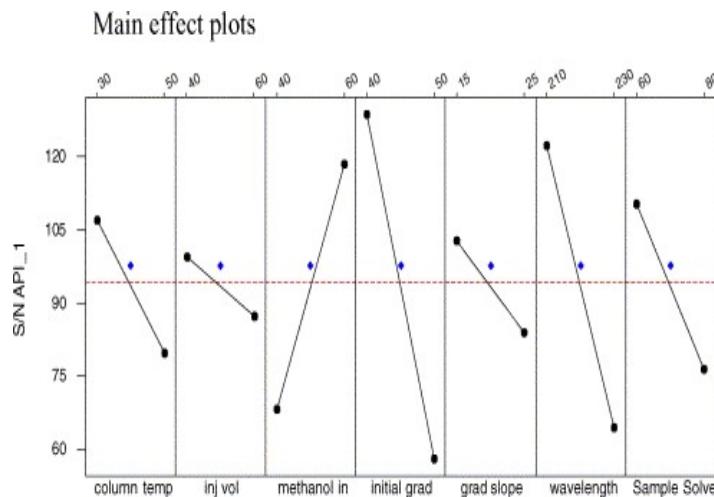
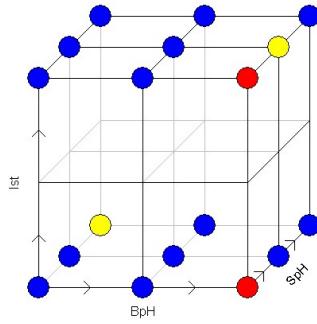


Pierre Chaminade

Lip(Sys)<sup>2</sup>/Chimie Analytique Pharmaceutique  
Faculté de Pharmacie Paris Saclay

# Plans d'expériences



# Plans d'expériences

- Permettent de planifier des conditions expérimentales (analyse, procédé) de façon à relier mathématiquement les paramètres de l'essai (facteurs) et la réponse (résultat attendu).
  - Double intérêt**
    - **organisation des essais**
    - **Modélisation du résultat**

⇒ *On ne fait pas toutes les expériences mais seulement certaines et on connaît le résultat de ce que l'on n'a pas fait !*
- Font appel au calcul matriciel
- Un même nom pour deux buts:
  - Explorer un procédé** → **analyse d'effets**
  - Mettre au point le procédé** → **optimisation**

# Applications

(repris de [www.coryent.com](http://www.coryent.com))

- Recherche et développement, fabrication, analyse, formulation de mélange.
- Production : réduction des coûts et des délais.
- Optimisation de caractéristiques, rendement, rapport qualité/prix.
- Conception de produits ou processus robustes.
- Identification des paramètres influents, et quantification de leur influence.
- Modèles prévisionnels.
- Qualité totale : Identification des causes de défauts, recherche du juste nécessaire, réduction du gaspillage (rebuts, retouches).
- Amélioration continue, réglage de machines, etc...

---

# Plan du cours

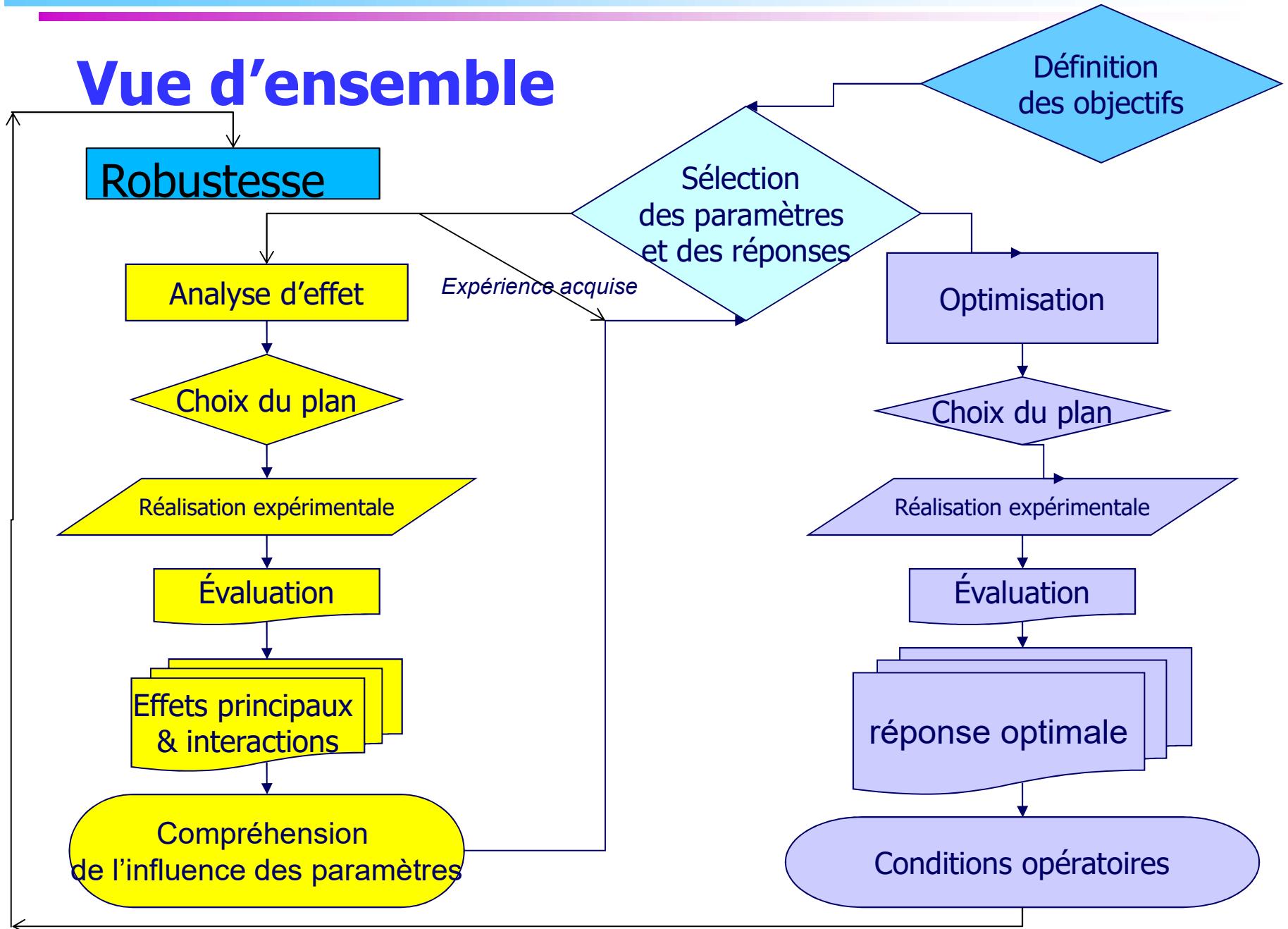
## 1 Analyse d'effets

- **Plans factoriels complets**
- **Interprétation**
- **Plans fractionnaires (Plackett-Burman)**

## 2 Optimisation

- **Exemples de plans**
- **interprétation**

# Vue d'ensemble



# Analyse d'effets

- Deux types de plans
  - Plan factoriel complet
    - $2^n$  essais pour  $n$  variables
    - Facile à calculer (Excel suffit).
    - Calcul des effets principaux et des interactions
  - Plans fractionnaires.
    - Nombre d'essais réduits
    - Plus difficile à interpréter.

# Plans factoriels complets

- Comportent toutes les combinaisons possibles entre les variables.
- Nombre d'essais  
 $= 2^{(\text{nbre de variables})}$
- Les variables peuvent être
  - Quantitatives (température, % d'un composant...)
  - Qualitatives (nature d'un composant....)

Essai	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>
1	+	+	+
2	-	+	+
3	+	-	+
4	-	-	+
5	+	+	-
6	-	+	-
7	+	-	-
8	-	-	-

# Calcul de l'effet des variables

- Nécessite une fonction de réponse.

*Exemple: Factors affecting the encapsulation of thioguanine in dehydratation rehydratation vesicles liposomes (Casals, Int. J. Pharm, 1996)*

Encapsulation de la thioguanine dans des liposomes contenant différentes quantités de cholestérol, de lipides anionique (-1) ou cationique (+1), sous différentes conditions de pH et de durée de sonication

Exp Name	pH	Cholesterol content	Charge of lipid	Time of Sonication	Encapsulation
N1	4.7	0	-1	30	9.84
N2	7.4	0	-1	30	8.63
N3	4.7	50	-1	30	3.12
N4	7.4	50	-1	30	6.13
N5	4.7	0	1	30	7.03
N6	7.4	0	1	30	6.11
N7	4.7	50	1	30	3.95
N8	7.4	50	1	30	9.86
N9	4.7	0	-1	120	7.25
N10	7.4	0	-1	120	6.35
N11	4.7	50	-1	120	10.88
N12	7.4	50	-1	120	5.74
N13	4.7	0	1	120	12.85
N14	7.4	0	1	120	12.4
N15	4.7	50	1	120	8.1
N16	7.4	50	1	120	14.07

# Calcul des effets principaux

- Effet du paramètre j pour n essais
- $y(+)$  réponse quand le paramètre est à l'état (+)
- $y(-)$ , quand il est à l'état (-)

$$E(j) = \frac{\sum_{i=1}^{n/2} y_i(+) - \sum_{i=1}^{n/2} y_i(-)}{n/2}$$

$$\begin{aligned} E(\text{pH}) &= \\ &[(8.63 + 6.13 + 6.11 + 9.86 + 6.35 \\ &+ 5.74 + 12.4 + 14.07) - (9.84 + \\ &3.12 + 7.03 + 3.95 + 7.25 + 10.88 \\ &+ 12.85 + 8.1)] / (16/2) \\ &= 6.27 / 8 \\ &= 0.78 \end{aligned}$$

