du modèle :

$$y_{th} = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2 + \varepsilon$$

En utilisant :  $\hat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y$  avec la matrice du modèle :

$$\begin{array}{cc}
 x_1 & x_2 \\
 1 & 0 \\
 0,5 & \sqrt{3}/2 \\
 -0,5 & \sqrt{3}/2 \\
 -1 & 0 \\
 -0,5 & -\sqrt{3}/2 \\
 0,5 & -\sqrt{3}/2 \\
 0 & 0
 \end{array}
 \longrightarrow
 X = \begin{array}{c}
 \begin{array}{cccccc}
 x_0 & x_1 & x_2 & x_1x_2 & x_1^2 & x_2^2 \\
 1 & 1 & 0 & 0 & 1 & 0 \\
 1 & 0,5 & \sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \\
 1 & -0,5 & \sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\
 1 & -1 & 0 & 0 & 1 & 0 \\
 1 & -0,5 & -\sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \\
 1 & 0,5 & -\sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\
 1 & 0 & 0 & 0 & 0 & 0
 \end{array}
 \end{array}$$



Donc

$$\hat{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1/3 & 1/6 & -1/6 & -1/3 & -1/6 & 1/6 & 0 \\ 0 & 1/2\sqrt{3} & 1/2\sqrt{3} & 0 & -1/2\sqrt{3} & -1/2\sqrt{3} & 0 \\ 0 & 1/\sqrt{3} & -1/\sqrt{3} & 0 & 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 0,5 & 0 & 0 & 0,5 & 0 & 0 & -1 \\ -1/6 & 1/3 & 1/3 & -1/6 & 1/3 & 1/3 & -1 \end{bmatrix} \cdot \begin{Bmatrix} 3,1 \\ 2,2 \\ 3,5 \\ 3,8 \\ 2,7 \\ 3,1 \\ 4,4 \end{Bmatrix} = \begin{Bmatrix} 4,4 \\ -0,383 \\ -0,029 \\ -0,981 \\ -0,950 \\ -1,717 \end{Bmatrix} \quad \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \\ a_{11} \\ a_{22} \end{Bmatrix}$$

# Réseaux de Doehlert – Application

Déterminer les coordonnées de l'optimum et prédire sa réponse.

Pour un modèle à deux facteurs, nous devons résoudre le système à 2 équations et 2 inconnues :

$$\begin{cases} \frac{\partial y_{th}}{\partial x_1} = a_1 + a_{12}x_2 + 2a_{11}x_1 = 0 \\ \frac{\partial y_{th}}{\partial x_2} = a_2 + a_{12}x_1 + 2a_{22}x_2 = 0 \end{cases}$$

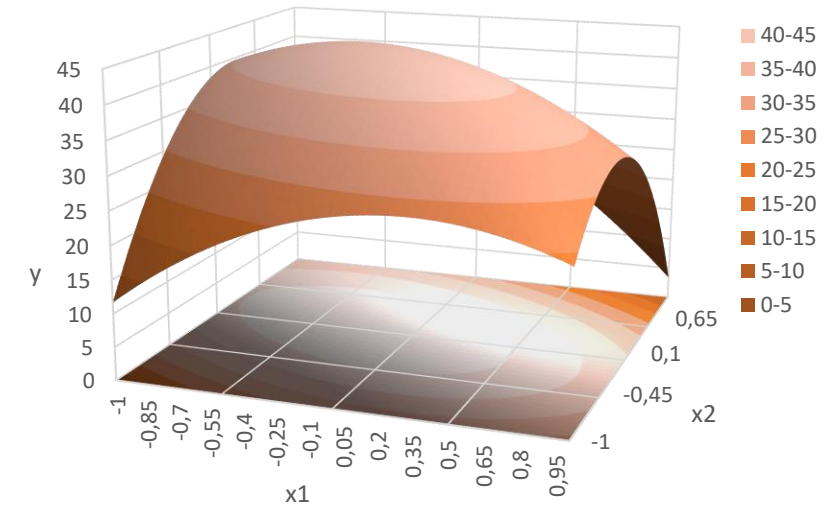
D'où

$$\begin{cases} x_1 = \frac{a_2 a_{12} - 2a_1 a_{22}}{4a_{11}a_{22} - a_{12}^2} = -0,232 \\ x_2 = \frac{a_1 a_{12} - 2a_2 a_{11}}{4a_{11}a_{22} - a_{12}^2} = 0,058 \end{cases}$$

Connaissant le modèle théorique et les coordonnées de l'optimum :

$$y_{th} = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2 + \varepsilon$$

$$\hat{y}_{th,max} = 4,4 + 0,383 \times 0,232 - 0,029 \times 0,058 + 0,981 \times 0,232 \times 0,058 - 0,950 \times 0,232^2 - 1,717 \times 0,058^2 = 4,44 \text{ mm}$$



$$\hat{A} = \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \\ a_{11} \\ a_{22} \end{Bmatrix} = \begin{Bmatrix} 4,4 \\ -0,383 \\ -0,029 \\ -0,981 \\ -0,950 \\ -1,717 \end{Bmatrix}$$

