Construction et Analyse d'un plan asymétrique de résolution IV

A. Kobilinsky[1]**, E.Mettler**[2]**, B. Carpentier**[2]

(1) Laboratoire de Biométrie. INRA, 78026 Versailles
(2) CNEVA, LERPAC, Paris.

RESUME. Dans une étude sur le nettoyage des surfaces, 16 facteurs susceptibles d'influer sur la qualité du nettoyage ont été initialement répertoriés. Certains de ces facteurs doivent impérativement être étudiés à plus de deux niveaux. Après avoir exposé les principes qui ont permis le choix du plan, une fraction 1/64 de résolution 4 d'un $4^3 2^6$, on en présente le mode d'analyse un peu particulier. Ce plan a été réalisé dans le cadre d'un projet CNEVA/INRA/UNIR intitulé "*Mise au point d'un test pour évaluer la nettoyabilité des surfaces colonisées par des biofilms*".

1 Méthodologie de l'étude

L'objectif de l'étude dans laquelle s'inscrit le plan d'expérience décrit ici est de mettre au point des techniques de laboratoire pour évaluer l'efficacité de procédures de nettoyage sur différents matériaux utilisés dans l'industrie agroalimentaire. L'intérêt d'une telle approche est clairement décrit dans Habert P. (1994).

Les échantillons de surface à comparer sont d'abord encrassés, puis soumis à un nettoyage et une désinfection. L'encrassement résiduel est mesuré par détachement aux ultrasons et dénombrement sur boites de pétri (Carpentier, Kobilinsky 1993). Une première liste de facteurs susceptibles de jouer est donnée au tableau 1. Pour plusieurs de ces facteurs, il est souhaitable d'étudier d'emblée plus que deux niveaux pour être représentatif de la diversité des situations possibles.

Pour le nettoyage, les échantillons sont placés par groupe de 8 dans un appareil Gardner-Erichsen (classiquement utilisé pour déterminer la lessivabilité de peintures). Les 8 échantillons de chaque groupe, manipulés simultanément dans des conditions aussi semblables que possible, constituent ce que nous appellerons un *bloc*. Un nombre global de 64 unités expérimentales correspondant à 8 tels blocs est jugé réalisable dans une première phase.

L'un des facteurs étudié est la température de nettoyage, égale soit à la température ambiante de 20°C, soit à la température de 4°C trouvée dans certains ateliers de travail de la viande. Cette température ne peut être variée à l'intérieur d'un bloc. C'est une contrainte dont on tiendra compte dans l'établissement du plan.

2 Principes généraux pour le choix du plan

Dans ce qui suit, le modèle d'analyse de variance est noté $y = X\beta + \varepsilon$. La reparamétrisation choisie est supposée de plein rang, c'est à dire telle que X est de plein rang lorsque le plan factoriel est complet. On suppose en outre que les paramètres dans β sont des effets factoriels, i.e. des effets principaux des facteurs ou des effets d'interactions.

En l'absence d'indications précises sur ces effets, une manière commode de formaliser l'objectif visé est la notion de *résolution*. Les plus utiles sont les résolutions 3, 4, 5.

		nb. de niveaux initialement requis	facteurs retenus	nb. de niv. retenu	code
— caractéristiques de la surface					
1	nature du matériau	4	⋆	4	mat
2	usure (par traitement de simulation)	2	⋆	2	us
— caractéristiques du biofilm					
3	souche (gram+; gram-)	4	⋆	2	sou
4	nature du milieu de culture	2			
5	concentration milieu de culture	2	⋆	2	mil
6	température milieu de culture	2			
7	age du biofilm	2			
— caractéristiques du nettoyage					
8	composition du détergent	4	⋆	4	dét
9	concentration du détergent	2			
10	température de nettoyage	2	⋆	2	T-net
11	durée du nettoyage	2	⋆	2	d-net
12	action mécanique (brosse, nb. de cycle, poids)	4	⋆	2	Pbros
— caractéristiques de la désinfection					
13	désinfectant	4	⋆	4	dés
14	temps de contact	2			
15	durée du rinçage	2			
16	température du rinçage	2			
AUTRES CONTRAINTES					
— Blocs de taille 8			⋆	8	bl
— Température de nettoyage (10) constante sur chaque bloc.					