Exp Name	pH	Cholesterol content	Charge of lipid	Time of Sonication	Encapsulation
N1	-1	-1	-1	-1	9.84
N2	1	-1	-1	-1	8.63
N3	-1	1	-1	-1	3.12
N4	1	1	-1	-1	6.13
N5	-1	-1	1	-1	7.03
N6	1	-1	1	-1	6.11
N7	-1	1	1	-1	3.95
N8	1	1	1	-1	9.86
N9	-1	-1	-1	1	7.25
N10	1	-1	-1	1	6.35
N11	-1	1	-1	1	10.88
N12	1	1	-1	1	5.74
N13	-1	-1	1	1	12.85
N14	1	-1	1	1	12.4
N15	-1	1	1	1	8.1
N16	1	1	1	1	14.07

# Calcul des effets principaux

- Effet du paramètre j pour n essais
- $y(+)$  réponse quand le paramètre est à l'état (+)
- $y(-)$ , quand il est à l'état (-)

$$E(j) = \frac{\sum_{i=1}^{n/2} y_i(+) - \sum_{i=1}^{n/2} y_i(-)}{n/2}$$

	pH	Cholesterol content	Charge of lipid	Time of Sonication
Effet	0.78	-1.08	2.05	2.87

Exp Name	pH	Cholesterol content	Charge of lipid	Time of Sonication	Encapsulation
N1	-1	-1	-1	-1	9.84
N2	1	-1	-1	-1	8.63
N3	-1	1	-1	-1	3.12
N4	1	1	-1	-1	6.13
N5	-1	-1	1	-1	7.03
N6	1	-1	1	-1	6.11
N7	-1	1	1	-1	3.95
N8	1	1	1	-1	9.86
N9	-1	-1	-1	1	7.25
N10	1	-1	-1	1	6.35
N11	-1	1	-1	1	10.88
N12	1	1	-1	1	5.74
N13	-1	-1	1	1	12.85
N14	1	-1	1	1	12.4
N15	-1	1	1	1	8.1
N16	1	1	1	1	14.07

# Calcul des interactions

- Pour compléter le tableau il suffit de calculer le produit des signes affectés aux paramètres

pH	Cholesterol content	Charge of lipid	Time of Sonication	pH x CHO	pH x Charge	pH x T sonic	CHO x Charge	CHO x T sonic	Charge x T sonic	pH x CHO x Charge	pH x CHO x T sonic	pHx Charge x Tsonic	CHO x Charge x Tsonic	pH x CHO x Charge x T sonic
A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
-	-	-	-	+	+	+	+	+	-	-	-	-	-	+
+	-	-	-	-	-	-	+	+	+	+	+	+	-	-
-	+	-	-	-	+	+	-	-	+	+	+	-	+	-
+	+	-	-	+	-	-	-	-	-	-	-	+	+	+
-	-	+	-	+	-	+	-	+	-	+	-	+	+	-
+	-	+	-	-	+	-	-	+	-	-	+	-	+	+
-	+	+	-	-	-	+	+	-	-	-	+	+	-	+
+	+	+	-	+	+	-	+	-	-	+	-	-	-	-
-	-	-	+	+	+	-	+	-	-	-	+	+	+	-
+	-	-	+	-	-	+	+	-	-	+	+	-	-	+
-	+	-	+	-	-	+	-	+	-	-	-	+	-	+
+	+	-	+	+	-	+	-	+	-	-	+	-	-	-
-	-	+	+	+	-	-	-	-	-	-	-	-	-	+
+	-	+	+	-	-	-	-	-	-	-	-	-	-	-
-	+	+	+	-	-	-	-	-	-	-	-	-	+	-
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+

# Calcul des effets pour les interactions.

- Même formule de calcul que la précédente

$$E(j) = \frac{\sum_{i=1}^{n/2} y_i(+) - \sum_{i=1}^{n/2} y_i(-)}{n/2}$$

pH	Cholesterol content	Charge of lipid	Time of Sonication		
0.78	-1.08	2.05	2.87		
pH x CHO	pH x Charge	pH x T sonic	CHO x Charge	CHO x T sonic	Charge x T sonic
1.65	1.84	-0.91	0.47	1.06	2.25
pH x CHO x Charge	pH x CHO x T sonic	pHx Charge x Tsonic	CHO x Charge x Tsonic		
1.66	-1.11	1.05	-2.00		
pH x CHO x Charge x T sonic					
1.01					

Donc, facile, mais comment savoir lesquelles de ces valeurs sont significatives ?

## Incertitude sur les effets (1).

- Estimation par réPLICATION
  - À partir d'essais dupliqués la variance poolée  $s^2$  est égale à la somme des carrés des différences de réponse /  $2n$
  - La variance des effets est donné par :  
$$s^2_{\text{effets}} = (2 s^2) / n$$

# Incertitude des effets par réPLICATION

pH	Cholesterol content	Charge of lipid	Time of Sonication			Encapsulation (1)	Encapsulation (2)	d	$d^2$
0.78	-1.08	2.05	2.87			9.84	11.46	1.62	2.6244
						8.63	9.16	0.53	0.2809
						3.12	3.02	0.1	0.01
						6.13	7.15	1.02	1.0404
						7.03	6.08	0.95	0.9025
						6.11	6.82	0.71	0.5041
						3.95	3.58	0.37	0.1369
						9.86	10.57	0.71	0.5041
						7.25	6.9	0.35	0.1225
						6.35	5.56	0.79	0.6241
						10.88	12.19	1.31	1.7161
						5.74	5.98	0.24	0.0576
						12.85	11.65	1.2	1.44
						12.4	13.3	0.9	0.81
						8.1	8.34	0.24	0.0576
						14.07	15.38	1.31	1.7161
1.01									
						$s^2$	0.39		
						$s^2$ effets	0.05		
						s effets	0.22		

$$\sigma_T^2 = \frac{\sum_{i=1}^n d_i^2}{2n}$$

$$\sigma_E = \sqrt{\left(\frac{1}{n} + \frac{1}{n}\right) \sigma_T^2}$$

## Incertitude sur les effets (2).

- Incertitude sur les effets en sacrifiant les interactions d'ordre 3 et plus

pH	Cholesterol content	Charge of lipid	Time of Sonication		
0.78	-1.08	2.05	2.87		
pH x CHO Charge	pH x T sonic	CHO x Charge	CHO x T sonic	Charge x T sonic	
1.65	1.84	-0.91	0.47	1.06	2.25
pH x CHO x Charge	pH x CHO x T sonic	pHx Charge x Tsonic	CHO x Charge x Tsonic		
1.66	-1.11	1.05	-2.00		
pH x CHO x Charge x T sonic					
1.01					

$$\sigma_{effets} = \sqrt{\frac{\sum_{ABC} E^2}{5}}$$

Ici  $\sigma_{effet} = 1.42$

# Calcul des effets par régression

- Revient à calculer un modèle linéaire avec des termes exprimant les interactions:  
$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + e$$
- Les coefficients ( $b$ ) représentent la mesure de l'effet des variables ( $x$ ) et/ou de leurs interactions ( $b_{12}$ )

# Notation matricielle

- Revient à résoudre
- Si le modèle est

$$y = X \cdot b$$

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + e$$

- Alors  $y$ ,  $X$  et  $b$  sont de la forme

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_{1(1)} & x_{2(1)} & x_{1(1)}x_{2(1)} \\ 1 & x_{1(2)} & x_{2(2)} & x_{1(2)}x_{2(2)} \\ 1 & x_{1(3)} & x_{2(3)} & x_{1(3)}x_{2(3)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1(n)} & x_{2(n)} & x_{1(n)}x_{2(n)} \end{bmatrix} \quad b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{12} \end{bmatrix}$$

# Estimation des paramètres de régression

- Revient à calculer  $b$  tel que  $y=X.b$
- D'où  $X^T X b = X^T y \Rightarrow b = (X^T X)^{-1} X^T y$
- La variance des coefficients du modèle se trouvent sur la matrice de variance-covariance  $V(b) = \sigma_{pe}^2 (X^T X)^{-1} = (1/N) \sigma_{pe}^2$

La significativité des coefficients est proportionnelle à l'erreur expérimentale

```
S2pe =
290.6044
>> V=S2pe*inv(X'*X)
=
18.1628      0      0      0      0      0      0
      0  18.1628      0      0      0      0      0
      0      0  18.1628      0      0      0      0
      0      0      0  18.1628      0      0      0
      0      0      0      0  18.1628      0      0
      0      0      0      0      0  18.1628      0
      0      0      0      0      0      0  18.1628
```

$X^T X$  est diagonale car les variables ne sont pas corrélées

# Estimation de la qualité de régression

- Variance totale:
  - $n$  nbre d'observations
- Variance de régression
  - $p$  nbre de coefficients
- Variance résiduelle
  - $n-k$  =nbre de réplications
  - Erreur d'ajustement
    - Erreur expérimentale

$$\sigma_T^2 = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} (y_{i,j} - \bar{y})^2}{n-1} = \frac{y^T y - n\bar{y}^2}{n-1}$$

$$\sigma_R^2 = \frac{\sum_i n_i (\hat{y} - \bar{y})^2}{p-1} = \frac{b^T X^T y - n\bar{y}^2}{p-1}$$

$$\sigma_{res}^2 = \frac{\sum_i \sum_j (y_{i,j} - \hat{y}_i)^2}{n-p} = \frac{e^T e}{n-p}$$

$$\sigma_{LOF}^2 = \frac{\sum_i n_i (\bar{y}_i - \hat{y}_i)^2}{k-p} = \frac{y^T y - b^T X^T y - SS_{PE}}{k-p}$$

