

PLANOR: programme de génération automatique de plans d'expériences réguliers

Version 2.2 pour Windows

A. Kobilinsky

Octobre 1999

Unité de Biométrie (BIA).
INRA
F78352. JOUY-EN-JOSAS Cedex.
FRANCE

Table des matières

1	Introduction	1
2	Présentation de la méthode	2
2.1	Exemple avec des facteurs à 2 niveaux	2
2.1.1	Définition et propriétés du plan	2
2.1.2	Principe de la recherche d'un plan par PLANOR	4
2.1.3	Randomisation	4
2.1.4	Variantes pour l'écriture du modèle	5
2.2	Exemple avec des facteurs à 3 niveaux	6
2.2.1	Définition et propriétés	6
2.2.2	Recherche par PLANOR de plusieurs solutions et sélection par l'étude des alias	8
2.2.3	Randomisation du plan en bloc précédemment défini	11
2.3	Exemple d'un mélange de facteurs à 6, 4 et 2 niveaux	12
2.3.1	Définition et propriétés. Décomposition en pseudofacteurs	12
2.3.2	Etude des alias	12
3	Exemples de plans pour le robot ARILAIT	15
3.1	Création d'un plan factoriel complet intégrant les contraintes de manipulation	15
3.1.1	Objectif et contraintes	15
3.1.2	Création par juxtaposition de deux sous-plans	16
3.1.3	Fusion des deux plans	20
3.1.4	Randomisation prenant en compte la structure des blocs	20
3.2	Plans pour une plaque	26
3.2.1	Plan de résolution 4 pour 8 facteurs	27
3.2.2	Plan de résolution 5 pour 5 facteurs	31
3.3	Plans pour deux plaques	32
3.3.1	Plan $2^6/2$ de résolution 5	35
3.3.2	Plan $2^8/8$ de résolution 4	35

3.3.3	Plan $4 \times 2^4/2$ de résolution 5	40
3.3.4	Plan $4 \times 2^7/16$ de résolution 4	40
4	Plans pour mélange de facteurs à 2 et 4 niveaux	45
4.1	Facteurs qualitatifs	45
4.1.1	Exemple	45
4.1.2	Plans de résolution 4	45
4.1.3	Plans de résolution 3, doublables en résolution 4	48
4.1.4	Prise en compte des blocs	50
4.2	Mélange de facteurs quantitatifs et qualitatifs	55
4.2.1	Effets polynomiaux et pseudofactoriels	55
4.2.2	Plans réguliers orthogonaux pour le modèle polynomial	59
4.2.3	Plans réguliers non orthogonaux pour le modèle polynomial	60
4.2.3.1	Fraction 1/2 d'un 4×4 pour deux facteurs quantitatifs	60
4.2.3.2	Fraction 1/16 d'un $4^3 2^4$ avec deux facteurs quantitatifs	65
5	Utilisation du programme	68
5.1	Condition d'emploi	68
5.2	Généralités	68
5.3	Description succincte des différents modules	68
5.4	Création ou modification d'un plan régulier	70
5.4.1	Ecran 1 (fig. 2)	70
5.4.1.1	nom	70
5.4.1.2	nb. d'unités	70
5.4.1.3	définition des facteurs de base	70
5.4.1.4	type de décomposition des facteurs	71
5.4.1.5	Recherche backtrack.	71
5.4.1.6	inclusion facteurs dans ensemble inéligible	73
5.4.1.7	Commentaire	74
5.4.2	Ecran 2 (fig. 3)	74
5.4.2.1	Facteurs de base et facteurs définis	74

5.4.2.2	Modèles, parties à estimer, parties de modèle : généralités.	74
5.4.2.3	Modèles, parties à estimer, parties de modèle : syntaxe. . .	75
5.4.2.4	Utilisation des parties de modèle	76
5.4.2.5	Suppression de termes	77
5.4.2.6	Les Hiérarchies.	77
5.4.2.7	Exemple avec plusieurs modèles et hiérarchies.	77
5.4.2.8	Facteurs prédéterminés	79
5.5	Randomisation	79
5.6	Recodage, sélection de facteurs, tri,	84
5.6.1	Création de nouveaux facteurs (fig. 4)	85
5.7	Etude des alias	85
5.8	Contenu des fichiers .REG du répertoire actif.	85
5.9	Initialisation	87
5.10	Contenu des fichiers REG, PS, PR, HIS, HIR	88
5.10.1	Fichiers .REG	88
5.10.2	Fichiers PS et PR	91
5.10.3	Fichiers HIS et HIR	92

Table des figures

1	Position des éprouvettes sur la plaque	15
2	Plan ROBOT1A, écran 1	17
3	Plan ROBOT1A, écran 2	18
4	Plan ROBOT1B, Redéfinition des niveaux de <i>pl</i> . Ecran 4	22
5	Randomisation du plan ROBOT1, écran 3	22
6	Définition de systèmes de blocs (<i>col1</i> , <i>col2</i> , <i>lig1</i> , <i>lig2</i>)	27
7	Exemple de structure en bloc (extrait de [3])	82
8	Sélection du modèle pour l'étude des alias. Ecran 6	86
9	Initialisation, écran 5	87

Liste des tableaux

1	Exemple 2.1: 4 facteurs et 2^3 unités expérimentales.	2
2	Effets confondus dans l'exemple 2.1.	3
3	Exemple 2.2: 4 facteurs à 3 niveaux sur 3 blocs de 9 unités	6
4	Effets confondus dans l'exemple 2.2.	7
5	Randomisation d'un plan en bloc	11
6	Décomposition en pseudofacteurs dans l'exemple 2.3.	12
7	Confusion pour des effets faisant intervenir des pseudofacteurs à 2 et 3 niveaux	14
8	Matrices clé des plans ROBOT1A et ROBOT1B	19
9	Plans ROBOT1, 1A, 1B avant randomisation (fichiers .PS)	21
10	Plan ROBOT1 randomisé, avant et après tri, édité par <i>écrit n0</i>	24
11	Plan ROBOT1 randomisé, avant et après tri. Edition par <i>écriture</i>	25
12	Plans de résolution 4 pour 16 unités et 6 à 8 facteurs	26
13	Plan ROBPL1R4 pour 1 plaque, 8 facteurs traitement	28
14	Randomisation du plan ROBPL1R4.PS (sorties intermédiaires)	30
15	Ensembles d'effets confondus dans le plan ROBPL1R4	31
16	Plan ROBPL1R5 pour 1 plaque, 5 facteurs traitement	33
17	Confusion avec les effets blocs dans ROBPL1R5	34
18	Recherche nb. maxi. de facteurs pour une fraction de taille 32 d'un 4×2^n	34
19	Plan ROBPL2R5 pour 2 plaques, 6 facteurs traitement	36
20	Confusion avec les effets blocs dans ROBPL2R5	37
21	Plan ROBPL2R4 pour 2 plaques, 8 facteurs traitement	38
22	Confusion d'effets dans trois solutions type d'un $2^8/8$ de résolution 4	39
23	Plan ROP2F4R5 pour 2 plaques, un $4 \times 2^4/2$ de résolution 5	41
24	Recherche de différentes solutions pour un $4 \times 2^7/16$	42
25	Structure d'alias dans un $4 \times 2^7/16$ de résolution 4	43
26	Plan ROP2F4R4 pour 2 plaques, un $4 \times 2^8/32$ de résolution 4	44
27	Plans de résolution 4 pour mélange de facteurs qualitatifs à 4 et 2 niveaux	47

28	Majorant de n_2 dans une fraction de taille 64, résolution 4 d'un $4^{n_4}2^{n_2}$. . .	47
29	2^{7-3} de résolution 4 obtenu par doublement d'un 2^{7-4} de résolution 3 . . .	49
30	Plan résol. 3 doublable en résol. 4 avec fact. à 4 et 2 niveaux	49
31	Fraction de résolution 4 d'un $4^3 \times 2^7$ en 8 blocs de 8 unités	51
32	Structure d'alias pour le plan du tableau 31	52
33	Modif. recherche tab. 31 pour estimer la variance inter-bloc	53
34	Fraction de résolution 4 d'un $4^3 \times 2^4$ en 8 blocs de 8 unités	54
35	Structure d'alias pour le plan du tableau 34	54
36	Fraction résol. 3 doublable en résol. 4 d'un $4^4 \times 2^{11}$ en 8 blocs de 8 unités .	56
37	Plan double du plan du tableau 36, de résolution 4	57
38	Décomposition en pseudofacteurs d'un facteur A quantitatif	58
39	Relations entre effets principaux polynomiaux et pseudofactoriels	58
40	Fraction 1/4 d'un $4 \times 4 \times 2^4$, adaptée à des fact. à 4 niveaux quantitatifs .	61
41	Expression des effets pseudofactoriels dans un modèle de degré 2	62
42	Fraction 1/2 d'un 4×4 , avec des facteurs à 4 niveaux quantitatifs	63
43	Estimations des effets polynomiaux dans la fraction $4^2/2$	64
44	Estimations des moindres carrés des effets linéaires dans la fraction $4^2/2$.	65
45	Fraction 1/16 d'un 4^32^4 incluant 2 fact. quant. à 4 niv.	66
46	Efficacités globales pour les 6 solutions	67
47	Efficacités pour les 3 meilleures solutions pour la trace	67
48	Menu général	68
49	Décomposition en pseudofacteurs d'un facteur A à 24 niveaux	71
50	Solution inadéquate pour une répartition en 3 systèmes de blocs croisés . .	78

1 Introduction

PLANOR génère des fractions de plan avec éventuellement un ou plusieurs systèmes de blocs. Il utilise une méthode dérivée de la méthode des matrices clés [23], décrite de façon détaillée dans [19], [20], [16], et de façon plus simplifiée dans [6]. Cette méthode conduit à des plans dit *réguliers* où les effets sont soit estimables indépendamment, soit complètement confondus.

Dans le cas le plus simple, l'utilisateur doit donner le modèle d'analyse de variance, préciser les termes qu'il veut estimer dans ce modèle. PLANOR cherche un ou plusieurs plans répondant à la demande de l'utilisateur parmi les plans constructibles par cette méthode. Comme il n'y a pas toujours de solution, le programme indique l'endroit où il s'est arrêté dans la construction du plan pour permettre à l'utilisateur de reformuler une demande permettant d'obtenir une solution.

PLANOR peut prendre en compte les contraintes de hiérarchie entre facteurs. Il permet aussi d'introduire plusieurs modèles et les familles de termes à estimer correspondantes. Cette possibilité est notamment utile dans un plan en blocs si un facteur est contraint de rester constant à l'intérieur de chaque bloc. Son effet est alors inestimable intra-bloc, c'est à dire dans un modèle comportant l'effet bloc. Mais il peut être estimable inter-bloc si on demande son estimation dans un modèle sans effet bloc: on évite ainsi de le confondre avec un autre effet traitement. Lorsqu'il y a plusieurs systèmes de blocs, on peut utiliser cette possibilité pour imposer l'estimation de certains effets dans des *strates* précisées à l'avance.

A l'origine, PLANOR a été réalisé pour élaborer des plans d'expériences permettant de piloter un robot. Ce robot, mis au point au Laboratoire du Génie de l'Hygiène et des Procédés Alimentaires de l'INRA Massy (LGHPA) dans le cadre d'un contrat avec l'Association pour le développement de la Recherche dans l'Industrie LAITIÈRE (ARILAIT), sert à tester des procédures de nettoyage et désinfection de surfaces. C'est pourquoi une partie importante de la notice est consacrée aux plans pour ce robot, qui illustrent d'ailleurs très bien les possibilités du programme.

L'analyse et l'interprétation des plans obtenus par PLANOR est généralement simple à condition de bien connaître les techniques d'analyse de variance. En particulier, lorsqu'il y a des blocs et que certains facteurs restent constants dans les blocs, l'analyse fait appel aux notions de *strates inter et intra-bloc*. L'effet principal d'un facteur constant dans chaque bloc est testé par comparaison à la variance *inter-bloc*, qui diffère de la variance d'erreur *intra-bloc* utilisée pour tester les autres effets. Un autre cas où l'analyse est un peu délicate est celui où le nombre de *degrés de liberté* de l'erreur est très réduit ou nul. On trouvera dans [18] et [16] quelques indications supplémentaires et des références sur ces sujets.

La présentation adoptée dans cette notice évite le recours au formalisme algébrique. Le texte y perd un peu en rigueur, mais est rendu ainsi accessible à un public beaucoup plus large. Les passages un peu ardu sont imprimés en caractères de petite taille et marqués au début par un astérisque. Ils peuvent être sautés sans nuire à la compréhension du reste du texte.

2 Présentation de la méthode

Les unités expérimentales sont repérées par les niveaux d'un certain nombre de facteurs, appelés *facteurs de base*. A partir de ces facteurs éventuellement décomposés, le programme définit si possible de nouveaux facteurs respectant les conditions imposées. Ces facteurs sont des combinaisons linéaires des facteurs de base ou des pseudofacteurs résultant de leur décomposition. Les quelques exemples suivants illustrent bien la méthode et ses propriétés.

2.1 Exemple avec des facteurs à 2 niveaux

2.1.1 Définition et propriétés du plan

Il y a 8 unités repérées par les combinaisons de niveaux de 3 facteurs traitement A , B , C ayant chacun 2 niveaux codés 0 et 1. A partir de ces trois facteurs de base, on définit un nouveau facteur D en posant $D = A + B + C \pmod{2}$. Ce même plan peut être aussi défini par la relation $D = ABC$ si l'on code 1 et -1 les niveaux. Pour distinguer les deux codages et présentations associées, on parle de notation *additive* dans le premier cas, de notation *multiplicative* dans le second. Le passage de la notation additive à la notation multiplicative s'opère en remplaçant chaque niveau α par le niveau $(-1)^\alpha$. Le tableau 1 présente le plan sous ses deux formes.

$D = A + B + C$				$D = ABC$				
A	B	C	D	A	B	C	D	ind.rép.
0	0	0	0	1	1	1	1	0
0	0	1	1	1	1	-1	-1	2
0	1	0	1	1	-1	1	-1	7
0	1	1	0	1	-1	-1	1	1
1	0	0	1	-1	1	1	-1	6
1	0	1	0	-1	1	-1	1	5
1	1	0	0	-1	-1	1	1	4
1	1	1	1	-1	-1	-1	-1	3

TAB. 1 – Exemple 2.1: 4 facteurs et 2^3 unités expérimentales.

Les effets factoriels étudiés sont les effets principaux A , B , C , D des facteurs et leurs interactions AB , AC , \dots , $ABCD$. Ces effets sont notés aussi de façon additive: $e(A)$ pour l'effet principal de A , $e(A+B)$, $e(A+C)$, \dots , $e(A+B+C+D)$ pour les interactions. La notation fonctionnelle $e()$ est avec cette notation additive indispensable pour distinguer une somme d'effets telle que $e(A+B)+e(C)$ de l'interaction correspondante $e(A+B+C)$. Elle sera aussi parfois utilisée en notation multiplicative pour distinguer un effet du produit de facteurs correspondant (égal à -1 ou 1).

La définition précise des effets est simple en notation multiplicative. Si l'on note $\tau(A,B,C,D)$ la réponse moyenne –en terme plus statistiques, l'espérance de la réponse–

terme ajouté	égalité déduite	effets confondus
	$0 = A + B + C + D$	$e(0), e(A + B + C + D)$
A	$A = B + C + D$	$e(A), e(B + C + D)$
B	$B = A + C + D$	$e(B), e(A + C + D)$
C	$C = A + B + D$	$e(C), e(A + B + D)$
$A + B$	$A + B = C + D$	$e(A + B), e(C + D)$
$A + C$	$A + C = B + D$	$e(A + C), e(B + D)$
$B + C$	$B + C = A + D$	$e(B + C), e(A + D)$
$A + B + C$	$A + B + C = D$	$e(A + B + C), e(D)$

TABLE 2 – Effets confondus dans l'exemple 2.1.

pour le traitement (A, B, C, D) , l'effet principal de A et l'interaction AB sont par exemple définis par

$$\begin{aligned}
e(A) &= \frac{1}{16} \sum_{A,B,C,D} A \tau(A, B, C, D) = \frac{1}{2} \sum_A A \tau(A, \dots, \dots) \\
&= \frac{1}{2} \left(\tau(1, \dots, \dots) - \tau(-1, \dots, \dots) \right) \\
e(AB) &= \frac{1}{16} \sum_{A,B,C,D} AB \tau(A, B, C, D) = \frac{1}{4} \sum_{A,B} AB \tau(A, B, \dots, \dots) \\
&= \frac{1}{4} \left(\tau(1, 1, \dots, \dots) - \tau(1, -1, \dots, \dots) - \tau(-1, 1, \dots, \dots) + \tau(-1, -1, \dots, \dots) \right).
\end{aligned}$$

Chaque point (\dots) indique que l'on a fait la moyenne sur la lettre correspondante. Par exemple :

$$\tau(A, B, \dots, \dots) = \frac{1}{4} \left(\tau(A, B, 1, 1) + \tau(A, B, 1, -1) + \tau(A, B, -1, 1) + \tau(A, B, -1, -1) \right)$$

La relation de définition $D = A + B + C \pmod{2}$ se réécrit sous la forme

$$A + B + C + D = 0 \pmod{2}. \tag{1}$$

En additionnant à cette égalité les sommes $A, \dots, A + B + C$ formées à partir des trois facteurs de base, on trouve les égalités figurant dans la seconde colonne du tableau 2. La règle simple suivante permet alors d'obtenir les ensembles d'effets confondus figurant dans la troisième colonne du tableau.

Règle 1 Les effets correspondant à deux sommes α et β égales sont confondus.

Noter que $e(0)$ est par définition la moyenne générale. L'interaction $e(A + B + C + D)$ est donc confondue avec cette moyenne générale.

En l'occurrence, la confusion de deux effets, par exemple $e(A + B)$ et $e(C + D)$ se traduit par le fait qu'on ne peut estimer que leur somme $e(A + B) + e(C + D)$. Les

effets confondus sont les mêmes si on rajoute 1 dans la définition de D , c'est à dire si $D = A + B + C + 1 \pmod{2}$, ou encore

$$A + B + C + D = 1 \pmod{2}, (ABCD = -1 \text{ en notation multiplicative}). \quad (2)$$

Dans ce cas ce sont les différences comme $e(A + B) - e(C + D)$ qui sont estimables et non les sommes.

L'examen de la troisième colonne du tableau 2 montre que les effets principaux se confondent dans cet exemple avec des interactions de 3 facteurs et que les interactions de deux facteurs sont confondues entre elles. Cela vient de ce que la relation de définition (1) comporte 4 facteurs. Un tel plan est dit de *résolution 4*. Si les interactions de 3 ou 4 facteurs sont négligées, chacune des sommes $e(A) + e(BCD), \dots, e(ABC) + e(D)$ contenant un effet principal peut être assimilée à cet effet. Les effets principaux sont alors estimables.

2.1.2 Principe de la recherche d'un plan par PLANOR

Dans ce qui précède, les propriétés du plan sont déduites de sa définition. PLANOR procède généralement en sens inverse. Il recherche le plan à partir du modèle et des termes à estimer. Ainsi dans cet exemple, après avoir précisé les facteurs de base A, B, C et le facteur à définir D , on introduit le modèle et les termes à estimer sous la forme symbolique suivante

$$\begin{array}{ll} \text{modèle} & : A + B + C + D + A.B + A.C + A.D + B.C + B.D + C.D \\ \text{termes à estimer} & : A + B + C + D. \end{array}$$

Le programme explore les possibilités pour D . Il élimine les choix qui ne permettent pas d'estimer les effets A, B, C, D dans le cadre du modèle considéré. Par exemple le choix $D = AB$ est éliminé parce qu'il conduit à une confusion des effets principaux A, B, D avec les interactions BD, AD, AB respectivement. Ici le seul choix qui peut être retenu est $D = \pm ABC$. En notation additive, $D = A + B + C$ ou $D = A + B + C + 1$. Après choix au hasard ou par l'utilisateur de la constante 0 ou 1 ajoutée à $A + B + C$ pour obtenir D , PLANOR construit le plan dans l'ordre systématique qui apparait au tableau 1. Ce plan est stocké dans un fichier suffixé par .PS. Les lettres PS sont les initiales de *Plan Systématique*.

2.1.3 Randomisation

L'affectation des traitements aux unités expérimentales (parcelles en agriculture, animaux en zootechnie, numéro de manipulation dans une expérience de laboratoire, ...) s'effectue généralement au hasard. Dans le cas présent, il n'y a pas de blocs et cette affectation au hasard, appelée *randomisation*, revient à tirer au hasard le numéro de l'unité affectée à chacun des 8 traitements. Le numéro tiré figure sous l'appellation *indice de répétition* dans le plan randomisé. Un possible résultat de ce tirage est donné dans le tableau 1.

Pour obtenir à partir du tableau 1 un plan facilement utilisable, il faut remplacer les numéros de niveaux par les niveaux réels, trier dans un ordre convenable, etc ... Ces opérations auxiliaires peuvent être effectuées en sélectionnant l’option *recodage, sélection de facteurs, tri, ...* du menu général (tableau 48).

2.1.4 Variantes pour l’écriture du modèle

PLANOR complète automatiquement le modèle par ajout de la constante et des termes “inclus” dans une des interactions. On peut donc abrégé l’écriture du modèle précédent en omettant les effets principaux puisque chacun d’eux est inclu dans une interaction:

$$\text{modèle: } A.B + A.C + A.D + B.C + B.D + C.D$$

Lorsqu’il y a beaucoup de facteurs, l’écriture du modèle peut devenir fastidieuse. Pour la raccourcir, on peut utiliser des parenthèses. On développe alors en supprimant les facteurs redondants dans chaque *effet* puis les effets redondants. Ainsi

$$\begin{aligned} (A + B + C + D)(A + B + C + D) &\rightarrow A.A + A.B + A.C + A.D + B.A + B.B + B.C + B.D \\ &\quad + C.A + C.B + C.C + C.D + D.A + D.B + D.C + D.D \\ &\rightarrow A + A.B + A.C + A.D + B.A + B + B.C + B.D \\ &\quad + C.A + C.B + C + C.D + D.A + D.B + D.C + D \\ &\rightarrow A + A.B + A.C + A.D + B + B.C + B.D + C + C.D + D \end{aligned}$$

donc le modèle précédant peut encore s’écrire sous la forme

$$(A + B + C + D)(A + B + C + D).$$

On peut éviter de réécrire deux fois l’expression $A + B + C + D$, en définissant cette expression comme une *partie de modèle* à laquelle est affectée un libellé qui peut être réutilisé dans le modèle ainsi que dans les termes à estimer.

partie de modèle	$P : A + B + C + D$
modèle	$P.P$
termes à estimer	P

Noter que les points séparant les facteurs dans chaque terme peuvent être remplacés par des blancs: $P P$ est équivalent à $P.P$.

Si l’on veut que le modèle contiennent toutes les interactions sauf une ou deux, disons $B.C$, $B.D$, il peut être commode d’utiliser une expression telle que $P.P \sim B.C + B.D$ utilisant le signe \sim pour soustraire les deux interactions $B.C$, $B.D$ au modèle $P.P$. La partie à droite de \sim peut être aussi écrite $B(C + D)$. Mais contrairement à la partie $P.P$ figurant à gauche de \sim , cette partie à droite du signe \sim n’est pas complétée par les sous-termes, en l’occurrence les effets principaux B , C , D . Seules les interactions $B.C$, $B.D$ sont donc ici supprimées du modèle.

$D = A + B + C$ $Bl = A + B + 2C$					$D = ABC$ $Bl = ABC^2$					Randomisation (cf tableau 5)	
A	B	C	D	Bl	A	B	C	D	Bl	Bl_0	ind-rép
0	0	0	0	0	1	1	1	1	1	2	2
0	0	1	1	2	1	1	j	j	j^2	0	6
0	0	2	2	1	1	1	j^2	j^2	j	1	6
0	1	0	1	1	1	j	1	j	j	1	0
0	1	1	2	0	1	j	j	j^2	1	2	1
0	1	2	0	2	1	j	j^2	1	j^2	0	1
0	2	0	2	2	1	j^2	1	j^2	j^2	0	7
0	2	1	0	1	1	j^2	j	1	j	1	2
0	2	2	1	0	1	j^2	j^2	j	1	2	5
1	0	0	1	1	j	1	1	j	j	1	8
1	0	1	2	0	j	1	j	j^2	1	2	7
1	0	2	0	2	j	1	j^2	1	j^2	0	8
1	1	0	2	2	j	j	1	j^2	j^2	0	2
1	1	1	0	1	j	j	j	1	j	1	7
1	1	2	1	0	j	j	j^2	j	1	2	0
1	2	0	0	0	j	j^2	1	1	1	2	4
1	2	1	1	2	j	j^2	j	j	j^2	0	4
1	2	2	2	1	j	j^2	j^2	j^2	j	1	4
2	0	0	2	2	j^2	1	1	j^2	j^2	0	3
2	0	1	0	1	j^2	1	j	1	j	1	5
2	0	2	1	0	j^2	1	j^2	j	1	2	8
2	1	0	0	0	j^2	j	1	1	1	2	6
2	1	1	1	2	j^2	j	j	j	j^2	0	0
2	1	2	2	1	j^2	j	j^2	j^2	j	1	3
2	2	0	1	1	j^2	j^2	1	j	j	1	1
2	2	1	2	0	j^2	j^2	j	j^2	1	2	3
2	2	2	0	2	j^2	j^2	j^2	1	j^2	0	5

TAB. 3 – Exemple 2.2: 4 facteurs à 3 niveaux sur 3 blocs de 9 unités

2.2 Exemple avec des facteurs à 3 niveaux

2.2.1 Définition et propriétés

Il y a 27 unités définies par 3 facteurs traitement A , B , C à 3 niveaux codés 0, 1, 2. On définit un nouveau facteur traitement D et un facteur bloc Bl en posant $D = A+B+C$ et $Bl = A + B + 2C \pmod{3}$.

Le plan résultant est donné dans le tableau 3. La notation multiplicative figurant au centre de ce tableau utilise la racine troisième $j = \exp(2\pi/3)$ de l'unité, qui remplace le -1 utilisé dans le cas des facteurs à 2 niveaux.

Les effets factoriels de l'effet principal de A sont A , A^2 , ceux de l'interaction entre A et B sont AB , A^2B^2 , AB^2 , A^2B . On énumère de même les effets factoriels appartenant aux autres effets principaux et interactions. En notation additive ces mêmes effets se notent $e(A)$, $e(2A)$, $e(A + B)$, $e(2A + 2B)$, $e(A + 2B)$, $e(2A + B)$, etc ... La définition précise de ces effets qui correspondent à une décomposition en contrastes orthogonaux, on dit aussi en degrés de liberté orthogonaux, n'est nullement indispensable pour comprendre et utiliser PLANOR.

En ajoutant $2D$ à l'égalité $D = A + B + C$ et en multipliant ensuite le résultat par 2 $\pmod{3}$, on obtient:

$$0 = A + B + C + 2D = 2A + 2B + 2C + D . \quad (3)$$

L'ajout des 27 combinaisons linéaires entre les facteurs de base A , B , C permet alors d'obtenir les égalités figurant à gauche du tableau 4. Les effets bloc Bl et $2Bl$ égaux respectivement à $A + B + 2C$ et $2A + 2B + C$ ont été rajoutés entre crochets (attention: $2 \times 2 C = 4C = C \pmod{3}$).

La règle 1 permet de déduire immédiatement de ces égalités les groupes d'effets confondus, qui apparaissent à droite du tableau 4.

A partir des égalités (3), on déduit directement les propriétés suivantes du plan (que l'on peut vérifier dans le tableau 4).

- Un effet principal ne se confond qu'avec des interactions de trois facteurs ou plus.
- Parmi les quatre degrés de liberté d'une interaction entre deux facteurs, deux se confondent avec une autre interaction de deux facteurs, et deux ne se confondent qu'avec des interactions de trois facteurs ou plus.
- Il n'y a pas d'ensemble de trois effets confondus qui ne contienne que des effets d'interactions de trois ou quatre facteurs.