TAB. 1 – *Liste initiale des facteurs a étudier*

- Un plan est dit de résolution 3 s'il permet d'estimer tous les effets principaux des facteurs sous l'hypothèse qu'il n'y a pas d'interaction de deux facteurs ou plus.
- Un plan est dit de résolution 4 s'il permet d'estimer tous les effets principaux dans un modèle comprenant les interactions de deux facteurs, mais aucune interaction de trois facteurs ou plus.
- Un plan est dit de résolution 5 s'il permet d'estimer tous les effets principaux et interactions de deux facteurs sous l'hypothèse qu'il n'y a pas d'interaction de trois facteurs ou plus.

L'estimabilité de la moyenne générale est habituellement aussi incluse dans la définition de ces plans.

Pour un nombre d'unités expérimentales donné, le nombre de facteurs qu'il est possible d'étudier diminue de façon nette lorsque la résolution augmente. Le choix de la résolution et du nombre de facteurs étudiés dépend du type de problème posé. Dans un état de grande ignorance des phénomènes, l'examen d'un nombre important de facteurs au moyen d'un plan de résolution 3 permet de déceler rapidement les quelques facteurs actifs dont les effets doivent être ultérieurement précisés par un plan de résolution supérieure. Dans cette perspective, il est même possible d'augmenter encore le nombre de facteurs en utilisant des plans dits hypersaturés où le nombre de facteurs dépasse le nombre d'unités (Lin 1995). Ces plans ne sont cependant adaptés que lorsqu'on pense que seuls quelques facteurs ont un effet prépondérant au sein d'un ensemble beaucoup plus vaste sur lequel on opère un criblage. Si cette hypothèse est erronée, l'analyse fait émerger un nombre important d'effets significatifs et il devient difficile de conclure compte tenu des confusions entre les effets principaux et les interactions.

Lorsque l'avancement des connaissances permet de préciser un petit nombre de facteurs actifs, on vise généralement la résolution 5 ou, si les interactions peuvent être précisées, un plan plus spécifiquement adapté à l'étude de ces interactions.

Dans l'étude décrite ici, les facteurs recensés ont a priori de bonnes chances d'être actifs et d'interagir. Le plan de résolution 3 semble donc contre-indiqué. La résolution 5 amène une réduction trop drastique du nombre de facteurs étudiés. La résolution 4 offre un bon compromis qui fournit une estimation fiable des effets principaux et permet, si le plan est régulier, d'apporter nombre de conclusions sur les interactions de deux facteurs.

Un plan est dit *régulier* s'il peut être défini par des relations de définition comme celles apparaissant au tableau 5. La régularité entraîne que les effets factoriels sont soit totalement confondus, soit orthogonaux. Dans le premier cas, les colonnes de X associées sont colinéaires. Dans le second cas, elles sont orthogonales. Un tel plan permet donc d'estimer orthogonalement les combinaisons linéaires disjointes des paramètres associés aux paquets de colonnes de X colinéaires. Les paramètres figurant dans une telle combinaison sont dit *confondus* ou *aliasés*. Dans le cas du plan de résolution 4 illustré au paragraphe 4, ces combinaisons contiennent soit un unique effet principal, soit un ensemble d'interactions. Ces ensembles d'interactions sont disjoints, ce qui facilite l'interprétation.

Pour une résolution donnée, le nombre de facteurs que l'on peut introduire dépend évidemment des nombres de niveaux de ces facteurs. Le cas où tous les facteurs ont deux niveaux est bien connu. Les tables de Connors & Young (1961) fournissent des fractions non orthogonales, mais efficaces, pour l'étude en résolution 5 d'un mélange de facteurs à 2 et 3 niveaux. Mais il n'existe pas à ce jour de tables ou programmes permettant d'obtenir de semblables fractions en résolution 4.