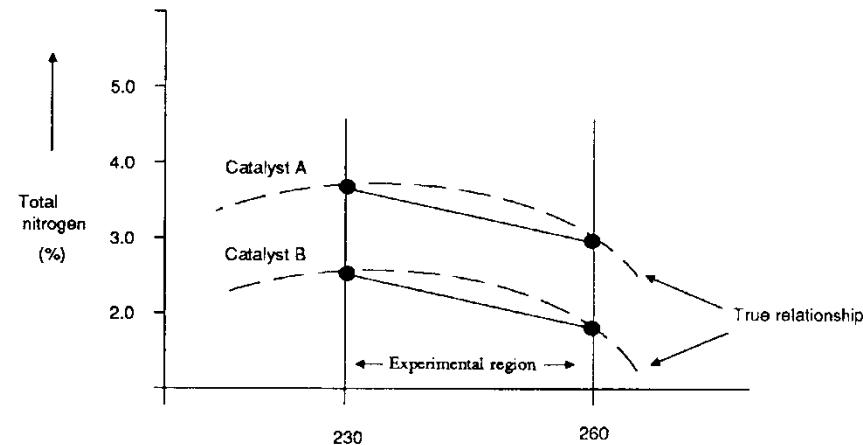
$$SS_{PE} = \sum_i \sum_j (y_{i,j} - \bar{y}_i)^2 \quad \sigma_{PE}^2 = \frac{SS_{PE}}{(n-k)}$$

# Analyse de la validité de la régression

- Si  $\sigma^2_T \approx \sigma^2_R$  alors  $\sigma^2_{\text{res}}$  est faible
  - Plus  $F = \sigma^2_R / \sigma^2_{\text{res}}$  est grand et plus on peut penser que le modèle décrit correctement les données.
- Si les points sont répliqués, on peut accéder à  $\sigma^2_{\text{pe}}$  et calculer  $\sigma^2_{\text{lof}}$ 
  - Si  $\sigma^2_{\text{lof}} < \sigma^2_{\text{pe}}$  l'ajustement est valide.
- Mais, si on utilise que deux niveaux + et -, il n'est pas possible de juger de l'ajustement

# Test de courbure et point au centre.

- L'analyse d'effets suppose que la réponse est linéaire.
  - On peut le tester en ajoutant un point central
  - Ce point peut être répliqué pour estimer l'erreur expérimentale et la significativité des variables.



essai	A	B
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5.1	0	0
5.2	0	0
5.3	0	0
5.4	0	0
5.5	0	0

# Randomisation

- Il est fortement déconseillé de réaliser le plan dans l'ordre des essais
- Randomisation pour éviter qu'un facteur externe ne se superpose à l'effet d'une variable.
- Exemple: tous les P3 (+) le matin et les P3 (-) l'après midi si les essais durent  $\frac{1}{2}$  h / 1 h.

N°	P1	P2	P3
1	+	+	+
2	-	+	+
3	+	-	+
4	-	-	+
5	+	+	-
6	-	+	-
7	+	-	-
8	-	-	-

# Retour sur Casals 1996

Exp No	Exp Name	Run Order	Incl/Excl	Cholesterol content		Charge of lipid	Time of Sonication	Encapsulation
				pH				
1	N1	1	Incl	4.7	0	-1	30	9.84
2	N2	10	Incl	7.4	0	-1	30	8.63
3	N3	12	Incl	4.7	50	-1	30	3.12
4	N4	8	Incl	7.4	50	-1	30	6.13
5	N5	9	Incl	4.7	0	1	30	7.03
6	N6	7	Incl	7.4	0	1	30	6.11
7	N7	15	Incl	4.7	50	1	30	3.95
8	N8	13	Incl	7.4	50	1	30	9.86
9	N9	4	Incl	4.7	0	-1	120	7.25
10	N10	19	Incl	7.4	0	-1	120	6.35
11	N11	18	Incl	4.7	50	-1	120	10.88
12	N12	17	Incl	7.4	50	-1	120	5.74
13	N13	6	Incl	4.7	0	1	120	12.85
14	N14	5	Incl	7.4	0	1	120	12.4
15	N15	14	Incl	4.7	50	1	120	8.1
16	N16	3	Incl	7.4	50	1	120	14.07
17	N17	2	Incl	6.1	33	0	60	7.01
18	N18	11	Incl	6.1	33	0	60	7.24
19	N19	16	Incl	6.1	33	0	60	6.92

La matrice d'expérience présentée dans l'article présente une estimation de l'erreur expérimentale au centre en plus d'une randomisation.

# Anova Casals 1996

Encapsulati on	DF	SS	MS (variance)	F	p	SD
Total	19	1394.3	73.3841			
Constant	1	1239.8	1239.8			
Total						
Corrected	18	154.502	8.58346			2.92975
Regression	10	113.417	11.3417	2.20842	<b>0.137</b>	3.36774
Residual	8	41.0853	5.13566			2.2662
Lack of Fit (Model Error)	6	41.0308	6.83847	251.107	<b>0.004</b>	2.61505
Pure Error (Replicate Error)	2	0.0544665	0.0272332			0.165025
N = 19	Q2 =	-1.539	Cond. no. =	1.109		
DF = 8	R2 =	0.734	Y-miss =	0		
	R2 Adj. =	0.402	RSD =	2.266		

- Le modèle montre une faible différence entre variance de régression et résiduelle associée à un manque d'ajustement.

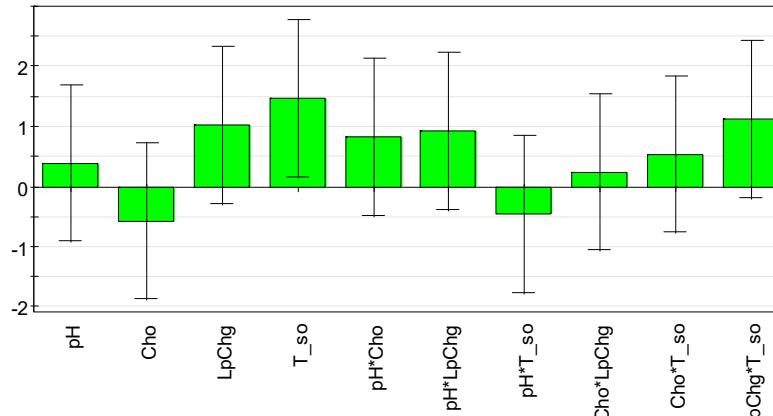
# Coefficient du modèle et effets

Seul le coefficient du temps de sonication est significatif ( $\neq 0$ )

Le résultat est le même pour les effets.

Investigation: Casals 1996 (MLR)

Scaled & Centered Coefficients for Encapsulation



N=19  
DF=8

R2=0.734  
Q2=-1.539

R2 Adj.=0.402  
RSD=2.266

Conf. lev.=0.95

Encapsulation	Effect	Conf. int(±)
T_so	2.93	2.59
LpChg*T_so	2.25	2.61
LpChg	2.05	2.61
pH*LpChg	1.84	2.61
pH*Cho	1.65	2.61
Cho	-1.13	2.59
Cho*T_so	1.08	2.61
pH*T_so	-0.91	2.61
pH	0.78	2.61
Cho*LpChg	0.47	2.61

N = 19	Q2 =	-1.539	Cond. no. =	1.109
DF = 8	R2 =	0.734	Y-miss =	0
	R2 Adj. =	0.402	RSD =	2.266
			Conf. lev. =	0.95

## Exemple 2

- E. Gonzales-Mira, Colloids & Surfaces B, 2010.

*Design and ocular tolerance of flurbiprofen loaded ultrasound-engineered nanostructured lipid carriers*

*Plan complet à 4 variables pour la fabrication des nanoparticules:*

*Flurbiprofen (FB)*

*Tween 80*

*%SA/L pourcentage d'acide stéarique p/r aux lipides totaux (huile de ricin)*

*Température de stockage des nanoparticules*

<b>FB</b>	FB	wt%	Quantitative	Controlled	0.05 to 0.15
<b>Tween 80</b>	T80	wt%	Quantitative	Controlled	1.6 to 2.6
<b>%SA/L</b>	%SA/L	%	Quantitative	Controlled	50 to 60
<b>storage temp</b>	sto_T	°C	Quantitative	Controlled	4 to 25

## Exemple 2

Plan complet randomisé à 2 niveaux par variable avec 3 mesures de la réponse (taille de nanoparticules) (3x16)

Exp No	Run Order	FB	Tween 80	%SA/L	storage temp	Mean Size
1	43	0.05	1.6	50	4	415.963
17	4	0.05	1.6	50	4	384.553
33	36	0.05	1.6	50	4	400.775
9	1	0.05	1.6	50	25	303.567
25	47	0.05	1.6	50	25	302.408
41	10	0.05	1.6	50	25	272.536
5	23	0.05	1.6	60	4	301.823
21	5	0.05	1.6	60	4	350.921
37	12	0.05	1.6	60	4	296.685
13	33	0.05	1.6	60	25	360.76
29	27	0.05	1.6	60	25	353.937
45	42	0.05	1.6	60	25	335.102
3	16	0.05	2.6	50	4	310.915
19	39	0.05	2.6	50	4	321.491
35	41	0.05	2.6	50	4	308.794
11	35	0.05	2.6	50	25	336.534
27	2	0.05	2.6	50	25	337.851
43	38	0.05	2.6	50	25	344.886
7	44	0.05	2.6	60	4	420.974
23	14	0.05	2.6	60	4	415.504
39	20	0.05	2.6	60	4	424.448
15	13	0.05	2.6	60	25	494.826
31	29	0.05	2.6	60	25	477.684
47	28	0.05	2.6	60	25	442.579
2	17	0.15	1.6	50	4	373.464
18	40	0.15	1.6	50	4	365.823
34	15	0.15	1.6	50	4	371.73
10	24	0.15	1.6	50	25	365.413
26	32	0.15	1.6	50	25	365.873
42	25	0.15	1.6	50	25	368.9
6	22	0.15	1.6	60	4	375.737
22	9	0.15	1.6	60	4	377.198
38	6	0.15	1.6	60	4	378.885
14	8	0.15	1.6	60	25	379.014
30	3	0.15	1.6	60	25	382.419
46	7	0.15	1.6	60	25	382.927
4	21	0.15	2.6	50	4	316.896
20	18	0.15	2.6	50	4	314.563
36	48	0.15	2.6	50	4	318.873
12	26	0.15	2.6	50	25	305.587
28	30	0.15	2.6	50	25	302.347
44	11	0.15	2.6	50	25	305.434
8	46	0.15	2.6	60	4	369.174
24	34	0.15	2.6	60	4	370.72
40	37	0.15	2.6	60	4	371.743
16	45	0.15	2.6	60	25	399.319
32	19	0.15	2.6	60	25	400.424
48	31	0.15	2.6	60	25	391.718