La dernière propriété montre que toute autre définition de Bl confondrait l'effet bloc avec au moins une interaction de deux facteurs. Cependant le choix effectué ici rend inestimables les effets d'interaction $e(C + D)$ et $e(2C + 2D)$ même si les interactions de trois ou quatre facteurs sont supposées nulles. Un choix plus judicieux serait $Bl = C + 2D$. Ce choix permettrait de confondre l'effet bloc avec les effets $e(C + 2D)$ et $e(2C + D)$ qui sont déjà confondus avec d'autres interactions de 2 facteurs et de toute façon inestimables.

	égalités induites par (3)	effets confondus
	$0 = A + B + C + 2D = 2A + 2B + 2C + D$	$e(0), e(A + B + C + 2D), e(2A + 2B + 2C + D)$
	$C = A + B + 2C + 2D = 2A + 2B + D$	$e(C), e(A + B + 2C + 2D), e(2A + 2B + D)$
	$2C = A + B + 2D = 2A + 2B + C + D$	$e(2C), e(A + B + 2D), e(2A + 2B + C + D)$
	$B = A + 2B + C + 2D = 2A + 2C + D$	$e(B), e(A + 2B + C + 2D), e(2A + 2C + D)$
	$B + C = A + 2B + 2C + 2D = 2A + D$	$e(B + C), e(A + 2B + 2C + 2D), e(2A + D)$
	$B + 2C = A + 2B + 2D = 2A + C + D$	$e(B + 2C), e(A + 2B + 2D), e(2A + C + D)$
	$2B = A + C + 2D = 2A + B + 2C + D$	$e(2B), e(A + C + 2D), e(2A + B + 2C + D)$
	$2B + C = A + 2C + 2D = 2A + B + D$	$e(2B + C), e(A + 2C + 2D), e(2A + B + D)$
	$2B + 2C = A + 2D = 2A + B + C + D$	$e(2B + 2C), e(A + 2D), e(2A + B + C + D)$
	$A = 2A + B + C + 2D = 2B + 2C + D$	$e(A), e(2A + B + C + 2D), e(2B + 2C + D)$
	$A + C = 2A + B + 2C + 2D = 2B + D$	$e(A + C), e(2A + B + 2C + 2D), e(2B + D)$
	$A + 2C = 2A + B + 2D = 2B + C + D$	$e(A + 2C), e(2A + B + 2D), e(2B + C + D)$
	$A + B = 2A + 2B + C + 2D = 2C + D$	$e(A + B), e(2A + 2B + C + 2D), e(2C + D)$
	$A + B + C = 2A + 2B + 2C + 2D = D$	$e(A + B + C), e(2A + 2B + 2C + 2D), e(D)$
$[Bl]$	$A + B + 2C = 2A + 2B + 2D = C + D$	$[e(Bl)], e(A + B + 2C), e(2A + 2B + 2D), e(C + D)$
	$A + 2B = 2A + C + 2D = B + 2C + D$	$e(A + 2B), e(2A + C + 2D), e(B + 2C + D)$
	$A + 2B + C = 2A + 2C + 2D = B + D$	$e(A + 2B + C), e(2A + 2C + 2D), e(B + D)$
	$A + 2B + 2C = 2A + 2D = B + C + D$	$e(A + 2B + 2C), e(2A + 2D), e(B + C + D)$
	$2A = B + C + 2D = A + 2B + 2C + D$	$e(2A), e(B + C + 2D), e(A + 2B + 2C + D)$
	$2A + C = B + 2C + 2D = A + 2B + D$	$e(2A + C), e(B + 2C + 2D), e(A + 2B + D)$
	$2A + 2C = B + 2D = A + 2B + C + D$	$e(2A + 2C), e(B + 2D), e(A + 2B + C + D)$
	$2A + B = 2B + C + 2D = A + 2C + D$	$e(2A + B), e(2B + C + 2D), e(A + 2C + D)$
	$2A + B + C = 2B + 2C + 2D = A + D$	$e(2A + B + C), e(2B + 2C + 2D), e(A + D)$
	$2A + B + 2C = 2B + 2D = A + C + D$	$e(2A + B + 2C), e(2B + 2D), e(A + C + D)$
	$2A + 2B = C + 2D = A + B + 2C + D$	$e(2A + 2B), e(C + 2D), e(A + B + 2C + D)$
$[2Bl]$	$2A + 2B + C = 2C + 2D = A + B + D$	$[e(2Bl)], e(2A + 2B + C), e(2C + 2D), e(A + B + D)$
	$2A + 2B + 2C = 2D = A + B + C + D$	$e(2A + 2B + 2C), e(2D), e(A + B + C + D)$

TAB. 4 – Effets confondus dans l'exemple 2.2.

Le plan permet donc d'estimer tous les effets principaux dans un modèle incluant les interactions de deux facteurs et l'effet bloc. Les spécifications permettant d'obtenir un

tel type de plan avec PLANOR sont:

$$\begin{array}{ll}
 \text{partie de modèle} & P : A + B + C + D \\
 \text{modèle} & Bl + P.P \\
 \text{termes à estimer} & P
 \end{array} \tag{4}$$

2.2.2 Recherche par PLANOR de plusieurs solutions et sélection par l'étude des alias

On peut dans cet exemple de taille réduite demander l'ensemble des solutions respectant la demande définie par (4). Il est facile de montrer qu'il y en a 144. Sous l'hypothèse de nullité des interactions de plus de trois facteurs, un tiers de ces solutions permettent d'estimer la moitié des degrés de liberté de chaque interaction entre deux facteurs. Les deux autres tiers confondent en outre avec les blocs les deux autres degrés de liberté de certaines interactions. Il n'est pas possible de spécifier au programme que l'on recherche de préférence un plan du bon tiers, mais on peut étudier a posteriori les solutions trouvées par l'option *étude des alias* du menu général (tableau 48).

Dans la suite de ce paragraphe, nous donnons le résultat de l'étude des alias pour une solution de chaque type. La solution appartenant au bon tiers est la solution 1 qui laisse, pour chacune des 6 interactions de deux facteurs, 2 degrés de liberté non confondus. La solution appartenant au mauvais tiers est la numéro 34 qui confond totalement deux interactions et pour laquelle il n'y a donc que 4 interactions bénéficiant de 2 degrés de liberté non confondus.

Les égalités définissant chaque plan sont présentées sous forme d'une matrice clé, avec les facteurs de base en ligne. A chaque facteur figurant soit dans le modèle, soit dans les termes à estimer, soit dans les hiérarchies ou parmi les facteurs prédéfinis, est associé une colonne de cette matrice précisant la combinaison linéaire des facteurs de base qui permet de définir ce facteur.

En fait plusieurs plans disjoints aux propriétés similaires sont obtenus en rajoutant à chaque facteur, modulo 3, un entier compris entre 0 et le coefficient "cR" donné en haut de la colonne correspondante.

L'étude des effets confondus, aussi appelés *alias*, commence par l'édition d'un tableau synthétique à partir duquel les listes d'effets confondus s'obtiennent simplement. Ce tableau est un intermédiaire technique n'intéressant que des utilisateurs très informés. L'utilisateur ordinaire pourra sans gêne sauter les paragraphes en petits caractères contenant ce tableau et les commentaires qui s'y rapportent.

Le modèle proposé de façon standard pour l'étude des alias est le modèle introduit initialement pour la recherche du plan. Il peut être modifié sans difficulté juste avant l'étude des alias.

Sorties de l'étude des alias pour les deux solutions retenues

Solution 1 (premier 3)

Pour obtenir les niveaux d'un facteur en tête d'une colonne, on multiplie les niveaux des facteurs de base figurant en tête de ligne par les coeffs dans la colonne et on ajoute un entier naturel fixé à l'avance et inférieur à cR. Calculs de niveaux effectués modulo 3.

		<i>blocs</i>				*
	<i>cR</i>	2	0	2	0	0
		3	3	3	3	3
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>Bl</i>
3	<i>A</i>	1	0	0	1	1
3	<i>B</i>	0	1	0	1	1
3	<i>C</i>	0	0	1	1	0

Etude des effets confondus.

★ Les colonnes des matrices ci-dessous donnent les relations de définition du plan (noyau de la matrice clé) d'où l'on déduit les effets confondus.

- Les colonnes de la première matrice engendrent tous les effets traitement confondus avec la moyenne générale. Par addition des vecteurs à coord. $\leq cT$, on peut en déduire tous les ensembles d'effets traitement confondus.
- Les combinaisons linéaires à coeffs $\leq cBT$ des colonnes de la seconde matrice donnent les effets blocs confondus et pour chacun d'eux un des effets traitement avec lequel il est confondu. Les autres effets traitement confondus avec cet effet bloc s'obtiennent en ajoutant les effets traitement confondus avec la moyenne.
- Les ensemble d'effets traitement confondus entre eux mais pas avec un effet bloc peuvent s'obtenir directement en ajoutant un des vecteurs non nul à coord. $\leq cTB$ aux ensembles d'effets traitement confondus avec les blocs.
- Les effets blocs non confondus s'obtiennent en ajoutant un vecteur non nul à coord. $\leq cB$ aux effets blocs confondus.
- Dans les ensembles d'effets confondus obtenus de la façon décrite ci-dessus, l'analyse détaillée des alias ne retient que ceux qui figurent effectivement dans le modèle (ce qui peut entraîner des différences dans la ventilation des effets).

						<i>cBT</i>	2
				<i>ordre</i>	3		3
<i>cB</i>	<i>cT</i>	<i>cTB</i>	<i>bl</i>				
0	2	2		3	<i>A</i>	1	<i>A</i> 0
0	2	0		3	<i>B</i>	1	<i>B</i> 0
0	2	2		3	<i>C</i>	1	<i>C</i> 2
0	0	0		3	<i>D</i>	2	<i>D</i> 1
0	0	0	*	3	<i>Bl</i>	0	<i>Bl</i> 2
						↑	↑
				matrice n^0	1	n^0	2

- Liste des effets traitement confondus avec la moyenne

$$; ABCD^2; A^2B^2C^2D;$$

- Ensembles d'effets confondus du modèle. Si l'ensemble contient un effet bloc, celui

ci est indiqué entre crochets.

$[Bl]; C^2D; AB;$
 $[Bl^2]; CD^2; A^2B^2;$
 $AC; B^2D;$
 $A^2D; BC;$
 $AD^2; B^2C^2;$
 $A^2C^2; BD^2;$

– liste des effets traitement non confondus

$A; A^2B; B^2; A^2; B; AB^2; C; D; C^2D^2;$
 $AD; B^2C; A^2C; BD; C^2; CD; D^2; AC^2;$
 $B^2D^2; BC^2; A^2D^2;$

– liste des effets blocs non confondus : vide

Solution 34 (premier 3)

Calculs de niveaux effectués modulo 3.

		<i>blocs</i>				*	
		<i>cR</i>	2	0	2	0	0
		3	3	3	3	3	3
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>Bl</i>	
3	<i>A</i>	1	0	0	2	2	
3	<i>B</i>	0	1	0	1	1	
3	<i>C</i>	0	0	1	1	2	

Etude des effets confondus.

						<i>cBT</i>		2
				ordre		3		3
*	<i>cB</i>	<i>cT</i>	<i>cTB</i>	<i>bl</i>				
	0	2	2	3	<i>A</i>	1	<i>A</i>	0
	0	2	2	3	<i>B</i>	2	<i>B</i>	0
	0	2	0	3	<i>C</i>	2	<i>C</i>	2
	0	0	0	3	<i>D</i>	1	<i>D</i>	2
	0	0	0	*	3	<i>Bl</i>	0	<i>Bl</i>
				3			0	1

– Liste des effets traitement confondus avec la moyenne

$; AB^2C^2D; A^2BCD^2;$

– Ensembles d'effets confondus du modèle. Si l'ensemble contient un effet bloc, celui

ci est indiqué entre crochets.

$$\begin{aligned}
 & [Bl^2]; C^2D^2; \\
 & [Bl]; CD; \\
 & A^2C; B^2D; \\
 & A^2D^2; B^2C^2; \\
 & A^2B; C^2D; \\
 & AB^2; CD^2; \\
 & BC; AD; \\
 & BD^2; AC^2;
 \end{aligned}$$

– liste des effets traitement non confondus

$$\begin{aligned}
 & A; BD; BC^2; A^2; B^2C; B^2D^2; B; AC; \\
 & AD^2; AB; C; D^2; B^2; A^2D; A^2C^2; D; C^2; A^2B^2;
 \end{aligned}$$

– liste des effets blocs non confondus

2.2.3 Randomisation du plan en bloc précédemment défini

La randomisation s’effectue ici en deux temps. Les vrais blocs d’unités expérimentales sont numérotés et on choisit au hasard, pour chaque numéro figurant dans le plan systématique, le numéro du vrai bloc associé. Puis on tire au hasard dans chaque bloc le numéro d’unité affecté à chaque traitement. Le tableau 5 donne un résultat possible de cette randomisation qui conduit aux colonnes intitulées Bl_0 et *ind-rép* du tableau 3. La colonne Bl_0 donne le numéro du vrai bloc, *ind-rép* celui de l’unité dans le bloc. Ces colonnes figurent après randomisation dans un fichier suffixé par .PR – Plan Randomisé – déduit du plan systématique .PS en remplaçant dans la colonne Bl_0 les numéros initiaux par les numéros, tirés au hasard, des vrais blocs et en adjoignant la colonne *ind-rép*. Comme dans l’exemple 2.1, l’obtention à partir de ce fichier .PR d’un fichier “prêt à l’emploi” demande quelques transformations annexes comprenant en particulier le tri sur Bl_0 et *ind-rép*.

no. de bloc dans le plan systématique	0	1	2
no. du véritable bloc	2	1	0

Bl	no. des unités affectées aux traitements du bloc								
0	2	1	5	7	0	4	8	6	3
1	6	0	2	8	7	4	5	3	1
2	6	1	7	8	2	4	3	0	5

TAB. 5 – Randomisation d’un plan en bloc

2.3 Exemple d'un mélange de facteurs à 6, 4 et 2 niveaux

2.3.1 Définition et propriétés. Décomposition en pseudofacteurs

Il y a 144 unités réparties en 6 blocs – facteur Bl –, sur lesquelles on expérimente 4 facteurs traitement A, B, C, D ayant respectivement 6, 6, 4 et 2 niveaux. A part D qui n'a que 2 niveaux, chacun de ces facteurs est décomposé en 2 *pseudofacteurs* de la façon indiquée dans le tableau 6. Noter que le programme utilise le caractère “_” pour indiquer un indijage: A_{-1} pour A_1 , B_{-2} pour B_2 , etc ...

A	A_1	A_2	B	B_1	B_2	Bl	Bl_1	Bl_2	C	C_1	C_2
1	1	2	1	1	2	1	1	2			
2	1	1	2	1	1	2	1	1	1	1	1
3	1	0	3	1	0	3	1	0	2	1	0
4	0	2	4	0	2	4	0	2	3	0	1
5	0	1	5	0	1	5	0	1	4	0	0
6	0	0	6	0	0	6	0	0			

TAB. 6 – Décomposition en pseudofacteurs dans l'exemple 2.3.

Les facteurs A, B, C sont pris comme facteurs de base. On définit alors Bl et D par les égalités suivantes:

$$\begin{aligned}
 Bl_1 &= A_1 + B_1 + C_1 && (\text{mod } 2) \\
 Bl_2 &= A_2 + 2B_2 && (\text{mod } 3) \\
 D &= A_1 + B_1 + C_1 + C_2 && (\text{mod } 2)
 \end{aligned}
 \tag{5}$$

2.3.2 Etude des alias

Pour réaliser l'étude des alias de ce plan précis, les égalités de définition (5) sont introduites dans le cadre “*facteurs prédéfinis*” de l'écran 2 (cadre figurant à droite et en bas dans l'exemple de la figure 3) sous la forme:

$$\begin{aligned}
 Bl_{-1} &: A_{-1} + B_{-1} + C_{-1} \\
 Bl_{-2} &: A_{-2} + 2B_{-2} \\
 D &: A_{-1} + B_{-1} + C_{-1} + C_{-2}
 \end{aligned}$$

On spécifie dans le cadre *facteurs définis* de ce même écran que Bl est un facteur bloc.

On suppose que le modèle comprend un effet bloc et un effet traitement incluant les effets principaux et interactions de deux facteurs. Compte tenu du fait qu'il est automatiquement complété par PLANOR, ce modèle, peut s'écrire:

$$Bl + A.B + A.C + A.D + B.C + B.D + C.D ,$$

ou encore en employant une partie de modèle

$$\begin{array}{ll}
 \text{partie de modèle} & P : A + B + C + D \\
 \text{modèle} & Bl + P.P
 \end{array}$$

Partant de ce modèle introduit dans l'écran 6 (cf § 5.7 et figure 8), l'étude des alias donne les résultats qui suivent. Dans la version actuelle, ces résultats sont donnés séparément pour les premiers 2 et 3 qui divisent le nombre d'unités.

Premier 2

- Liste des effets traitement confondus avec la moyenne

$$; A_1 B_1 C_1 C_2 D;$$

- Ensembles d'effets confondus du modèle. Si l'ensemble contient un effet bloc, celui ci est indiqué entre crochets.

$$\begin{aligned} & [Bl_1]; C_2 D; \\ & A_1 B_1; C_1 C_2 D; \\ & A_1 D; B_1 C_1 C_2; \\ & B_1 D; A_1 C_1 C_2; \end{aligned}$$

- liste des effets traitement non confondus

$$A_1; B_1 C_1; B_1; A_1 C_1; C_1; C_2; D; A_1 C_2; B_1 C_2; C_1 D; C_1 C_2;$$

- liste des effets blocs non confondus: vide.

Premier 3

- Liste des effets traitement confondus avec la moyenne: vide.
- Ensembles d'effets confondus du modèle. Si l'ensemble contient un effet bloc, celui ci est indiqué entre crochets.

$$\begin{aligned} & [Bl_2^2]; A_2^2 B_2; \\ & [Bl_2]; A_2 B_2^2; \end{aligned}$$

- liste des effets traitement non confondus

$$; A_2; B_2; A_2^2 B_2^2; A_2^2; A_2 B_2; B_2^2;$$

- liste des effets blocs non confondus: vide

On peut facilement déduire de ces études menées séparément pour les premiers 2 et 3 les effets confondus dans le plan global. La règle générale est simple. Tout effet se décompose sous forme d'un produit $\alpha_2 \alpha_3$ où α_2 et α_3 sont exprimables à partir des pseudofacteurs à 2 et 3 niveaux respectivement. Par exemple $\alpha_2 = A_1 D$ et $\alpha_3 = A_2 B_2^2$ pour l'effet $A_1 A_2 B_2^2 D$. Soient alors $\alpha_2 \alpha_3$ et $\beta_2 \beta_3$ deux effets ainsi décomposés.

Règle 2 $\alpha_2 \alpha_3$ et $\beta_2 \beta_3$ sont confondus si et seulement si α_2 est confondu avec β_2 et α_3 est confondu avec β_3 .

En particulier, si ni α_2 , ni α_3 ne sont confondus, l'effet $\alpha_2\alpha_3$ n'est pas confondu. Dans cet exemple par exemple A_1 , A_2 et A_2^2 ne sont pas confondus, et les produits A_1A_2 , $A_1A_2^2$ ne le sont donc pas non plus. Les 5 degrés de liberté de A sont donc estimables. On voit de même que les autres effets principaux sont estimables dans le modèle qui comprend toutes les interactions de 2 facteurs traitement et l'effet bloc.

Examinons maintenant une interaction telle que AB . Les 25 effets correspondants apparaissent tous dans le tableau 7 avec, dans la même case, les autres effets avec lesquels ils sont confondus. Trois degrés de liberté seulement sont confondus avec des interactions non négligeables: A_1B_1 confondu avec C_1C_2D , $A_2B_2^2$, $A_2^2B_2$ confondus avec Bl_2 , Bl_2^2 .

	A_1	B_1	A_1B_1	C_1C_2D
A_2	A_1A_2	B_1A_2	$A_1B_1A_2$	$A_2C_1C_2D$
A_2^2	$A_1A_2^2$	$B_1A_2^2$	$A_1B_1A_2^2$	$A_2^2C_1C_2D$
B_2	A_1B_2	B_1B_2	$A_1B_1B_2$	$B_2C_1C_2D$
B_2^2	$A_1B_2^2$	$B_1B_2^2$	$A_1B_1B_2^2$	$B_2^2C_1C_2D$
A_2B_2	$A_1A_2B_2$	$B_1A_2B_2$	$A_1B_1A_2B_2$	$A_2B_2C_1C_2D$
$A_2^2B_2^2$	$A_1A_2^2B_2^2$	$B_1A_2^2B_2^2$	$A_1B_1A_2^2B_2^2$	$A_2^2B_2^2C_1C_2D$
$A_2B_2^2$	$A_1A_2B_2^2$	$B_1A_2B_2^2$	$A_1B_1A_2B_2^2$	$A_2B_2^2C_1C_2D$
Bl_2	A_1Bl_2	B_1Bl_2	$A_1B_1Bl_2$	$C_1C_2DBl_2$
$A_2^2B_2$	$A_1A_2^2B_2$	$B_1A_2^2B_2$	$A_1B_1A_2^2B_2$	$A_2^2B_2C_1C_2D$
Bl_2^2	$A_1Bl_2^2$	$B_1Bl_2^2$	$A_1B_1Bl_2^2$	$C_1C_2DBl_2^2$

TAB. 7 – Confusion pour des effets faisant intervenir des pseudofacteurs à 2 et 3 niveaux

3 Exemples de plans pour le robot ARILAIT

Pour mesurer l'efficacité du nettoyage et de la désinfection de surfaces ouvertes, le laboratoire LGHPA et l'association ARILAIT, déjà mentionnés dans l'introduction, ont mis au point un robot pouvant salir, puis nettoyer et désinfecter une surface dans des conditions reproductibles [5]. La surface est une plaque en acier inoxydable comportant 16 éprouvettes circulaires de 5 cm de diamètre, disposées sur la plaque comme indiqué en figure 1.

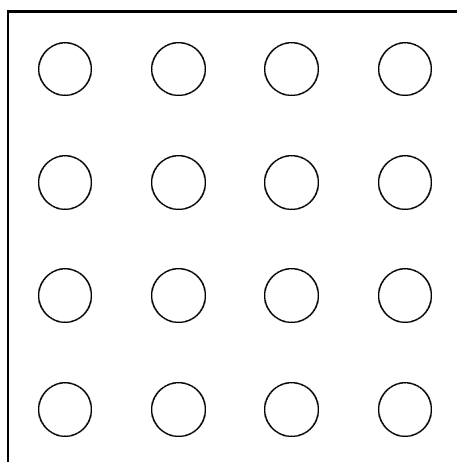


FIG. 1 – Position des éprouvettes sur la plaque

Les facteurs susceptibles d'influencer l'efficacité du nettoyage et de la désinfection sont multiples. Ce sont par exemple la nature du matériel constituant l'éprouvette, sa rugosité, le type et la quantité de souillure, la concentration des bactéries dans cette souillure, la nature, la concentration, la température d'application, le temps d'action de chacun des produits utilisés pour le nettoyage et la désinfection.

Le maniement du robot implique certaines contraintes. Il se déplace par ligne et par colonne et les niveaux des facteurs ne peuvent varier ad libitum d'une éprouvette à l'autre. D'autre part, l'efficacité du nettoyage et de la désinfection est évaluée en comparant l'éprouvette nettoyée et désinfectée à une éprouvette témoin encrassée de la même façon mais non nettoyée. Pour chaque traitement, on doit donc avoir, au moins à l'encrassement, deux éprouvettes traitées de façon identique.

3.1 Création d'un plan factoriel complet intégrant les contraintes de manipulation

3.1.1 Objectif et contraintes

Dans l'exemple de ce paragraphe, un unique produit commercial, alcalin chloré, est appliqué sous forme de mousse pour nettoyer et désinfecter. Les facteurs étudiés sont

n-sou :	la nature de la souillure	caillé Saint-Paulin.
q-sou :	la quantité de souillure déposée sur les éprouvettes, variée en ajoutant un poids sur le bras du robot	0,01 g/éprouvette 0,10 g/éprouvette.
rug :	la rugosité des éprouvettes	0,25 μ m 0,73 μ m.
conc :	la concentration du produit de nettoyage et de désinfection	1% 3% (v/v)
T-act :	le temps d'action du produit.....	15mn 30mn.

Il n'y a pas dans cette expérience de cloison étanche permettant d'éviter la diffusion de la mousse d'une partie de la plaque à l'autre. Les deux derniers facteurs, concentration et temps d'action du produit, restent donc obligatoirement constants au cours d'une même manipulation de la plaque. Pour étudier les 4 combinaisons (1%, 15mn), (1%, 30mn), (3%, 15mn), (3%, 30mn) de ces deux facteurs, il faut au minimum 4 telles manipulations, soit 4 plaques correspondant à 4 semaines d'expérimentation. Pour étudier la variabilité d'une plaque à l'autre, on répète deux de ces combinaisons (3%,15mn), (1%, 30mn), ce qui conduit à un total de 6 plaques.

Sur chacune des 6 plaques, on peut expérimenter les 8 combinaisons de niveaux des trois autres facteurs: (caillé, 0.01g/e, 0.25 μ m), (caillé, 0.10g/e, 0.25 μ m), (caillé, 0.01g/e, 0.73 μ m), (caillé, 0.10g/e, 0.73 μ m), (StPaul, 0.01g/e, 0.25 μ m), (StPaul, 0.10g/e, 0.25 μ m), (StPaul, 0.01g/e, 0.73 μ m), (StPaul, 0.10g/e, 0.73 μ m), et disposer pour chacun de ces traitements d'une éprouvette traitée et du témoin associé.

Pour simplifier les manipulations, l'encrassement est effectué colonne par colonne et la nature et la quantité de souillure ne sont modifiées que lors des changements de colonne. On dispose dans chaque colonne deux éprouvettes de chaque rugosité, l'une servant de témoin pour l'autre. Au nettoyage, les éprouvettes traitées sont remises aux mêmes emplacements et les éprouvettes témoins sont remplacées par d'autres éprouvettes qui ne servent qu'à boucher les trous.

Le plan est donc dans cet exemple clairement défini et la recherche algorithmique inutile. Il est cependant intéressant d'examiner comment procéder pour créer pratiquement ce plan et le randomiser.

3.1.2 Création par juxtaposition de deux sous-plans

Comme il a été indiqué, les 6 plaques sont formées d'un plan de 4 plaques comprenant l'ensemble des 4 \times 8 traitements et d'un second plan de 2 plaques comprenant une moitié de ces traitements.

Ces deux plans, notés respectivement ROBOT1A et ROBOT1B sont obtenus séparément puis fusionnés. Pour chacun d'eux, les facteurs de base sont les numéros de plaque *pl*, de colonne dans la plaque *col*, d'unité dans la colonne *u*. Les facteurs définis sont les facteurs

traitement : $n\text{-sou}$, $q\text{-sou}$, Rug , $conc$, $T\text{-act}$. Les figures 2 et 3 donnent les deux écrans permettant de définir la recherche pour ROBOT1A. Le contenu de ces deux écrans est détaillé au § 5.4.