Ceci a amené à retenir ici 4 niveaux, plutôt que 3, pour les facteurs étudiés à plus que 2 niveaux. En décomposant chaque facteur à 4 niveaux en deux *pseudofacteurs* à 2 niveaux comme cela est indiqué pour le facteur matériau (*mat*) dans le tableau 2, on a alors la possibilité d'utiliser les techniques de construction de fractions de plans pour des facteurs à 2 niveaux. On doit toutefois tenir compte de ce qu'un effet tel que $mat_1.mat_2$ est un effet principal du facteur matériau au même titre que mat_1 et mat_2 et qu'il n'y a donc pas symétrie entre les facteurs et pseudofacteurs. Cette méthodologie, décrite plus en détail dans Cliquet & *al* (1994), donne une grande souplesse combinatoire pour l'obtention de plans réguliers et permet d'avoir recours à des programmes de recherche automatique des relations de définition tels que PLANOR (Kobilinsky 1994), FACTEX (SAS 1989), KEYFINDER (Zemroch 1992).

		facteur mat	pseudofacteurs mat$_1$	mat$_2$
	0	carreau de céramique	1	1
	1	résines acrylique	1	-1
(1)	2	résine époxy	-1	1
	3	résine polyurethane	-1	-1

TAB. 2 – *Décomposition en pseudofacteurs du facteur matériau*

(1) pour des raisons expliquées dans la suite du texte, ce matériau a été remplacé par de la résine acrylique

La connaissance du nombre maximum de facteurs étudiables en résolution 5, 4, 3 est déterminante pour le choix du plan. Le paragraphe suivant précise ce nombre dans le cas dit *asymétrique* où tous les facteurs n'ont pas le même nombre de niveaux. On s'intéresse plus précisément au cas des mélanges de facteurs à 4 et 2 niveaux.

3 Limite pour le nombre de facteurs dans le cas asymétrique

En résolution 5 ou 3 tous les paramètres du modèle doivent être estimés. Le nombre de paramètres doit donc être inférieur au nombre d'unités expérimentales. Cette propriété permet d'obtenir, pour un nombre d'unités donné, un majorant au nombre de facteurs introduisibles. Un résultat de Margolin (1969) permet d'obtenir un majorant analogue en résolution 4.

3.1 Résolution 5

Notons n_2 et n_4 les nombres respectifs de facteurs à 2 et 4 niveaux. Le nombre p de paramètres du modèle incluant les effets principaux et interactions de deux facteurs, donné par la formule

$$p = 1 + n_2 + 3n_4 + \frac{n_2(n_2-1)}{2} + 9\frac{n_4(n_4-1)}{2} + 3n_2n_4 \ ,$$

est reporté dans le tableau 3 pour $n_4 \leq 4$ et $n_2 \leq 11$. De l'examen de ce tableau il ressort qu'avec 64 unités, on a obligatoirement $n_2 \leq 2$ si $n_4 = 3$, $n_2 \leq 5$ si $n_4 = 2$, $n_2 \leq 8$ si $n_4 = 1$ et $n_2 \leq 10$ si $n_4 = 0$.

Mais si l'on recherche un plan régulier, les limites accessibles sont inférieures. On sait que si $n_4 = 0$, la régularité impose $n_2 \leq 8$. Elle impose de même les inégalités $n_2 \leq 0$ si $n_4 = 3$, $n_2 \leq 3$ si $n_4 = 2$ et $n_2 \leq 6$ si $n_4 = 1$. La théorie des plans fractionnaires (John 1971, Kobilinsky et Monod 1995) permet de montrer très simplement les deux premières inégalités. La troisième s'obtient à l'aide d'un programme tel que PLANOR.

n_4 \ n_2	0	1	2	3	4	5	6	7	8	9	10	11
0	1	2	4	7	11	16	22	29	37	46	$\underline{56}$	67
1	4	8	13	19	26	34	43	53	$\underline{64}$	76	89	103
2	16	23	31	40	50	$\underline{61}$	73	86	100	115	131	148
3	37	47	$\underline{58}$	70	83	97	112	128	145	163	182	202
4	67	80	94	109	125	142	160	179	199	220	242	265

TAB. 3 – *nb. p de paramètres du modèle en résolution* 5

Dans le cas de la résolution 5, les plans réguliers possèdent, outre l'orthogonalité, une propriété de robustesse importante: les estimateurs des effets principaux y restent non biaisés si certaines interactions de trois facteurs sont non nulles.