# Réseaux de Doehlert – Application

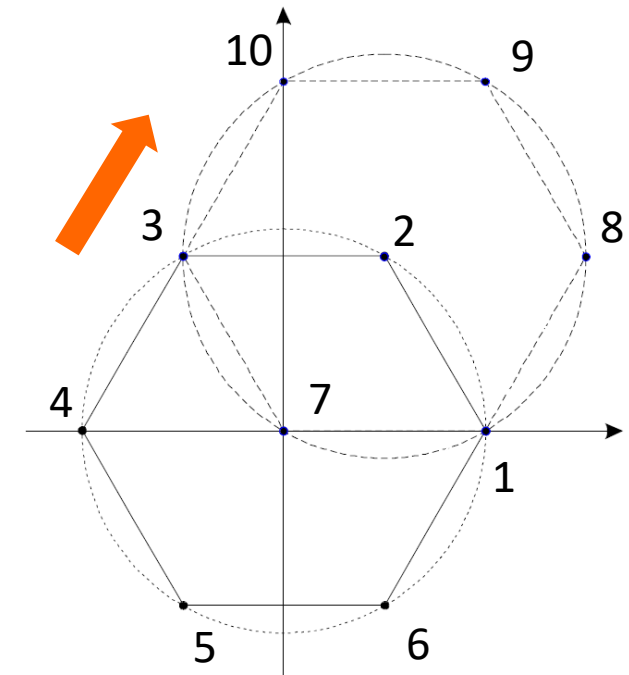
On souhaite plutôt **minimiser la réponse**. Comme  $a_1 < 0$  et  $a_2 < 0$ , la réponse tend à diminuer avec les niveaux supérieurs. On décide donc de faire glisser le plan vers des niveaux de  $x_1$  et de  $x_2$  supérieurs (couche plus épaisse, température plus haute).

4 points sont conservés (1, 2, 3 et 7) et 3 autres sont ajoutés (8, 9 et 10) :

$$Y = \begin{Bmatrix} y_8 \\ y_9 \\ y_{10} \\ y_3 \\ y_7 \\ y_1 \\ y_2 \end{Bmatrix} = \begin{Bmatrix} 2,5 \\ 3,2 \\ 3,1 \\ 3,5 \\ 4,4 \\ 3,1 \\ 2,2 \end{Bmatrix}$$

Le calcul reste le même en recentrant les variables sur le point n° 2 :

$$\hat{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1/3 & 1/6 & -1/6 & -1/3 & -1/6 & 1/6 & 0 \\ 0 & 1/2\sqrt{3} & 1/2\sqrt{3} & 0 & -1/2\sqrt{3} & -1/2\sqrt{3} & 0 \\ 0 & 1/\sqrt{3} & -1/\sqrt{3} & 0 & 1/\sqrt{3} & -1/\sqrt{3} & 0 \\ 0,5 & 0 & 0 & 0,5 & 0 & 0 & -1 \\ -1/6 & 1/3 & 1/3 & -1/6 & 1/3 & 1/3 & -1 \end{bmatrix} \cdot \begin{Bmatrix} 2,5 \\ 3,2 \\ 3,1 \\ 3,5 \\ 4,4 \\ 3,1 \\ 2,2 \end{Bmatrix} = \begin{Bmatrix} 2,2 \\ -0,533 \\ -0,346 \\ 0,808 \\ 0,800 \\ 1,400 \end{Bmatrix}$$



# Réseaux de Doehlert – Application

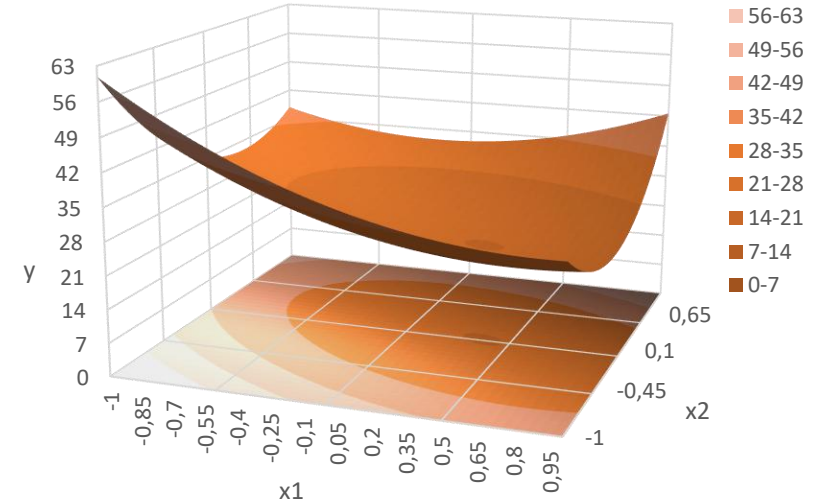
La détermination de l'optimum se fait de la même manière :