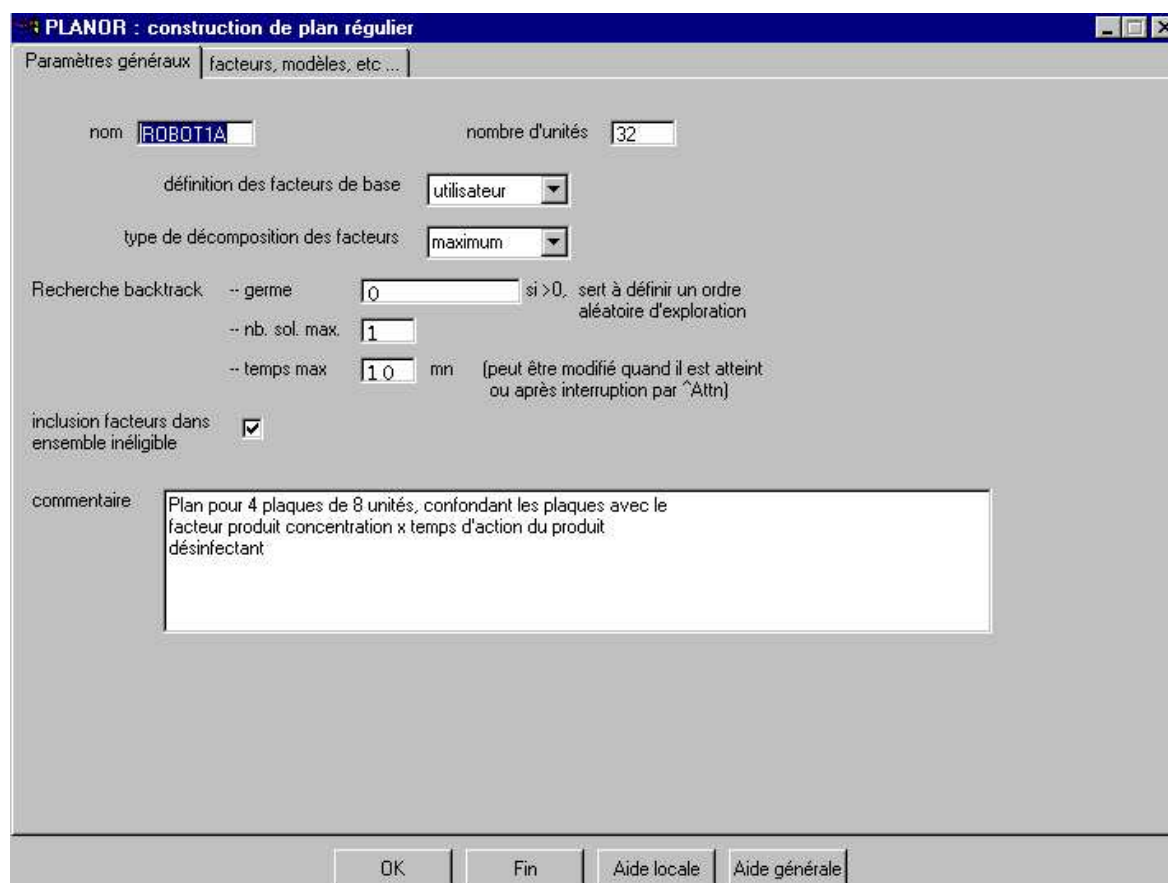


FIG. 2 – Plan ROBOT1A, écran 1

La zone des hiérarchies, dans l'écran 2 (fig. 3), précise que $conc$, $T\text{-act}$ sont définis à partir du numéro de plaque pl , et $n\text{-sou}$, $q\text{-sou}$ à partir du numéro de plaque pl et du numéro de colonne dans la plaque col . Ainsi $conc$ et $T\text{-act}$ sont obligatoirement constants sur chaque plaque et $n\text{-sou}$ et $q\text{-sou}$ constants sur chaque colonne du plan trouvé.

Le plan ROBOT1A doit être complet, c'est à dire comporter chacun des 2^5 traitements définis par $n\text{-sou}$, $q\text{-sou}$, Rug , $conc$, $T\text{-act}$. Pour être certain qu'il en soit ainsi, on introduit un modèle comprenant l'interaction des 5 facteurs dans la zone modèle de l'écran 2. Le programme complète automatiquement ce modèle en introduisant tous les termes inclus dans cette interaction, de la constante aux interactions entre quatre de ces facteurs.

La partie à estimer associée à ce modèle est laissée vide. Le programme inclut alors automatiquement la moyenne générale dans cette partie. Cette dernière ne peut donc être confondue avec aucun des effets figurant dans le modèle et cette condition implique que les 2^5 traitements sont effectivement présents dans le plan.

Le fichier ROBOT1A.REG peut être utilisé comme point de départ pour construire le plan ROBOT1B. Dans l'écran 1, on remplace le nom par ROBOT1B, le nombre d'unités

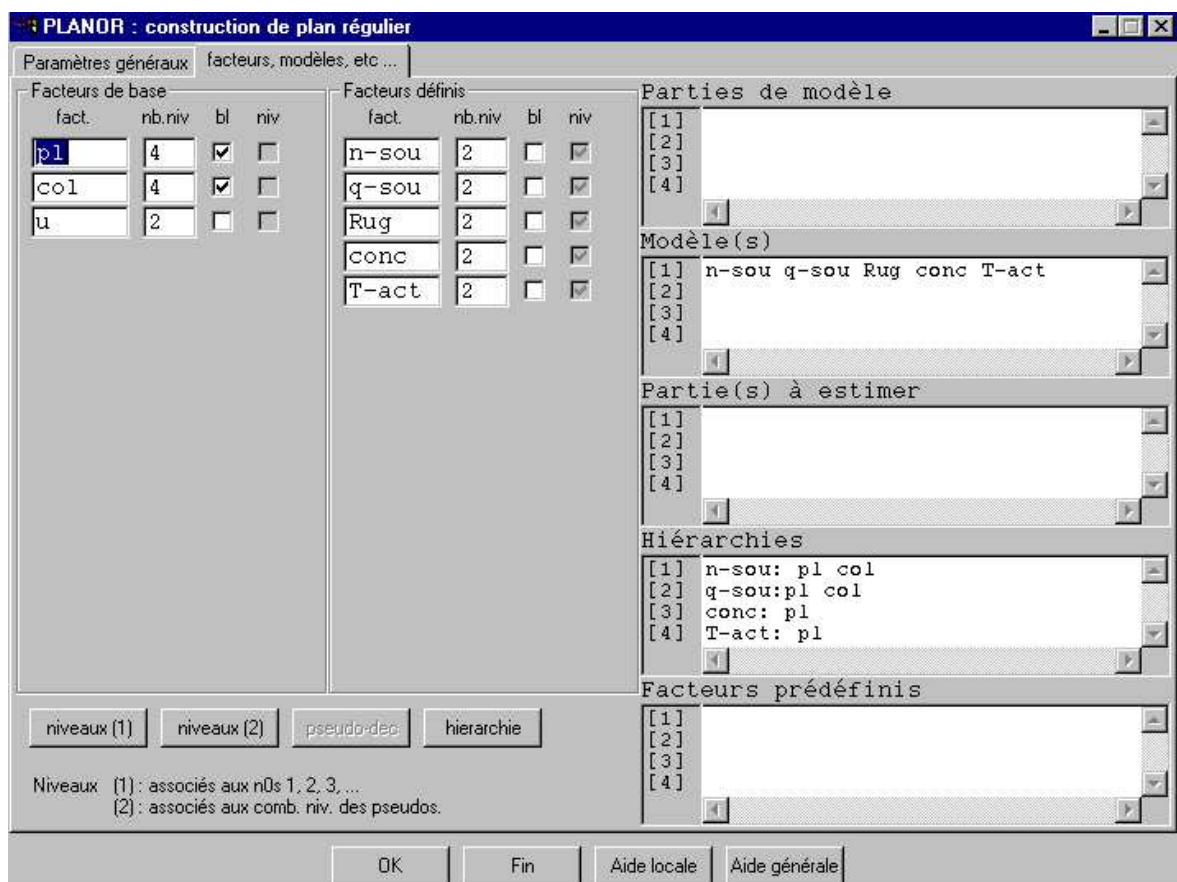


FIG. 3 – Plan ROBOT1A, écran 2

par 16. Dans l'écran 2, on change en 2 le nombre de niveaux de pl et on supprime par exemple $T-act$ du modèle. Le plan obtenu comprend alors toutes les combinaisons de niveaux des quatre facteurs restant dans le modèle.

Mais $T-act$ ne figure pas dans le modèle et rien n'interdit qu'il ne reste constant sur ce second plan. Pour l'obliger à prendre ses deux niveaux, on peut constituer un second modèle contenant seulement $T-act$, associé à une partie à estimer vide. Cette dernière précaution est toutefois inutile si les facteurs, $T-act$ en particulier, sont déclarés d'office *inéligibles* dans le champs correspondant de l'écran 1 (fig. 2, au dessus de la zone commentaire), ce qui signifie qu'aucun des effets principaux ne peut être confondu avec la moyenne générale, ou de façon équivalente que chacun des facteurs prend l'ensemble de ses niveaux sur le plan.

Les résultats de la recherche de PLANOR sont enregistrés dans un fichier ASCII que l'utilisateur peut ensuite imprimer ou consulter par n'importe quel éditeur. Le nom proposé pour ce fichier en standard est le nom de l'expérience suivi du suffixe OUT. Après les étapes de création des deux plans on obtient donc des fichiers ROBOT1A.OUT et ROBOT1B.OUT qui contiennent les valeurs des paramètres déterminant la recherche et les matrices clés. Ces matrices clés sont données dans le tableau 8.

		Plan ROBOT1A								
		2	2	2	2	2	2	2	2	2
		pl ₁	pl ₂	col ₁	col ₂	n-sou	q-sou	Rug	conc	T-act
2	pl ₁	1	0	0	0	0	0	0	1	0
2	pl ₂	0	1	0	0	0	0	0	0	1
2	col ₁	0	0	1	0	1	0	0	0	0
2	col ₂	0	0	0	1	0	1	0	0	0
2	u	0	0	0	0	0	0	1	0	0

		Plan ROBOT1B								
		2	2	2	2	2	2	2	2	2
		pl	col ₁	col ₂	n-sou	q-sou	Rug	conc	T-act	
2	pl	1	0	0	0	0	0	1	1	
2	col ₁	0	1	0	1	0	0	0	0	
2	col ₂	0	0	1	0	1	0	0	0	
2	u	0	0	0	0	0	1	0	0	

TAB. 8 – Matrices clé des plans ROBOT1A et ROBOT1B

Ces matrices conduisent à prendre ici $n-sou=col_1$, $q-sou=col_2$, $Rug=u$, $conc=pl_1$, $T-act=pl_2$ pour construire ROBOT1A et $n-sou=col_1$, $q-sou=col_2$, $Rug=u$, $conc=pl$, $T-act=pl$ pour ROBOT2. col_1 , col_2 , notés col_1 , col_2 dans le programme, sont les pseudofacteurs résultant de la décomposition du facteur col à 4 niveaux, et de même pl_1 , pl_2 , notés pl_1 , pl_2 , ceux qui résultent de la décomposition du facteur pl dans le plan ROBOT1A.

Il apparait dans la figure 2, champs *nb. sol. max.*, qu'une seule solution est demandée. Dans ce cas après obtention de l'unique matrice clé, le programme se poursuit automatiquement par la construction du plan qui est stocké dans le fichier PS associé, ROBOT1A.PS ou ROBOT1B.PS. Le contenu de ce fichier peut être ensuite édité ou manipulé en utilisant l'option *recodage, sélection de facteurs, tri, ...* du menu general

(tab. 48). Cette option fait apparaître le menu situé à gauche de l'écran 4 (fig. 4)

La sous-option *écrit n0* permet en particulier d'éditer ces plans, tels qu'ils apparaissent dans le tableau 9. Les niveaux sont dans ce tableau donnés par leur numéro, qui commence à 0. Il est possible d'obtenir une édition plus explicite par la sous option *écriture* qui remplace, pour les facteurs préalablement recodés par l'utilisateur, les numéros par les niveaux introduits. Un exemple d'édition de ce type est donné dans le tableau 11.

L'obtention du fichier ROBOT1A.PS est suivie dans ce cas de l'appel du module de randomisation. Pour éviter cette randomisation, ici inutile puisque ROBOT1A n'est qu'une partie du plan, ce module est interrompu par un *esc*.

3.1.3 Fusion des deux plans

En l'absence de recodage, les niveaux donnés à un facteur qui a p niveaux sont assimilés à leurs numéros $0, \dots, p - 1$. Ainsi, si le facteur pl n'a pas été recodé, ses niveaux sont 0, 1, 2, 3 dans le plan ROBOT1A et 0, 1 dans le plan ROBOT1B. La fusion conduit alors à un plan inadapté: le facteur pl a 4 niveaux au lieu de 6 et il y a 16 unités au lieu de 8 pour chacune des plaques 0 et 1.

Il est donc indispensable avant d'opérer la fusion de recoder les niveaux du facteur pl , ce qui peut se faire de deux façons:

- lors de la création du plan, en positionnant le curseur sur le facteur à recoder et en cliquant sur l'un des deux boutons "niveaux (1)", "niveaux (2)" (fig. 3)
- en modifiant un fichier PS déjà créé par l'option *recodage, sélection de facteurs, tri, ...* du menu général et la sous option *nouv. fact..* On indique que l'on veut redéfinir pl en tapant $pl: pl$ dans la fenêtre ad hoc, puis on clique sur le bouton "niveaux" (fig. 4).

Après recodage, le choix *fusion* dans l'option *recodage, sélection de facteurs, tri, ...* permet de fusionner les fichiers. Dans le fichier résultant, appelé ici ROBOT1.PS (tab. 9), ne sont retenus que les facteurs communs aux deux fichiers. Les pseudofacteurs pl_1, pl_2 présents seulement dans ROBOT1A sont donc d'office éliminés. Les niveaux retenus après fusion sont ceux qui figurent dans l'un ou l'autre des fichiers. Ils sont renumérotés de 0 au nombre total de niveaux.

3.1.4 Randomisation prenant en compte la structure des blocs

Pour être effectivement mis en place, ce plan ROBOT1 doit être randomisé par l'option *Randomisation* du menu général (tab. 48). Cette randomisation doit permuter les plaques entre elles pour que concentration et temps d'action soient toujours constants par plaque après la randomisation. De même elle doit permuter les colonnes de chaque plaque entre elles. Cela est indiqué au programme par le modèle de randomisation qui comprend les termes pl et $pl.col$ (fig.5). La permutation aléatoire des 6 plaques et les 6 permutations aléatoires des 4 colonnes de chaque plaque sont complétées par les 24 permutations des 2 unités de chaque colonne. En sortie, les niveaux tirés au hasard par

Plan ROBOT1B											Plan ROBOT1									
	$\frac{p}{l}$	$\frac{p}{l}$	$\frac{c}{o}$	$\frac{c}{o}$		$\frac{n}{s}$	$\frac{q}{s}$		$\frac{c}{o}$	$\frac{T}{a}$		$\frac{c}{o}$	$\frac{n}{c}$	$\frac{T}{a}$						
	1	2	1	2	$\frac{p}{l}$	$\frac{c}{o}$	$\frac{u}{o}$	$\frac{u}{o}$	$\frac{R}{u}$	$\frac{g}{c}$		$\frac{p}{l}$	$\frac{c}{o}$	$\frac{s}{o}$	$\frac{q}{o}$	$\frac{R}{u}$	$\frac{c}{o}$	$\frac{n}{c}$	$\frac{T}{a}$	
1			0	0	*	*	0	0	0	0										
2			0	0	0	0	0	0	0	0		*	*							
3			1	0	0	2	1	0	0	0		1	0	0	0	0	0	0	0	0
4			1	0	1	2	1	0	0	1		2	1	0	0	0	1	1	1	1
5			0	1	0	1	0	1	0	0		1	0	1	0	0	0	0	0	0
6			0	1	1	1	0	1	0	1		1	1	1	0	0	1	1	1	1
7			1	1	0	3	1	1	0	0		0	1	0	1	0	0	0	0	0
8			1	1	1	3	1	1	0	1		1	1	0	1	0	1	1	1	1
9			0	0	0	0	0	0	0	1		0	1	0	0	0	0	0	0	0
10			0	0	1	0	0	0	0	1		1	1	1	0	0	1	1	1	1
11			1	0	0	2	1	0	1	0		0	1	0	0	1	0	0	0	0
12			1	0	1	2	1	0	1	1		1	1	1	1	1	1	1	1	1
13			0	1	0	1	0	1	1	0		0	1	0	1	0	1	0	0	0
14			0	1	1	1	0	1	1	1		1	1	1	1	1	1	1	1	1
15			1	1	0	3	1	1	1	0		0	0	0	1	1	0	0	0	0
16			1	1	1	3	1	1	1	1		1	1	1	1	1	1	1	1	1
17			0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
18			0	0	0	1	0	0	0	0		0	0	0	0	0	0	0	0	0
19			0	0	0	2	1	0	0	0		0	0	0	0	0	0	0	0	0
20			0	0	1	0	0	0	0	0		0	0	0	0	0	0	0	0	0
21			0	0	1	0	2	1	0	0		0	1	0	0	0	0	0	0	0
22			0	1	0	1	2	1	0	0		0	0	0	1	0	0	0	0	0
23			0	1	1	0	1	2	1	0		0	0	0	1	0	0	0	0	0
24			1	1	1	0	3	2	1	0		1	1	1	1	1	1	1	1	1
25			0	0	0	1	0	1	0	1		1	1	0	0	0	0	0	0	0
26			1	0	0	1	2	1	0	1		1	1	1	1	1	1	1	1	1
27			0	1	0	1	1	0	1	1		0	1	0	1	0	1	0	0	0
28			1	1	0	1	3	1	0	1		1	1	1	1	1	1	1	1	1
29			0	0	1	1	0	3	1	1		1	1	0	0	0	0	0	0	0
30			1	0	1	1	2	3	1	1		1	1	1	1	1	1	1	1	1
31			0	1	1	1	1	3	1	1		1	1	1	0	0	0	0	0	0
32			1	1	1	1	3	3	1	1		1	1	1	1	1	1	1	1	1
33			0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
34			0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
35			0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0	0
36			0	0	0	2	0	0	0	1		0	0	0	0	0	0	0	0	0
37			2	2	1	0	0	0	0	1		0	0	0	0	0	0	0	0	0
38			4	2	1	0	1	1	1	1		0	1	0	1	1	1	1	1	1
39			3	2	1	0	2	1	0	0		0	1	0	1	0	1	0	1	0
40			5	2	1	0	2	1	0	1		0	1	0	1	1	1	1	1	1
41			2	1	0	1	2	1	0	1		0	1	0	1	1	0	0	0	0
42			4	1	0	3	2	1	0	1		1	1	1	1	1	1	1	1	0
43			3	1	0	1	0	1	1	0		0	1	1	0	0	0	0	0	0
44			5	1	0	1	2	1	0	1		1	1	1	1	1	1	1	1	1
45			2	3	1	1	0	1	1	0		1	1	1	0	0	0	0	0	0
46			4	3	1	1	1	0	1	1		1	1	1	1	1	1	1	1	0
47			3	3	1	1	1	3	1	1		1	1	1	1	0	0	0	0	0
48			5	3	1	1	1	3	1	1		1	1	1	1	1	1	1	1	1

TAB. 9 – Plans ROBOT1, 1A, 1B avant randomisation (fichiers .PS)

1. Dans cette édition des plans, obtenue par l'option *écrit n0*, les niveaux sont donnés par leur numéros.
2. les étoiles au dessous de *pl* et *col* indiquent que ces facteurs ont été définis comme des facteurs blocs.

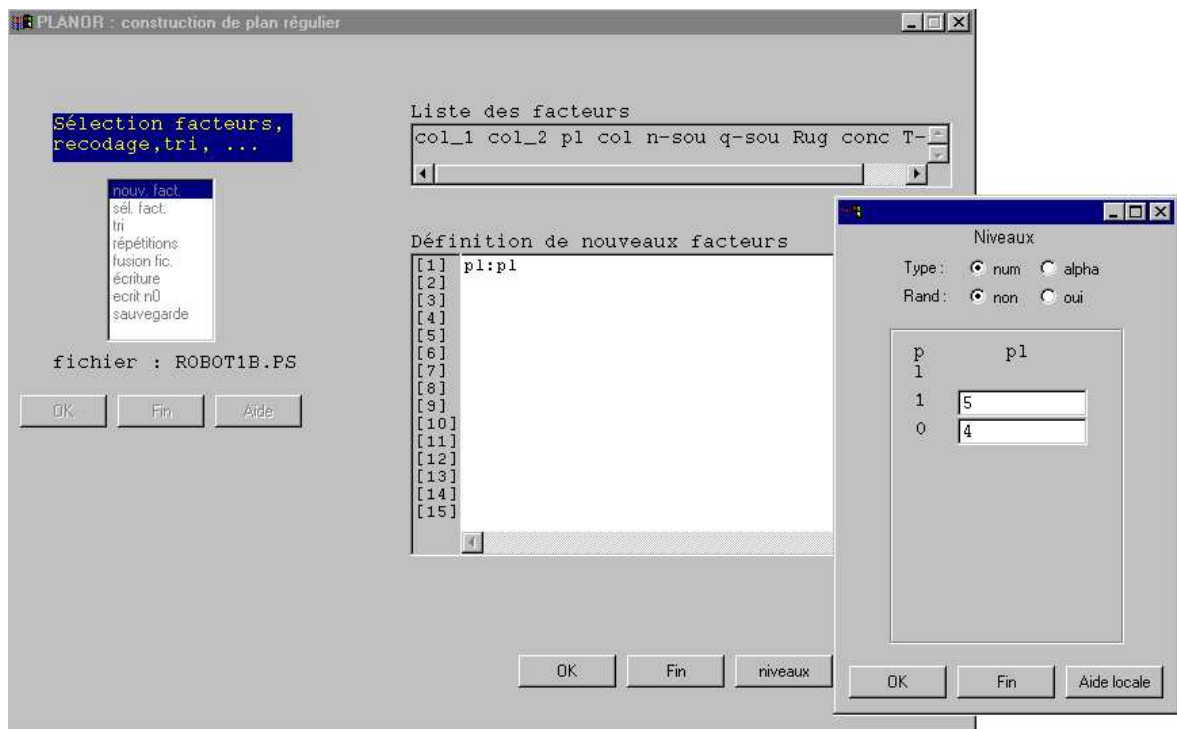


FIG. 4 – Plan ROBOT1B, Redéfinition des niveaux de *p1*. Ecran 4

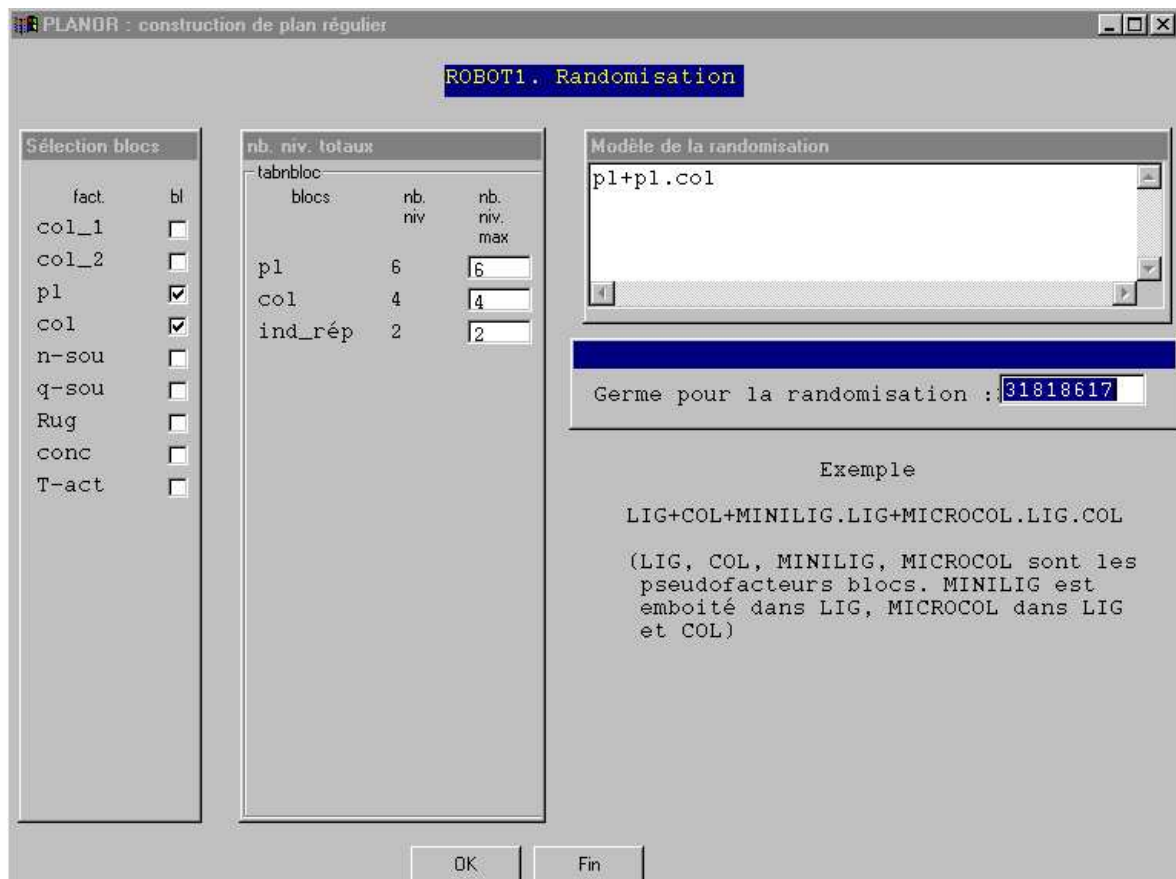


FIG. 5 – Randomisation du plan ROBOT1, écran 3

la randomisation viennent remplacer les niveaux du plan systématique dans les colonnes *pl* et *col*. Une colonne supplémentaire *ind-rep*, l'indice de répétition, permet de repérer l'unité dans chaque colonne. Les traitements gardent donc le même ordre que dans le plan systématique mais les niveaux des blocs qui leur sont associés font référence aux unités vraies tirées par la randomisation. Le plan randomisé est stocké dans le fichier ROBOT1.PR et est également édité dans un fichier ASCII, de nom ROBOT1.OUT en standard.

Les unités du plan randomisé peuvent être réordonnées d'une façon adaptée à la manipulation par le robot en utilisant le *tri* de l'option *recodage, sélection de facteurs, tri, ...* (fig.4), avec pour clés de tri d'abord les plaques, ensuite les colonnes, enfin l'indice de répétition. L'ordre utilisé dans ce tri pour les niveaux d'un facteur est l'ordre croissant naturel, alphabétique ou numérique selon le type de ces niveaux (dans le cas alphabétique, les lettres majuscules sont classées avant toutes les lettres minuscules).

Les tableaux 10 et 11 donne le fichier randomisé avant et après tri, édité avec les numéros de niveaux d'une part (*écrit n0*), les niveaux introduits par l'utilisateur d'autre part (*écriture*). Noter que le tri est opéré sur les niveaux introduits par l'utilisateur, non sur les numéros de ces niveaux.