3.2 Résolution 4

Proposition 3.1 (Margolin) *Dans un plan de résolution* 4*, les colonnes de X associées à la constante, aux effets principaux et aux interactions avec un facteur détermniné sont indépendantes et leur nombre est donc inférieur au nombre d'unités.*

Ce résultat est valable quels que soient les nombres de niveaux des différents facteurs du plan.

Dans le cas symétrique où il y a n facteurs a 2 niveaux, le nombre p de colonnes considéré dans la proposition est $1+n+(n-1)=2n$. Avec 2^k unité, on trouve donc que n est inférieur à 2^{k-1}, majorant que l'on sait pouvoir atteindre avec un plan régulier.

Dans le cas particulier d'un mélange de n_4 facteurs à 4 niveaux et n_2 facteurs à 2 niveaux, le nombre de colonnes de X associées aux effets principaux est n_2+3n_4. Si le facteur déterminé a 4 niveaux, ses interactions avec les autres facteurs à 4 niveaux produisent $9(n_4-1)$ colonnes et ses interactions avec les facteurs à 2 niveaux en produisent $3n_2$. Le nombre p de colonnes de la proposition est donc, si $n_4>0$

$$p = 1 + n_2 + 3n_4 + 3(n_2 + 3(n_4 - 1)) \ .$$

Ce nombre est reporté dans le tableau 4 pour $1 \leq n_4 \leq 4$ et $n_2 \leq 15$. Pour que ce nombre

n_4 \ n_2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	4	8	12	16	20	24	28	32	36	40	44	48	52	56	60	$\underline{64}$
2	16	20	24	28	32	36	40	44	48	52	56	60	$\underline{64}$	68	72	76
3	28	32	36	40	44	48	52	56	60	$\underline{64}$	68	72	76	80	84	88
4	40	44	48	52	56	60	$\underline{64}$	68	72	76	80	84	88	92	96	100

TAB. 4 – *nb. p de paramètres du modèle en résolution* 4

soit inférieur au nombre d'unités 64, on doit avoir $n_2 \leq 6$ pour $n_4=4$, $n_2 \leq 9$ pour $n_4=3$, $n_2 \leq 12$ pour $n_4=2$, $n_2 \leq 15$ pour $n_4=1$. Si l'on impose en outre la régularité, les limites trouvées par programme descendent à $n_2 \leq 4$ pour $n_4=4$, $n_2 \leq 7$ pour $n_4=3$, $n_2 \leq 12$ pour $n_4=2$ et $n_4 \leq 15$ pour $n_4=1$. Ces dernières limites sont celles atteintes par le programme après un temps de recherche limité. Noter cependant que quand $n_4=1$ ou $n_4=2$, ces limites coincident avec la borne de Margolin. La borne sur n_2 donnée par la proposition 3.1 est donc atteinte avec un plan régulier de 64 unités lorsque $n_4=1$ ou 2 de même que lorsque $n_4=0$.

Ce résultat rejoint ceux donnés par Margolin (1969), Agrawal & Dey (1983), qui sont repris et développés dans un chapitre du livre de Dey (1985). Les constructions de ces auteurs, basées sur des matrices d'Hadamard, permettent de construire lorsque $n_4 = 1$ ou 2 des plans de résolution 4 orthogonaux pour les effets principaux et comportant le nombre maximum de facteurs à deux niveaux autorisé par la proposition 3.1. L'orthogonalité pour les effets principaux signifie que les colonnes de X associées aux effets principaux sont orthogonales entre elles et orthogonales aux colonnes associées aux interactions de deux facteurs.

Ces plans existent pour des nombres d'unités qui ne sont pas nécessairement des puissances de 2. Ainsi si le nombre d'unités N est multiple de 16, $N = 16m$, la technique de Margolin permet de construire à partir d'une matrice d'Hadamard d'ordre $4m$ un tel plan pour $n_4 = 1$ et $n_2 = 4m - 1$. Elle redonne lorsque $m = 4$ la limite $n_2 = 15$ donnée précédemment mais permet aussi en prenant $m = 3$ de trouver un plan à 48 unités ayant le nombre maximum $n_2 = 11$ facteurs à 2 niveaux. Ce dernier plan n'est cependant pas défini par des relations de définition et n'est pas régulier. Il ne permet pas d'exploiter les résultats concernant les interactions aussi simplement que le plan régulier considéré dans la suite de cet article.