# ANOVA

Mean Size	DF	SS	MS (variance)	F	p	SD
Total	48	6.39E+06	133184			
Constant	1	6.29E+06	6.29E+06			
Total						47.7766
Corrected	47	107283	2283			
Regression	10	86979	8698	15.8506	<b>0.000</b>	93.2625
Residual	37	20303	549			23.4253
Lack of Fit	5	15335	3067	19.7552	<b>0.000</b>	55.3812
(Model Error)						12.4601
Pure Error	32	4968	155			
(Replicate Error)						
N = 48	Q2 =	0.681	Cond. no. =	1		
DF = 37	R2 =	0.811	Y-miss =	0		
	R2 Adj. =	0.76	RSD =	23.43		

Attention, quand ce test est significatif +++++

Celui-là n'a pas beaucoup de sens.

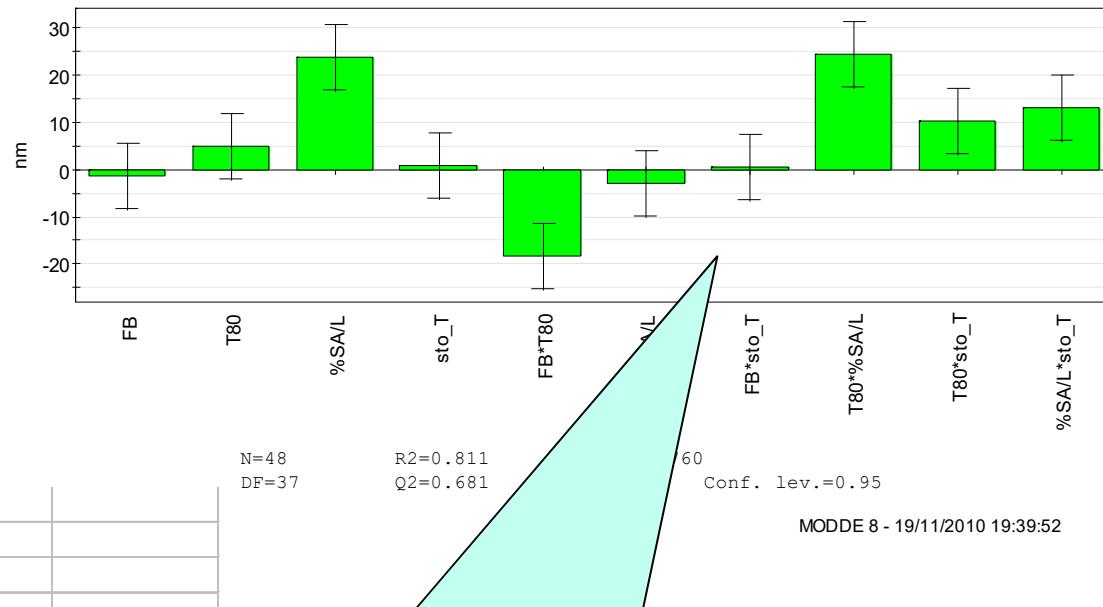
# Coefficients et effets

Certaines variables n'ont pas d'effet propre mais sont importantes à travers les interactions.

Mean Size	Effect	Conf. int(±)
T80*%SAL	48.77	13.70
%SAL	47.47	13.70
FB*T80	-36.59	13.70
%SAL*sto_T		
	26.64	13.70
T80*sto_T	20.66	13.70
T80	9.87	13.70
FB*%SAL	-5.44	13.70
FB	-2.56	13.70
sto_T	2.27	13.70
FB*sto_T	1.45	13.70
N = 48	Q2 =	0.681
DF = 37	R2 =	0.811
	R2 Adj. =	0.76
	Y-miss =	0
	RSD =	23.43
	Conf. lev. =	0.95

Investigation: Gonzalez-Mira 2010 (2) (MLR)

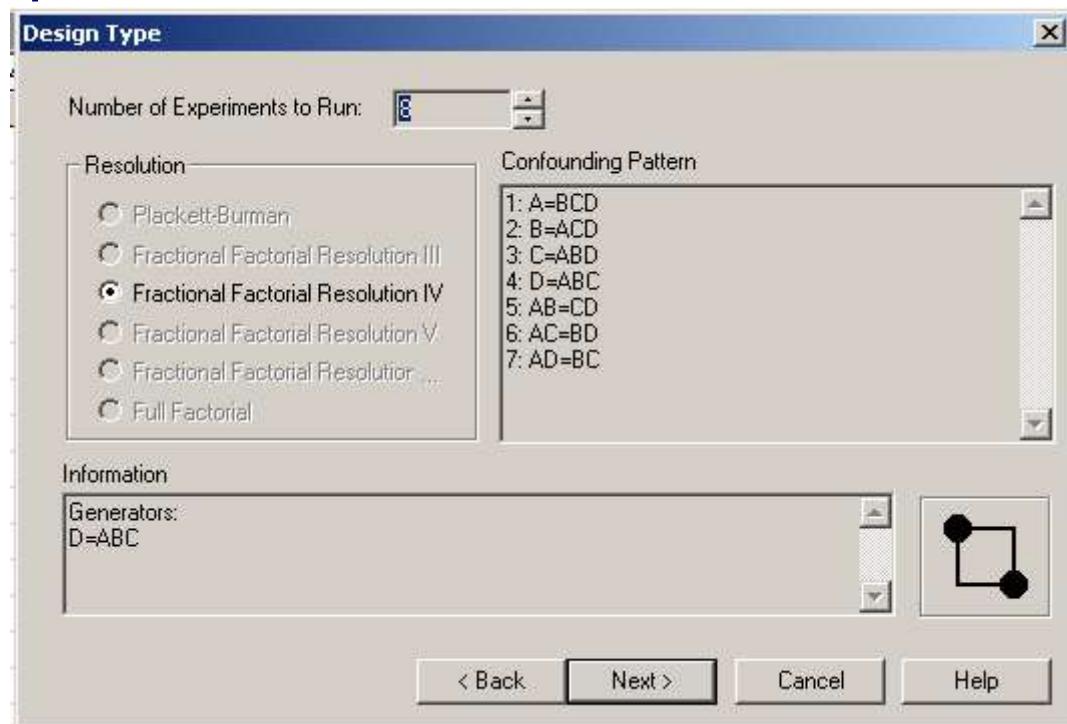
Scaled & Centered Coefficients for Mean Size



On peut encore optimiser le modèle en supprimant les termes NS et en cherchant à transformer les Y (réponse) pour les linéariser

# Plans fractionnaires.

- Permettent de réduire le nombre d'essais
  - Mais les effets principaux sont confondus avec les interactions.
  - Il faut refaire un plan sur les variables significatives



# Exemple de travail, process de sphéronisation

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	y
	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	-1	-1
	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	4.2
	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	-1	4.8
	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1	6.1
	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1	3.7
	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1	6.4
	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-1	4.7
	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	3.7
	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	4.4
	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	6.5
	-1	1	1	1	-1	-1	-1	1	1	-1	-1	-1	-1	1	-1	6.6
	1	1	-1	1	1	-1	1	-1	1	-1	-1	-1	1	-1	-1	3.4
	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	3.9
	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1	6.7
	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1	5.4
	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	1	-1	1	8.3
	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	6.3
Effets (E)	-2.26	0.49	0.79	-0.89	-0.41	-0.36	-0.14	0.09	-0.24	-0.19	0.14	-0.09	-0.19	-0.39	0.41	
	$\sigma_{effets} = \sqrt{\frac{\sum_{ABC} E^2}{5}}$															0.28

- Process: sphéronisation A= quantité d'eau (250-325 ml), B = extrusion (39-59 rpm), C=tamis (0,8-1,5 mm), D = sphéroniseur (700-1010 rpm)
- Réponse (y) dureté des comprimés.