Remarques:

- Le fait que *col* n'apparaisse dans le modèle de randomisation que lié à *pl* indique une relation de hiérarchie et implique que les permutations aléatoires des numéros de colonne se font indépendamment d'une plaque à l'autre. De même l'indice de répétition, considéré comme hiérarchisé par tous les autres facteurs, est randomisé séparément dans chaque colonne de chaque plaque. Un choix de modèle différent pourrait être fait si la plaque se trouvait en position verticale, avec par exemple la colonne 1 en haut et la colonne 4 en bas. Il faudrait alors tenir compte, aussi bien dans la construction du plan que dans sa randomisation, d'un éventuel effet colonne traduisant des différences dans l'encrassement ou le nettoyage en fonction de la hauteur. Le facteur colonne serait alors croisé avec le facteur plaque et non plus hiérarchisé par lui.
- Le type de randomisation effectuée, étudiée de façon approfondie dans [3], conduit normalement à un modèle pour l'analyse où apparaissent un certain nombre de termes, dits *termes ancestraux*, qui sont déduits de façon automatique du modèle de la randomisation. Dans cet exemple, ces termes, qui figurent dans le fichier ASCII créé par la randomisation, sont *pl, pl.col, pl.col.ind-rép*. Les effets associés dans le modèle d'analyse de variance sont aléatoires et l'analyse est normalement effectuée par une procédure prenant en compte ces effets aléatoires (cf [2], [18]).
- L'introduction des facteurs *pl* et *col* est rendue ici indispensable par les contraintes de manipulation du robot. Cependant, l'effet plaque est apparu négligeable dans les premiers essais, sans doute à cause du couplage de chaque éprouvette traitée avec une éprouvette témoin et il est naturel de penser que l'effet colonne, sur la plaque maintenue horizontale, est encore plus négligeable. En d'autre terme, l'excellente reproductibilité des opérations effectuées par le robot conduit légitimement à supposer que la variabilité entre deux unités de colonnes différentes est du même ordre

Avant tri									Après tri									
			<i>n</i>	<i>q</i>		<i>c</i>	<i>T</i>	<i>i</i>				<i>n</i>	<i>q</i>		<i>c</i>	<i>T</i>	<i>i</i>	
			—	—		—	—	<i>n</i>				—	—		—	—	<i>n</i>	
	<i>c</i>	<i>s</i>	<i>s</i>	<i>R</i>	<i>o</i>	<i>u</i>	<i>a</i>	<i>r</i>				<i>c</i>	<i>s</i>	<i>s</i>	<i>R</i>	<i>o</i>	<i>a</i>	
	<i>p</i>	<i>o</i>	<i>o</i>	<i>o</i>	<i>u</i>	<i>n</i>	<i>c</i>	<i>e</i>				<i>p</i>	<i>o</i>	<i>o</i>	<i>o</i>	<i>u</i>	<i>n</i>	
	<i>l</i>	<i>l</i>	<i>u</i>	<i>u</i>	<i>g</i>	<i>c</i>	<i>t</i>	<i>p</i>				<i>l</i>	<i>l</i>	<i>u</i>	<i>u</i>	<i>g</i>	<i>c</i>	
	*	*						*				*	*					*
1	0	2	0	0	0	0	0	1				1	2	0	1	1	0	0
2	3	0	0	0	0	1	1	0				2	2	0	1	1	0	0
3	0	0	1	0	0	0	0	0				3	2	1	0	0	1	0
4	3	2	1	0	0	1	1	1				4	2	1	0	0	0	0
5	0	3	0	1	0	0	0	1				5	2	2	1	0	0	0
6	3	3	0	1	0	1	1	0				6	2	2	1	0	1	0
7	0	1	1	1	0	0	0	1				7	2	3	0	1	0	0
8	3	1	1	1	0	1	1	0				8	2	3	0	1	1	0
9	0	2	0	0	1	0	0	0				9	3	0	0	0	0	1
10	3	0	0	0	1	1	1	1				10	3	0	0	0	1	1
11	0	0	1	0	1	0	0	1				11	3	1	1	1	0	1
12	3	2	1	0	1	1	1	0				12	3	1	1	1	1	1
13	0	3	0	1	1	0	0	0				13	3	2	1	0	1	1
14	3	3	0	1	1	1	1	1				14	3	2	1	0	0	1
15	0	1	1	1	1	0	0	0				15	3	3	0	1	0	1
16	3	1	1	1	1	1	1	1				16	3	3	0	1	1	1
17	2	1	0	0	0	0	0	1				17	4	0	1	1	0	0
18	5	2	0	0	0	1	0	1				18	4	0	1	1	1	0
19	4	1	0	0	0	0	1	1				19	4	1	0	0	1	0
20	1	2	0	0	0	1	1	1				20	4	1	0	0	0	1
21	2	2	1	0	0	0	0	0				21	4	2	1	0	1	0
22	5	3	1	0	0	1	0	0				22	4	2	1	0	0	1
23	4	2	1	0	0	0	1	1				23	4	3	0	1	1	0
24	1	1	1	0	0	1	1	0				24	4	3	0	1	0	0
25	2	3	0	1	0	0	0	0				25	5	0	0	1	0	1
26	5	0	0	1	0	1	0	0				26	5	0	0	1	1	0
27	4	3	0	1	0	0	1	1				27	5	1	1	1	0	1
28	1	3	0	1	0	1	1	1				28	5	1	1	1	1	0
29	2	0	1	1	0	0	0	1				29	5	2	0	0	1	1
30	5	1	1	1	0	1	0	0				30	5	2	0	0	0	1
31	4	0	1	1	0	0	1	0				31	5	3	1	0	0	1
32	1	0	1	1	0	1	1	0				32	5	3	1	0	1	0
33	2	1	0	0	1	0	0	0				33	0	0	1	0	0	0
34	5	2	0	0	1	1	0	0				34	0	0	1	0	1	0
35	4	1	0	0	1	0	1	0				35	0	1	1	1	0	0
36	1	2	0	0	1	1	1	0				36	0	1	1	1	0	0
37	2	2	1	0	1	0	0	1				37	0	2	0	0	1	0
38	5	3	1	0	1	1	0	1				38	0	2	0	0	0	0
39	4	2	1	0	1	0	1	0				39	0	3	0	1	1	0
40	1	1	1	0	1	1	1	1				40	0	3	0	1	0	0
41	2	3	0	1	1	0	0	1				41	1	0	1	1	0	1
42	5	0	0	1	1	1	0	1				42	1	0	1	1	1	1
43	4	3	0	1	1	0	1	0				43	1	1	1	0	0	1
44	1	3	0	1	1	1	1	0				44	1	1	1	0	1	1
45	2	0	1	1	1	0	0	0				45	1	2	0	0	1	1
46	5	1	1	1	1	1	0	1				46	1	2	0	0	0	1
47	4	0	1	1	1	0	1	1				47	1	3	0	1	1	1
48	1	0	1	1	1	1	1	1				48	1	3	0	1	0	1

TAB. 10 – Plan ROBOT1 randomisé, avant et après tri, édité par *écrit n0*

Avant tri								Après tri									
pl	col	n-sou	q-sou	Rug	conc	T-act	ind- rep	pl	col	n-sou	q-sou	Rug	conc	T-act	ind- rep		
6	4	2	2	2	2	2	2	6	4	2	2	2	2	2	2		
1	4	2	caillé	10mg	0.25	1%	15mn	1	1	0	0	St-Paul	100mg	0.73	1%	15mn	0
2	1	0	caillé	10mg	0.25	3%	30mn	0	2	0	0	St-Paul	100mg	0.25	1%	15mn	1
3	4	0	St-Paul	10mg	0.25	1%	15mn	0	3	0	1	caillé	10mg	0.73	1%	15mn	0
4	1	2	St-Paul	10mg	0.25	3%	30mn	1	4	0	1	caillé	10mg	0.25	1%	15mn	1
5	4	3	caillé	100mg	0.25	1%	15mn	1	5	0	2	St-Paul	10mg	0.25	1%	15mn	0
6	1	3	caillé	100mg	0.25	3%	30mn	0	6	0	2	St-Paul	10mg	0.73	1%	15mn	1
7	4	1	St-Paul	100mg	0.25	1%	15mn	1	7	0	3	caillé	100mg	0.25	1%	15mn	0
8	1	1	St-Paul	100mg	0.25	3%	30mn	0	8	0	3	caillé	100mg	0.73	1%	15mn	1
9	4	2	caillé	10mg	0.73	1%	15mn	0	9	1	0	caillé	10mg	0.25	3%	30mn	0
10	1	0	caillé	10mg	0.73	3%	30mn	1	10	1	0	caillé	10mg	0.73	3%	30mn	1
11	4	0	St-Paul	10mg	0.73	1%	15mn	1	11	1	1	St-Paul	100mg	0.25	3%	30mn	0
12	1	2	St-Paul	10mg	0.73	3%	30mn	0	12	1	1	St-Paul	100mg	0.73	3%	30mn	1
13	4	3	caillé	100mg	0.73	1%	15mn	0	13	1	2	St-Paul	10mg	0.73	3%	30mn	0
14	1	3	caillé	100mg	0.73	3%	30mn	1	14	1	2	St-Paul	10mg	0.25	3%	30mn	1
15	4	1	St-Paul	100mg	0.73	1%	15mn	0	15	1	3	caillé	100mg	0.25	3%	30mn	0
16	1	1	St-Paul	100mg	0.73	3%	30mn	1	16	1	3	caillé	100mg	0.73	3%	30mn	1
17	0	1	caillé	10mg	0.25	1%	15mn	1	17	2	0	St-Paul	100mg	0.25	1%	30mn	0
18	3	2	caillé	10mg	0.25	3%	15mn	1	18	2	0	St-Paul	100mg	0.73	1%	30mn	1
19	2	1	caillé	10mg	0.25	1%	30mn	1	19	2	1	caillé	10mg	0.73	1%	30mn	0
20	5	2	caillé	10mg	0.25	3%	30mn	1	20	2	1	caillé	10mg	0.25	1%	30mn	1
21	0	2	St-Paul	10mg	0.25	1%	15mn	0	21	2	2	St-Paul	10mg	0.73	1%	30mn	0
22	3	3	St-Paul	10mg	0.25	3%	15mn	0	22	2	2	St-Paul	10mg	0.25	1%	30mn	1
23	2	2	St-Paul	10mg	0.25	1%	30mn	1	23	2	3	caillé	100mg	0.73	1%	30mn	0
24	5	1	St-Paul	10mg	0.25	3%	30mn	0	24	2	3	caillé	100mg	0.25	1%	30mn	1
25	0	3	caillé	100mg	0.25	1%	15mn	0	25	3	0	caillé	100mg	0.25	3%	15mn	0
26	3	0	caillé	100mg	0.25	3%	15mn	0	26	3	0	caillé	100mg	0.73	3%	15mn	1
27	2	3	caillé	100mg	0.25	1%	30mn	1	27	3	1	St-Paul	100mg	0.25	3%	15mn	0
28	5	3	caillé	100mg	0.25	3%	30mn	1	28	3	1	St-Paul	100mg	0.73	3%	15mn	1
29	0	0	St-Paul	100mg	0.25	1%	15mn	1	29	3	2	caillé	10mg	0.73	3%	15mn	0
30	3	1	St-Paul	100mg	0.25	3%	15mn	0	30	3	2	caillé	10mg	0.25	3%	15mn	1
31	2	0	St-Paul	100mg	0.25	1%	30mn	0	31	3	3	St-Paul	10mg	0.25	3%	15mn	0
32	5	0	St-Paul	100mg	0.25	3%	30mn	0	32	3	3	St-Paul	10mg	0.73	3%	15mn	1
33	0	1	caillé	10mg	0.73	1%	15mn	0	33	4	0	St-Paul	10mg	0.25	1%	15mn	0
34	3	2	caillé	10mg	0.73	3%	15mn	0	34	4	0	St-Paul	10mg	0.73	1%	15mn	1
35	2	1	caillé	10mg	0.73	1%	30mn	0	35	4	1	St-Paul	100mg	0.73	1%	15mn	0
36	5	2	caillé	10mg	0.73	3%	30mn	0	36	4	1	St-Paul	100mg	0.25	1%	15mn	1
37	0	2	St-Paul	10mg	0.73	1%	15mn	1	37	4	2	caillé	10mg	0.73	1%	15mn	0
38	3	3	St-Paul	10mg	0.73	3%	15mn	1	38	4	2	caillé	10mg	0.25	1%	15mn	1
39	2	2	St-Paul	10mg	0.73	1%	30mn	0	39	4	3	caillé	100mg	0.73	1%	15mn	0
40	5	1	St-Paul	10mg	0.73	3%	30mn	1	40	4	3	caillé	100mg	0.25	1%	15mn	1
41	0	3	caillé	100mg	0.73	1%	15mn	1	41	5	0	St-Paul	100mg	0.25	3%	30mn	0
42	3	0	caillé	100mg	0.73	3%	15mn	1	42	5	0	St-Paul	100mg	0.73	3%	30mn	1
43	2	3	caillé	100mg	0.73	1%	30mn	0	43	5	1	St-Paul	10mg	0.25	3%	30mn	0
44	5	3	caillé	100mg	0.73	3%	30mn	0	44	5	1	St-Paul	10mg	0.73	3%	30mn	1
45	0	0	St-Paul	100mg	0.73	1%	15mn	0	45	5	2	caillé	10mg	0.73	3%	30mn	0
46	3	1	St-Paul	100mg	0.73	3%	15mn	1	46	5	2	caillé	10mg	0.25	3%	30mn	1
47	2	0	St-Paul	100mg	0.73	1%	30mn	1	47	5	3	caillé	100mg	0.73	3%	30mn	0
48	5	0	St-Paul	100mg	0.73	3%	30mn	1	48	5	3	caillé	100mg	0.25	3%	30mn	1

TAB. 11 – Plan ROBOT1 randomisé, avant et après tri. Edition par *écriture*

que la variabilité entre les unités d'une même colonne. Dans ces conditions, on peut s'autoriser à ne pas prendre en compte l'effet colonne lors de l'analyse.

3.2 Plans pour une plaque

Pour la réalisation des plans décrits ci-après, l'introduction de cloisons étanches amovibles permet de séparer soit les colonnes, soit les lignes de la plaque. Par suite, les traitements peuvent être modulés d'une colonne à une autre, ou d'une ligne à une autre, à chacune des étapes de l'expérimentation.

Par ailleurs, l'encrassement se fait en deux passes. A l'issue de chacune d'elles sont donc obtenues 8 éprouvettes *actives* et les 8 éprouvettes témoins associées. Dans les étapes suivantes de l'expérimentation, nettoyage et désinfection, les $16 = 2 \times 8$ éprouvettes actives sont disposées simultanément aux 16 emplacements de la plaque.

Avec 16 unités, il est possible d'étudier jusqu'à 5 facteurs à 2 niveaux en résolution 5, et jusqu'à 8 en résolution 4. Pour un nombre intermédiaire 6 ou 7, il n'est pas possible de faire mieux que la résolution 4. Plus précisément, si le modèle inclut tous les effets principaux et les interactions de deux facteurs, il n'est pas possible d'estimer en plus des effets principaux ne fut ce qu'une interaction de 2 facteurs.

On pourrait en déduire qu'il est toujours préférable d'utiliser 8 facteurs plutôt que 7 ou 6, mais ce n'est pas exact car la taille des paquets d'interactions confondues augmente avec le nombre de facteurs (tableau 12). Si les tests correspondant à certains de ces paquets sont significatifs, l'interprétation est certainement d'autant plus délicate que le nombre de facteurs est élevé (cf [18] pour un exemple d'interprétation de plan de ce type).

facteurs de base A, B, C, D		facteurs définis $E = ABC, F = ABD, G = ACD, H = BCD$	
Confusions entre interactions			
avec les 6 facteurs A, B, C, D, E, F	avec les 7 facteurs A, B, C, D, E, F, G	avec les 8 facteurs A, B, C, D, E, F, G, H	
$AB; CE; DF$	$AB; DF; CE$	$AB; DF; GH; CE$	
$AC; BE$	$AC; DG; BE$	$AC; FH; DG; BE$	
$BC; AE$	$BC; FG; AE$	$BC; DH; FG; AE$	
$AD; BF$	$AD; CG; BF$	$AD; EH; CG; BF$	
$BD; AF$	$BD; AF; EG$	$BD; EG; AF; CH$	
$CD; EF$	$CD; AG; EF$	$CD; EF; AG; BH$	
$DE; CF$	$BG; CF; DE$	$BG; CF; AH; DE$	

TAB. 12 – Plans de résolution 4 pour 16 unités et 6 à 8 facteurs

Par ailleurs, si certaines interactions sont supposées négligeables, ou si l'estimation de certains effets principaux déjà connus n'est pas requise, l'estimation de certaines interactions devient possible. Dans ce cas le nombre 6, 7 ou 8 de facteurs retenus peut jouer de façon primordiale.

Les considérations précédentes montrent le type de dispositif que l'on peut espérer mettre en place en expérimentant avec une seule plaque. En pratique, il faut tenir compte des contraintes imposées par la manipulation du robot. Les deux exemples qui suivent montrent que, si les effets colonnes et lignes sont négligés, l'adaptation à ces contraintes peut s'effectuer sans réduction supplémentaire du nombre de facteurs étudiés.

3.2.1 Plan de résolution 4 pour 8 facteurs

Les paramètres à introduire en entrée du programme pour cet exemple, ainsi que la matrice clé obtenue en sortie, figurent au tableau 13. La forme de ce tableau ne correspond pas à celle des écrans de saisie, mais plutôt au récapitulatif qui figure dans le premier fichier de sortie des résultats, ou dans le fichier obtenu par l'option *Contenu des fichiers .REG* du menu général (tableau 48).

Pour définir ce plan, on divise la plaque en deux macrocolonnes de deux colonnes (facteurs $col1$, $col2$) et deux macrolignes de deux lignes ($lig1$, $lig2$) comme indiqué sur la figure 6.

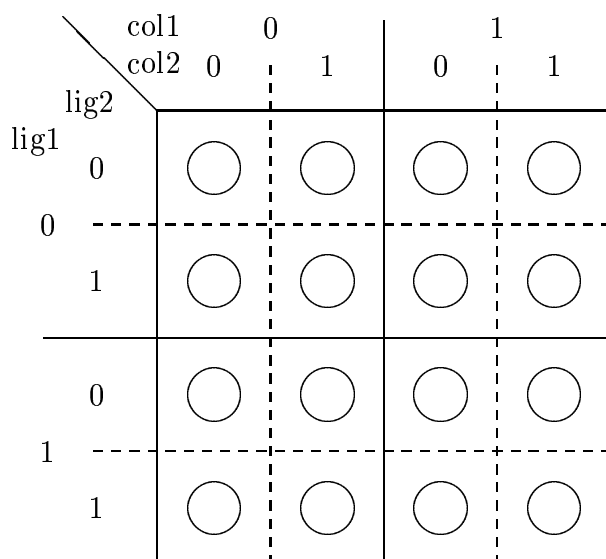


FIG. 6 – Définition de systèmes de blocs ($col1$, $col2$, $lig1$, $lig2$)

Dans chaque macroligne, on dispose au premier encrassement les éprouvettes actives sur l'une des lignes ($lig2=0$), les témoins sur l'autre ($lig2=1$). On échange le rôle des deux lignes au second encrassement ($lig2=0$: témoins, $lig2=1$: éprouvettes actives). Dans la suite des opérations chacune des éprouvettes actives est remise à l'emplacement qu'elle occupait à l'encrassement.

Pour faciliter les manipulations, l'encrassement s'effectue par colonne et le nettoyage par ligne. On évite à l'encrassement de changer de souillure au sein d'une même colonne. La nature de la souillure ($n-sou$) et sa concentration en bactérie ($c-bat$) doivent donc être identiques pour les deux éprouvettes actives encrassées simultanément. Par suite les niveaux de ces deux facteurs sont définis à partir des niveaux des pseudofacteurs $col1$, $col2$, $lig2$ ce qui est indiqué dans les deux premières lignes de la zone des hiérarchies.

De même, la concentration (*conc*) et le temps d'action (*T-act*) du produit de nettoyage ne peuvent être modifiés sur une même ligne. Les niveaux de ces facteurs sont donc définis à partir de *lig1* et *lig2*, ce que l'on indique aussi dans la zone des hiérarchies.

On a admis dans ce cas la possibilité de changer le poids sur le bras du robot au milieu de la colonne pour faire varier la quantité de souillure *q-sou*. Aucune contrainte de hiérarchie n'est donc à prendre en compte pour ce facteur.

L'intensité du brossage *bross* est de même modifiée par l'ajout d'un poids sur le bras du robot. Pour éviter des changements trop fréquents, on n'autorise sa modification que toutes les deux unités, c'est à dire seulement en milieu et fin de chaque ligne. C'est pour prendre en compte cette contrainte qu'est introduite la dernière ligne de la zone hiérarchie qui fait dépendre le niveau de *bross* de la ligne (*lig1*, *lig2*) et de la macrocolonne (*col1*).

La principale requête formulée est définie par la paire *modèle—partie à estimer* numéro 1. Le développement du terme *p.p* dans le modèle 1 donne tous les effets principaux et interactions entre deux des 8 facteurs traitement. La partie à estimer (*p*) dans le cadre de ce modèle contient tous les effets principaux. Le plan résultant, donné par sa matrice clé et ses relations de définition au bas du tableau 13, est donc de résolution 4. En outre, l'inclusion, dans le modèle 1, du facteur *lig2* correspondant au numéro d'encrassement permet d'assurer, même en présence d'un effet additif de l'encrassement, l'estimabilité de tous les effets principaux traitement.

En revanche, les autres facteurs colonne ou ligne *col1*, *col2*, *lig1* ne sont pas pris en compte dans ce modèle 1, pour la simple raison que les contraintes de hiérarchie ne le permettent pas. Ainsi si le modèle inclut le terme *lig1.lig2*, il ne peut y avoir de solution puisque *T-act* et *conc* sont constants dans chaque ligne.

Dans le cadre d'un modèle comprenant les effets lignes et colonnes, il n'est donc pas possible d'estimer tous les effets principaux. Mais on peut s'assurer de l'estimabilité de certains d'entre eux. Ainsi dans cet exemple, le rajout de la paire *modèle—partie à estimer* numéro 2 permet de s'assurer de l'estimabilité des effets principaux rugosité et nature du matériau constituant l'éprouvette (*rug*, *nat*), même si le modèle inclue outre les effets ligne et colonne les interactions *colonne*×*encrassement* (*col1.col2.lig2*) et *macrocolonne*×*ligne* (*col1.lig1.lig2*). Dans le dispositif résultant, deux éprouvettes actives situées dans la même demi-ligne diffèrent à la fois par leur rugosité et par leur nature et il en est de même pour deux éprouvettes actives encrassées simultanément dans une même colonne.

La randomisation du plan doit bien entendu respecter la structure en colonnes et lignes, mais aussi la structure en macrocolonnes et macrolignes pour éviter des changements intempestifs du poids sur le bras du robot, à l'encrassement ou au nettoyage. La randomisation ad-hoc est obtenue en prenant comme modèle de la randomisation *col1+col1.col2+lig1+lig1.lig2*. Cette formule indique que la randomisation est définie par :

- une permutation aléatoire des macrocolonnes (*col1*),
- une permutation aléatoire des macrolignes (*lig1*),
- pour chacune des macrocolonnes, une permutation aléatoire des deux colonnes la constituant (*col2* pour *col1* fixé),

SORTIE	COMMENTAIRE
Classe 0, pseudofacteurs <i>col1</i>]0)= , pseudofacteurs associés : fonction fi : 1 0	Randomisation des macrocolonnes Pas de pseudofacteurs hiérarchisant permutation : 0 \mapsto 1, 1 \mapsto 0
Classe 1, pseudofacteurs <i>lig1</i>]1)= , pseudofacteurs associés : fonction fi : 1 0	Randomisation des macrolignes Pas de pseudofacteurs hiérarchisant permutation : 0 \mapsto 1, 1 \mapsto 0
Classe 2, pseudofacteurs <i>col2</i>]2)=0 , pseudofacteurs associés : <i>col1</i> fonction fi : 0 0 1 1 1 0	randomisation des colonnes dans chaque macrocolonne permutation macrocolonne 0 : 0 \mapsto 0, 1 \mapsto 1 macrocolonne 1 : 0 \mapsto 1, 1 \mapsto 0
Classe 3, pseudofacteurs <i>lig2</i>]3)=1 , pseudofacteurs associés : <i>lig1</i> fonction fi : 0 0 1 1 0 1	randomisation des lignes dans chaque macroligne permutation macroligne 0 : 0 \mapsto 0, 1 \mapsto 1 macroligne 1 : 0 \mapsto 0, 1 \mapsto 1

TAB. 14 – Randomisation du plan ROBPL1R4.PS (sorties intermédiaires)

- pour chacune des macrocolonnes, une permutation aléatoire des deux lignes la constituant (*lig2* pour *lig1* fixé).

Les permutations aléatoires des colonnes sont menées indépendamment dans les deux macrocolonnes. De même les permutations des lignes sont menées indépendamment dans les deux macrolignes. Les permutations choisies peuvent être explicitées (tableau 14) en demandant les impressions intermédiaires dans l'écran 5 (fig. 9) obtenu par l'option *initialisation* du menu général.

Quel est l'intérêt de cette randomisation ? Les expériences passées semblent montrer que l'utilisation des témoins rend négligeables l'effet *position de l'éprouvette sur la plaque*. Il n'est cependant pas certain qu'il en soit toujours ainsi et la prudence impose d'éviter l'utilisation systématique du plan très régulier obtenu avant randomisation.

En principe, l'étude théorique de ce type de randomisation conduit à retenir comme modèle pour l'analyse statistique un modèle incluant tous les effets blocs qui respectent

modèle : $col1.col2.lig1.lig2+p.p$

[lig2]; n-sou c-bat ; rug nat ; q-sou bross ; T-act conc ;
 [col2]; n-sou ;
 [lig2 col2]; c-bat ;
 [lig1]; T-act ;
 [lig1 lig2]; conc ;
 [lig1 col2]; q-sou nat ; n-sou T-act ; bross rug ; c-bat conc ;
 [lig1 lig2 col2]; q-sou rug ; c-bat T-act ; bross nat ; n-sou conc ;
 [col1]; q-sou ;
 [lig2 col1]; bross ;
 [col1 col2]; n-sou q-sou ; T-act nat ; c-bat bross ; conc rug ;
 [lig2 col1 col2]; q-sou c-bat ; T-act rug ; n-sou bross ; conc nat ;
 [lig1 col1]; n-sou nat ; c-bat rug ; q-sou T-act ; conc bross ;
 [lig1 lig2 col1]; c-bat nat ; n-sou rug ; T-act bross ; q-sou conc ;
 [lig1 col1 col2]; nat ;
 [lig1 lig2 col1 col2]; rug ;

TAB. 15 – Ensembles d’effets confondus dans le plan ROBPL1R4

les hiérarchies, c’est à dire ne dissocient jamais la colonne $col2$ de la macrocolonne $col1$ et de même la ligne $lig2$ de la macroligne $lig1$. Les termes de ce modèle :

$col1$; $lig1$; $lig1.col1$; $col2.col1$; $col2.col1.lig1$; $lig2.lig1$; $lig2.lig1.col1$; $lig2.lig1.col2.col1$

sont les *termes ancestraux*. Le programme les déduit automatiquement du modèle de la randomisation et en fournit la liste sur le fichier de sortie du module de randomisation. Ces effets blocs sont rendus aléatoires par la randomisation et, en toute rigueur, l’analyse doit prendre cela en compte pour tester chaque effet traitement dans la bonne “*strate*”.

Dans cet exemple, chaque effet bloc est confondu avec un ou plusieurs effets traitement de la façon indiquée au tableau 15, tableau que l’on obtient par l’étude des alias avec le modèle $col1.col2.lig1.lig2+p.p$. Il est donc hors de question d’effectuer l’analyse avec décomposition en strate évoquée ci-dessus. On est obligé de se reposer pour l’analyse sur l’hypothèse formulée d’absence d’effet position sur la plaque. Il n’y a d’ailleurs pas non plus de degrés de liberté résiduels permettant d’estimer une variance d’erreur. On doit donc pour analyser faire l’hypothèse que certaines des combinaisons linéaires d’effets traitement apparaissant dans le tableau 15 sont nulles et détecter les autres par une procédure du type de celles décrites dans [16] ou [18].

3.2.2 Plan de résolution 5 pour 5 facteurs

Le tableau 16 décrit les paramètres d’entrée pour obtenir un plan de résolution 5 pour 5 facteurs. Le facteur intensité du brossage n’est pas étudié dans ce plan, aussi n’est-il pas nécessaire de décomposer les colonnes en deux macrocolonnes comme dans l’exemple précédent.

La requête principale est ici définie par la paire *modèle—partie à estimer* 1, qui indique que l’on veut estimer tous les termes dans le modèle $p.p$ incluant les effets principaux et les interactions de deux facteurs.

Il est impossible d'estimer les 16 paramètres du modèle $p.p$ dans un modèle qui inclurait en outre un effet bloc tel que l'effet encrassement $lig2$. En revanche on peut imposer comme dans l'exemple précédent que tous les effets principaux soient estimables dans un modèle incluant l'effet encrassement ($lig2$) et que la rugosité soit en outre estimable intra-colonne, c'est à dire en présence d'un terme col dans le modèle. Ces deux dernières contraintes correspondent aux paires *modèle—partie à estimer* 2 et 3. La substitution de $col.lig2$ à col dans la paire 3 produirait le diagnostic suivant, qui prouve qu'on ne peut pas imposer exactement la même contrainte à la rugosité que dans l'exemple précédent:

ECHEC DE LA RECHERCHE BACKTRACK

Recherche arreteee sur le facteur rug

Ordre d'introduction des pseudofacteurs :

lig1	lig2	col_1	col_2	n-sou	c-bat	T-act	conc	rug
2	2	2	2	2	2	2	2	2
								*

La solution obtenue est reportée au bas du tableau 16 et le tableau 17 précise les confusions avec les effets blocs obtenues par l'étude des alias avec le modèle $col.lig1.lig2+p.p$. Il n'y a pas de degrés de liberté pour estimer une variance résiduelle et l'analyse statistique doit donc utiliser une méthode de détection des effets influents de même type que dans l'exemple précédent.