Après résolution :

$$\begin{cases} x_1 = \frac{a_2 a_{12} - 2a_1 a_{22}}{4a_{11}a_{22} - a_{12}^2} = 0,317 \\ x_2 = \frac{a_1 a_{12} - 2a_2 a_{11}}{4a_{11}a_{22} - a_{12}^2} = 0,032 \end{cases}$$

Et on en déduit :

$$\hat{y}_{th,min} = a_0 + a_1 x_1 + a_2 x_2 + a_{12} x_1 x_2 + a_{11} x_1^2 + a_{22} x_2^2 = 2,11 \text{ mm}$$



$$\hat{A} = \begin{Bmatrix} a_0 \\ a_1 \\ a_2 \\ a_{12} \\ a_{11} \\ a_{22} \end{Bmatrix} = \begin{Bmatrix} 2,2 \\ -0,533 \\ -0,346 \\ 0,808 \\ 0,800 \\ 1,400 \end{Bmatrix}$$

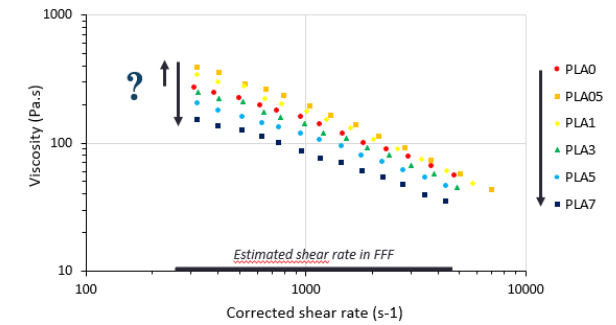
Les premiers calculs réalisés, les itérations deviennent simples car basées sur le premier plan.

La précision du modèle reste cependant contrainte par les erreurs de prédiction. L'application de ces calculs n'est donc valide que si le modèle théorique correspond à l'expérimentation. Le cas contraire, la prédiction serait trop éloignée de la réalité et le modèle n'aurait plus aucun sens.



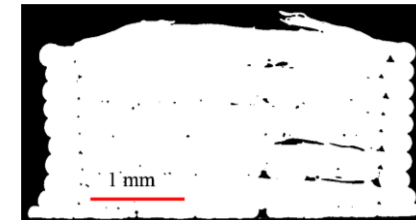
# Réseaux de Doehlert – Application

Optimisation matériaux x comportement rhéologique x propriétés



$$CY : \left( \eta(\dot{\gamma}) = \eta_0 (1 + (\lambda \dot{\gamma})^a)^{\frac{n-1}{a}} \right)$$

[Ginoux 2020]

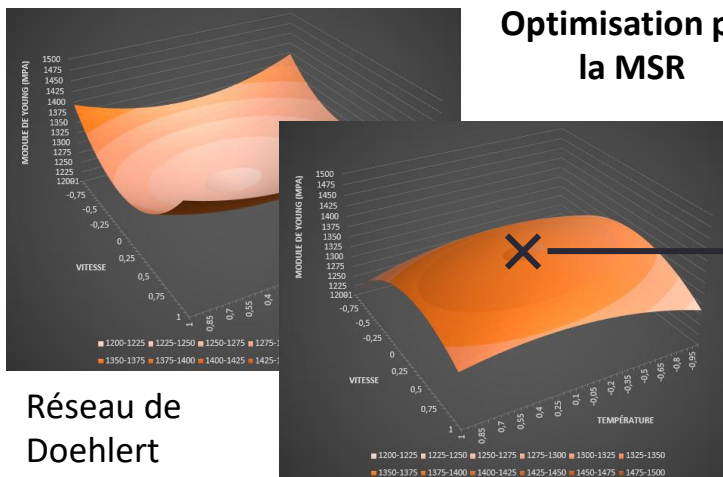


[Ginoux 2021b]

$$\eta(\dot{\gamma}, \bar{M}_w, T, \varphi_v) = \eta_{0, \bar{M}_{w0}, T_0}(\varphi_v) \left( \frac{\bar{M}_w}{\bar{M}_{w0}} \right)^\alpha \exp \left[ \frac{E_{aV} \left( \begin{cases} \varphi_v \leq \Phi_c \\ \varphi_v > \Phi_c \end{cases} \right) \left( \frac{1}{T} - \frac{1}{T_0} \right)}{R} \right]$$

$$\times \left[ 1 + \left( \lambda_{\bar{M}_{w0}, T_0}(\varphi_v) \left( \frac{\bar{M}_w}{\bar{M}_{w0}} \right)^\alpha \dot{\gamma} \exp \left[ \frac{E_{aH} \left( \begin{cases} \varphi_v \leq \Phi_c \\ \varphi_v > \Phi_c \end{cases} \right) \left( \frac{1}{T} - \frac{1}{T_0} \right)}{R} \right] \right)^a \right]^{\frac{n_2 \varphi_v^2 + n_1 \varphi_v + n_0 - 1}{a}}$$