# Relation plan fractionnaire / plan complet

Essai	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	y
2	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	4.8
3	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	6.1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3.7
8	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	1	4.4
12	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	3.9
13	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	-1	1	6.7
15	-1	1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	8.3
16	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	6.3
1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	4.2
4	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	1	-1	-1	3.7
5	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	-1	1	-1	6.4
6	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	1	-1	4.7
9	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	6.5
10	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	-1	1	-1	6.6
11	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	-1	3.4
14	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	-1	5.4

- Exemple « sphéronisation »
- Plan trié en fonction de la colonne ABCD
- Si on ne garde que la partie supérieure telle que  $ABCD=1$  on crée un plan demi - factoriel

# Fusion effets principaux et interactions.

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD	y
2	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	4.8
3	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	1	6.1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3.7
8	1	1	-1	-1	1	-1	-1	-1	1	-1	-1	1	1	1	1	4.4
12	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	3.9
13	-1	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	1	1	6.7
15	-1	1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	8.3
16	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	6.3

- En fait  $A^2=B^2=C^2=D^2=ABCD=1$

–  $A=BCD$ ,  $B=ACD$ ,  $C=ABD$ ,  $D=ABC$

–  $AB=CD$ ,  $AC=BD$ ,  $AD=BC$

$$E(j) = \frac{\sum_{i=1}^{n/2} y_i(+) - \sum_{i=1}^{n/2} y_i(-)}{n/2} \text{ avec } n=8$$

	A BCD	B+ ACD	C+ ABD	D+ ABC	AB+ CD	AC+ BD	AD+ BC	y
2	1	-1	1	-1	-1	1	-1	4.8
3	-1	-1	-1	-1	-1	1	1	6.1
7	1	1	1	1	1	1	1	3.7
8	1	1	-1	-1	1	-1	-1	4.4
12	1	-1	-1	1	-1	-1	1	3.9
13	-1	-1	1	1	1	-1	-1	6.7
15	-1	1	1	-1	-1	-1	1	8.3
16	-1	1	-1	1	-1	1	-1	6.3

Effets (E)	-2.65	0.30	0.70	-0.75	-0.60	-0.60	-0.05	

## Comparaison $2^n$ et $2^{(n-1)}$

	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	ACD	BCD	ABCD
Effets (E)	-2.26	0.49	0.79	-0.89	-0.41	-0.36	-0.14	0.09	-0.24	-0.19	0.14	-0.09	-0.19	-0.39	0.41
	A+	B+	C+	D+	AB+	AC+	AD+								
	BCD	ACD	ABD	ABC	CD	BD	BC								
Effets (E)	-2.65	0.30	0.70	-0.75	-0.60	-0.60	-0.05								

- On perd nécessairement de l'information

# Création à partir du plan n-1

- Choisir le générateur et étendre le plan
- Ex D=ABC donc  $D^2=ABCD=1$ ...calculer les effets confondus (alias).

A	B	C	AB	AC	BC	ABC
-1	-1	-1	1	1	1	-1
-1	-1	1	1	-1	-1	1
-1	1	-1	-1	1	-1	1
-1	1	1	-1	-1	1	-1
1	-1	-1	-1	-1	1	1
1	-1	1	-1	1	-1	-1
1	1	-1	1	-1	-1	-1
1	1	1	1	1	1	1

A	B	C	D
-1	-1	-1	-1
-1	-1	1	1
-1	1	-1	1
-1	1	1	-1
1	-1	-1	1
1	-1	1	-1
1	1	-1	-1
1	1	1	1

# Création d'un plan quart –factoriel ( $2^{n-2}$ )

- Partir du plan complet n-2
  - Choisir un premier générateur  $D=ABC$ 
    - $A=BCD, B=ACD, C=ABD, D=ABC$
    - $AB=CD, BD=AC$
  - Choisir un second générateur  $E=BC$ 
    - $A=BCD=ED=ABCE$
    - $B=ACD=CE=ABDE$
    - $C=ABD=BE=ACDE$
    - $D=ABC=BCDE=AE$
    - $E=ABCDE=BC=AD$
    - ...

A	B	C	D	E
-1	-1	-1	-1	1
-1	-1	1	1	-1
-1	1	-1	1	-1
-1	1	1	-1	1
1	-1	-1	1	1
1	-1	1	-1	-1
1	1	-1	-1	-1
1	1	1	1	1

## Plackett-Burman

- Plans de screening basés sur les matrices d'Hadamard (matrice carrée dont les coefficients sont tous 1 ou -1 et dont les lignes sont toutes orthogonales entre elles).
- Permettent d'évaluer les effets principaux de N variables en N+1 essais.
- Mêmes modalités de calcul que les plans précédents

# Établissement d'un plan de P. B.

- N=8        + + + - - -
- N=12       + + - + + + - - -
- N=16       + + + + - + + + - + - -
- N=20       + + - + + + + - + - - + + -
- N=24       + + + + + - + + + - + - + - - -
  
- La première ligne est décalée pour former les lignes suivantes
- La dernière ligne ne contient que des -
- Calcul des interactions signe inversé.
- On peut laisser des variables « vides » pour apprécier la significativité des effets

# Exemple Plackett-Burman

*Fabrication of starch-based microparticles by an emulsification-crosslinking method Bing-Zheng Li, Journal of Food Engineering, 2009*

Experimental field definition for the Plackett-Burman design

Symbol code	Factors	Experimental values	
		Low level (-1)	High level (+1)
$X_1$	OSA starch concentration (g/g)	0.22	0.25
$X_2$	TSTP concentration (g/g)	0.015	0.020
$X_3$	PEG molecular weight (g/mol)	10,000	20,000
$X_4$	PEG concentrations (g/g)	0.25	0.30
$X_5$	Volume ratio of $w_1/w_2$ (mL/mL)	10:100	15:95
$X_6$	Agitation rate (rpm)	250	400
$X_7$	Crosslinking time (h)	8	12
$X_8, X_9, X_{10}, X_{11}$	Dummy factors	-	-

**Table 2**  
Plackett-Burman design matrix with response value

Run	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$	$X_{10}$	$X_{11}$	$Y(D[4,3]^a)$ ( $\mu\text{m}$ )
1	1	1	1	-1	-1	-1	1	-1	1	1	-1	39.1
2	-1	-1	1	-1	1	1	-1	1	1	1	-1	39.7
3	1	1	-1	-1	-1	1	-1	1	1	-1	1	51.3
4	1	1	-1	1	1	1	-1	-1	-1	1	-1	49.1
5	1	-1	1	1	-1	1	1	1	-1	-1	-1	19.7
6	1	-1	1	1	1	-1	-1	-1	1	-1	1	55.8
7	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	48.8
8	-1	1	1	-1	1	1	1	-1	-1	-1	1	31.8
9	-1	1	1	1	-1	-1	-1	1	-1	1	1	19.7
10	1	-1	-1	-1	1	-1	1	1	-1	1	1	66.8
11	-1	1	-1	1	1	-1	1	1	1	-1	-1	35.6
12	-1	-1	-1	1	-1	1	1	-1	1	1	1	30.9

<sup>a</sup>  $D[4,3]$ , the volume weighted mean diameter of OSA-MPs.

# Exemple Plackett-Burman (2)

**Table 3**

Analysis of variance for Plackett-Burman factorial model

Source	Sum of squares	Degrees of freedom	Mean square	F-value	p-value Prof > F
Model	2129.29	7	304.18	10.55	0.0190
$X_1$	472.51	1	472.51	16.40	0.0155
$X_2$	102.67	1	102.67	3.56	0.1321
$X_3$	490.24	1	490.24	17.01	0.0146
$X_4$	370.74	1	370.74	12.86	0.0230
$X_5$	400.21	1	400.21	13.89	0.0204
$X_6$	156.24	1	156.24	5.42	0.0804
$X_7$	136.69	1	136.69	4.74	0.0950
Residual	115.28	4	28.82		
Corrected total	2244.57	11			

# Plans pour l'optimisation

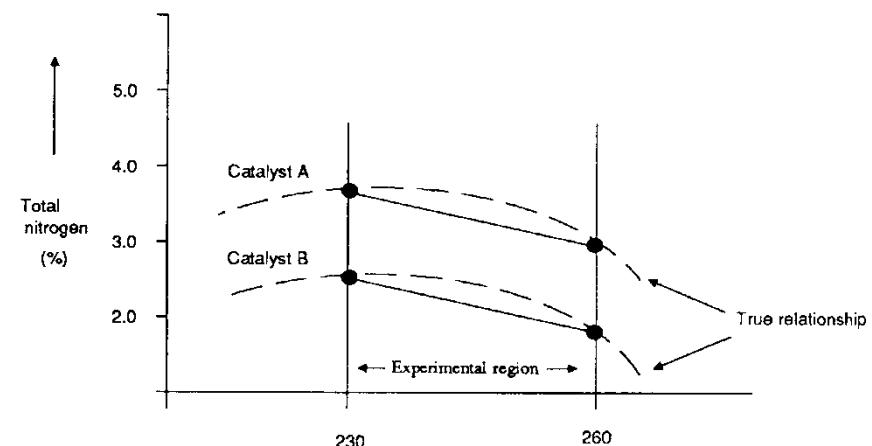
Surfaces de réponse.

## Plans d'expériences pour l'optimisation.

- Modèle mathématique capable de décrire un optimum
  - Modèle de degré 2 avec interactions

$$y = a_0 + a_{11} x_1^2 + a_1 x_1 + a_{22} x_2^2 + a_2 x_2 + a_{12} x_1 x_2 + e$$

variables	2	3	4	5	...	8
Coefficients	6	10	15	21	...	45



# Calcul par régression

$$y = X.b$$

$$b = (X^T X)^{-1} X^T y$$

$$V(b) = \sigma_{pe}^2 (X^T X)^{-1}$$

- Ne pas oublier que le plan suivi ( $X$ ) influence l'incertitude sur les coefficients du modèle  $V(b)$ .