3.3 Plans pour deux plaques

On sait qu'avec 32 unités, il est possible d'étudier jusqu'à 6 facteurs à 2 niveaux en résolution 5 et jusqu'à 16 en résolution 4.

Si l'un des facteurs étudiés est qualitatif à 4 niveaux, les autres ayant toujours 2 niveaux, la situation est assez différente. En résolution 5, il peut y avoir jusqu'à 4 facteurs à 2 niveaux en plus du facteur à 4 niveaux. En résolution 4, il peut y avoir jusqu'à 7 facteurs à 2 niveaux en plus du facteur à 4 niveaux.

★ Le plafond de 7 s'obtient aisément par la recherche définie dans le tableau 18. Le résultat de cette recherche montre l'impossibilité d'aller au delà du 7ième facteur.

Cette recherche introduit comme facteurs de base le facteur A à 4 niveaux et trois autres facteurs à 2 niveaux. On peut prouver simplement que ce choix n'est pas restrictif en utilisant le résultat suivant : si des facteurs ne peuvent être choisis comme facteurs de base, il existe entre eux une relation de définition. Ainsi si aucun des deux ensembles $\{A, B, C, D\}$, $\{A, B, C, E\}$ ne peut être utilisé comme ensemble de base, on peut former deux relations de la forme $A_1^\alpha A_2^\beta BCD = 1$, $A_1^\gamma A_2^\delta BCE = 1$. Mais la multiplication de ces deux relations donne une relation entre les trois facteurs A, D, E qui ne peut exister en résolution 4.

Pour la résolution 5, l'impossibilité de dépasser 4 facteurs à 2 niveaux résulte plus simplement d'un calcul élémentaire du nombre de *degrés de liberté*. Avec 5 facteurs à 2 niveaux, il y aurait à estimer outre la constante, les 3 + 5 paramètres associés aux effets principaux et les $3 \times 5 + 5 \times 4/2 = 25$ paramètres associés aux interactions, soit un total de 34 paramètres, ce qui excède le nombre d'unités expérimentales.

nom	:	ROBPL1R5.REG
nb. d'unités	:	16
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	1
- germe	:	876986
Inclusion des facteurs dans l'ensemble inéligible	:	oui
Commentaire		
1 plaque avec deux encrassements, résolution 5		

Fact. de base			Facteurs à définir				Parties de modèle
fac.	nb.	b	fac.	nb.	type	niveaux	
	niv.	l		niv.	niv.		
lig1	2	←	n-sou	2	lit.	caillé	1 p.p
lig2	2	←				St-Paul	2 lig2
col	4	←	c-bat	2	lit.	3%	3 col
						6%	
			T-act	2	lit.	15mn	
						30mn	
			conc	2	lit.	1%	
						3%	
			rug	2	num.	0.25	
						0.75	
							Parties à estimer
							1 p p
							2 p
							3 rug
							Hierarchies
							1 n-sou: lig2 col
							2 c-bat: lig2 col
							3 T-act: lig1 lig2
							4 conc: lig1 lig2

Définition du plan						
matrice clé						relations de définition
	n-sou	c-bat	T-act	conc	rug	
lig1	0	0	1	1	0	n-sou = lig2 col ₁ col ₂
lig2	1	1	1	0	1	c-bat = lig2 col ₁
col ₁	1	1	0	0	0	T-act = lig1 lig2
col ₂	1	0	0	0	1	conc = lig1
						rug = lig2 col ₂

TAB. 16 – Plan ROBPL1R5 pour 1 plaque, 5 facteurs traitement

Les pseudofacteurs indicés (col_1 , col_2) sont ceux qui résultent d'une décomposition automatique d'un facteur et sont notés avec un – dans le programme.

modèle : col.lig1.lig2 + p.p

- [col₁]; n-sou rug ;
- [lig2]; T-act conc ;
- [lig2 col₁]; c-bat ;
- [col₂]; n-sou c-bat ;
- [col₁ col₂]; c-bat rug ;
- [lig2 col₂]; rug ;
- [lig2 col₁ col₂]; n-sou ;
- [lig1]; conc ;
- [lig1 col₁]; c-bat T-act ;
- [lig1 lig2]; T-act ;
- [lig1 lig2 col₁]; c-bat conc ;
- [lig1 col₂]; T-act rug ;
- [lig1 col₁ col₂]; n-sou T-act ;
- [lig1 lig2 col₂]; conc rug ;
- [lig1 lig2 col₁ col₂]; n-sou conc ;

TAB. 17 – Confusion avec les effets blocs dans ROBPL1R5

nom	:	ESSAI
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	1
- germe	:	0
Inclusion des facteurs dans l'ensemble inéligible	:	oui

Fact. de base		Facteurs à définir		Parties de modèle	
fac.	nb. niv.	fac.	nb. niv.		
A	4	E	2	1 P:A+B+C+D+E+F+G+H+I+J+K+L	
B	2	F	2	Modèles	
C	2	G	2	1 P.P	
D	2	H	2	Parties à estimer	
		I	2	1 P	
		J	2		
		K	2		
		L	2		

ECHEC DE LA RECHERCHE BACKTRACK

Recherche arrêtée sur le facteur I

Ordre d'introduction des pseudofacteurs :

A1 A2 B C D E F G H I J K L
2 2 2 2 2 2 2 2 2 2 2 2

*

TAB. 18 – Recherche nb. maxi. de facteurs pour une fraction de taille 32 d'un 4×2^n

Si l'unique facteur à 4 niveaux est quantitatif, les choses sont encore différentes. On se reportera pour l'analyse de ce cas à [16], [6] et au paragraphe 4 relatif aux plans pour mélanges de facteurs à 2 et 4 niveaux.

Les exemples qui suivent illustrent, dans le cas du robot, les cas considérés ci-dessus.

3.3.1 Plan $2^6/2$ de résolution 5

Le tableau 19 donne les paramètres d'entrée pour un tel plan et la solution sous forme de matrice clé.

L'estimation de l'effet plaque est requis (paire *modèle—partie à estimer* 1). On impose en outre que tous les effets principaux soient estimables même en présence d'un effet encrassement (paire 2) et que la rugosité varie intra-colonne pour les deux éprouvettes encrassées simultanément (paire 3). Les contraintes relatives aux hiérarchies sont du même type que dans les plans pour une plaque.

Le plan obtenu permet donc d'estimer la constante, l'effet plaque, les 6 effets principaux et 15 interactions de deux facteurs, sous l'hypothèse qu'il n'y a pas d'interactions de trois facteurs et plus. Il laisse pour l'estimation de la variance résiduelle $9 = 32 - (1 + 1 + 6 + 15)$ degrés de liberté, qui correspondent aux effets bloc non confondus et distincts de *pl* dans le tableau 20 obtenu par l'étude des alias du modèle *pl.col.lig1.lig2 + p.p.*

3.3.2 Plan $2^8/8$ de résolution 4

Le tableau 21 donne les paramètres de la recherche. Vingt solutions différentes sont requises avec un germe non nul. Dans ce cas, le programme repart à zero pour chaque recherche en réordonnant de façon aléatoire l'ensemble des choix possibles pour chaque colonne de la matrice clé à définir. Cela permet d'obtenir des solutions nettement différentes sans faire une recherche exhaustive.

Pour comparer les solutions, on utilise l'option *étude des alias* du menu général avec le modèle *p.p.* Le tableau 22 donne les sorties pour les solutions 3, 7, 9. Ces sorties illustrent en fait les trois types de résultats obtenus par l'étude des alias des vingt solutions.

Le premier type représenté par la solution 3 donne 7 paquets d'effets traitement confondus, dont un formé de 3 interactions de deux facteurs et les 6 autres de deux interactions seulement. Il conduit donc à 15 interactions confondues et $13 = C_8^2 - 15$ interactions non confondues. Ce type de plan permet donc d'estimer (orthogonalement) 8 effets principaux, 13 interactions isolées et 7 combinaisons linéaires d'interactions. La présence de *pl* dans la partie à estimer 1 permet aussi de s'assurer de l'estimabilité de l'effet plaque. Après estimation de ces $29 = 8 + 13 + 7 + 1$ effets et de la constante, il reste donc seulement $2 = 32 - 30$ degrés de liberté pour l'estimation de la variance résiduelle, degrés de liberté qui correspondent aux deux effets bloc non confondus et distincts de *pl* obtenus avec une analyse des alias du modèle *pl.col1.col2.lig1.lig2 + p.p.* (bas du tableau 22). Ce faible nombre de degrés de liberté oblige, soit à combiner cette estimation avec une

nom	:	ROBPL2R5
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	1
- germe	:	8765869
Inclusion des facteurs dans l'ensemble inéligible	:	oui
Commentaire		
2 plaques avec encrassement en deux passes, résolution 5		

Fact. de base			Fact. définis		Parties de modèle
fac.	nb.	bloc	fac.	nb.	1 p:n-sou+q-sou+c-bat+T-act+conc+rug
	niv.			niv.	
pl	2	←	n-sou	2	
lig1	2	←	q-sou	2	Modèles
lig2	2	←	c-bat	2	1 pl+p.p
col	4	←	T-act	2	2 pl.lig2
			conc	2	3 pl.col.lig2
			rug	2	
					Parties à estimer
					1 pl + p.p
					2 p
					3 rug
					Hiérarchies
					1 n-sou: col lig2
					2 c-bat: col lig2
					3 T-act: lig1 lig2
					4 conc: lig1 lig2

Définition du plan							
matrice clé						relations de définition	
	n-sou	q-sou	c-bat	T-act	conc	rug	
pl	0	1	0	0	0	1	n-sou = lig2 col ₂
lig1	0	1	0	1	1	1	q-sou = pl lig1 lig2 col ₂
lig2	1	1	1	1	0	1	c-bat = lig2 col ₁
col ₁	0	0	1	0	0	1	T-act = lig1 lig2
col ₂	1	1	0	0	0	1	conc = lig1
							rug = pl lig1 lig2 col ₁ col ₂

TAB. 19 – Plan ROBPL2R5 pour 2 plaques, 6 facteurs traitement

[col₁]; q-sou rug ;
 [col₁ col₂]; n-sou c-bat ;
 [lig2]; T-act conc ;
 [lig2 col₁]; c-bat ;
 [lig2 col₂]; n-sou ;
 [pl col₂]; q-sou T-act ;
 [pl col₁ col₂]; T-act rug ;
 [pl lig2 col₂]; q-sou conc ;
 [pl lig2 col₁ col₂]; conc rug ;
 [lig1]; conc ;
 [lig1 col₁]; c-bat T-act ;
 [lig1 col₂]; n-sou T-act ;
 [lig1 lig2]; T-act ;
 [lig1 lig2 col₁]; c-bat conc ;
 [lig1 lig2 col₂]; n-sou conc ;
 [pl lig1 col₂]; c-bat rug ;
 [pl lig1 col₁ col₂]; q-sou c-bat ;
 [pl lig1]; n-sou q-sou ;
 [pl lig1 col₁]; n-sou rug ;
 [pl lig1 lig2 col₂]; q-sou ;
 [pl lig1 lig2 col₁ col₂]; rug ;

liste des effets blocs non confondus

col₂ ; lig2 col₁ col₂ ; pl ; pl col₁ ; pl lig2 ;
 > pl lig2 col₁ ; lig1 col₁ col₂ ; lig1 lig2 col₁ col₂ ;
 > pl lig1 lig2 ; pl lig1 lig2 col₁ ;

TAB. 20 – Confusion avec les effets blocs dans ROBPL2R5

Le > indique que la ligne se continue.

nom	:	ROBPL2R4
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	20
- germe	:	208877454
Inclusion des facteurs dans l'ensemble inéligible	:	oui
Commentaire		
2 plaques avec encrassement en deux passes, 8 facteurs, résolution 4		

Fact. de base			Fact. définis		Parties de modèle
fac.	nb.	bloc	fac.	nb.	1 p:n-sou+q-sou+c-bat+T-act +conc+bross+rug+nat
	niv.			niv.	
pl	2	←	n-sou	2	Modèles
lig1	2	←	q-sou	2	1 pl+p.p
lig2	2	←	c-bat	2	2 pl.lig2
col1	2	←	T-act	2	3 pl.col1.col2.lig2
col2	2	←	conc	2	Parties à estimer
			bross	2	1 pl+p
			rug	2	2 p
			nat	2	3 rug+nat
					Hiérarchies
					1 n-sou: pl col1 col2 lig2
					2 c-bat: pl col1 col2 lig2
					3 T-act: pl lig1 lig2
					4 conc : pl lig1 lig2
					5 bross: pl lig1 lig2 col1

Définition du plan									
matrice clé								relations de définition	
	n-sou	q-sou	c-bat	T-act	conc	bross	rug	nat	
pl	0	0	0	1	0	0	1	1	n-sou = col2
lig1	0	1	0	1	1	1	1	1	q-sou = lig1 col1
lig2	0	0	1	0	1	1	1	0	c-bat = lig2 col1 col2
col1	0	1	1	0	0	1	0	1	T-act = pl lig1
col2	1	0	1	0	0	0	1	1	conc = lig1 lig2
									bross = lig1 lig2 col1
									rug = pl lig1 lig2 col2
									nat = pl lig1 col1 col2

TAB. 21 – Plan ROBPL2R4 pour 2 plaques, 8 facteurs traitement

Ensembles d'effets traitement du modèle confondus entre eux		
<p>Solution 3, Germe : 798251298</p> <p>rug nat ; n-sou c-bat ; q-sou conc ; conc rug ; q-sou nat ; conc nat ; q-sou rug ; c-bat conc ; n-sou q-sou ; n-sou conc ; q-sou c-bat ; n-sou nat ; c-bat rug ; n-sou rug ; c-bat nat ;</p>	<p>Solution 7, Germe : 539794919</p> <p>q-sou rug ; T-act nat ; conc bross ; n-sou c-bat ; T-act rug ; q-sou nat ; q-sou T-act ; rug nat ; conc rug ; q-sou bross ; q-sou conc ; bross rug ; T-act conc ; bross nat ; conc nat ; T-act bross ; q-sou c-bat ; n-sou nat ; c-bat rug ; n-sou T-act ; n-sou q-sou ; c-bat nat ; n-sou rug ; c-bat T-act ;</p>	<p>Solution 9, Germe : 1636010972</p> <p>bross nat ; n-sou T-act ; q-sou c-bat ; q-sou nat ; c-bat bross ; T-act conc ; q-sou bross ; c-bat nat ; n-sou conc ; q-sou T-act ; n-sou c-bat ; conc nat ; n-sou q-sou ; c-bat T-act ; conc bross ; n-sou bross ; T-act nat ; q-sou conc ; n-sou nat ; T-act bross ; c-bat conc ;</p>
Effets traitement non confondus avec d'autres effets traitement		
<p>Solution 3</p> <p>n-sou ; conc bross ; T-act nat ; >c-bat ; q-sou bross ; T-act rug ; >q-sou T-act ; bross rug ; conc ; >bross nat ; T-act bross ; T-act ; >q-sou ; bross ; n-sou bross ; >T-act conc ; c-bat bross ; nat ; >n-sou T-act ; rug ; c-bat T-act ;</p>	<p>Solution 7</p> <p>c-bat ; n-sou ; rug ; q-sou ; >T-act ; nat ; c-bat bross ; >conc ; bross ; c-bat conc ; >n-sou bross ; n-sou conc ;</p>	<p>Solution 9</p> <p>n-sou ; T-act ; T-act rug ; >n-sou rug ; rug ; conc rug ; conc ; >c-bat ; q-sou ; q-sou rug ; >c-bat rug ; rug nat ; bross rug ; >bross ; nat ;</p>
Effets bloc non confondus (avec un effet traitement)		
<p>Solution 3</p> <p>pl ; lig1 ; pl lig1 lig2 col1 ;</p>	<p>Solution 7</p> <p>lig2 ; pl lig2 col2 ; col2 ; >lig2 col1 col2 ; pl ; >pl lig2 col1 ; pl lig1 lig2 col2 ; >pl lig1 col1 col2 ;</p>	<p>Solution 9</p> <p>lig2 ; pl ; pl lig1 ; >pl lig1 col2 ; pl lig2 col2 ; >lig2 col1 col2 ; lig1 lig2 col1 ; >pl col1 col2 ; pl lig1 col1 ;</p>

TAB. 22 – Confusion d'effets dans trois solutions type d'un $2^8/8$ de résolution 4
Le > indique que la ligne se continue.

estimation antérieurement obtenue, soit à utiliser une procédure de détection des effets influents du type déjà mentionné.

Le second type représenté par la solution 7 donne 11 ensembles d'effets traitement confondus dont deux ensembles de 3 interactions et 9 de deux, soit 24 interactions confondues et 4 qui ne le sont pas. Il permet donc d'estimer la constante, l'effet plaque, les 8 effets principaux, 4 interactions isolées, 11 combinaisons linéaires d'interactions et laisse $7 = 32 - (1 + 1 + 8 + 4 + 11)$ degrés de liberté pour l'estimation de la variance résiduelle. Ces 7 degrés de liberté correspondent aux effets bloc non confondus et distincts de pl obtenus avec une analyse des alias du modèle $pl.col1.col2.lig1.lig2 + p.p$ (bas du tableau 22).

Finalement, le troisième type représenté par la solution 9 permet d'estimer 7 interactions isolées, 7 combinaisons linéaires de trois interactions et laisse 8 degrés de liberté pour estimer la variance résiduelle.

Le premier type de plan permet d'estimer d'avantage d'interactions et est donc préférable aux deux autres, sauf peut-être dans le cas où l'on s'attend à ce que de nombreux effets soit significatifs et où on tient à ce que l'expérience fournisse un estimateur correct de la variance résiduelle.

Le choix précis du plan à l'intérieur du premier type est sans importance si on accorde la même importance a priori à toutes les interactions de deux facteurs. Nous avons retenu pour l'exemple la solution 3, correspondant au germe 798251298, dont la matrice clé et les relations de définition sont données en bas du tableau 21. La construction du plan correspondant s'obtient par l'option *Création à partir d'une matrice obtenue précédemment* du menu général, dans laquelle on spécifie le numéro de la solution retenue. On peut aussi réobtenir cette solution isolée en refaisant une recherche d'une seule solution avec le germe 798251298.

3.3.3 Plan $4 \times 2^4/2$ de résolution 5

Le tableau 23 donne les paramètres d'entrée permettant d'obtenir un tel plan et la solution obtenue. Le modèle contient la constante, l'effet pl , les trois paramètres de l'effet principal du facteur qualitatif $n-sou$ à 4 niveaux, les 4 effets principaux des facteurs à 2 niveaux et leurs 6 interactions, les 12 paramètres d'interaction entre le facteur $n-sou$ et chacun des 4 autres facteurs. Il contient donc au total $27 = 1 + 1 + 3 + 4 + 6 + 12$ paramètres et laisse pour estimer la variance d'erreur $5 = 32 - 27$ degrés de liberté. Ceux-ci correspondent aux effets bloc non confondus et distincts de pl apparaissant au bas du tableau 23. Ils ont été obtenus avec l'analyse des alias du modèle $pl.col.lig1.lig2 + p.p$.

3.3.4 Plan $4 \times 2^7/16$ de résolution 4

Avant d'examiner l'adaptation d'un tel plan au robot, on peut faire une recherche pour savoir quels types de plans de cette forme on peut obtenir. L'examen des alias pour les 20 solutions obtenues par la recherche définie dans le tableau 24 montre une unique structure d'alias et suggère qu'il n'y a, à une permutation près des facteurs à 2 niveaux d'une part, et des trois pseudofacteurs A_1, A_2, A_1A_2 qu'une seule solution. Le tableau 25

nom	:	ROP2F4R5
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	1
- germe	:	6794539
Inclusion des facteurs dans l'ensemble inéligible	:	oui
Commentaire		
2 plaques avec encrassement en deux passes, résolution 5		
1 facteur qualitatif à 4 niveaux, 4 facteurs à 2 niveaux		

Fact. de base			Fact. définis		Parties de modèle	
fac.	nb.	b	fac.	nb.		
	niv.	l		niv.		
pl	2	←	n-sou	4	1 p:n-sou+c-bat+T-act+conc+rug	
lig1	2	←	c-bat	2	Modèles	
lig2	2	←	T-act	2	1 pl+p.p	
col	4	←	conc	2	2 pl.lig2	
			rug	2	Parties à estimer	
					1 pl+p p	
					2 c-bat+T-act+conc+rug	
					Hiérarchies	
					1 n-sou: pl col lig2	
					2 c-bat: pl col lig2	
					3 T-act: pl lig1 lig2	
					4 conc: pl lig1 lig2	

Définition du plan							
matrice clé						relations de définition	
	n-sou ₁	n-sou ₂	c-bat	T-act	conc	rug	
pl	0	0	1	1	0	0	n-sou ₁ = col ₁
lig1	0	0	0	1	1	0	n-sou ₂ = lig2 col ₁
lig2	0	1	0	0	1	1	c-bat = pl col ₁ col ₂
col ₁	1	1	1	0	0	0	T-act = pl lig1
col ₂	0	0	1	0	0	1	conc = lig1 lig2
							rug = lig2 col ₂

liste des effets blocs non confondus
pl ; pl col ₁ ; pl lig1 col ₂ ; pl lig1 col ₁ col ₂ ; lig1 lig2 col ₂ ; lig1 lig2 col ₁ col ₂ ;

TAB. 23 – Plan ROP2F4R5 pour 2 plaques, un $4 \times 2^4/2$ de résolution 5

précise les confusions pour l'une des solutions. On voit qu'il y a 18 effets non confondus et 13 paquets d'interactions. Un tel dispositif ne laisse donc aucun degré de liberté résiduel et ne permet pas d'estimer un éventuel effet bloc supplémentaire.

nom	:	ESSAI1
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de decomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	20
- germe	:	456934
Inclusion des facteurs dans l'ensemble inéligible	:	oui

Fact. de base		Facteurs à définir		Parties de modèle	
fac.	nb. niv.	fac.	nb. niv.		
A	4	E	2	1 P:A+B+C+D+E+F+G+H	
B	2	F	2	Modèles	
C	2	G	2	1 P.P	
D	2	H	2	Parties à estimer	
				1 P	

TAB. 24 – Recherche de différentes solutions pour un $4 \times 2^7/16$

Le tableau 26 définit un plan de ce type pour le robot. L'impossibilité d'estimer pl impose de ne pas le faire figurer dans la partie à estimer.

Solution 4, Germe : 208449859
liste détaillée des ensembles d'effets confondus du modèle
$A_1 B ; E F ;$ $A_1 A_2 B ; D H ; C G ;$ $A_1 C ; E H ;$ $A_1 A_2 C ; D F ; B G ;$ $B C ; F H ; A_1 A_2 G ;$ $D E ; A_1 G ;$ $A_1 D ; E G ;$ $A_1 A_2 D ; B H ; C F ;$ $B D ; F G ; A_1 A_2 H ;$ $A_1 H ; C E ;$ $C D ; G H ; A_1 A_2 F ;$ $A_1 F ; B E ;$ $C H ; B F ; A_1 E ; D G ;$
liste des effets non confondus
$; A_1 ; A_2 ; A_1 A_2 ; B ; A_2 B ; C ; A_2 C ; A_2 G ; G ;$ $>D ; A_2 D ; A_2 H ; H ; A_2 F ; F ; A_2 E ; A_1 A_2 E ; e ;$

TAB. 25 – Structure d'alias dans un $4 \times 2^7/16$ de résolution 4

nom	:	ROP2F4R4
nb. d'unités	:	32
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	10 mn
- nb. sol. max.	:	1
- germe	:	777754354
Inclusion des facteurs dans l'ensemble inéligible	:	oui
Commentaire		
2 plaques avec encrassement en deux passes, résolution 4		
1 facteur qualitatif à 4 niveaux, 7 facteurs à 2 niveaux		

Fact. de base			Fact. définis		Parties de modèle
fac.	nb.	bloc	fac.	nb.	1 p:n-sou+q-sou+c-bat+T-act +conc+bross+rug+nat
		niv.			
pl	2	←	n-sou	4	Modèles
lig1	2	←	q-sou	2	1 pl+p.p
lig2	2	←	c-bat	2	Parties à estimer
col1	2	←	T-act	2	1 p
col2	2	←	conc	2	Hiérarchies
			bross	2	1 n-sou: pl col1 col2 lig2
			rug	2	2 c-bat: pl col1 col2 lig2
			nat	2	3 T-act: pl lig1 lig2
					4 conc : pl lig1 lig2
					5 bross: pl lig1 lig2 col1

Définition du plan										
matrice clé										
	n-sou ₁	n-sou ₂	q-sou	c-bat	T-act	conc	bross	rug	nat	relations de définition
pl	1	0	0	0	0	1	0	1	1	n-sou ₁ = pl lig2 col1 col2
lig1	0	0	1	0	1	1	1	0	0	n-sou ₂ = col1 col2
lig2	1	0	0	1	1	1	1	0	1	q-sou = lig1 col1
col1	1	1	1	1	0	0	1	0	0	c-bat = lig2 col1 col2
col2	1	1	0	1	0	0	0	1	1	T-act = lig1 lig2
										conc = pl lig1 lig2
										bross = lig1 lig2 col1
										rug = pl col2
										nat = pl lig2 col2

liste des effets blocs non confondus
pl ; pl col ₁ ; pl lig1 col ₂ ; pl lig1 col ₁ col ₂ ; lig1 lig2 col ₂ ; lig1 lig2 col ₁ col ₂ ;

TAB. 26 – Plan ROP2F4R4 pour 2 plaques, un $4 \times 2^8/32$ de résolution 4

4 Plans pour mélange de facteurs à 2 et 4 niveaux

4.1 Facteurs qualitatifs

4.1.1 Exemple

Dans une étude sur l'efficacité du nettoyage et de la désinfection de surfaces, on a initialement répertorié 16 facteurs susceptibles d'agir, dont quatre facteurs qualitatifs importants qu'on souhaiterait étudier à plus de deux modalités, comme le type de produit désinfectant.

Les expériences de comparaison des différents traitements sont menées sur des échantillons de surfaces traités par bloc de 8. Le facteur *température de nettoyage* T , étudié à 2 niveaux 4° et 20° , est obligatoirement constant sur chaque bloc.

Pour déterminer les facteurs les plus influents, 64 traitements correspondant à 8 blocs peuvent être testés dans une première phase. Pour minimiser le nombre d'essais, on décide de n'étudier qu'à deux niveaux les facteurs autres que les quatre déjà mentionnés. Pour ces derniers, on retient quatre niveaux plutôt que trois. Cela permet de couvrir un spectre de traitements plus large, et enrichit beaucoup les possibilités de création de fractions de plan.

Dans un premier temps, les blocs sont ignorés. On cherche le nombre maximum de facteurs à 2 niveaux rajoutables aux 4 facteurs à 4 niveaux en limitant la résolution à 3 ou 4, car la résolution 5 est inatteignable avec ces quatre facteurs à 4 niveaux.

★ Il existe obligatoirement des relations de définition liant ces quatre facteurs car le plan, de taille 64, ne peut contenir qu'une fraction des $256 = 4^4$ combinaisons de niveaux de ces quatre facteurs.

Une première recherche d'un plan de résolution 4 trouve en quelques secondes des relations permettant de définir 4 facteurs E, F, G, H à deux niveaux en plus des 4 facteurs A, B, C, D à 4 niveaux. La recherche d'un 5ième facteur I n'a toujours pas abouti après 5 mn et est abandonnée (tab. 27).

Deux possibilités se présentent pour étudier d'avantage de facteurs : réduire le nombre de facteurs à 4 niveaux ou limiter la résolution à 3.

4.1.2 Plans de résolution 4

Un plan de résolution 4 est un plan où tous les effets principaux sont estimables dans un modèle incluant, outre la constante et les effets principaux, toutes les interactions de deux facteurs. En utilisant une reparamétrisation convenable, par exemple celle décrite dans [20], on prouve facilement le résultat suivant, où X est la matrice du modèle linéaire.

Proposition 4.1 *Dans un plan de résolution 4, les colonnes de X associées à la constante, aux effets principaux et aux interactions entre un facteur déterminé et chacun des autres facteurs sont indépendantes.*

Ce résultat a été prouvé par Margolin [21] en utilisant la reparamétrisation associée aux polynômes orthogonaux. Il fournit un minorant au nombre d'unités N permettant de construire un plan de résolution 4 avec des facteurs fixés à l'avance : N doit être supérieur au nombre de colonnes indépendantes données par la proposition.