De même, la méthode d'Agrawal & Dey donne pour $N = 32m$, $n_4 = 2$, des plans comportant le nombre maximum $n_2 = 8m - 4$ de facteurs à deux niveaux. Elle permet donc de trouver des plans non régulier lorsque par exemple $N = 96$. Agrawal et Dey proposent aussi une méthode de construction pour le cas $n_4 = 3$, $N = 64m$ unités. Mais le nombre $n_2 = 16m - 12$ de facteurs à deux niveaux correspondant est alors inférieur à la borne fournie par la proposition 3.1 et également assez nettement inférieur dans le cas $m = 1$ à la borne 7 obtenue par programme.

3.3 Résolution 3

On se reportera pour ces plans très connus et utilisés aux ouvrages de Dey (1985), Collombier (1995), Benoist & al (1994).

4 Construction et analyse du plan

Les considérations précédentes ont conduit à la sélection d'un plan régulier comprenant les facteurs marqués d'un $*$ dans le tableau 1. Ce plan comprend $n_4 = 3$ facteurs à 4 niveaux et $n_2 = 6$ facteurs à deux niveaux. Le choix d'un n_2 inférieur au plafond de 7 diminue la taille des ensemble d'effets confondus et facilite en conséquence l'interprétation. Ce choix permet aussi d'imposer qu'un des degrés de liberté de l'effet bloc soit non confondu avec une interaction.

Les relations de définition du plan figurent à droite du tableau 5. Elles ont été trouvées par le programme PLANOR en introduisant les spécifications données à gauche de ce tableau. Rappelons que les facteurs indexés sont ceux qui résultent de la décomposition en pseudofacteurs des facteurs traitement à 4 niveaux (matériau, détergent, désinfectant) et du facteur bloc. L'effet bl_1 représente donc un degré de liberté de l'effet bloc.

La première ligne du tableau précise que l'on veut estimer bl_1 et tous les effets traitement à l'exception de *T-net* dans le cadre du modèle comprenant les effets principaux, les interactions entre deux facteurs et l'effet bloc.

L'effet *T-net* ne peut pas être estimé dans un modèle comprenant l'effet bloc, mais on demande en ligne 2 qu'il puisse être estimé dans le modèle obtenu en supprimant l'effet

estimer	P+bl$_1$	*dans le modèle*	bl+Q.Q
estimer	T-net	*dans le modèle*	Q.Q
où	P= mat+dét+dés+us+sou+mil+Pbros+d-net		
et	Q= P+T-net		
	T-net *est défini à partir de* bl		

		Facteurs de base
us	=	mat$_1$ mat$_2$ dét$_1$ dés$_1$
sou	=	mat$_2$ dét$_2$ dés$_1$
mil	=	mat$_2$ dét$_1$ dés$_2$
Pbros	=	mat$_1$ dét$_1$ dét$_2$ dés$_2$
d-net	=	mat$_1$ mat$_2$ dét$_1$ dét$_2$ dés$_1$ dés$_2$
T-net	=	mat$_1$ mat$_2$ dét$_2$ dés$_2$
bl$_1$	=	mat$_1$ dét$_2$ dés$_1$ dés$_2$
bl$_2$	=	mat$_1$ mat$_2$ dés$_1$ dés$_2$
bl$_3$	=	mat$_1$ dés$_2$

TAB. 5 – *Formulation de la demande et relations de définition trouvées par PLANOR*

bloc. On est ainsi certain qu'il ne se confond pas avec un autre effet principal ou une interaction de deux facteurs et qu'il peut être estimé inter-bloc.

Les 64 unités du plan peuvent être repérées par les combinaisons de niveaux -1 et $+1$ des 6 pseudofacteurs mat_1, mat_2, $dét_1$, $dét_2$, $dés_1$, $dés_2$, appelés facteurs de base. Les niveaux des autres facteurs sont obtenus en effectuant les produits indiqués. Pour simplifier on a systématiquement affecté un signe positif à ces produits. On aurait pu prendre l'opposé de certains de ces produits, par exemple définir l'usure par *us*= -mat_1 mat_2 $dét_1$ $dés_1$. Un tel changement de signe qui revient à intervertir les niveaux -1 et $+1$ ne change pas la nature des confusions d'effets,

Le tableau 6 donne certaines des 63 combinaisons linéaires estimables d'effets et leurs estimations classées par valeurs absolues décroissantes, pour la variable efficacité de la désinfection y. Grace à la régularité, ces combinaison sont disjointes, c'est à dire que deux d'entre elles ne peuvent inclure le même effet.