---

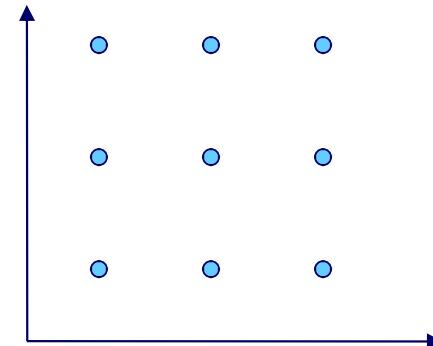
## Plans classiques

- Plan factoriel à trois niveaux
- Plan central composite
- Plan de Box-Behnken
- Plan de Doehlert

## Plan factoriel à 3 niveaux.

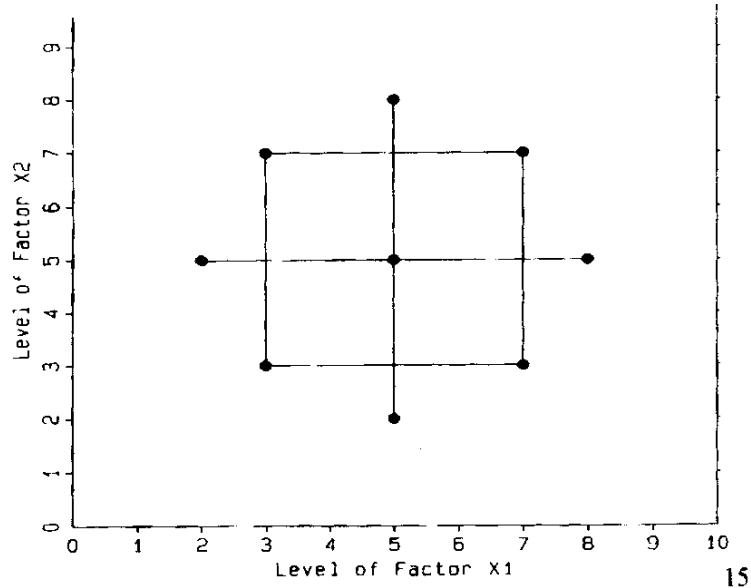
- Plan en  $3^k$
- Progression rapide du nombre d'essais en fonction du nombre de variables
- Couverture assez médiocre de l'espace paramétrique

	A	B
1	-1	-1
2	-1	0
3	-1	1
4	0	-1
5	0	0
6	0	1
7	1	-1
8	1	0
9	1	1



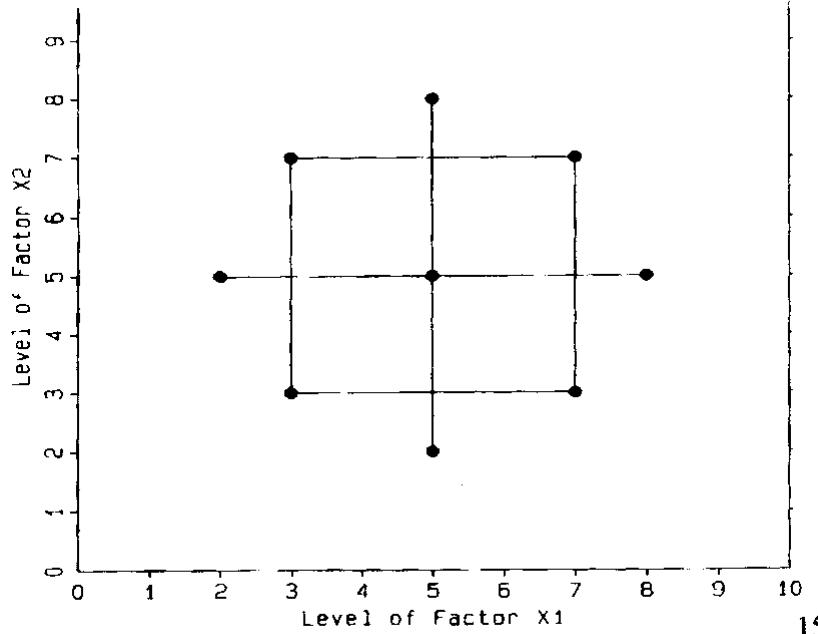
# Central composite

- Permet de couvrir l'espace paramétrique avec 5 niveaux pour chaque paramètre.



- Composé d'un plan factoriel à 2 niveaux (+1 et -1)
  - auquel on ajoute
    - Un point central (répliqué) au niveau 0
    - Une « étoile » à  $+ \alpha$  et  $- \alpha$

# Plan central composite

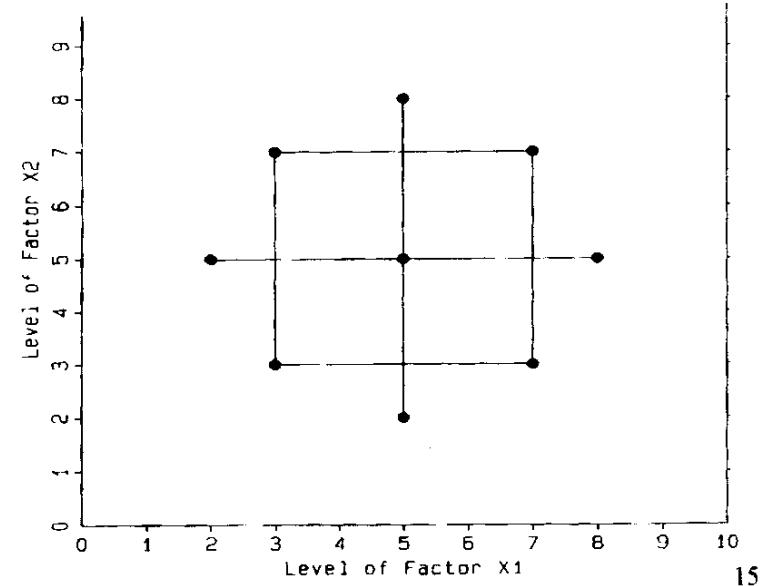


	A	B
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5	-1.4	0
6	1.4	0
7	0	-1.4
8	0	1.4
9	0	0

variables	$\alpha$
2	1.40
3	1.68
4	2.00
5	2.38
6	2.83

# Central composite

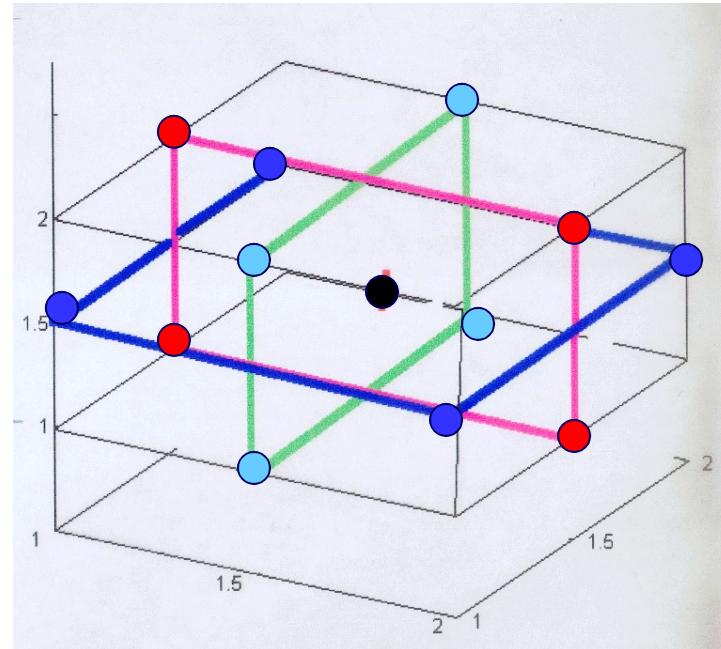
- Son efficacité est bonne jusqu'à 4 variables.



Facteurs	Paramètres	Essais	Efficacité p/e
1	3	5	0.60
2	6	9	0.67
3	10	15	0.67
4	15	25	0.60
5	21	43	0.49
6	28	77	0.36
7	36	143	0.25
8	45	273	0.16
9	55	531	0.10
10	66	1045	0.06

# Box Behnken

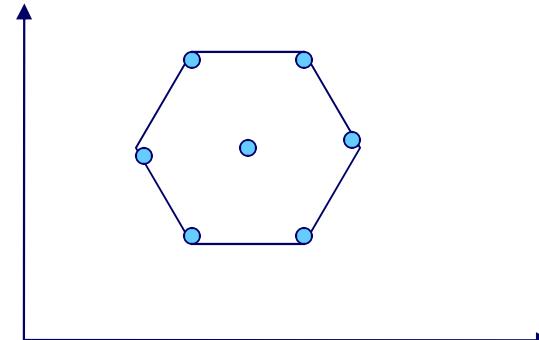
- Moins bonne couverture de l'espace paramétrique que le plan central composite.
- Intéressant pour 5 ou 6 variables.



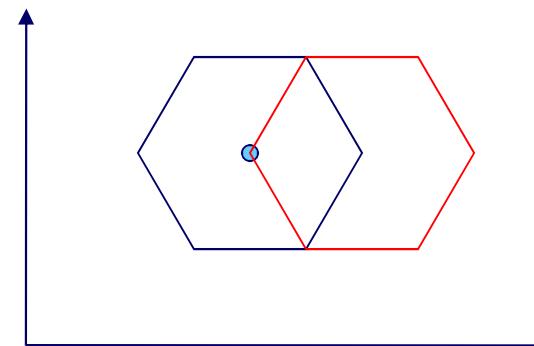
Facteurs	Paramètres	Essais	
		Central Composite	Box-Behnken
1	3	5	
2	6	9	
3	10	15	13
4	15	25	25
5	21	43	41
6	28	77	49
7	36	143	
8	45	273	
9	55	531	
10	66	1045	

# Plans de Doehlert

variables	coefficients	central composite		Doehlert	
		exp	efficacité	exp	efficacité
2	6	9	0.67	7	0.86
3	10	15	0.67	13	0.77
4	15	25	0.6	21	0.71
5	21	43	0.49	31	0.68
8	45	273	0.16	73	0.62



Possède une meilleure efficacité que le central composite.  
Permet de déplacer le plan quand les bornes apparaissent mal choisies



# Exemple optimisation

«Reversible protein precipitation to ensure stability during encapsulation within PLGA microspheres » de A. Giteau et al. parue dans European Journal of Pharmaceutics and Biopharmaceutics (2008).