[Ginoux 2021a]



Réseau de Doehlert

Optimisation par la MSR

	Vitesse (mm/s)	Température (°C)	Viscosité (Pa.s)
PLA0	70	211	60,3
PLA05	60	221	46,7
PLA1	95	239	12,9
PLA3	81	217	21,8
PLA5	72	192	56,1

Variation similaire par rapport à la rhéologie capillaire

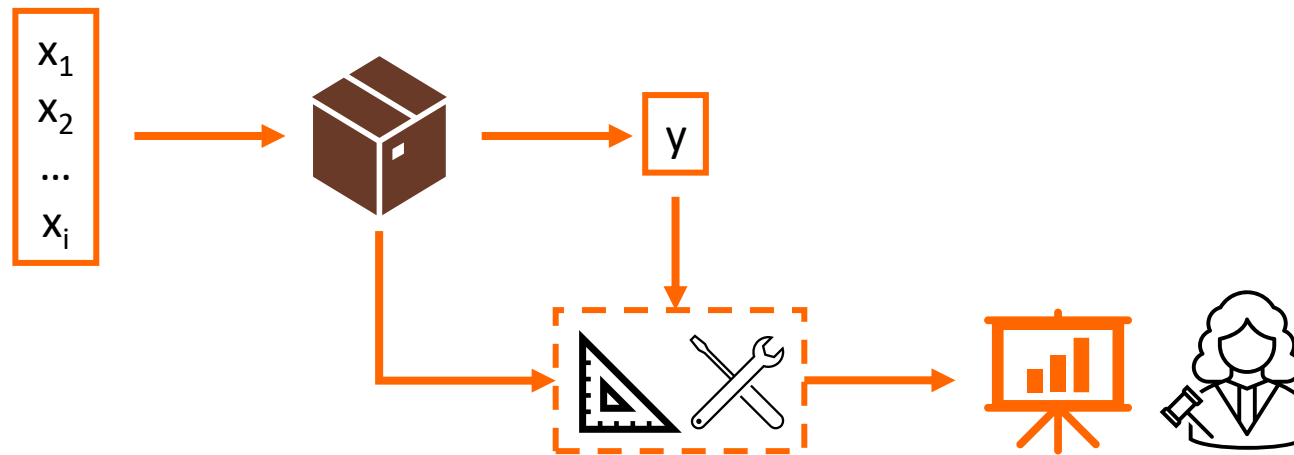
Température paramétrée vs. Température matière ? Transfert thermique ?

# Mieux interpréter avec les outils statistiques

## Limites des plans d'expériences :

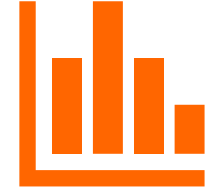
Les plans d'expériences sont des méthodes efficaces pour ordonner les essais et prédire un résultat. Cependant, la compréhension du système qui en découle et l'acquisition des connaissances nécessitent une bonne compréhension des résultats qui peuvent être sujette à l'**appréciation subjective de l'expérimentateur**.

Pour **éviter les erreurs de jugement**, les **outils statistiques** sont un moyen de **rendre la méthode plus objective** puisqu'ils permettent **une interprétation plus quantitative que qualitative**. Les outils statistiques sont une aide décisionnelle pour les **choix opérateur**.



# Statistiques – Introduction

Les **outils statistiques** regroupent un ensemble de **méthodes mathématiques** servant à expliciter les résultats. Il en existe de nombreuses dont l'**ANOVA** (*ANalysis Of VAriance*) qui reste la méthode la plus répandue pour la méthodologie des plans d'expériences.



---

## Dans quel but ?

- Déterminer les **effets influents** de ceux non ou peu influents
- Mieux interpréter les résultats
- **Aider à la prise de décision**