Il n'y a pas de répétitions permettant de dégager d'emblée des degrés de liberté pour estimer la variance d'erreur. On ne peut pas non plus utiliser à cette fin les interactions d'ordre élevé car, à l'exception de q, toutes les combinaisons contiennent soit un effet principal, soit une interaction de 2 facteurs traitement. Pour déterminer à partir de quel seuil ces sommes diffèrent significativement de 0, on les représente par une demi-droite de Henry (Half Normal probability plot) dans la figure 1. On trouvera des références sur cette représentation dans Kobilinsky, Fliss & Carpentier (1995).

Ce graphique fait apparaître un saut vers la valeur 0.14 qu'il est naturel d'utiliser comme seuil de détection. On remarque que les effets température de nettoyage *T-net*, usure du matériau *us*, pression de brossage *Pbros*, représentés respectivement par les symboles 0, 5, s, sont parmi les effets négligeables. Ceci justifie que l'on fasse une analyse en supprimant ces facteurs du modèle, y compris les interactions où ils apparaissent.

En fait, suite à des difficultés de dernière minute rencontrées dans l'obtention du biofilm sur l'un des matériaux, la résine époxy, ce matériau a été remplacé par l'un des trois autres, la résine acrylique, qui se trouve de ce fait deux fois plus représenté dans le plan que les deux autres matériaux. Le déséquilibre ainsi introduit est pris en compte dans la définition des effets en donnant un poids double à la résine acrylique dans le calcul des effets n'impliquant pas l'effet matériau. De cette façon, bien que le plan perde sa régularité, les combinaisons linéaires estimables d'interactions restent assez simples pour permettre une exploitation des résultat allant au delà des seuls effets principaux. Le tableau 7 donne ces combinaisons après suppression des facteurs *T-net*, *us*, *Pbros*. Cette suppression et la disparition des effets où apparaît mat_2 (sans mat_1) permettent de dégager 10 degrés de liberté pour estimer la variance d'erreur. L'examen du tableau 2 permets de comprendre pourquoi mat_2 disparaît : si les matériaux 1 et 2 sont identiques, les contrastes définis par les colonnes mat_1 et mat_2 coincident tous deux avec la différence entre les matériaux 0 et 3.

	y	
v	0.58	mat$_1$ mat$_2$
H	0.53	dét$_1$
d	−0.45	dét$_1$ sou
B	−0.44	dés$_1$
Y	−0.37	mil
R	−0.33	[bl$_2$ bl$_3$] + Pbros d-net + mat$_2$ dés$_1$ + dét$_2$ sou
A	−0.31	dés$_2$
W	−0.29	us Pbros + dés$_2$ sou
z	0.28	dés$_2$ T-net + mil Pbros + mat$_1$ mat$_2$ dét$_2$
V	0.26	sou
w	−0.24	mat$_1$ mat$_2$ dés$_2$ + dét$_2$ T-net
o	−0.23	mat$_1$ mat$_2$ mil + dét$_2$ Pbros + sou d-net
2	−0.21	dét$_1$ d-net + dés$_1$ T-net
b	−0.21	mat$_2$ dét$_1$ dét$_2$
P	0.19	mat$_2$
N	−0.19	dét$_1$ dét$_2$ dés$_1$
8	−0.18	dét$_1$ T-net + dés$_1$ d-net + mat$_2$ Pbros
I	0.17	dét$_1$ dés$_2$ + mat$_2$ mil
e	−0.17	mat$_1$ d-net
n	−0.16	mat$_1$ dét$_1$
f	0.15	mat$_1$
j	0.15	mat$_1$ dét$_2$
$\star$	0.15	d-net
g	0.11	[bl$_3$] + mat$_1$ dés$_2$ + dét$_1$ dét$_2$ Pbros
		...
0	0.09	[bl$_1$ bl$_2$ bl$_3$] + T-net
l	−0.08	mil d-net + mat$_1$ mat$_2$ sou
		...
s	−0.07	Pbros
U	0.06	mat$_1$ T-net + dét$_1$ dét$_2$ mil + dés$_1$ dés$_2$ sou
		...
5	−0.05	us
r	−0.05	dés$_2$ Pbros + mil T-net + mat$_1$ dét$_1$ dét$_2$ + us sou
		...
q	0.03	
		...
F	0.00	[bl$_1$ bl$_3$] + dét$_2$ dés$_1$ + mat$_2$ sou
C	0.00	dés$_1$ dés$_2$