- L' objectif: trouver les conditions de précipitation de protéines pour assurer leur stabilité et permettre leur encapsulation dans des microsphères de PLGA (poly(D,L-lactide-co-glycolide)).
- Des particules de protéines sont formées par addition de chlorure de sodium et d'un solvant miscible à l'eau (glycofurool ou Tetrahydrofurfuryl-polyethylenglycol) à des solutions de protéines.
- La stabilité des protéines ainsi précipitées est vérifiée par mesure de leur activité enzymatique ceci après réhydratation des particules de protéines précédemment formées.
-

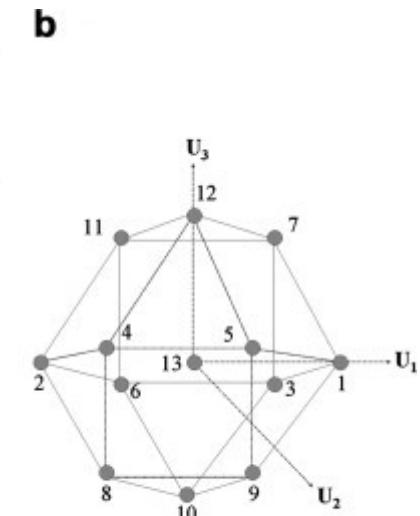
# Exemple optimisation

- Trois facteurs influençant la précipitation des protéines sont évalués :
- – La force ionique exprimée par la concentration en chlorure de sodium (en mole/l)
- – Le rapport entre le volume de phase aqueuse et le glycofurool exprimé en volume de phase aqueuse (en µl)
- – La quantité de protéines en mg

Plan à 5 niveaux pour le premier facteur, 7 pour le second et 3 pour le troisième.

Le plan comprend 13 essais dont le N° 13, situé au centre du plan est répété 3 fois.

N°Exp	Ionic strength U <sub>1</sub> (M)	Aqueous phase volume U <sub>2</sub> (µl)	Protein quantity U <sub>3</sub> (mg)
1	0.59	90.0	0.50
2	0.01	90.0	0.50
3	0.44	155.0	0.50
4	0.16	25.0	0.50
5	0.44	25.0	0.50
6	0.16	155.0	0.50
7	0.44	111.7	0.90
8	0.16	68.3	0.10
9	0.44	68.3	0.10
10	0.30	133.4	0.10
11	0.16	111.7	0.90
12	0.30	46.6	0.90
13	0.30	90.0	0.50
14	0.30	90.0	0.50
15	0.30	90.0	0.50

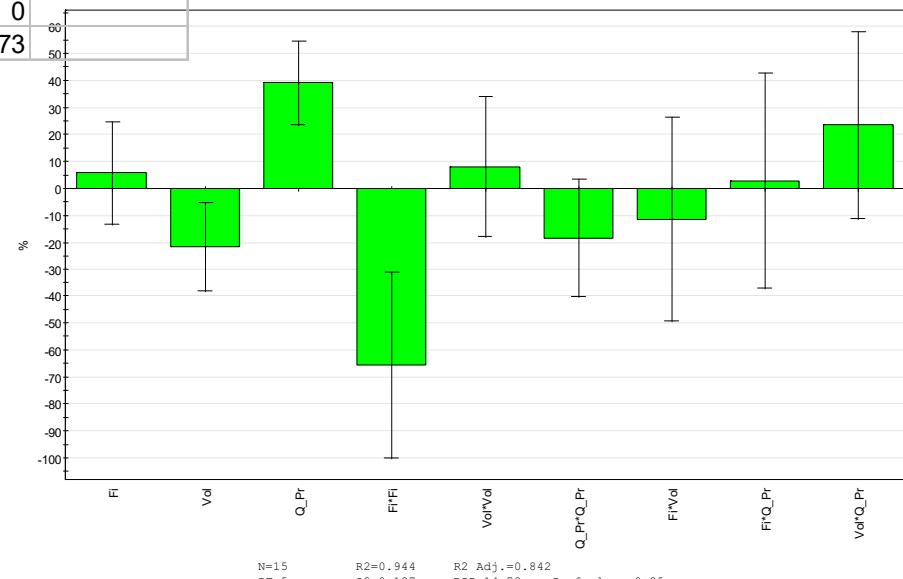


# Exemple optimisation: lysozyme

Lysozyme	DF	SS	MS (variance)	F	p	SD
Total	15	78930	5262			
Constant	1	59661.1	59661.1			
Total						
Corrected	14	19268.9	1376.35			37.0992
Regression	9	18184.4	2020.49	9.31471	<b>0.012</b>	44.9498
Residual	5	1084.57	216.913			14.728
Lack of Fit	3	1046.57	348.856	18.3608	<b>0.052</b>	18.6777
(Model Error)						
Pure Error (Replicate Error)	2	38	19			4.3589
N = 15	Q2 =	0.127	Cond. no. =	5.39		
DF = 5	R2 =	0.944	Y-miss =	0		
	R2 Adj. =	0.842	RSD =	14.73		

Investigation: Exemple Geteau JPB 2008 (MLR)

Scaled & Centered Coefficients for Lysozyme

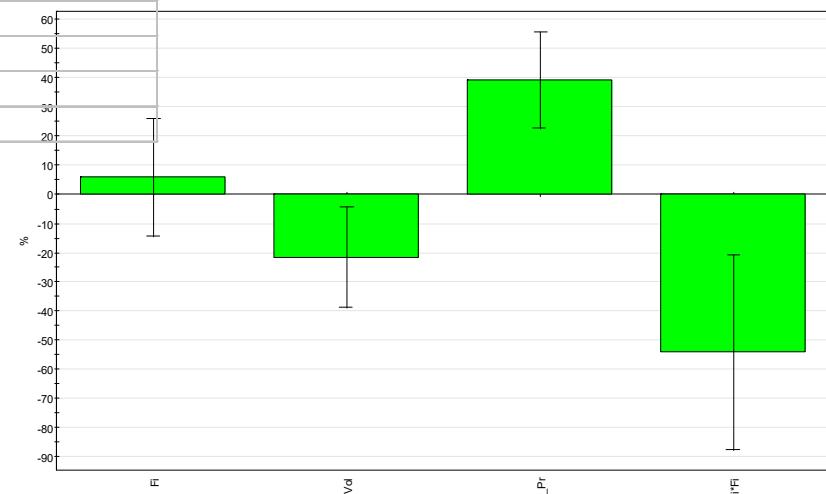


# Amélioration de l'ANOVA par élimination des termes ns du modèle

Lysozyme	DF	MS		F	p	SD
		SS	(variance)			
Total	15	78930	5262			
Constant	1	59661.1	59661.1			
Total						
Corrected	14	19268.9	1376.35			37.0992
Regression	4	16042.1	4010.54	12.4289	<b>0.001</b>	63.3288
Residual	10	3226.79	322.679			17.9633
Lack of Fit (Model Error)	8	3188.79	398.599	20.9789	<b>0.046</b>	19.9649
Pure Error (Replicate Error)	2	38	19			4.3589
N = 15	Q2 =	0.455	Cond. no. =	3.487		
DF = 10	R2 =	0.833	Y-miss =	0		
	R2 Adj. =	0.766	RSD =	17.96		

Investigation: Exemple Gateau JPB 2008 (MLR)

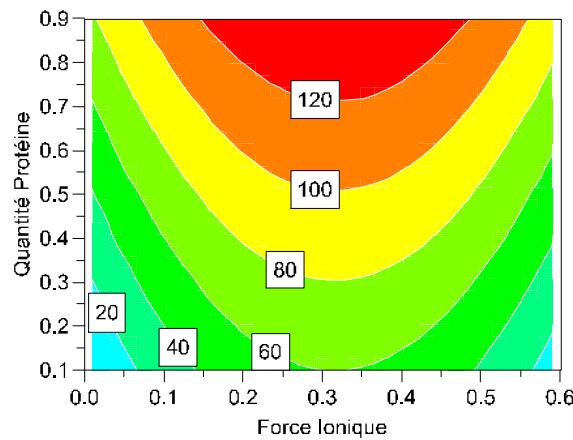
Scaled & Centered Coefficients for Lysozyme



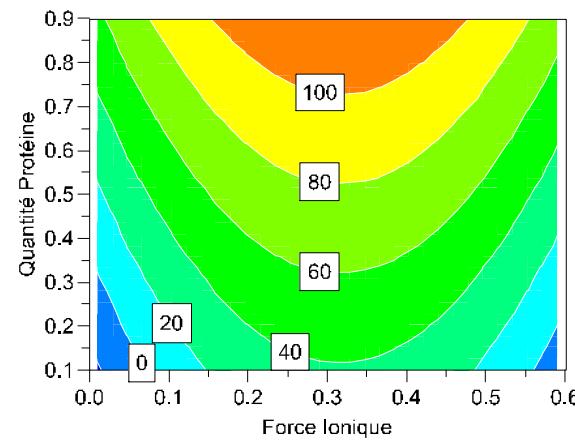
N=15 R2=0.833 R2 Adj.=0.766  
DF=10 Q2=0.455 RSD=17.96 Conf. lev.=0.95

# Surface de réponse

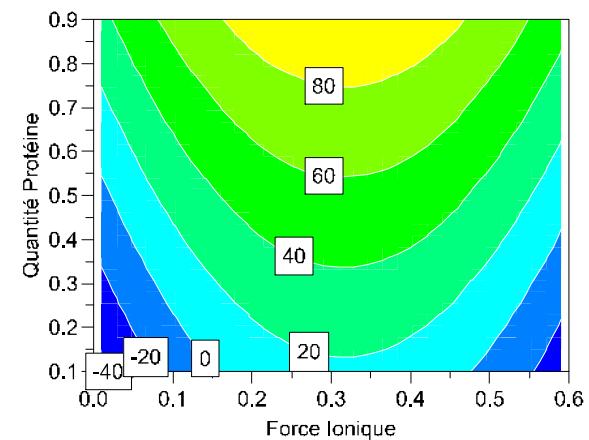
Investigation: Exemple Gateau JPB 2008 (MLR)  
4D Contour of Lysozyme