---

## Dans quel cas ?

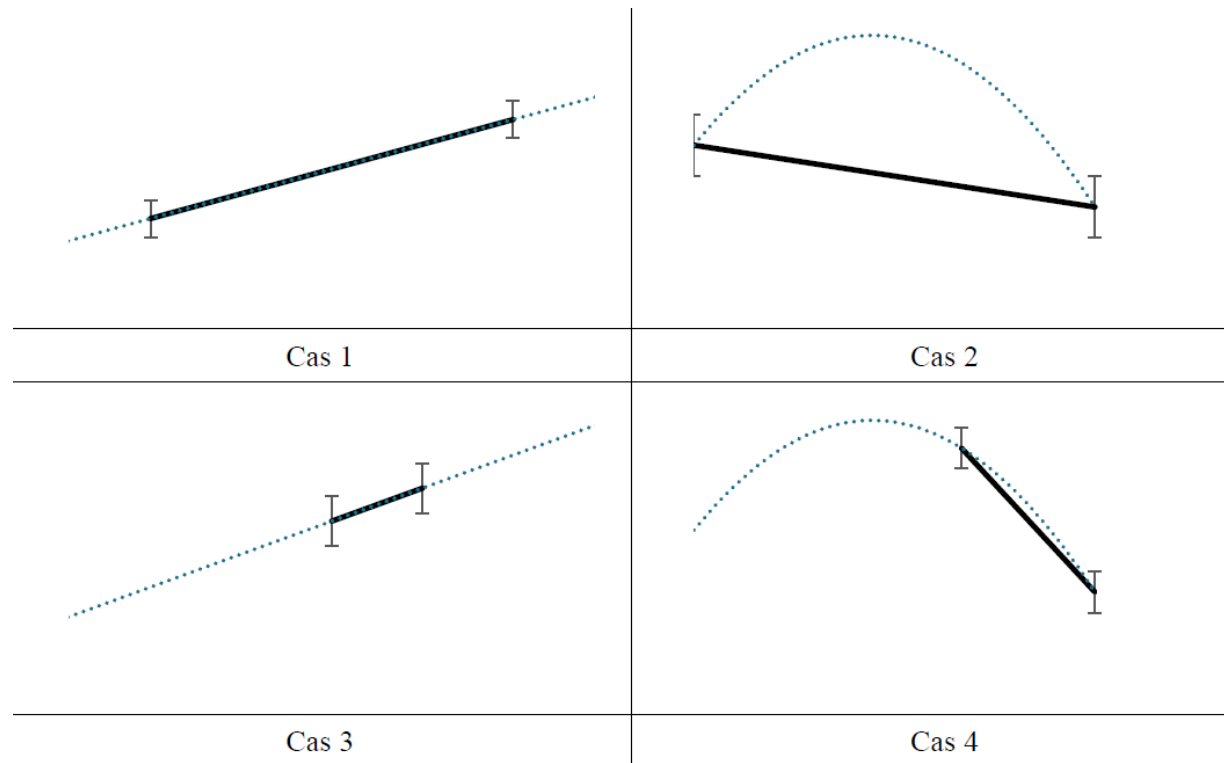
- Faibles écarts entre les effets et entre les effets et les erreurs expérimentales
- Interactions complexes
- Écart entre la prédiction et l'expérience



# Statistiques – Importance des choix opérateur

L'erreur expérimentale est souvent l'objet d'**erreurs dues à la mesure**, pouvant être relativement fixes (erreur systématique). Il convient alors de prendre des **valeurs de niveaux assez éloignées pour observer un effet suffisant** par rapport à l'erreur expérimentale, mais **pas trop éloignées pour respecter l'hypothèse de linéarité** (en particulier pour le criblage).

**Cas idéal** : Effet linéaire et supérieur à l'erreur expérimentale



Effet non linéaire, non représentatif de la réalité

Effet linéaire, mais l'erreur expérimentale dépasse l'effet

**Cas réaliste et préférable** : Effet linéaire dans le domaine expérimental et supérieur aux erreurs

# Statistiques – Coefficients de détermination

Le coefficient de détermination est un **indicateur de la conformité du modèle** supposé vis-à-vis des expériences et donc du système étudié. Il est défini comme la variance expliquée par la régression sur la variance des mesures :

$$R^2 = \frac{SC_{mod}}{SC_{mes}} = \frac{\sum_{i=1}^n (y_{th,i} - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Avec  $SC_{mod}$  la somme des carrés du modèle issu de la régression  $y_{th}$  corrigée de la moyenne  $\bar{y}$  et  $SC_{mes}$  la somme des carrés des expériences mesurées  $y$  corrigées de la moyenne  $\bar{y}$ .

Pour un résidu suffisamment faible, on peut montrer que :

$$SC_{mes} = SC_{mod} + SC_r$$

Avec  $SC_r$  la somme des carrés des résidus.

Le coefficient de détermination peut alors s'écrire :

$$R^2 = 1 - \frac{SC_r}{SC_{mes}} = 1 - \frac{\sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Dans le cas des **plans orthogonaux**, les résidus sont nuls et  $R^2 = 1$  : la précision du modèle est au plus juste.

# Statistiques – Coefficients de détermination

Une augmentation du nombre de variables entraîne au passage celle du coefficient de détermination au détriment de la robustesse du modèle. La **considération du nombre de variables explicatives** permet alors de **s'affranchir du biais** causé par elles. Ainsi, il est possible de **comparer deux modèles** ayant un nombre de variables différent pour les mêmes expériences.

Le coefficient de détermination ajusté est alors utilisé et défini comme :

$$R_{aj}^2 = 1 - \frac{CM_r}{CM_{mes}}$$

Avec  $CM_r$  le carré moyen des résidus et  $CM_{mes}$  le carré moyen des mesures expérimentales corrigées de la moyenne tels que  $CM = SC/ddl$ .