TAB. 6 – *Combinaisons linéaires estimables des effets*

FIG. 1 – *Graphique des quantiles*
Les lettres sont celles qui apparaissent à gauche du tableau 6

	y
CM erreur (10 ddl)	0.3002
ecart-type	0.5479

	effet	$F(1{,}10)$	S	$\hat{e}$
1	constante			0.358
2	mil	128.5	$***$	-0.365
3	mat_1 mat_2	72.6	$***$	0.583
4	$dét_1$	59.0	$***$	0.526
5	$dét_1$ sou	43.3	$***$	-0.450
6	$dés_1$	41.4	$***$	-0.440
7	$dés_2$	20.4	$**$	-0.309
8	$dés_2$ sou	18.4	$**$	-0.294
9	mat_1 mat_2 $dét_2$	16.6	$**$	0.279
10	sou	13.9	$**$	0.255
11	mat_1	12.5	$**$	0.171
12	bl_2 bl_3 + $dét_2$ sou	12.4	$**$	-0.340
13	mat_1 mat_2 $dés_2$	12.0	$**$	-0.237
14	mat_1 mat_2 mil + sou d-net	10.9	$**$	-0.226
15	$dét_1$ d-net	9.7	$*$	-0.212
16	$dét_1$ $dét_2$ $dés_1$	7.4	$*$	-0.185
17	mat_1 $dét_1$ $dét_2$	6.9	$*$	-0.127
18	$dés_1$ d-net	6.6	$*$	-0.175
19	mat_1 d-net	6.5	$*$	-0.123
20	mat_1 $dét_1$	5.8	$*$	-0.164
21	bl_3 - $dét_1$ mil	5.1	$*$	0.217
22	mat_1 $dét_2$	4.6		0.146
23	d-net	4.6		0.146
24	bl_1 bl_2 + $dés_1$ sou	4.2		-0.197
25	mat_1 $dés_2$ + $dét_1$ mil	2.4		-0.105
26	$dét_1$ $dés_1$	2.4		0.106
	...			
54	$dés_2$ d-net	0.0		0.005

TAB. 7 – *Estimation et test des effets dans le modèle simplifié*

Les combinaisons d'effets apparaissant au tableau 7 ont été obtenues par une technique classique, décrite par exemple dans Kobilinsky (1989). On remarque que certaines de ces combinaisons ne sont plus disjointes. Par exemple l'effet $dét_1.mil$ apparait dans les combinaisons linéaires 21 : bl_3 - $dét_1.mil$ et 25 : $mat_1.dés_2{+}dét_1.mil$.

L'interprétation des résultats significatifs ne pose pas de problème pour les combinaisons d'effets réduites à un terme. Par ailleurs les arguments suivants amènent à assimiler les sommes 12 : $bl_2.bl_3+$ $dét_2.sou$ et 14 : $mat_1.mat_2.mil$ + $sou.d\text{-}net$ à $dét_2.sou$ et $mat_1.mat_2.mil$ respectivement.

La première fait intervenir un effet bloc $bl_2.bl_3$. Or l'effet bl_1 non confondu est négligeable, et 4 autres effets blocs apparaissent dans des sommes non significativement différentes de 0. Il paraît donc raisonnable de faire l'hypothèse qu'il n'y a pas d'effet bloc et que l'estimateur obtenu pour la somme $bl_2.bl_3+$ $dét_2.sou$ estime correctement $dét_2.sou$.

Pour la seconde somme, on est tenté d'attribuer son effet significatif à l'interaction $mat_1.mat_2.mil$ car les effets principaux associés $mat_1.mat_2$ et mil sont beaucoup plus significatifs que ceux des deux autres facteurs sou et $d\text{-}net$.