Ainsi dans le cas qui nous préoccupe, si il y a n_4 facteurs à 4 niveaux, n_2 facteurs à deux niveaux, le nombre de colonnes associées aux effets principaux est $n_2 + 3n_4$. Si le facteur déterminé a 4 niveaux, le nombre de colonnes associées aux interactions entre ce facteur et chacun des $n_2 + n_4 - 1$ autres est $3(n_2 + 3(n_4 - 1))$. On doit donc avoir

$$N \geq 1 + n_2 + 3n_4 + 3(n_2 + 3(n_4 - 1)) .$$

Si N et n_4 sont fixés, cette formule donne le majorant suivant au nombre n_2 de facteurs à deux niveaux introductibles en résolution 4 :

$$n_2 \leq \frac{N}{4} - 3n_4 + 2 . \quad (6)$$

Dans les cas $\{n_4 = 1, N = 16k\}$ et $\{n_4 = 2, N = 8k\}$, on sait construire des plans de résolution 4 ayant ce nombre maximum n_2 de facteurs à 2 niveaux [21], [1]. Ces constructions, basées sur des matrices d'Hadamard, ne garantissent pas la régularité et donc la simplicité des confusions entre interactions de deux facteurs, mais elles permettent de former des plans ayant un nombre d'unités qui n'est pas puissance de 2.

En revanche pour $n_4 = 3$, on ne connaît pas de méthode générale comparable et seul un algorithme tel que celui de PLANOR permet de former des plans orthogonaux comportant un nombre appréciable de facteurs à deux niveaux.

Les résultats obtenus par PLANOR avec $N = 64$ et n_4 compris entre 1 et 4 sont récapitulés dans le tableau 27. Ce tableau précise les facteurs de base utilisés dans chaque recherche.

★ Les choix opérés pour ces facteurs de base ne sont pas restrictifs. Cela est clair lorsque les facteurs de base sont les facteurs A, B, C à quatre niveaux puisqu'en résolution 4 il ne peut y avoir de relation de définition liant ces trois facteurs. Lorsqu'il y a seulement deux facteurs A et B à 4 niveaux, on peut certainement leur adjoindre C . La résolution 4 interdit toute relation de définition entre ces trois facteurs et entraîne donc que les 32 combinaisons de leurs niveaux sont présentes. Si on ne pouvait y adjoindre encore un autre facteur à 2 niveaux, cela signifierait que tous les autres facteurs sont définis à partir de A, B, C et que le plan est formé de deux répétitions des mêmes 32 traitements, ce qui doit être évité. Un raisonnement analogue vaut dans le cas où il n'y a qu'un seul facteur à 4 niveaux.

Les nombres maximum de facteurs à 2 niveaux donnés par ce tableau 27 sont à comparer aux majorants provenant de (6) qui sont explicités pour n_4 compris entre 1 et 6 dans le tableau 28.

Les nombres maximum de facteurs à 2 niveaux atteints dans un délai raisonnable sont légèrement au dessous des plafonds fournis par le tableau 28 lorsque $n_4 = 4$ ou $n_4 = 3$. Ils sont égaux à ces plafonds dans les cas $n_4 = 2$ et $n_4 = 1$.

La résolution 4 donne des estimations correctes des effets principaux même en présence d'interactions. Elle permet en conséquence de tirer des conclusions généralement assez fiables. Mais le tableau 27 montre que l'obtention de cette résolution oblige à

nb. d'unités	: 64
définition des facteurs de base	: utilisateur
type de décomposition des facteurs	: maximum

Partie de modèle	
PM: $A + B + C + D + E + F + G + H + I + J + K + L + M + N + O + P + Q$	
Modèle PM.PM	Partie à estimer PM

nb. fact. 4 niv.	nb. fact. 2 niv. maxi	facteurs de base	facteurs définis et pt. d'arrêt à 5 mn
4	4	$A_1 A_2 B_1 B_2 C_1 C_2$	$D_1 D_2 E F G H I \star J K L M N O P Q$
3	7	$A_1 A_2 B_1 B_2 C_1 C_2$	$D E F G H I J \star K L M N O P Q$
2	12	$A_1 A_2 B_1 B_2 C D$	$E F G H I J K L M N \star O P Q$
1	15	$A_1 A_2 B C D E$	$F G H I J K L M N O P \star Q$

TAB. 27 – Plans de résolution 4 pour mélange de facteurs qualitatifs à 4 et 2 niveaux

L'étoile indique le facteur que le programme était en train d'essayer de définir après 5 mn de recherche. En général ce facteur est atteint en une fraction de minute. Le nombre figurant dans la seconde colonne est le nombre de facteurs à 2 niveaux à gauche de \star , c'est à dire le nombre maximum que l'on peut arriver à définir en un temps raisonnable

n_4	1	2	3	4	5	6
n_2 max	15	12	9	6	3	0

TAB. 28 – Majorant de n_2 dans une fraction de taille 64, résolution 4 d'un $4^{n_4}2^{n_2}$

restreindre beaucoup les ambitions initialement affichées quant au nombre de facteurs étudiés.

4.1.3 Plans de résolution 3, doublables en résolution 4

Si on suspecte l'existence d'un petit nombre de facteurs vraiment actifs sans avoir d'indications plus précises, une bonne stratégie peut être d'augmenter le nombre de facteurs étudiés en ramenant à 3 la résolution.

★ Le nombre maximum de facteurs à 2 niveaux ajoutables aux facteurs à 4 niveaux en résolution 3 est facile à trouver. La condition requise est de ne pas avoir de relation de définition liant deux facteurs seulement. Par exemple, si A, B, C, D sont les facteurs à quatre niveaux, E, F, \dots ceux à deux niveaux, on ne doit pas avoir de relations telles que $B_1 = A_1A_2, E = A_1A_2, E = F$. En d'autres termes les éléments $A_1, A_2, A_1A_2, B_1, B_2, B_1B_2, C_1, C_2, C_1C_2, D_1, D_2, D_1D_2, E, F, \dots$ doivent être tous différents. Comme il y a au plus 63 éléments différents et non nuls formés à partir des facteurs de base, il reste une fois déterminés les facteurs à 4 niveaux $51 = 63 - 12$ éléments utilisables pour définir les facteurs à 2 niveaux. Plus généralement, si s_4 est le nombre de facteurs à 4 niveaux, le nombre maximum de facteurs à 2 niveaux ajoutables est $s_2 = 63 - 3 \times s_4$.

La théorie donne par ailleurs le nombre s_4 maximum de facteur à 4 niveaux que l'on peut introduire préalablement : $21 = 63/3$ (nombre de sous-espaces vectoriels de dimension 1 dans l'espace vectoriel F_4^3 de dimension 3 sur le corps de Galois à 4 éléments).

Les chances de détecter les facteurs actifs augmentent avec le nombre des facteurs étudiés, mais l'abaissement de la résolution fragilise beaucoup les conclusions qui peuvent être tirées. Un tel plan de résolution 3 n'est donc généralement envisagé que comme une étape intermédiaire permettant une détection rapide des facteurs les plus importants. Il peut se poursuivre par un autre plan restreint aux facteurs trouvés actifs pour confirmer ou infirmer les résultats trouvés et étudier les interactions. Une autre continuation possible, utile quand le nombre d'effets significatifs empêche d'interpréter à cause des confusions possibles avec les interactions, est un *doublément du plan* permettant d'avoir au total un plan de résolution 4.

Cette possibilité de doublement est bien connue lorsque tous les facteurs ont deux niveaux : le second plan est égal au produit du premier par -1 . Le tableau 29 donne un exemple classique, pour 7 facteurs A, B, C, D, E, F, G à 2 niveaux.

Si certains facteurs A, B, \dots sont à 4 niveaux, cette même procédure de doublement par le plan opposé, appliquée aux facteurs à 2 niveaux et aux pseudofacteurs $A_1, A_2, B_1, B_2, \dots$ issus des facteurs à 4 niveaux, ne conduit pas dans tous les cas à un plan de résolution 4. Pour qu'elle y conduise, il faut que le plan initial de résolution 3 ait en outre la propriété de permettre l'estimation des effets principaux de la forme A_1A_2, B_1B_2, \dots dans le cadre du modèle incluant toutes les interactions entre deux facteurs.

Le tableau 30 illustre la recherche correspondante lorsqu'il y a 4 facteurs à 4 niveaux. La contrainte supplémentaire est introduite sous forme de la paire *Modèle-Partie à estimer* numéro 2.

En bas du tableau 30 sont donnés les nombres maximum de facteurs à 2 niveaux pouvant être rajoutés aux facteurs à 4 niveaux avec cette contrainte supplémentaire, pour

Plan initial (N=1)								Complément : $-1 \times$ plan initial (N = -1)									
	N	A	B	C	D	E	F	G		N	A	B	C	D	E	F	G
1	1	1	1	1	1	1	1	1	9	-1	-1	-1	-1	-1	-1	-1	-1
2	1	1	1	-1	1	-1	-1	-1	10	-1	-1	-1	1	-1	1	1	1
3	1	1	-1	1	-1	1	-1	-1	11	-1	-1	1	-1	1	-1	1	1
4	1	1	-1	-1	-1	-1	1	1	12	-1	-1	1	1	1	1	-1	-1
5	1	-1	1	1	-1	-1	1	-1	13	-1	1	-1	-1	1	1	-1	1
6	1	-1	1	-1	-1	1	-1	1	14	-1	1	-1	1	1	-1	1	-1
7	1	-1	-1	1	1	-1	-1	1	15	-1	1	1	-1	-1	1	1	-1
8	1	-1	-1	-1	1	1	1	-1	16	-1	1	1	1	-1	-1	-1	1

TAB. 29 – 2^{7-3} de résolution 4 obtenu par doublement d'un 2^{7-4} de résolution 3

un nombre de facteurs à 4 niveaux égal à 2, 3, 4. Dans chaque cas la recherche a été arrêtée sur un temps limite de quelques minutes. La contrainte supplémentaire garantit que le doublement du plan résultant par son opposé donne un plan global de résolution 4.

nb. d'unités	: 64
définition des facteurs de base	: utilisateur
type de decomposition des facteurs	: maximum

Partie de modèle	
PM: $A + B + C + D + E + F + G + H + I + J + K + L + M + N + O + P + Q + R + S$	
Modèle	Partie à estimer
1 PM	1 PM
2 PM.PM	2 $A_1A_2 + B_1B_2 + C_1C_2 + D_1D_2$
***** RECHERCHE INTERROMPUE à 5 mn	
Recherche arretée sur le facteur S	
Ordre d'introduction des pseudofacteurs:	
$A_1 A_2 B_1 B_2 C_1 C_2 D_1 D_2 E F G H I J K L M N O P Q R S$ ★	

nb. fact. 4 niv.	nb. fact. 2 niv. maxi
2	26
3	20
4	14

TAB. 30 – Plan résol. 3 doublable en résol. 4 avec fact. à 4 et 2 niveaux

★ La propriété de passage à la résolution 4 par doublement du plan dans le cas des facteurs à deux niveaux est vraie et facile à prouver dans un contexte beaucoup plus général que celui considéré dans cette notice. N'importe quel plan de résolution 3 pour facteurs à 2 niveaux, par exemple le plan de Plackett et Burman pour 11 facteurs et 12 unités dérivé de la matrice d'Hadamard d'ordre 12, donne en le doublant par le plan opposé un plan de résolution 4, c'est à dire un plan où l'estimation usuelle des effets principaux comme demi-différence des moyennes entre les niveaux 1 et -1 est non biaisée même si certains couples

de facteurs présentent des interactions.

Si le plan initial est un plan régulier constructible par les méthodes considérées ici, il en est de même du plan doublé. Ainsi dans l'exemple du tableau 29, les relations de définition du plan initial à 8 unités sont $D = AB$, $E = AC$, $F = BC$, $G = ABC$. Il est alors immédiat, si N est le pseudofacteur prenant la valeur 1 pour le premier plan, -1 pour le second, que le plan global est défini à partir des 4 facteurs de base N , A , B , C par les relations $D = ABN$, $E = ACN$, $F = BCN$, $G = ABC$. Il est facile de voir que les relations de définition du plan initial qui restent valides au niveau global sont les relations impliquant un nombre pair des lettres A à G . En particulier les relations de trois lettres vérifiées par le plan initial ne sont plus vérifiées sur le plan global. Par exemple la relation $ABD = 1$ vérifiée sur le plan initial ne peut être valide globalement puisque $ABD = -1$ sur le plan complémentaire. Les relations de définition du plan global incluent donc au minimum 4 lettres et ce plan est bien de résolution 4.

Dans le cas où certains facteurs A , B , ... ont quatre niveaux, il peut exister des relations de définition du plan initial dans lesquelles figurent trois facteurs seulement, mais un nombre pair de pseudofacteurs : $A_1A_2EF = 1$, $A_1A_2B_1C_1 = 1$ par exemple. Ces relations sont vérifiées par le plan complémentaire et donc par le plan global. La résolution de ce dernier n'excède alors pas 3 puisqu'il confond avec la moyenne des interactions de trois facteurs.

Ces relations gênantes contiennent trois lettres distinctes dont une au moins qui apparaît deux fois. Elles contiennent donc obligatoirement un produit tel que A_1A_2 , B_1B_2 de deux pseudofacteurs issus d'un même facteur à 4 niveaux. En imposant en plus de la résolution 3 l'estimabilité de ces produits A_1A_2 , B_1B_2 , etc ... dans un modèle incluant toutes les interactions de deux facteurs, on évite ces relations gênantes et le plan généré donne en le doublant par l'opposé un plan régulier de résolution 4.

4.1.4 Prise en compte des blocs

L'introduction des blocs et de la contrainte de hiérarchie pour la température s'opère sans difficulté.

Le tableau 31 précise les paramètres de la recherche d'un plan de résolution 4 comprenant 3 facteurs à 4 niveaux et 7 facteurs, dont la température, à 2 niveaux. Parmi les 20 solutions trouvées, c'est la 6ième qui a été sélectionnée parce qu'elle donne un nombre maximum de 26 effets non confondus dans le modèle incluant l'effet bloc et les interactions de deux facteurs. Ces effets non confondus, qui incluent les $15 = 3 \times 3 + 6$ effets principaux des facteurs différents de la température, apparaissent dans le tableau 32, qui donne aussi les 37 paquets d'effets confondus. Une fois estimés les $63 = 26 + 37$ effets non confondus et combinaisons linéaires d'effets confondus, il ne reste plus de degrés de liberté pour estimer la variance d'erreur. De même on constate que les effets bloc sont tous confondus avec des interactions. Il n'est donc pas possible de dégager des degrés de liberté pour l'estimation de la variance *inter-bloc* par comparaison à laquelle l'effet principal du facteur température T confondu avec les blocs est normalement testé.

En fait, les 19 autres solutions trouvées dans cette recherche ne permettent pas non plus de dégager de degrés de liberté pour l'estimation des variances inter ou intra-bloc. On est donc contraint, avec ce type de plan comportant un nombre maximum de facteurs à 2 niveaux, d'effectuer l'analyse par des techniques comme celles décrites et référencées dans [16], [18], avec le problème supplémentaire créé par la présence de deux strates inter et intra-bloc.

Une alternative est de diminuer le nombre de facteurs à 2 niveaux en essayant de

nom	:	HYGIEN1.REG
nb. d'unités	:	64
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	30 mn
- nb. sol. max.	:	20
- germe	:	398723
Inclusion des facteurs dans l'ensemble inéligible	:	oui

Fact. de base		Fact. définis			Parties de modèle	
fac.	nb. niv.	fac.	nb. niv.	bloc		
A	4	Bl	8	←	1	S1:A+B+C+D+E+F+G+H+I
B	4	T	2		2	S:S1+T
C	4	D	2		Modèles	
		E	2		1	Bl+S.S
		F	2		2	S.S
		G	2		Parties à estimer	
		H	2		1	S1
		I	2		2	T
					Hiérarchies	
					1	T : Bl

Définition du plan, solution 6, germe 520571533											
blocs	matrice clé										relations de définition
	*	*	*	T	D	E	F	G	H	I	
	Bl_1	Bl_2	Bl_3								$Bl_1 = A_2C_1$
A_1	0	1	0	1	1	0	1	0	1	0	$Bl_2 = A_1A_2B_1C_1$
A_2	1	1	0	0	1	1	1	1	0	1	$Bl_3 = B_1B_2C_1$
B_1	0	1	1	0	1	1	1	1	1	0	$T = A_1B_2C_1$
B_2	0	0	1	1	0	1	1	0	0	1	$D = A_1A_2B_1C_2$
C_1	1	1	1	1	0	0	1	1	1	1	$E = A_2B_1B_2C_2$
C_2	0	0	0	0	1	1	0	0	1	1	$F = A_1A_2B_1B_2C_1$
											$G = A_2B_1C_1$
											$H = A_1B_1C_1C_2$
											$I = A_2B_2C_1C_2$

TAB. 31 – Fraction de résolution 4 d'un $4^3 \times 2^7$ en 8 blocs de 8 unités

liste détaillée des ensembles d'effets confondus du modèle	
$[Bl_2]; C_1 C_2 D; B_2 F; A_1 G;$ $[Bl_3]; B_1 B_2 C_1; A_1 A_2 F;$ $[Bl_2 Bl_3]; A_1 A_2 B_2;$ $[Bl_1]; DH; A_2 C_1; B_1 G;$ $[Bl_1 Bl_2]; A_1 B_1; C_1 C_2 H;$ $[Bl_1 Bl_3]; C_2 E; A_2 B_1 B_2;$ $[Bl_1 Bl_2 Bl_3]; T;$ $DI; B_1 T; A_2 F;$ $C_1 C_2 I; A_2 B_2;$ $B_1 B_2 F; A_1 A_2 C_1;$ $C_1 F; HI; A_1 A_2 B_1 B_2; TG;$ $A_1 T; B_2 C_1;$ $B_1 B_2 T; C_2 H; A_1 A_2 G;$ $C_1 C_2 E; A_1 F; B_2 G;$ $DE; C_1 T; A_1 B_2; FG;$ $A_1 A_2 B_1; C_2 D;$ $A_2 T; B_1 F; EH;$ $EI; B_1 C_1; A_2 G;$	$A_1 C_1; B_2 T;$ $A_2 B_1; TF; C_1 G;$ $A_1 A_2 T; C_2 I; B_1 B_2 G;$ $TE; C_1 D; A_2 H;$ $B_2 I; A_2 C_1 C_2;$ $FI; C_1 H; A_2 D;$ $C_2 T; B_1 B_2 H; A_1 A_2 I;$ $B_1 B_2 I; A_1 A_2 H; C_2 G;$ $B_1 E; C_1 I; FH;$ $B_1 C_2; A_1 A_2 D;$ $A_1 E; C_1 C_2 F; B_2 D;$ $DF; B_2 C_1 C_2; A_2 I; EG;$ $A_2 C_2; B_1 B_2 E;$ $TD; B_1 I; C_1 E;$ $A_2 E; B_1 B_2 C_2; TH; GI;$ $A_1 A_2 C_2; TI; B_1 D; GH;$ $B_1 C_1 C_2; A_1 H;$ $A_1 C_1 C_2; EF; B_1 H; DG;$ $A_1 D; B_2 E; C_1 C_2 G;$
liste des effets non confondus	
$A_1; G; B_1; A_2; C_1; B_1 B_2; A_1 A_2; F; B_2;$ $> A_1 B_1 B_2; C_2; B_1 B_2 C_1 C_2; B_1 B_2 D; E; A_1 C_2;$ $> B_2 H; A_1 A_2 C_1 C_2; H; C_1 C_2 T; C_1 C_2; D; A_1 I;$ $> C_2 F; B_2 C_2; A_1 A_2 E; I;$	
liste des effets blocs non confondus (vide)	

TAB. 32 – Structure d'alias pour le plan du tableau 31

dégager des degrés de liberté pour estimer les deux erreurs. Pour l'erreur inter-bloc, on sait qu'un des effets bloc est confondu avec la température. On pourrait espérer imposer que les six autres effets soient non confondus, mais il est facile de montrer quand il y a trois facteurs A, B, C à quatre niveaux qu'au moins un effet de chacune des interactions AB, AC, BC est confondu avec les blocs.

★ Considérons par exemple l'interaction BC . Dans les expressions définissant les sept effets blocs en fonction des pseudofacteurs de base, A ne peut apparaître que sous les trois formes A_1, A_2, A_1A_2 . Il y a donc au moins un de ces effets bloc où il est absent car s'il apparaît dans deux d'entre eux sous une même forme, il ne figure pas dans le produit de ces deux. Cet effet bloc s'exprime donc à partir de B et C et les deux doivent apparaître puisqu'il ne peut être confondu avec un effet principal.

Il y a donc dans ce cas au moins quatre effets blocs confondus et au plus trois effets blocs non confondus. Pour être certain d'obtenir un dispositif donnant le nombre maximum d'effets blocs non confondus, on rajoute trois effets blocs à la partie à estimer numéro 0 du tableau 31, soit sous la forme $Bl_1, Bl_2, Bl_1.Bl_2$ de trois effets liés, soit sous la forme Bl_1, Bl_2, Bl_3 de trois effets indépendants (tab. 33). Dans les deux cas, les 4

	modèle		partie à estimer
1	$Bl+S.S$	ou	$Bl_1+Bl_2+Bl_1.Bl_2+S1$ $Bl_1+Bl_2+Bl_3+S1$
2	$S.S$		T

TAB. 33 – Modif. recherche tab. 31 pour estimer la variance inter-bloc

premiers facteurs à deux niveaux, dont T , sont trouvés très rapidement et la recherche s'éternise sur le 5ième facteur. On retient donc 4 facteurs à 2 niveaux en plus des trois à 4 niveaux et du facteur bloc à 8 niveaux.

Pour chacun des deux choix de partie à estimer 1, vingt solutions sont recherchées avec un germe non nul.

Lorsque la partie à estimer 1 est $Bl_1+Bl_2+Bl_1.Bl_2$, l'analyse des alias des 20 solutions fait apparaître une même structure d'alias (ce qui laisse penser qu'il n'y a en fait qu'une seule solution, à des permutations près des facteurs ou pseudofacteurs). Cette structure d'alias comprend 26 effets non confondus et 32 paquets d'effets confondus. Elle laisse donc $5 = 63 - (26 + 32)$ degrés de liberté pour estimer la variance résiduelle et comporte par construction trois degrés de liberté pour estimer la variance inter-bloc.

Lorsque la partie à estimer 1 est $Bl_1+Bl_2+Bl_3$, l'analyse des alias des 20 solutions fait apparaître plusieurs structures. Outre les trois effets bloc non confondus, ces structures comprennent en général 32 ou 33 effets traitement non confondus, 24 ou 25 paquets d'effets confondus et laissent 3 ou 4 degrés de liberté pour estimer la variance résiduelle. Cependant l'une d'entre elle, la solution 6, se distingue nettement des 19 autres: elle comprend 42 effets traitement non confondus et 17 paquets d'effets. Elle laisse donc seulement 1 degré de liberté pour la résiduelle, mais apparaît par ailleurs nettement supérieure aux autres solutions. Les tableaux 34 et 35 précisent cette solution et donne sa structure d'alias.

Rien n'interdit bien sûr de dupliquer un bloc d'un tel dispositif. Cela donne un degré de liberté supplémentaire pour la variance inter et 7 pour la variance résiduelle intra-bloc. Ces degrés de liberté donnent des variances d'erreur "*pures*" en ce sens qu'elles ne sont pas gonflées par d'éventuelles interactions de trois facteurs ou plus. Cependant l'analyse statistique du dispositif non régulier ainsi obtenu est plus complexe et l'intérêt pratique de ce genre de répétition faible dans un contexte de criblage des facteurs influents, où l'absence d'effets de plusieurs facteurs permet a contrario de dégager sans peine les facteurs actifs.

La prise en compte des blocs dans les plans de résolution 3 considéré au tableau 30 diminue le nombre de facteurs à 2 niveaux qu'il est possible d'introduire. Avec 4 facteurs à 4 niveaux, la recherche du 12ième facteur à 2 niveaux n'a toujours pas débouché après 10 mn. L'objectif initial d'étude de 16 facteurs, 4 à 4 niveaux et 12 à 2 niveaux, ne semble donc pas accessible en résolution 3 avec la contrainte d'estimabilité des effets de type A_1A_2 garantissant l'atteinte de la résolution 4 par doublement du plan.

Le tableau 36 donne les paramètres de la recherche et une solution pour 11 facteurs à 2 niveaux. La paire modèle-partie à estimer 1 impose que T , qui figure dans le modèle $B1+S$, ne soit confondu avec aucun des autres facteurs figurant dans $S1$. Il est donc inutile de rajouter une paire (S, T) analogue à la paire $(S.S, T)$ introduite dans les recherches de plans de résolution 4.

Le tableau 37 montre comment obtenir le plan doublé en introduisant directement ses relations de définition dans la zone des *facteurs prédéfinis*. Ce plan comporte un pseudofacteur bloc supplémentaire $B14$ prenant le niveau 1 sur le premier plan, -1 sur le second. Les relations de définition de ce plan doublé se déduisent immédiatement de celles du plan initial en ajoutant $B14$ à toutes les sommes d'un nombre pair de pseudofacteurs.

4.2 Mélange de facteurs quantitatifs et qualitatifs

4.2.1 Effets polynomiaux et pseudofactoriels

Supposons maintenant qu'au moins un des facteurs à 4 niveaux, disons A , est quantitatif. Il y a alors dissymétrie entre les trois effets A_1, A_2, A_1A_2 associés et il est possible de jouer sur cette dissymétrie pour sélectionner le plan. Nous décrivons ici de façon sommaire la façon de procéder. On trouvera dans [16], [6] une description plus détaillée et en particulier la définition précise des effets polynomiaux.