D'où :

$$R_{aj}^2 = 1 - \frac{SC_r/ddl_r}{SC_{mes}/ddl_{mes}} = 1 - \frac{n-1}{n-p} \times \frac{\sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Avec  $ddl_r$  le degré de liberté des résidus,  $ddl_{mes}$  le degré de liberté des mesures corrigées de la moyenne,  $n$  le nombre de configurations et  $p$  le nombre de coefficient du modèle.

# Statistiques – ANOVA

L'**ANOVA** compare la **variance des coefficients** ou du modèle par rapport à celle des résidus. Elle permet d'évaluer l'**importance des coefficients** du modèle et la **significativité du modèle**. Il existe plusieurs tests dont le **test de Student** et le **test de Fisher**. Les *valeurs*–*t* (Student) et les *valeurs*–*F* (Fisher) peuvent être calculées puis comparées à celles dans les tables pour en déterminer la significativité (si **valeur calculée > valeur table**).

Pour les coefficients :

$$\text{valeur-t} = \frac{a_i}{\sigma_a}$$

Avec  $a_i$  la valeur du coefficient à étudier et  $\sigma_a$  l'écart-type des coefficients défini comme étant la racine du rapport des carrés moyens des résidus sur le nombre d'expériences ( $\sigma_a = \sqrt{CM_r/n}$ ). Soit :

$$\text{valeur-t} = \sqrt{\frac{n(n-p)}{\sum_{i=1}^n \varepsilon_i^2}} \cdot a_i$$

Pour le modèle :

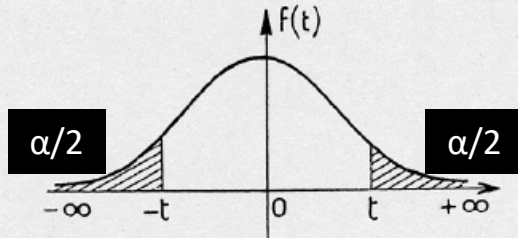
$$\text{valeur-F} = \frac{CM_{mod}}{CM_r}$$

Soit :

$$\text{valeur-F} = \frac{n-p}{p-1} \cdot \frac{\sum_{i=1}^n (y_{th,i} - \bar{y})^2}{\sum_{i=1}^n \varepsilon_i^2}$$



# Statistiques – Table de Student



## Test de Student

La *valeur-t* de la table est obtenue à partir de la probabilité  $P = 1 - \alpha$  que la valeur soit significative qui est fonction du degré de liberté.

Pour l'ANOVA, la *valeur-t* se calcule à partir du **ddl des résidus**. Il s'agit alors de lire la valeur de la table à partir du  $ddl_r$  au niveau de significativité choisi et de la comparer à la valeur calculée.

Si  $t_{calculé} > t_{table}$ , alors le coefficient est supposé significatif, et donc influent sur la réponse.

$\alpha$ ddl	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,478	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291



# Statistiques – Tables de Fisher-Snedecor

TABLE DE F (FISHER-SNEDECOR),  $P = 0,95$

		Degrés de liberté du numérateur : $v_1$											
		1	2	3	4	5	6	7	8	9	10	12	14
Degrés de liberté du dénominateur : $v_2$	1	161	200	216	225	230	234	237	239	241	242	244	245
	2	18,5	19	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4
	3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,71
	4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6	5,96	5,91	5,87
	5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,64
	6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,1	4,06	4	3,96
	7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,53
	8	5,32	4,46	4,07	3,84	3,69	3,58	3,5	3,44	3,39	3,35	3,28	3,24
	9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,03
	10	4,96	4,1	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,86
	11	4,84	3,98	3,59	3,36	3,2	3,09	3,01	2,95	2,9	2,85	2,79	2,74
	12	4,75	3,89	3,49	3,26	3,11	3	2,91	2,85	2,8	2,75	2,69	2,64
	13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,6	2,55
	14	4,6	3,74	3,34	3,11	2,96	2,85	2,76	2,7	2,65	2,6	2,53	2,48
	15	4,54	3,68	3,29	3,06	2,9	2,79	2,71	2,64	2,59	2,54	2,48	2,42
	16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,37
	17	4,45	3,59	3,2	2,96	2,81	2,7	2,61	2,55	2,49	2,45	2,38	2,33
	18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,29
	19	4,38	3,52	3,13	2,9	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,26
	20	4,35	3,49	3,1	2,87	2,71	2,6	2,51	2,45	2,39	2,35	2,28	2,22
	21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,2
	22	4,3	3,44	3,05	2,82	2,66	2,55	2,46	2,4	2,34	2,3	2,23	2,17
	23	4,28	3,42	3,03	2,8	2,64	2,53	2,44	2,37	2,32	2,27	2,2	2,15
	24	4,26	3,4	3,01	2,78	2,62	2,51	2,42	2,36	2,3	2,25	2,18	2,13
	25	4,24	3,39	2,99	2,76	2,6	2,49	2,4	2,34	2,28	2,24	2,16	2,11
	26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,09
	27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,2	2,13	2,08
	28	4,2	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,06
	29	4,18	3,33	2,93	2,7	2,55	2,43	2,35	2,28	2,22	2,18	2,1	2,05
	30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,04
	32	4,15	3,29	2,9	2,67	2,51	2,4	2,31	2,24	2,19	2,14	2,07	2,01
	34	4,13	3,28	2,88	2,65	2,49	2,38	2,29	2,23	2,17	2,12	2,05	1,99
	36	4,11	3,26	2,87	2,63	2,48	2,36	2,28	2,21	2,15	2,11	2,03	1,98
	38	4,1	3,24	2,85	2,62	2,46	2,35	2,26	2,19	2,14	2,09	2,02	1,96
	40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2	1,95
	50	4,03	3,18	2,79	2,56	2,4	2,29	2,2	2,13	2,07	2,03	1,95	1,89
	60	4	3,15	2,76	2,53	2,37	2,25	2,17	2,1	2,04	1,99	1,92	1,86
	70	3,98	3,13	2,74	2,5	2,35	2,23	2,14	2,07	2,02	1,97	1,89	1,84
	80	3,96	3,11	2,72	2,49	2,33	2,21	2,13	2,06	2	1,95	1,88	1,82
	90	3,95	3,1	2,71	2,47	2,32	2,2	2,11	2,04	1,99	1,94	1,86	1,8
	100	3,94	3,09	2,7	2,46	2,31	2,19	2,1	2,03	1,97	1,93	1,85	1,79
	$\infty$	3,84	3	2,6	2,37	2,21	2,1	2,01	1,94	1,88	1,83	1,75	1,69