Vol. Phase Aqueuse = 25



Vol. Phase Aqueuse = 90

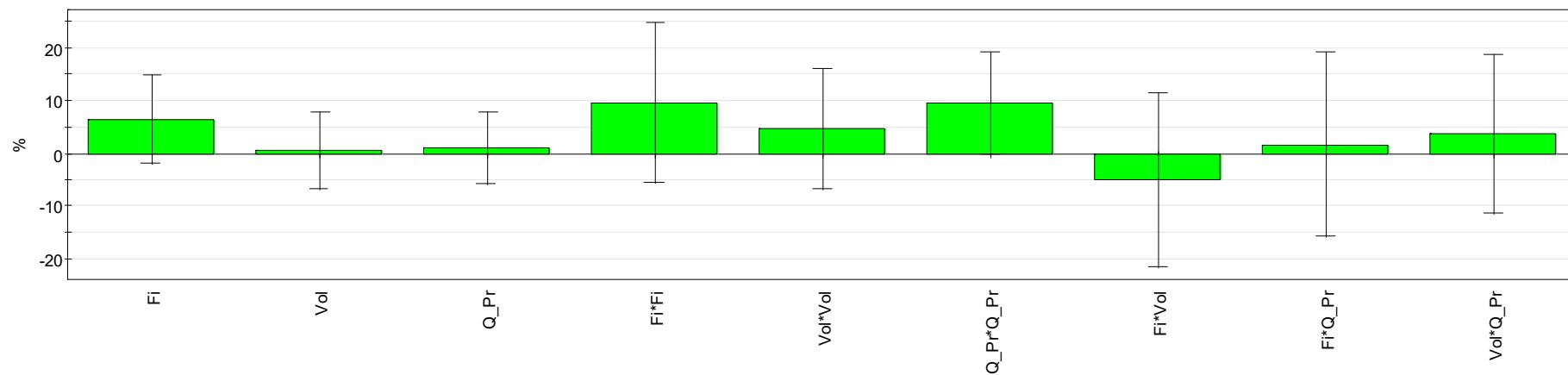


Vol. Phase Aqueuse = 155

MODDE 8 - 22/11/2010 19:11:04

# Exemple optimisation: $\beta$ galactosidase

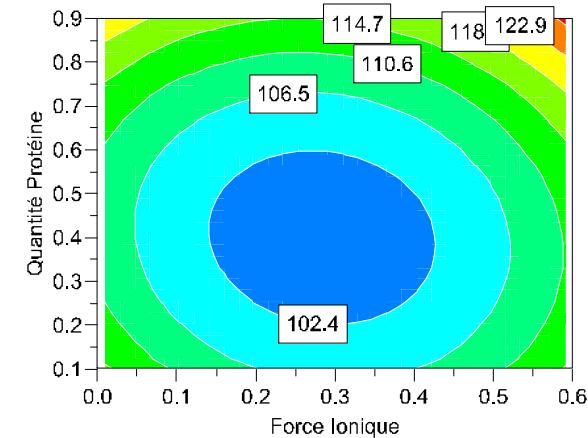
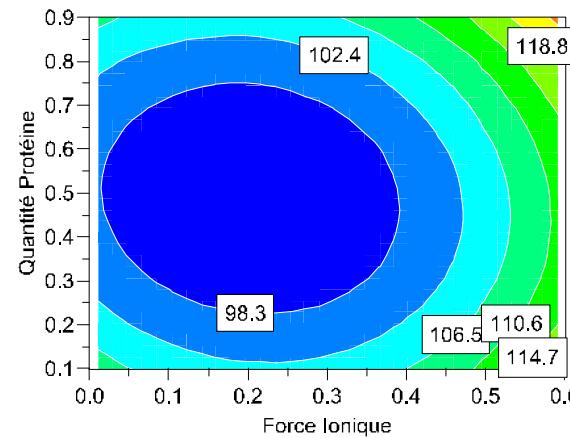
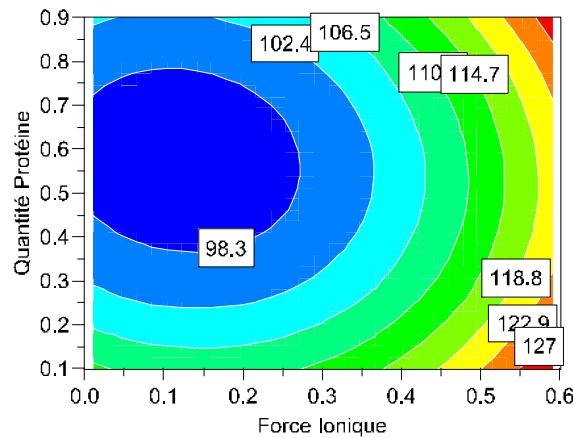
B-Galactosidase	DF	SS	MS (variance)	F	p	SD
Total	15	161117	10741.1			
Constant	1	160373	160373			
Total						
Corrected	14	743.594	53.1138			7.28792
Regression	4	222.775	55.6937	1.06935	<b>0.421</b>	7.46282
Residual	10	520.819	52.0819			7.21678
Lack of Fit (Model Error)	8	480.152	60.019	2.95176	<b>0.278</b>	7.7472
Pure Error (Replicate Error)	2	40.6667	20.3333			4.50925
N = 15	Q2 =	-0.501	Cond. no. =	3.487		
DF = 10	R2 =	0.3	Y-miss =	0		
	R2 Adj. =	0.019	RSD =	7.217		



# Exemple optimisation: $\beta$ galactosidase

Investigation: Exemple Gateau JPB 2008 (MLR)

4D Contour of B-Galactosidase



Vol. Phase Aqueuse = 25

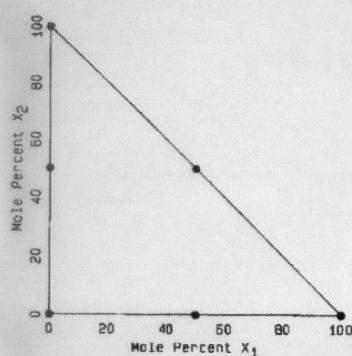
Vol. Phase Aqueuse = 90

Vol. Phase Aqueuse = 155

MODDE 8 - 22/11/2010 19:16:33

# Plans de mélanges

## Plans de mélanges.



$$\sum_{i=1}^n \% x_i = \% x_1 + \% x_2 + \dots + \% x_n = 100$$

$$\sum_{i=1}^3 \% x_i = \% x_1 + \% x_2 + \% x_3 = 100 \Rightarrow \% x_3 = 1 - (\% x_1 + \% x_2)$$

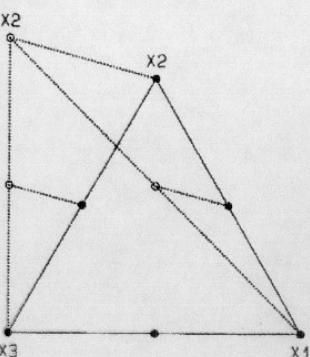
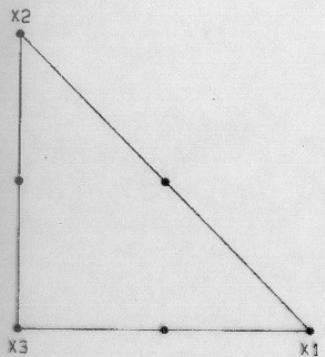
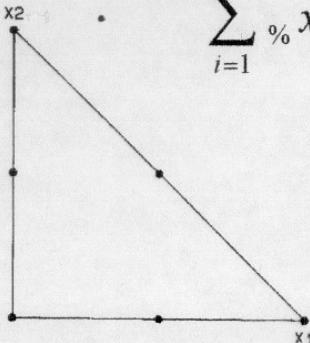


Figure 12.29 Conversion of an orthogonal mixture space to an equilateral triangular mixture space. See text for details.

# Plan de mélange

- Limonene encapsulation in freeze-drying of gum Arabic–sucrose–gelatin systems Vikas Kaushika and Yrjö H. Roos , Food Science and Technology, 2006

Exp No	Arabic Gum	Sucrose	Gelatin	ENCAPS
1	1	0	0	75
2	0	1	0	4.2
3	0	0	1	66
4	0.5	0.5	0	57
5	0.5	0	0.5	54
6	0	0.5	0.5	36
7	0.666667	0.166667	0.166667	64.2
8	0.166667	0.666667	0.166667	47.4
9	0.166667	0.166667	0.666667	67.7
10	0.333333	0.333333	0.333333	50.9
11	0.333333	0.333333	0.333333	51.3
12	0.333333	0.333333	0.333333	49.8

ENCAPS	DF	SS	MS (variance)	F	p	SD
Total	12	36113.9	3009.49			
Constant	1	32396	32396			
Total						
Corrected	11	3717.85	337.986			18.3844
Regression	5	3399.31	679.862	12.806	<b>0.004</b>	26.0742
Residual	6	318.536	53.0893			7.28624
Lack of Fit (Model Error)	4	317.329	79.3322	131.489	<b>0.008</b>	8.90686
Pure Error (Replicate Error)	2	1.20667	0.603335			0.776747
N = 12	Q2 =	0.268	Cond. no. =	14.54		
DF = 6	R2 =	0.914	Y-miss =	0		
	R2 Adj. =	0.843	RSD =	7.286		

# Plan de mélange