Par *effets polynomiaux* nous désignons ici tous les effets, quelles que soient les natures qualitatives ou quantitatives des facteurs qu'ils contiennent, où les facteurs quantitatifs apparaissent sous forme polynomiale. Par exemple, si A est quantitatif à 4 niveaux, B qualitatif à 2 niveaux, les effets polynomiaux sont les effets principaux $B, \text{lin } A, \text{quad } A, \text{cub } A$ et les interactions $\text{lin } A.B, \text{quad } A.B, \text{cub } A.B$. Ils sont à distinguer des *effets pseudofactoriels* associés aux produits entre pseudofacteurs: $B, A_1, A_2, A_1A_2, A_1.B, A_2.B, A_1A_2.B$. Lorsqu'il y a des facteurs quantitatifs, ce sont les effets polynomiaux qui ont vraiment du sens, servent à formaliser les hypothèses et doivent être estimés.

nom	:	HYGIEN3.REG
nb. d'unités	:	64
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	15 mn
- nb. sol. max.	:	3
- germe	:	6648718
Inclusion des facteurs dans l'ensemble inéligible	:	oui

Fact. de base		Fact. définis			Parties de modèle	
fac.	nb. niv.	fac.	nb. niv.	bloc		
A	4	Bl	8	←	1 S1:A+B+C+D+E+F+G+H+I+J+K+L+M+N	
B	4	T	2		2 S:S1+T	
C	4	D	4		Modèles	
		E	2		1 Bl+S	
		F	2		2 S.S	
		G	2		Parties à estimer	
		H	2		1 S1	
		I	2		2 A ₁ A ₂ +B ₁ B ₂ +C ₁ C ₂ +D ₁ D ₂	
		J	2		Hiérarchies	
		K	2		1 T : Bl	
		L	2			
		M	2			
		N	2			

Définition du plan, solution 6, Germe 520571533																	
matrice clé														relations de définition			
blocs	*	*	*	T	D ₁	D ₂	E	F	G	H	I	J	K	L	M	N	
	Bl ₁	Bl ₂	Bl ₃														$Bl_1 = A_1 C_1 C_2$
A ₁	1	0	1	1	0	1	0	0	0	1	1	0	1	1	1	0	$Bl_2 = B_1 B_2 C_2$
A ₂	0	0	0	0	1	0	1	0	1	0	0	1	0	1	1	0	$Bl_3 = A_1 B_1 C_2$
B ₁	0	1	1	0	0	1	1	1	1	1	0	0	1	0	1	0	$T = A_1 B_2$
B ₂	0	1	0	1	1	0	0	0	1	1	0	0	0	1	0	1	$D_1 = A_2 B_2 C_1$
C ₁	1	0	0	0	1	1	0	0	1	1	1	0	1	1	1	1	$D_2 = A_1 B_1 C_1 C_2$
C ₂	1	1	1	0	0	1	0	1	0	1	0	1	0	1	0	0	$E = A_2 B_1$
																	$F = B_1 C_2$
																	$G = A_2 B_1 B_2 C_1$
																	$H = A_1 B_1 B_2 C_1 C_2$
																	$I = A_1 C_1$
																	$J = A_2 C_2$
																	$K = A_1 B_1 C_1$
																	$L = A_1 A_2 B_2 C_1 C_2$
																	$M = A_1 A_2 B_1 C_1$
																	$N = B_2 C_1$

TAB. 36 – Fraction résol. 3 doublable en résol. 4 d'un $4^4 \times 2^{11}$ en 8 blocs de 8 unités

nom	:	HYGIEN4.REG
nb. d'unités	:	128
définition des facteurs de base	:	utilisateur

Parties de modèle	
1	S1 : A+B+C+D+E+F+G+H+I+J+K+L+M+N
2	S : S1+T
Modèles	
1	B1.B14+S.S

Fact. de base			Fact. définis			Facteurs prédéfinis	
fac.	nb. bloc niv.		fac.	nb. bloc niv.			
A	4		B1	8	←	1	$B1_1 : A_1 + C_1 + C_2$
B	4		T	2		2	$B1_2 : B_1 + B_2 + C_2$
C	4		D	2		3	$B1_3 : A_1 + B_1 + C_2$
B14	2	←	...			4	$T : A_1 + B_2 + B14$
			N	2		5	$D_1 : A_2 + B_2 + C_1$
						6	$D_2 : A_1 + B_1 + C_1 + C_2 + B14$
						7	$E : A_2 + B_1 + B14$
						8	$F : B_1 + C_2 + B14$
						9	$G : A_2 + B_1 + B_2 + C_1 + B14$
						10	$H : A_1 + B_1 + B_2 + C_1 + C_2$
						11	$I : A_1 + C_1 + B14$
						12	$J : A_2 + C_2 + B14$
						13	$K : A_1 + B_1 + C_1$
						14	$L : A_1 + A_2 + B_2 + C_1 + C_2$
						15	$M : A_1 + A_2 + B_1 + C_1 + B14$
						16	$N : B_2 + C_1 + B14$

TAB. 37 – Plan double du plan du tableau 36, de résolution 4

La méthode de construction utilisée par PLANOR n'est pas directement adaptée aux facteurs quantitatifs et à l'estimation des effets polynomiaux. Cependant, si l'on choisit la correspondance entre les niveaux quantitatifs et les niveaux des pseudofacteurs d'une façon appropriée, les effets pseudofactoriels s'expriment en fonction des effets polynomiaux sous une forme simple, qui peut être exploitée pour chercher des relations de définition entre pseudofacteurs adaptées à la nature quantitative de certains facteurs.

Plus précisément, pour chaque facteur quantitatif à 4 niveaux, la correspondance entre niveaux ordonnés et niveaux des pseudofacteurs est choisie comme indiqué au tableau 38. Ce choix correspond au choix standard de PLANOR et il n'est donc pas nécessaire de passer par la redéfinition des niveaux pour l'obtenir.

Les niveaux ordonnés apparaissant dans le tableau 38 sont les quatre premiers entiers 0, 1, 2, 3, auxquels on peut toujours se ramener par changement d'origine et d'échelle. La correspondance du tableau 38 induit les relations figurant dans le tableau 39.

	notation additive		notation multiplicative	
A	A_1	A_2	A_1	A_2
0	1	1	-1	-1
1	1	0	-1	1
2	0	1	1	-1
3	0	0	1	1

TAB. 38 – *Décomposition en pseudofacteurs d'un facteur A quantitatif*

$$\begin{array}{ll}
 A_1 & = (2 \operatorname{lin} A - \operatorname{cub} A)/\sqrt{5} & \operatorname{lin} A & = (2A_1 + A_2)/\sqrt{5} \\
 A_1 A_2 & = \operatorname{quad} A & \operatorname{quad} A & = A_1 A_2 \\
 A_2 & = (\operatorname{lin} A + 2 \operatorname{cub} A)/\sqrt{5} & \operatorname{cub} A & = (-A_1 + 2A_2)/\sqrt{5}
 \end{array}$$

TAB. 39 – *Relations entre effets principaux polynomiaux et pseudofactoriels*

Ces relations sont utilisées pour exprimer un effet pseudofactoriel en fonction des effets polynomiaux ou réciproquement un effet polynomial en fonction des effets pseudofactoriels. Les produits obtenus après substitution des termes figurant à gauche des égalités dans le tableau 39 par ceux figurant à droite peuvent ensuite être développés de la façon usuelle. Par exemple, si B est aussi un facteur quantitatif à 4 niveaux décomposé de façon analogue à A et C un facteur à deux niveaux:

$$\begin{aligned}
 A_2 C &= (\operatorname{lin} A + 2 \operatorname{cub} A)C/\sqrt{5} = (\operatorname{lin} A.C + 2 \operatorname{cub} A.C)/\sqrt{5}, \\
 A_1 B_1 &= (2 \operatorname{lin} A - \operatorname{cub} A)(2 \operatorname{lin} B - \operatorname{cub} B)/5 \\
 &= (4 \operatorname{lin} A \operatorname{lin} B - 2 \operatorname{cub} A \operatorname{lin} B - 2 \operatorname{lin} A \operatorname{cub} B + \operatorname{cub} A \operatorname{cub} B)/5, \\
 A_1 A_2 B_1 &= \operatorname{quad} A(2 \operatorname{lin} B - \operatorname{cub} B)/\sqrt{5} \\
 &= (2 \operatorname{quad} A \operatorname{lin} B - \operatorname{quad} A \operatorname{cub} B)/\sqrt{5} \\
 \operatorname{lin} A \operatorname{lin} B &= (2A_1 + A_2)(2B_1 + B_2)/5 \\
 &= (4A_1 B_1 + 2A_2 B_1 + 2A_1 B_2 + A_2 B_2)/5.
 \end{aligned}$$

Ces relations et les hypothèses effectuées sur les effets polynomiaux permettent de

hiérarchiser les différents effets pseudofactoriels et suggèrent des stratégies de recherche de plans réguliers.

Lorsque tous les facteurs sont qualitatifs, le principal critère de classification d'un effet factoriel est le nombre de facteurs qui y apparaissent. Une hypothèse réaliste, lorsqu'on dispose de peu d'information a priori, est qu'un tel effet est d'autant plus faible que ce nombre est plus grand. En particulier, les interactions de trois facteurs et plus sont souvent supposées nulles, ce qui amène à rechercher des plans de résolution 5.

Ce nombre de facteurs figurant dans l'effet reste un élément important pour classer les effets polynomiaux dans le cas où il y a des facteurs quantitatifs. Mais un autre critère intervient dans ce cas: le degré de l'effet polynomial. Ce degré est 1 pour un facteur qualitatif et égal au degré polynomial pour un facteur quantitatif, c'est à dire à 1 pour un effet linéaire, 2 pour un effet quadratique, 3 pour un effet cubique. Le degré d'un effet faisant intervenir plusieurs facteurs est la somme des degrés en chaque facteur. Ainsi les degrés de lin A quad B , cub $A.C$ sont respectivement $3 = 1 + 2$ et $4 = 3 + 1$. Une hypothèse fréquemment formulée est que les effets polynomiaux de degré 3 ou plus sont nuls. Une hypothèse moins restrictive est que seuls sont nuls les effets polynomiaux de degré 3 n'appartenant pas aux effets principaux, c'est à dire impliquant au moins deux facteurs.

Deux approches peuvent être utilisées pour exploiter les hypothèses sur les effets polynomiaux. Dans la première [13], on s'appuie exclusivement sur les effets pseudofactoriels nuls pour trouver un plan régulier. Le plan résultant est orthogonal vis à vis du modèle polynomial et permet d'estimer les effets polynomiaux avec une efficacité de 1, c'est à dire avec une variance identique, à la correction pour la taille près, à celle d'un plan factoriel complet. Dans la seconde [8], une utilisation plus fine des relations du tableau 39 permet d'obtenir des plans ayant un nombre d'unités beaucoup plus faible. Ceux-ci ne sont plus orthogonaux, mais conservent néanmoins d'excellentes efficacités pour estimer les effets polynomiaux.

Les deux approches conduisent en général à des plans ayant des propriétés de distribution uniforme des points qui les rendent assez robustes vis à vis du modèle sélectionné et pour cette raison préférable dans de nombreux cas aux plans D -optimaux obtenus par des procédures algorithmiques. La description que nous donnons de ces approches ici a été reprise ultérieurement dans [16] auquel on pourra se reporter pour un complément d'information.

4.2.2 Plans réguliers orthogonaux pour le modèle polynomial

Sous l'hypothèse que les effets polynomiaux de degré 3 ou plus sont nuls, les effets pseudofactoriels s'expriment uniquement à partir d'effets polynomiaux de degré 3 ou plus sont également nuls. En incluant les autres effets pseudofactoriels, non nuls, dans le modèle et les termes à estimer, on est alors certain d'obtenir un plan permettant d'estimer tous les effets polynomiaux de degré 1 ou 2. Le plan obtenu est orthogonal vis à vis de ces effets polynomiaux qu'il estime avec la même efficacité qu'un plan factoriel complet, c'est à dire avec une efficacité de 1. Cette efficacité n'est pas toujours dans ce cas l'efficacité maximum, mais les plans ainsi construits ont une très bonne efficacité globale et aussi

une bonne robustesse vis à vis du modèle.

Considérons par exemple un cas où il y a 16 unités, deux facteurs quantitatifs A et B à 4 niveaux, un facteur qualitatif C à deux niveaux.

Si les effets polynomiaux de degré 3 et plus sont nuls, il en est de même de tous les effets pseudofactoriels contenant 3 symboles ou plus. Par exemple, $A_1A_2B_1 = (2 \text{ quad } A \text{ lin } B - \text{quad } A \text{ cub } B)/\sqrt{5}$ est nul puisque $\text{quad } A \text{ lin } B$ et $\text{quad } A \text{ cub } B$, de degrés 3 et 5 respectivement, sont tous deux nuls.

Le plan de résolution 5 défini par $A_1A_2B_1B_2C = 1$ permet alors d'estimer tous les effets pseudofactoriels à 1 ou 2 symboles, et par suite tous les effets polynomiaux de degré 1 ou 2. Il est donc adapté à la nature quantitative des facteurs A et B . Ce plan peut s'obtenir dans PLANOR en définissant le modèle et les termes à estimer par $P.P$ où P est la partie de modèle définie par $P : A_{-1} + A_{-2} + B_{-1} + B_{-2} + C$.

De façon générale, on obtient facilement le plus petit degré des effets polynomiaux en fonction desquels s'exprime un effet pseudofactoriel. Par exemple, si A, B sont quantitatifs à 4 niveaux, C qualitatif à 4 niveaux, ce plus petit degré est 4 pour l'effet $A_1B_1B_2C_1C_2$. On totalise pour l'obtenir le nombre de pseudofacteurs issus de facteurs quantitatifs et le nombre de facteurs qualitatifs, soit ici 3 pour les trois pseudofacteurs A_1, B_1, B_2 et 1 pour le facteur qualitatif C .

De ce mode de calcul, il résulte que si le dispositif est de résolution 5 lorsqu'on assimile à des facteurs les pseudofacteurs issus de facteurs quantitatifs, il permet d'estimer orthogonalement tous les effets d'un modèle polynomial de degré 2.

Par exemple avec 64 unités, il est possible d'étudier, en résolution 5, jusqu'à 8 facteurs à 2 niveaux. La substitution à certains facteurs de pseudofacteurs issus de facteurs quantitatifs à 4 niveaux permet d'obtenir des plans orthogonaux pour un modèle de degré 2 avec 1, 2, 3 ou 4 facteurs quantitatifs à 4 niveaux et respectivement 6, 4, 2 ou 0 facteurs à deux niveaux. Le tableau 40 donne un exemple avec 2 facteurs quantitatifs A, B à 4 niveaux, 4 facteurs C, D, E, F à deux niveaux. Un facteur bloc BL à 4 niveaux a été rajouté. Enfin on a inclus dans le modèle et les termes à estimer les effets principaux de A et B pour permettre l'estimation des effets cub A et cub B . On aurait pu, sans restreindre le choix, prendre A, B plus deux des facteurs à 2 niveaux comme facteurs de base. Mais c'est l'option standard qui a été ici retenue pour la définition des facteurs de base, afin d'illustrer les sorties correspondantes.

4.2.3 Plans réguliers non orthogonaux pour le modèle polynomial

4.2.3.1 Fraction 1/2 d'un 4×4 pour deux facteurs quantitatifs Considérons le cas où il y a deux facteurs quantitatifs A et B à 4 niveaux et un modèle supposé de degré 2. Ce modèle a 6 paramètres: la constante, lin A , lin B , quad A , quad B et lin A lin B . Mais il n'y a que 5 effets pseudofactoriels dont l'expression en fonction des effets polynomiaux ne fait apparaître aucun de ces paramètres et qui sont donc nuls: $A_1A_2B_1, A_1A_2B_2, A_1B_1B_2, A_2B_1B_2, A_1A_2B_1B_2$. La méthode décrite dans le paragraphe précédent ne permet pas dans ce cas de trouver une fraction 1/2 adéquate, car une telle

nom	:	QUANT1
nb. d'unités	:	64
définition des facteurs de base (qui définissent l'unité)	:	standard
type de décomposition des facteurs	:	maximum
...		

Facteurs à définir		
fac.	nb. niv.	bloc
A	4	
B	4	
C	2	
D	2	
E	2	
F	2	
BL	4	←

Parties de modèle	
P: A ₁ +A ₂ +B ₁ +B ₂ +C+D+E+F	
Modèles	
P.P+A+B+BL	
Parties à estimer	
P.P+A+B+BL	

Définition du plan											
matrice clé										relations de définition	
	A ₁	A ₂	B ₁	B ₂	C	D	E	F	BL ₁		BL ₂
2 ₁	1	0	0	0	1	0	0	1	1	0	C = A ₁ A ₂ B ₁ B ₂
2 ₂	0	1	0	0	1	0	0	1	0	1	F = A ₁ A ₂ DE
2 ₃	0	0	1	0	1	0	0	0	1	0	BL ₁ = A ₁ B ₁ D
2 ₄	0	0	0	1	1	0	0	0	0	1	BL ₂ = A ₂ B ₂ E
2 ₅	0	0	0	0	0	1	0	1	1	0	
2 ₆	0	0	0	0	0	0	1	1	0	1	

TAB. 40 – Fraction 1/4 d'un $4 \times 4 \times 2^4$, adaptée à des fact. à 4 niveaux quantitatifs

fraction devrait permettre d'estimer $11 = 16 - 5$ paramètres avec 8 unités seulement.

Compte tenu de la nullité supposée des effets polynomiaux de degré 3 ou plus, les effets pseudofactoriels s'expriment sous la forme donnée dans le tableau 41.

$$\begin{array}{ll}
 A_1 & = 2 \operatorname{lin} A / \sqrt{5} & B_1 & = 2 \operatorname{lin} B / \sqrt{5} \\
 A_1 A_2 & = \operatorname{quad} A & B_1 B_2 & = \operatorname{quad} B \\
 A_2 & = \operatorname{lin} A / \sqrt{5} & B_2 & = \operatorname{lin} B / \sqrt{5} \\
 \\
 & & A_1 B_1 & = 4 \operatorname{lin} A \operatorname{lin} B / 5 \\
 & & A_1 B_2 & = 2 \operatorname{lin} A \operatorname{lin} B / 5 \\
 & & A_2 B_1 & = 2 \operatorname{lin} A \operatorname{lin} B / 5 \\
 & & A_2 B_2 & = \operatorname{lin} A \operatorname{lin} B / 5
 \end{array}$$

TAB. 41 – Expression des effets pseudofactoriels dans un modèle de degré 2

Il est clair que $\operatorname{lin} A$ peut être estimé à partir d'un seul des deux effets A_1 ou A_2 . Ces deux effets n'apportent cependant pas la même information. A partir d'une estimation $\hat{e}(A_1)$ de variance σ^2/n (où n est la taille de la fraction), on obtient une estimation $\sqrt{5}\hat{e}(A_1)/2$ de $\operatorname{lin} A$ de variance $(5/4)\sigma^2/n$, tandis qu'à partir d'une estimation $\hat{e}(A_2)$ de même variance σ^2/n , on obtient une estimation $\sqrt{5}\hat{e}(A_2)$ de $\operatorname{lin} A$ de variance $5\sigma^2/n$ quatre fois supérieure. Il est donc préférable de confondre A_2 plutôt que A_1 sur la fraction.

Une même remarque s'applique à $\operatorname{lin} B$ et à $\operatorname{lin} A \operatorname{lin} B$. Ainsi ce dernier effet peut être estimé à partir de chacun des effets $A_1 B_1, A_1 B_2, A_2 B_1, A_2 B_2$. Mais la variance d'estimation, égale à $(25/16)\sigma^2/n$ si l'on part d'une estimation $\hat{e}(A_1 B_1)$ de variance σ^2/n , devient $(25/4)\sigma^2/n$ si l'on part de $A_1 B_2$ ou $A_2 B_1$ et $25\sigma^2/n$ si l'on part de $A_2 B_2$. Les informations apportées, par définition inversement proportionnelles aux variances, sont dans les rapports 16, 4, 4, 1. Lorsqu'on combine plusieurs estimations, les informations se cumulent. Par exemple si on dispose d'estimations indépendantes $\hat{e}(A_1 B_1), \hat{e}(A_1 B_2)$ de variance σ^2/n , l'estimation combinée de $\operatorname{lin} A \operatorname{lin} B$

$$\frac{16}{20} \frac{5}{4} \hat{e}(A_1 B_1) + \frac{4}{20} \frac{5}{2} \hat{e}(A_1 B_2)$$

a pour variance $(25/(16+4))\sigma^2/n$. Il faut donc confondre en priorité les effets pseudofactoriels apportant le minimum d'information, c'est à dire $A_2 B_2$ et si nécessaire $A_1 B_2, A_2 B_1$.

Ces considérations amène à rechercher une fraction 1/2 du plan factoriel complet de taille $16 = 4 \times 4$ ne confondant les effets pseudofactoriels importants $A_1, A_1 A_2, B_1, B_1 B_2, A_1 B_1$ qu'avec les 5 effets nuls. Le modèle et la partie à estimer correspondants apparaissent sous l'intitulé recherche n0 1 dans le tableau 42. Le modèle contient tous les effets non nuls, à savoir ici ceux où apparaissent au plus 2 pseudofacteurs. Cette recherche n'aboutit malheureusement pas.

Il faut donc assouplir la demande en autorisant la confusion des effets importants suscités avec des effets mineurs tels que $A_2 B_1, A_1 B_2, A_2 B_2$. Le modèle et la partie à estimer correspondants sont précisés sous le titre recherche n0 2 dans le tableau 42. L'étude des

nom	: QUANT2
nb. d'unités	: 8
définition des facteurs de base	: standard
type de décomposition des facteurs	: maximum
Recherche backtrack - temps maxi	: 10 mn
- nb. sol. max.	: 999
...	

Fact. définis		Recherche n0 1	Recherche n0 2
fac.	nb. niv.	Parties de modèle	Modèle
		P: $A_1+A_2+B_1+B_2$	$A_1.A_2+B_1.B_2+A_1.B_1$
A	4	Modèle	Partie à estimer
B	4	P.P	$A_1+A_1.A_2+B_1+B_1.B_2+A_1.B_1$
		Partie à estimer	Relation de définition
		$A_1+A_1.A_2+B_1+B_1.B_2+A_1.B_1$	$A_2B_2 = 1$
		ÉCHEC DE LA RECHERCHE	
		Recherche arrêtée sur le facteur B ₂	

TAB. 42 – Fraction 1/2 d'un 4×4 , avec des facteurs à 4 niveaux quantitatifs

alias des solutions obtenues montre qu'il n'y a en fait qu'une solution définie par la relation $A_2B_2 = 1$. Pour cette solution, les combinaisons linéaires estimables de base sont

$$\begin{aligned}
E(\langle y, \mathbf{1} \rangle / 8) &= e(\mathbf{1}) && + e(A_2.B_2) \\
E(\langle y, A_1 \rangle / 8) &= e(A_1) && + e(A_1.A_2.B_2) \\
E(\langle y, A_2 \rangle / 8) &= e(A_2) && + e(B_2) \\
E(\langle y, A_1.A_2 \rangle / 8) &= e(A_1.A_2) && + e(A_1.B_2) \\
E(\langle y, B_1 \rangle / 8) &= e(B_1) && + e(A_2.B_1.B_2) \\
E(\langle y, A_1.B_1 \rangle / 8) &= e(A_1.B_1) && + e(A_1.A_2.B_1.B_2) \\
E(\langle y, A_2.B_1 \rangle / 8) &= e(A_2.B_1) && + e(B_1.B_2) \\
E(\langle y, A_1.A_2.B_1 \rangle / 8) &= e(A_1.A_2.B_1) && + e(A_1.B_1.B_2),
\end{aligned} \tag{7}$$

où y est le vecteur des 8 observations et $\langle \rangle$ dénote le produit scalaire usuel de \mathbb{R}^8 . Les vecteurs $\mathbf{1}, A_1, \dots, A_1A_2, \dots$ figurant dans les produits scalaires sont les vecteurs de 1 et -1 naturellement associés en notation multiplicative aux pseudofacteurs ou produits de pseudofacteurs correspondants.

Compte tenu des égalités du tableau 41 et de la nullité des effets pseudofactoriels comprenant 3 ou 4 symboles, le système (7) se réécrit sous la forme

$$\begin{aligned}
E(\langle y, \mathbf{1} \rangle / 8) &= e(\mathbf{1}) && + \text{lin } A \text{ lin } B / 5 \\
E(\langle y, A_1 \rangle / 8) &= 2 \text{ lin } A / \sqrt{5} \\
E(\langle y, A_2 \rangle / 8) &= \text{lin } A / \sqrt{5} && + \text{lin } B / \sqrt{5} \\
E(\langle y, A_1.A_2 \rangle / 8) &= \text{quad } A && + 2 \text{ lin } A \text{ lin } B / 5 \\
E(\langle y, B_1 \rangle / 8) &= 2 \text{ lin } B / \sqrt{5} \\
E(\langle y, A_1.B_1 \rangle / 8) &= 4 \text{ lin } A \text{ lin } B / 5 \\
E(\langle y, A_2.B_1 \rangle / 8) &= 2 \text{ lin } A \text{ lin } B / 5 && + \text{quad } B \\
E(\langle y, A_1.A_2.B_1 \rangle / 8) &= 0
\end{aligned} \tag{8}$$

Les produits scalaires de y avec A_1 , B_1 , $A_1.B_1$ permettent d'estimer directement les effets linéaires $\text{lin } A$ et $\text{lin } B$ et l'interaction $\text{lin } A \text{ lin } B$. Les estimations de la moyenne générale et des deux effets quadratiques sont déduites de ces estimations et des produits scalaires avec $\mathbf{1}$, $A_1.A_2$, $A_2.B_1$.

Le tableau 43 donne les estimations ainsi obtenues, ainsi que leurs variances et efficacités associées. Le calcul des variances est basé sur le fait que les produits scalaires à gauche de (8), $\langle y, \mathbf{1} \rangle / 8$, $\langle y, A_1 \rangle / 8$, $\langle y, A_2 \rangle / 8$, ... sont non corrélés de variance $\sigma^2/8$. La comparaison des variances ainsi obtenues à la variance $\sigma^2/16$ des effets polynomiaux dans le plan factoriel complet donne, après correction pour tenir compte du fait qu'il y a moitié moins d'observations dans la fraction, les efficacités figurant à droite du tableau 43.

La notation $e()$ est utilisée pour représenter les effets polynomiaux dans ce tableau. Ainsi l'effet linéaire $\text{lin } A$, l'interaction $\text{lin } A \text{ lin } B$ y sont représentés par $e(\text{lin } A)$, $e(\text{lin } A \text{ lin } B)$. Cette notation, adoptée dans [6] pour distinguer les pseudofacteurs polynomiaux de leurs effets, est également commode pour représenter les estimations en mettant un chapeau ou un tilde sur le e .

	estimation	variance	efficacité
$\hat{e}(\text{lin } A)$	$= (\sqrt{5}/2) \langle y, A_1 \rangle / 8$	$(5/4) \sigma^2/8$	$4/5$
$\hat{e}(\text{lin } B)$	$= (\sqrt{5}/2) \langle y, B_1 \rangle / 8$	$(5/4) \sigma^2/8$	$4/5$
$\hat{e}(\text{lin } A \text{ lin } B)$	$= (5/4) \langle y, A_1 B_1 \rangle / 8$	$(25/16) \sigma^2/8$	$16/25$
$\hat{e}(\mathbf{1})$	$= \left(\langle y, \mathbf{1} \rangle - (1/4) \langle y, A_1 B_1 \rangle \right) / 8$	$(17/16) \sigma^2/8$	$16/17$
$\hat{e}(\text{quad } A)$	$= \left(\langle y, A_1 A_2 \rangle - (1/2) \langle y, A_1 B_1 \rangle \right) / 8$	$(5/4) \sigma^2/8$	$4/5$
$\hat{e}(\text{quad } B)$	$= \left(\langle y, A_2 B_1 \rangle - (1/2) \langle y, A_1 B_1 \rangle \right) / 8$	$(5/4) \sigma^2/8$	$4/5$

TAB. 43 – Estimations des effets polynomiaux dans la fraction $4^2/2$

L'information apportée par $\langle y, A_2 \rangle / 8$ sur $\text{lin } A$ and $\text{lin } B$ n'a pas été utilisée. Elle peut être utilisée pour obtenir les estimateurs des moindres carrés de ces deux paramètres. Il n'y a à prendre en compte pour faire le calcul que les lignes associées aux produits scalaires $\langle y, A_1 \rangle / 8$, $\langle y, A_2 \rangle / 8$, $\langle y, B_1 \rangle / 8$ qui se réécrivent sous la forme matricielle suivante

$$E \begin{bmatrix} \langle y, A_1 \rangle / 8 \\ \langle y, A_2 \rangle / 8 \\ \langle y, B_1 \rangle / 8 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 0 \\ 1 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \text{lin } A \\ \text{lin } B \end{bmatrix},$$

et conduisent aux estimations données dans le tableau 44.

Au lieu d'améliorer l'estimation de $\text{lin } A$ et $\text{lin } B$ avec l'information apportée par $\langle y, A_2 \rangle / 8$, on peut utiliser le pseudofacteur A_2 pour diviser le dispositif en deux blocs. Si on note C le facteur bloc, on a $C = A_2$ et la ligne associée à A_2 dans le système (8) devient

$$E (\langle y, A_2 \rangle / 8) = \text{lin } A / \sqrt{5} + \text{lin } B / \sqrt{5} + e(C).$$

estimation	variance	efficacité
$\tilde{e}(\text{lin } A) = (\sqrt{5}/12) \left(5 \langle y, A_1 \rangle + 2 \langle y, A_2 \rangle - \langle y, B_1 \rangle \right) / 8$	$(25/24)\sigma^2/8$	24/25
$\tilde{e}(\text{lin } B) = (\sqrt{5}/12) \left(- \langle y, A_1 \rangle + 2 \langle y, A_2 \rangle + 5 \langle y, B_1 \rangle \right) / 8$	$(25/24)\sigma^2/8$	24/25

TAB. 44 – Estimations des moindres carrés des effets linéaires dans la fraction $4^2/2$

L'information apportée par $\langle y, A_2 \rangle / 8$ peut donc être utilisée pour estimer l'effet bloc $e(C)$:

$$\hat{e}(C) = \langle y, A_2 \rangle / 8 - \left(\hat{e}(\text{lin } A) + \hat{e}(\text{lin } B) \right) / \sqrt{5}$$

Cette répartition en deux blocs est utile si la perte d'information sur les effets linéaires qu'elle induit est compensée par la diminution de la variance résiduelle σ^2 .