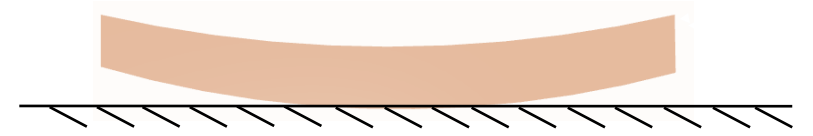
## Test de Fisher

La *valeur*—F de la table est obtenue à partir du ddl au numérateur ( $v_1$ ) et de celui au dénominateur ( $v_2$ ). Il existe plusieurs tables : chaque table correspond aux valeurs obtenues à une probabilité  $P = 1 - \alpha$  donnée. Le cas le plus courant est  $P = 0,95$ .

Pour l'ANOVA, la *valeur*—F se calcule à partir du ddl du modèle et du ddl des résidus. Il s'agit alors de lire la valeur de la table à partir du *ddl<sub>mod</sub>* au numérateur et du *ddl<sub>r</sub>* au dénominateur.

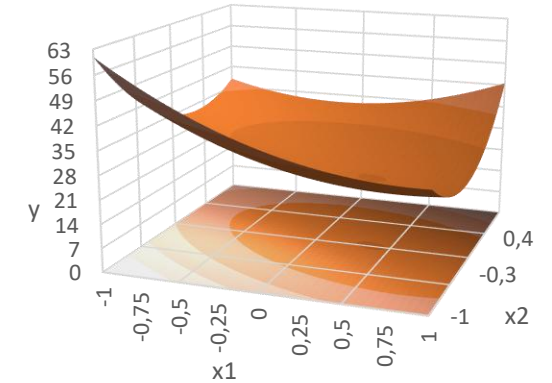
Si  $F_{calculé} > F_{table}$ , alors le modèle est supposé significatif, et donc apte à prédire la réponse.

# Statistiques – Application ANOVA



Reprenons le cas du réseau de Doehlert vu précédemment :

Pour  $Y = \begin{Bmatrix} 2,5 \\ 3,2 \\ 3,1 \\ 3,5 \\ 4,4 \\ 3,1 \\ 2,2 \end{Bmatrix}$  et  $y_{th} = a_0 + a_1x_1 + a_2x_2 + a_{12}x_1x_2 + a_{11}x_1^2 + a_{22}x_2^2 + \varepsilon$ ,



l'estimateur des coefficients était :  $\hat{A} = \begin{Bmatrix} 2,2 \\ -0,533 \\ -0,346 \\ 0,808 \\ 0,800 \\ 1,400 \end{Bmatrix}$ . On en déduit des réponses prédites théoriquement  $y_{th,i}$  :

$$\hat{Y} = X \cdot \hat{A} = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0,5 & \sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \\ 1 & -0,5 & \sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -0,5 & -\sqrt{3}/2 & \sqrt{3}/4 & 0,25 & 0,75 \\ 1 & 0,5 & -\sqrt{3}/2 & -\sqrt{3}/4 & 0,25 & 0,75 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{Bmatrix} 2,2 \\ -0,567 \\ -0,346 \\ 0,808 \\ 0,75 \\ 1,417 \end{Bmatrix} = \begin{Bmatrix} 2,47 \\ 3,23 \\ 3,07 \\ 3,53 \\ 4,37 \\ 3,13 \\ 2,20 \end{Bmatrix} \quad \varepsilon = 0,03$$

# Statistiques – Application ANOVA

Importance des coefficients :

*valeur-t* = valeur sur  
l'écart-type du coef.  
en valeur absolue

Coefficient	Valeur	Ecart-type	<i>valeur-t</i>	Significativité
$a_0$	2,20	0,3086	71,29	Oui
$a_1$	-0,533	0,3086	17,28	Oui
$a_2$	-0,346	0,3086	11,22	Non
$a_{12}$	-0,808	0,3086	26,20	Oui
$a_{11}$	-0,800	0,3086	25,92	Oui
$a_{22}$	-1,400	0,3086	45,37	Oui

> 12,706 ?