En définitive, on peut dans cet exemple faire des hypothèses réalistes pour estimer tous les paramètres du modèle. Cela rend possible l'estimation des moyennes associées aux paires de facteurs et permet même d'obtenir une prédiction de la variable étudiée pour chaque combinaison de niveaux des facteurs actifs.

REFERENCES

Agrawal V. and Dey A. (1983). Orthogonal Resolution-IV Designs for Some Asymmetrical Factorials. *Technometrics*, **25**, n0 2, 197-199.

Benoist D., Tourbier Y. et Germain-Tourbier S. (1994). Plans d'expériences : construction et analyse. Tec-Doc, Lavoisier.

Box G.E.P. and Wilson K.B. (1951). On the experimental attainment of optimum conditions. *J. Roy. Stat. Soc. Ser. B*, 13, 1-45. (reproduit dans *The collected Works of G.E.P Box*, Vol 2, Ed : G.C. Tiao, Wadsworth.)

Carpentier B.et Kobilinsky A. (1993). Outils pour l'étude de la nettoyabilité des surfaces utilisées dans l'industrie agro-alimentaire. *Actes de Contaminexpert 93.*

Cliquet S., Durier C. et Kobilinsky A. (1994). Principle of a fractional design for qualitative and quantitative factors; application to the production of *Bradyrhizobium japonicum* in culture and inoculation media. *Agronomie*, **14**, 569-587.

Collombier D. (1996). Plans d'expérience factoriels. Springer Verlag, *collection Mathématiques et applications*, **21**, Heidelberg.

Connor W.S.and Young S. (1961). Fractional factorial designs with factors at two and three levels. National Bureau of Standards. Applied Mathematics Series. **58**. (Reproduit dans McLean and Anderson, 1984, Marcel Dekker, New-York).

Dey A. (1985). Orthogonal Fractional Factorial Designs. *Wiley Eastern*, New Dehli.

Dey A. and Ramakrishna G.V.S. (1977). A Note on Orthogonal Main-Effect Plans. *Technometrics*, **19**, 511-512.

Draper N. and Mitchell T. (1970). Construction of a set of 512-run designs of resolution $\geq$ 5 and a set of even 1024-run designs of resolution $\geq$ 6. *The Annals of Mathematical Statistics*, **41**, n0 3, 876-887.

Habert P. (1994). Film bactérien : le défi de l'ultrapropreté. *La recherche*, **263**, 330-332.

John P.W.M. (1971). Statistical Design and Analysis of Experiments. Macmillan,

Kobilinsky A. (1989). Tactiques en analyse de variance et en régression. *Rev Modulad*,**1,** 25-58, Ed: INRIA.

Kobilinsky A. (1994). PLANOR: programme de génération automatique de plans d'expériences réguliers. Document interne, Labo. de Biométrie, INRA, Versailles.

Kobilinsky A., Fliss M. et Carpentier B. (1995). Méthodes et exemples de construction de plans factoriels fractionnaires. Dans : Stratégie expérimentale et procédés biotechnologiques. Collection *Récents progrès en génie des procédés*, Vol 9, **36**, p19-28, Ed: GFGP, Diffuseur: Lavoisier.

Kobilinsky A.and Monod H. (1991). Experimental design generated by group morphisms: an introduction. *Scand. J. Statist.*, **18**, 119-134.

Kobilinsky A. and Monod H. (1995). Juxtaposition of regular factorial designs and the complex linear model. *Scand. J. Statist* **22**, n° 2, 223-254.

Lin K.J. (1995). Generating Systematic Supersaturated Designs. *Technometrics*, **37**, n0 2, 213-225.

Margolin B.H. (1969). Resolution IV Fractional Factorial Designs. *J. Roy. Statist. Soc. Ser. B*, **31**, 514-523.

SAS/QC Software (1989). The FACTEX Procedure. Reference Version 6, First Edition. SAS Institute Inc.

Zemroch, P.J. (1992). KEYFINDER - A complete toolkit for generating fractinal-replicate and blocked factorial designs. In Dodge Y. and J. Whittaker (eds). Computational Statistics Vol 2, 263-268. Physica-Verlag.