On calcule d'abord  $\sigma_a = \sqrt{CM_r/n} = \sqrt{(n-1)\varepsilon^2/(n(n-p))}$  avec  $n = 7$  et  $p = 6$ , et  $\varepsilon = 0,03$ .

On calcule ainsi la *valeur-t* que l'on compare à la valeur de la table.

Pour  $ddl_r = n - p = 1$  et  $P = 1 - \alpha = 0,95$  : *valeur-t* = 12,706.

On en déduit que les effets de premier ordre des facteurs ne sont pas significatifs à 95 % de probabilité.

# Statistiques – Application ANOVA

Significativité du modèle :

Source var.	SC	ddl	CM	valeur-F	Significativité
Modèle	3,0105	5	0,6021	90,31	Non
Erreur	0,0067	1	0,0067		
Total	313,72				

> 230 ?

On calcule la somme des carrés pour chacune des sources de variation :

$$SC_{mod} = \sum_{i=1}^n (y_{th,i} - \bar{y})^2 \quad \text{et} \quad SC_r = (n - 1)\varepsilon^2$$

On calcule les degrés de liberté :

$$ddl_{mod} = p - 1 = 5 \quad \text{et} \quad ddl_r = n - p = 1$$

On calcule les carrés moyens comme étant le rapport  $SC/ddl$ .

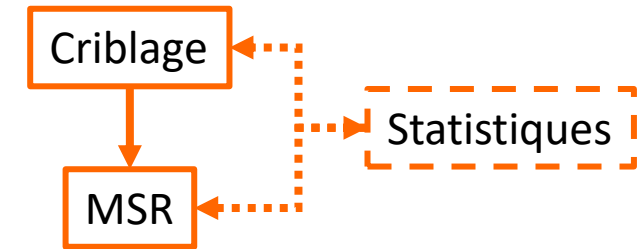
On calcule la  $valeur-F = CM_{mod}/CM_r$  que l'on compare à celle obtenue avec la table.

Pour  $ddl_{mod} = 5$ ,  $ddl_r = 1$  et  $P = 1 - \alpha = 0,95$  : **valeur-F = 230**.

**Le modèle n'est pas significatif à 95 % de probabilité.** Cependant, un rapide calcul sur tableur indique qu'il est significatif à 92 % de probabilité.

# Résumé de la méthodologie des plans d'expériences

- (1) Définir la problématique
- (2) Sélectionner le plan (type de plan, réponse(s), facteurs, niveaux)
- (3) Faire les essais
- (4) Analyser les résultats (calculs matriciels, Pareto, ANOVA, etc.)
- (5) Vérifier les résultats (expérience de confirmation)



Si beaucoup de facteurs : sélectionner un plan de **criblage** en (2) et **déterminer les facteurs les plus influents** en (4)

Dans le cas d'un **Taguchi**, l'étape (2) se décompose en :

1. Lister les facteurs et interactions possibles
2. Classer dans les groupes et construire le graphe linéaire
3. Sélectionner la table orthogonale la plus proche et ajuster l'expérimentation à celle-ci
4. Ordonner les expériences avec la table triangulaire correspondante

Si besoin d'**optimiser** : après détermination des **facteurs influents**

Dans le cas d'un **Doehlert**, itérations en boucle de (2) à (4) jusqu'à optimisation :

1. Placer le premier plan
2. Analyser les résultats et déterminer l'optimum
3. Glisser le réseau jusqu'à l'obtention de l'optimum souhaité

# Références

- Fisher RA. The design of experiments. 2nd ed. Edinburgh: Oliver and Boyd; 1937.
- Pillet M. Les plans d'expériences par la méthode Taguchi. 3rd ed. Editions d'Organisation; 1997.
- Vivier S. Stratégies d'optimisation par la méthode des plans d'expériences et application aux dispositifs électrotechniques modélisés par éléments finis. École Centrale de Lille et Université des Sciences et Technologies de Lille, 2002.
- Tinsson W. Plans d'expérience : constructions et analyses statistiques. 1st ed. Berlin, Heidelberg: Springer; 2010.
- Doehlert DH. Uniform shell designs. Appl Stat 1970;19:231–9. doi:10.2307/2346327.
- Goupy J, Creighton L. Introduction aux plans d'expériences. Paris: Dunod; 2006.