On peut aussi utiliser le pseudofacteur $A_1 A_2 B_1$ pour définir un second système de deux blocs croisé avec le premier. L'inconvénient est qu'il ne reste alors plus aucun degré de liberté pour estimer l'erreur.

L'intérêt pratique de ce petit exemple est limité, mais il montre bien la souplesse de cette méthode de construction. L'exemple du paragraphe suivant, extrait de [6], illustre l'emploi de cette méthode pour créer un plan de taille nettement plus importante.

4.2.3.2 Fraction $1/16$ d'un $4^3 2^4$ avec deux facteurs quantitatifs Pour optimiser le milieu de culture d'un Rhizobium symbiote du soja (*Bradyrhizobium japonicum*), on étudie 7 facteurs de composition de ce milieu, dont trois à 4 niveaux et 4 à 2 niveaux. Sur les trois facteurs à 4 niveaux, l'un A est qualitatif, les deux autres B et C quantitatifs. Le tableau 45 précise la recherche effectuée.

Le premier couple *modèle, partie à estimer* permet de s'assurer que les effets principaux et interactions majeures entre deux facteurs ne sont pas confondus entre eux. Sont considérées comme majeures toutes les interactions où n'apparaissent aucun des deux pseudofacteurs B_2 et C_2 .

Le second couple *modèle, partie à estimer* impose en outre au plan d'être de résolution 4, c'est à dire de permettre l'estimation de tous les effets principaux dans un modèle comportant toutes les interactions de deux facteurs. L'aspect quantitatif n'est pas pris en compte dans ce second couple, introduit pour assurer un maximum de robustesse à l'estimation des effets principaux de tous degrés.

La recherche exhaustive complète, requise en donnant 999 comme nombre maximum de solutions, conduit à un ensemble de 1152 solutions. L'examen de ces solutions, mené par un programme spécifique, a montré qu'elles étaient toutes dérivées de 6 solutions de base en permutant entre eux les facteurs ou pseudofacteurs ayant des rôles symétriques, c'est à dire D, E, F, G d'abord, ensuite B et C et finalement $A_1, A_2, A_1 A_2$. Ces 6 solutions de base sont données dans le tableau 46, où sont également reportées trois mesures globales d'efficacité permettant de comparer ces solutions du point de vue variance.

Le tableau 47 précise les efficacités, pour chaque effet, des trois plans les meilleurs

nom	:	RHIZO6
nb. d'unités	:	64
définition des facteurs de base (qui définissent l'unité)	:	utilisateur
type de décomposition des facteurs (en pseudofacteurs)	:	maximum
Recherche backtrack - temps maxi	:	20 mn
- nb. sol. max.	:	999
- germe	:	0
Inclusion des facteurs dans l'ensemble inéligible	:	oui

Fact. de base		Facteurs à définir		Parties de modèle	
fac.	nb. niv.	fac.	nb. niv.	1	2
A	4	D	2	p: A+B ₁ +C ₁ +D+E+F+G	
B	4	E	2	q: A+B+C+D+E+F+G	
C	4	F	2	Modèles	
		G	2	1	p.p
				2	q.q
				Parties à estimer	
				1	p.p
				2	q

TAB. 45 – Fraction 1/16 d'un $4^3 2^4$ incluant 2 fact. quant. à 4 niv.

pour l'efficacité globale. Dans le troisième, les interactions incluant A , qui ont chacune trois degrés de liberté, sont caractérisées par trois efficacités, appelées efficacités principales. La définition précise de ces efficacités est donnée dans [15] et [16]. La plus petite de ces 3 efficacités est la borne inférieure des efficacités pour tous les contrastes appartenant à cette interaction.

plan	relations de définition	trace	dét	val. pr. min.
1	$A_1B_1B_2C_1D = 1; A_2B_1B_2C_2E = 1$ $A_2B_1C_1C_2F = 1; A_1B_2C_1C_2G = 1$	0.948	0.976	0.434
2	$A_1B_1B_2C_1D = 1; A_2B_1B_2C_2E = 1$ $A_1B_1C_1C_2F = 1; A_1A_2B_2C_1C_2G = 1$	0.934	0.969	0.460
3	$A_1B_2C_1D = 1; A_2B_2C_2E = 1$ $A_1B_1C_1C_2F = 1; A_1A_2B_1B_2C_1C_2G = 1$	0.912	0.958	0.460
4	$A_1B_2C_1D = 1; A_2B_2C_2E = 1$ $A_1B_1C_1C_2F = 1; A_1A_2B_2C_1C_2G = 1$	0.900	0.956	0.330
5	$A_1B_2C_1D = 1; A_2B_1C_2E = 1$ $A_1B_1B_2C_2F = 1; A_2B_2C_1C_2G = 1$	0.883	0.946	0.400
6	$A_1B_2C_1D = 1; A_1B_1C_2E = 1$ $A_2B_1B_2C_2F = 1; A_1A_2B_2C_1C_2G = 1$	0.874	0.944	0.330

TAB. 46 – Efficacités globales pour les 6 solutions

Effet	Efficacité		
	Plan 1	Plan 2	Plan 3
A	1	1	1
lin B, lin C	1	1	1
quad B, quad C	1	1	1
D,E,F,G	1	1	1
A, lin B	1	1	1 0.840 0.800
A, lin C	1	1	1 0.960 0.840
AD	1	1	1 1 0.833
AE	1	1	1 1 0.952
AF,AG	1	1	1 1 0.800
lin B, lin C	0.640	0.960	0.800
lin B.D	0.840	0.800	0.840
lin B.G	0.840	0.800	0.800
lin B.E	0.840	0.800	0.960
lin B.F	0.840	0.960	0.800
lin C.D	0.840	0.960	0.840
lin C.E, lin C.F, lin C.G	0.840	0.800	0.800
DE	1	0.800	1
DF	1	0.960	1
DG	0.810	0.800	1
EG	1	0.667	0.800
EF	0.810	0.800	0.800
FG	1	0.800	0.667

TAB. 47 – Efficacités pour les 3 meilleures solutions pour la trace

5 Utilisation du programme

5.1 Condition d'emploi

PLANOR fonctionne sous Windows (95, 98, NT). La notice d'installation est dans le fichier *instal*.

5.2 Généralités

L'essentiel des indications utiles au maniement du programme apparaît à l'écran. L'utilisateur peut aussi facilement accéder à une aide en ligne, locale ou générale, en cliquant sur les boutons "Aide locale" ou "Aide générale".

En règle générale, l'utilisateur introduit l'information nécessaire au programme dans différentes zones. Pour les déplacements entre zones, modifications et validations, le programme suit les conventions usuelles de Windows. La frappe de *esc* permet un retour en arrière. La frappe de $\hat{B}reak$ ($\hat{A}ttn$ sur certains claviers) est utilisée pour interrompre une recherche de plan en cours et modifier le temps limite autorisé pour cette recherche (le symbole $\hat{\quad}$ représente la touche *Ctrl* que l'on doit maintenir appuyée en frappant la touche qui suit). Si ce temps est ramené à 0 mn, la recherche est effectivement interrompue, sinon elle est reprise en validant le nouveau temps par \leftarrow (Enter). Le $\hat{A}lt Del$ (appuyer sur les 3 touches à la fois) permet d'interrompre totalement le programme. On doit ensuite sélectionner le "gestionnaire des tâches" et sélectionner "fin de tache" deux fois de suite.

5.3 Description succincte des différents modules

L'appel de PLANOR fait apparaître le menu général du tableau 48. Ce paragraphe décrit les différents choix possible de ce menu.

Programme PLANOR	
<i>création d'un plan régulier</i>	§ 5.4
<i>Modification d'un plan régulier</i>	§ 5.4
<i>Création à partir d'une matrice obtenue précédemment</i>	§ 5.4
<i>Randomisation</i>	§ 5.5
<i>Recodage, sélection de facteurs, tri, ...</i>	§ 5.6
<i>Etude des alias</i>	§ 5.7
<i>Contenu des fichiers REG</i>	§ 5.8
<i>Initialisation</i>	§ 5.9

TAB. 48 – Menu général

Lors de la *création d'un plan régulier*, l'utilisateur introduit les paramètres définissant

sa recherche (fig. 2 et 3). Ces paramètres sont stockés dans un fichier suffixé par REG. Dans ce même fichier REG sont aussi stockés les résultats de la recherche lorsqu'elle aboutit. Ces résultats sont une ou plusieurs matrices clés.

La *modification d'un plan régulier* permet de reprendre et modifier les paramètres d'un plan régulier antérieurement défini. Dans une phase exploratoire on peut ainsi changer le modèle, la partie à estimer, introduire ou supprimer des facteurs définis. Cette option peut aussi être utilisée pour créer un nouveau plan par modification d'un ancien. On prendra soin dans ce cas de changer le nom du plan pour ne pas écraser le fichier REG de départ.

Si une seule matrice est demandée, le programme enchaîne automatiquement, lorsque la recherche aboutit, le module de création du plan puis le module de randomisation. Il construit ainsi un fichier suffixé par PS donnant le Plan dans un ordre Systématique, puis un fichier PR donnant le Plan Randomisé.

Dans le cas où plusieurs matrices sont cherchées, les solutions obtenues peuvent être examinées par *l'étude des alias*. Les plans PS et PR correspondant à la solution retenue sont alors construits en utilisant l'option *Création à partir d'une matrice obtenue précédemment*.

La randomisation peut être interrompue par un *esc* et reprise ultérieurement en sélectionnant l'option *randomisation*. Les paramètres définissant la randomisation – modèle donnant la structure des blocs et germe en particulier – sont enregistrés sur le fichier PS de départ. Si ce fichier est randomisé à nouveau, ces mêmes paramètres sont alors proposés de façon standard à l'utilisateur qui peut ainsi reproduire simplement la même randomisation.

Plusieurs opérations annexes s'avèrent nécessaires pour fabriquer à partir d'un ou de plusieurs plans réguliers le plan adapté à une situation particulière. Le module *Recodage, sélection de facteurs, tri, ...* permet d'effectuer simplement plusieurs de ces opérations: création d'un nouveau facteur produit à partir de pseudofacteurs, recodage des niveaux, élimination des facteurs inutiles, répétitions de certaines unités, fusion de plusieurs plans, tri. Le plan créé par ces opérations peut être stocké sous forme d'un nouveau fichier PS, ou édité dans un fichier ASCII servant de base à la réalisation du plan. Dans ce dernier cas, les numéros de niveaux utilisés pour le codage interne peuvent être remplacés par les niveaux explicites réels pour faciliter la lecture du plan.

A chaque fichier PS (resp. PR) est associé un fichier suffixé par HIS (resp. HIR) dans lequel sont enregistrées les modifications effectuées. Ce fichier permet d'indiquer, lors de l'édition du plan, la façon dont il a été obtenu.

Les modules de PLANOR créent, outre les fichiers spécifiques REG, PS, PR, HIS, HIR, des fichiers ASCII donnant de façon explicite les résultats des calculs. Un nom standard suffixé par OUT est proposé pour ces fichiers, mais l'utilisateur peut naturellement modifier ce nom à sa guise.

Les fichiers spécifiques sont en fait des fichiers APL dont le suffixe standard SF a été modifié. Les personnes disposant d'un interpréteur APL peuvent donc lire ces fichiers, dont le contenu est précisé au paragraphe 5.8.

5.4 Création ou modification d'un plan régulier

L'information nécessaire pour définir la recherche est donnée dans les écrans 1 et 2 (figure 2, 3).

5.4.1 Ecran 1 (fig. 2)

L'écran 1 contient les champs suivants.

5.4.1.1 nom Ce nom est utilisé comme préfixe des fichiers de sortie. Ainsi, si on inscrit EXEMPLE, les fichiers de sortie sont EXEMPLE.REG, EXEMPLE.OUT, etc ...

5.4.1.2 nb. d'unités : nombre d'unités expérimentales.

5.4.1.3 définition des facteurs de base . Choix *standard*, *utilisateur*.

Ce champs définit le mode de définition des facteurs de base, dont les combinaisons de niveaux repèrent les différentes unités expérimentales.

Ces facteurs de base peuvent être des *pseudofacteurs* utilisés uniquement pour repérer les unités, mais qui n'ont aucun sens physique réel et ne figurent dans aucun modèle. C'est en particulier le cas dans l'option *standard* où le programme introduit un pseudofacteur pour chaque premier figurant dans la décomposition en nombre premier du nombre d'unités. Le programme donne aussi un nom standard à ces pseudofacteurs, formé avec le premier en question suivi du trait de soulignement _ et du numéro d'ordre. Par exemple, si il y a 36 unités, les pseudofacteurs introduits dans l'option *standard* sont 2_1, 2_2, 3_1, 3_2.

Ce choix *standard* donne beaucoup de souplesse puisque il ne fixe au départ aucun des facteurs mais les détermine librement comme combinaison des pseudofacteurs de base. Mais pour cette même raison, c'est aussi le choix qui conduit aux recherches les plus longues. Aussi, dans bien des cas, utilise-t-on de préférence l'option *utilisateur* qui permet de sélectionner des facteurs de base parmi les facteurs réellement actifs correspondant aux traitements ou aux blocs et figurant dans le modèle ou les hiérarchies. Les facteurs de base sont alors introduits dans la table figurant à gauche de l'écran 2 (fig. 3) décrit en détail plus loin.

Par définition, toutes les combinaisons de niveaux des facteurs de bases figurent dans le plan et cela implique qu'il ne peut pas y avoir de relation de définition où ne figurent que ces facteurs. Le choix des facteurs de base par l'utilisateur interdit donc certaines relations de définition et est susceptible de restreindre l'ensemble des solutions atteignables.

Aussi essaye-t-on, lorsqu'on a recours à ce mode utilisateur, de choisir dès le départ des facteurs de base pour lesquels on veut voir figurer dans le plan toutes les combinaisons de niveaux. Le fait de choisir ces facteurs comme facteurs de base n'introduit alors pas de contrainte supplémentaire.

Par exemple, si les unités sont structurées par des systèmes de blocs (blocs, sous-bloc, lignes, colonnes, etc . . .), on peut introduire les facteurs blocs correspondants, éventuellement complétés par un pseudofacteur associé au numéro de répétition, comme facteurs de base. De même, dans la recherche d'une répartition en blocs d'un plan factoriel complet, on peut prendre comme facteurs de base les facteurs traitement, éventuellement complétés par un pseudofacteur associé au numéro de répétition si les traitements sont répétés.

Dans tous les cas, si une seule solution est requise, tout choix des facteurs de base pour lequel la recherche aboutit est acceptable. Ce n'est qu'en cas d'échec de la recherche qu'il faut parfois se demander si la raison de l'échec n'est pas un choix défectueux des facteurs de base.

5.4.1.4 type de décomposition des facteurs . Choix *maximum*, *minimum*, *au choix*.

Les facteurs dont le nombre de niveaux n'est pas premier peuvent être décomposés en produits de plusieurs pseudofacteurs. Dans tous les cas de figure, la décomposition doit aboutir à des pseudofacteurs dont le nombre de niveaux est une puissance d'un premier, mais il peut y avoir plusieurs décompositions de ce type comme le montre l'exemple du tableau 49 qui donne les décompositions d'un facteur A à 24 niveaux. La décomposition

	maximum	minimum	au choix
A_1	(2)	A_1 (8)	A_1 (4)
A_2	(2)	A_2 (3)	A_2 (2)
A_3	(2)		A_3 (3)
A_4	(3)		

TAB. 49 – Décomposition en pseudofacteurs d'un facteur A à 24 niveaux

la plus poussée est obtenue avec l'option *maximum* proposée de façon standard par le programme. Tous les pseudofacteurs obtenus dans cette décomposition ont un nombre de niveaux premier. L'option *minimum* est à l'opposé celle qui donne le nombre le plus restreint de pseudofacteurs. Les nombres de niveaux des pseudofacteurs obtenus dans cette option sont les plus grandes puissances de premier divisant le nombre de niveaux du facteur décomposé. L'option *au choix* permet de choisir librement et indépendamment pour chaque facteur la décomposition utilisée. Ce choix s'opère dans dans l'écran 2 (fig. 3), lorsqu'on est positionné sur le facteur à décomposer, en cliquant sur le bouton "pseudo-dec". Si on omet de cliquer, c'est la décomposition maximale qui est choisie.

D'un point de vue pratique, l'option *maximum* proposée en standard est celle qui donne le plus de flexibilité. On le constate empiriquement et cela a été démontré pour certaines classes de plans factoriels complets répartis en blocs [24], [25].

5.4.1.5 Recherche backtrack. Pour expliquer le sens des trois paramètres *germe*, *nb. de solutions*, *temps maximum* associés, nous décrivons ici succinctement cette recherche dans le cas où les nombres de niveaux sont tous puissances d'un même premier. Une description détaillée du cas général est donnée dans [17].

Le programme établit d'abord les listes $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_s$ de vecteurs possibles pour chacune des colonnes 1, 2, \dots s de la matrice clé.

Il cherche ensuite successivement les colonnes de la matrice clé de façon à respecter les contraintes imposées. Une fois sélectionnées des colonnes 1 à i compatibles avec les contraintes, il en déduit les vecteurs autorisables de la liste \mathcal{L}_{i+1} . Si l'ensemble de ces vecteurs est non vide, il choisit le premier d'entre eux comme colonne $i + 1$ et continue en séquence. Si cet ensemble est vide, il sélectionne dans la liste \mathcal{L}_i l'élément autorisable suivant s'il en existe. S'il n'en existe pas, il revient sur le choix de la colonne $i - 1$, et ainsi de suite.

Lorsqu'il n'y a pas de solution, l'arrêt de la procédure intervient lorsque la liste \mathcal{L}_1 est épuisée, ce qui signifie que tous les choix possibles ont été examinés. Comme cet examen exhaustif peut dans certains cas de figure prendre un temps considérable, un *temps limite* est donné au delà duquel le programme s'interrompt. L'utilisateur a alors la possibilité soit de poursuivre la recherche en augmentant le temps limite, soit d'arrêter effectivement.

Il est clair que l'ordre dans lequel sont classés les listes \mathcal{L}_i influe de façon déterminante sur le choix de la matrice clé dans les cas où on cherche une seule solution. Pour éviter d'obtenir systématiquement la même solution, on a la possibilité de réordonner au hasard ces listes en introduisant un *germe* différent de 0. La valeur de ce germe détermine en fait entièrement la façon aléatoire dont ces listes sont réordonnées et donc la solution obtenue. La même recherche conduite avec un même germe conduit donc toujours à la même solution.

Dans bien des cas, on est amené à chercher dans l'ensemble des solutions vérifiant les contraintes imposées la meilleure possible pour un critère non pris en compte dans la recherche. Par exemple, parmi toutes les fractions de résolution 4, on cherche celle qui confond le plus petit nombre possible d'interactions. Dans ce cas, il est souhaitable d'obtenir toutes les solutions possibles. On déclenche à cette fin une recherche exhaustive en mettant le plus grand nombre possible 999 dans le champs *nb. de solutions*. Pour trouver toutes les solutions, le programme poursuit alors la recherche après obtention d'une solution, en passant, dans la liste \mathcal{L}_s associée à la dernière colonne s de la matrice clé, à l'élément autorisable suivant. Si il n'y en a pas, il revient à la colonne $s - 1$ et sélectionne l'élément autorisable suivant, et ainsi de suite. Il s'arrête soit lorsque la liste \mathcal{L}_1 est épuisée, soit après atteinte du temps limite.

L'inconvénient de cette procédure exhaustive de recherche est qu'elle peut durer très longtemps (d'autant plus longtemps que dans cette version du programme, les symétries entre facteurs ne sont pas prises en compte pour raccourcir la recherche). Par ailleurs, les solutions obtenues consécutivement sont souvent très proches et ne diffèrent bien souvent que par une seule des colonnes de la matrice clé. L'arrêt prématuré des recherches conduit donc à un ensemble de solutions trop peu diversifié.

Pour obtenir un assortiment de solutions bien différentes sans effectuer la recherche exhaustive, il suffit alors de demander un nombre de solution réduit (strictement inférieur à 999 en tout cas) et d'introduire un germe aléatoire non nul. Dans ce cas on repart à zéro pour la recherche de chaque nouvelle solution, en réordonnant de façon aléatoire les listes

\mathcal{L}_i . Les solutions obtenues sont stockées dans le fichier .REG et après étude, on peut en sélectionner une par son numéro pour construire explicitement le plan. Le germe associé à chacune des solutions figure aussi sur le fichier de sortie .OUT et il est donc possible de réobtenir la solution sélectionnée dans une recherche isolée.

La recherche backtrack utilise donc les trois paramètres suivants.

- **germe**. Il détermine l'ordre dans lequel sont explorées les solutions a priori possibles pour chaque colonne de la matrice clé, ordre lexicographique si ce germe est nul et aléatoire si il est strictement positif. Dans ce dernier cas, si $1 < \text{nb. solutions} < 999$, cet ordre est redéfini à chaque recherche.
- **nb. sol. max.**. Si il est égal à 999, la recherche backtrack se poursuit jusqu'à ce que toutes les solutions soient obtenues ou jusqu'à atteinte du temps limite. Une telle recherche exhaustive n'a des chances d'aboutir que pour des problèmes de petite taille.
Si $1 < \text{nb. solutions} < 999$ et que le germe est non nul, la recherche redémarre à zéro chaque fois qu'une solution a été trouvée, avec un nouvel ordre aléatoire d'exploration des colonnes potentiellement utilisables.
- **temps maxi**. Rappelons qu'il est toujours possible d'interrompre la recherche en cours en frappant \hat{Attn} (\hat{Break}) et en ramenant à 0 mn ce temps maximum.
Si ce temps limite est atteint, le programme propose de le prolonger. Dans bien des cas cependant, cette atteinte du temps limite résulte de la non existence d'une solution. Dans ce cas, la poursuite de la recherche peut prendre un temps considérable car elle nécessite pour prouver l'absence de solution l'exploration de toutes les possibilités. En pratique, il est donc déconseillé de poursuivre trop longtemps une recherche.

5.4.1.6 inclusion facteurs dans ensemble inéligible

La donnée du modèle et des termes à estimer rend *inéligibles* certaines relations de définition, c'est à dire interdit que ces relations ne puissent être utilisées pour définir le plan. Cependant si l'effet principal d'un facteur ne figure pas dans la liste des termes à estimer, la recherche peut amener une relation de définition où ne figure que cet unique facteur dont seule une fraction des niveaux apparaît alors dans le plan. Ceci n'est dans certain cas pas acceptable, par exemple s'il s'agit d'un facteur bloc dont tous les niveaux doivent être présent.

L'inclusion systématique des facteurs dans l'ensemble inéligible, obtenu en cochant la case correspondante permet de rendre *inéligible* toute relation de définition où n'intervient qu'un seul facteur et de s'assurer ainsi que chaque facteur prend tous ses niveaux sur le plan.

Il est possible de n'imposer cette contrainte qu'à certains facteurs, en sélectionnant le choix *non* et incluant ces facteurs dans un modèle auquel est associé une partie à estimer vide. Celle-ci est interprétée comme une partie réduite à la moyenne générale. Les termes du modèle, ne pouvant être confondus avec la moyenne générale, sont inéligibles et les facteurs inclus dans ce modèle prennent obligatoirement tous leurs niveaux sur le plan.

5.4.1.7 Commentaire . Champs libre permettant d'apporter quelques précisions sur le plan considéré.

5.4.2 Ecran 2 (fig. 3)

Cet écran contient 7 zones de dialogue. A gauche les tables dans lesquelles sont introduits les listes de *facteurs de base* ou de *facteurs définis*. A droite 5 fenêtres d'édition pour introduire les *parties de modèle*, *modèles* et *parties à estimer* associées, les *hiérarchies* et enfin les *facteurs prédéfinis*.

5.4.2.1 Facteurs de base et facteurs définis

Les déplacements s'opèrent avec les flèches ou les tabulations de la façon ordinaire avec Windows. Pour introduire un nouveau facteur après (resp. avant) le facteur sur lequel le curseur est positionné, frapper *enter* \leftarrow (resp. \rightarrow). Pour supprimer le facteur sur lequel le curseur est positionné, frapper *Del*.

Dans la troisième colonne de ces deux tables, on peut indiquer (en cochant) quels sont les facteurs blocs. Cette indication n'est pas utilisée dans la recherche backtrack. Elle n'est prise en compte que pour distinguer effets bloc et effet traitement dans l'étude des alias et dans l'étape de randomisation. Si l'utilisateur a omis de préciser les facteurs bloc dans cet écran 2, il peut encore le faire avant la phase de randomisation, mais l'introduction dans l'écran 2 a l'avantage de clarifier dès le départ la situation et de faciliter l'étude des alias.

La table des facteurs de base n'apparaît que lorsqu'on a retenu l'option *utilisateur* dans l'écran 1 pour définir ces facteurs. Les considérations du paragraphe précédent donnent dans ce cas un fil conducteur pour la sélection de ces facteurs. Rappelons que le produit de leurs nombres de niveaux, introduits dans la seconde colonne de la table, doit être égal au nombre d'unités figurant dans l'écran 1.

Les niveaux des facteurs définis sont déduits de ceux des facteurs de base en utilisant les règles définies par la matrice clé. Cette matrice clé est cherchée par la procédure backtrack de façon à rendre estimables les termes à estimer dans les modèles associés et à respecter les éventuelles hiérarchies. Certains des facteurs peuvent être prédéterminés par l'utilisateur. Les colonnes de la matrice clé associées sont alors fixées dans la recherche.

La suite de ce paragraphe rappelle de façon brève les fonctions et syntaxes des informations entrées dans les fenêtres figurant à droite de l'écran 2. Une introduction plus pédagogique à ces notions au travers d'exemples figure dans le paragraphe 2 et des descriptions plus générales peuvent être trouvées dans [16], [17].

5.4.2.2 Modèles, parties à estimer, parties de modèle : généralités. On introduit en général un *modèle* unique donnant les effets factoriels non négligeables. On précise ensuite la *partie à estimer* de ce modèle. Pour simplifier l'écriture en évitant de réécrire plusieurs fois la même expression, on peut introduire aussi des *parties de modèle*.

On doit aussi pouvoir indiquer les contraintes de *hiérarchie* entre facteurs.

Il peut être nécessaire d'introduire plusieurs modèles et parties à estimer correspondantes, pour traiter par exemple des cas où les systèmes de blocs induisent plusieurs strates.

Un modèle est systématiquement complété par les termes inclus dans les termes y apparaissant. Ainsi si $A.B$ y apparaît, A , B et la moyenne générale sont systématiquement inclus par le programme. Ceci n'est pas vrai pour une partie à estimer. Si elle contient le terme $A.B$ mais pas les termes A et B , cela indique que l'on veut estimer l'interaction $A.B$, mais pas nécessairement les effets principaux A et B . Nous verrons que l'utilisation du facteur constant $\mathbf{1}$ permet de réintroduire facilement des effets principaux sans allourdir l'écriture du modèle.

Une partie à estimer vierge est interprétée comme étant la moyenne générale. Les termes du modèle complété associé ne peuvent donc être confondus avec la moyenne générale, ce qui interdit toute relation de définition basée seulement sur les facteurs d'un de ces termes. Ces relations de définition sont *inéligibles*.

L'utilisation d'une partie à estimer vierge permet en particulier de s'assurer que certains produits de facteurs prennent tous leurs niveaux sur le plan: il suffit de faire figurer ces produits dans le modèle associé.

De même, un modèle vierge est interprété comme étant la moyenne générale. Les termes de la partie à estimer associée ne peuvent donc être confondus avec la moyenne générale et conduisent à des relations de définition inéligibles. Compte tenu du fait que la partie à estimer n'est pas complétée comme le modèle, cette utilisation d'un modèle vierge donne plus de souplesse pour introduire des relations de définition inéligibles que n'en donne l'utilisation d'une partie à estimer vierge. Si l'on veut que l'ensemble des termes inéligibles soit réduit strictement à la partie à estimer, on laisse donc le modèle vierge et on décoche s'il y a lieu la case *inclusion facteurs dans ensemble inéligible* décrite au paragraphe 5.4.1.6.

Pour faciliter l'écriture du modèle ou de la partie à estimer, on a la possibilité aussi de supprimer certains termes. La suppression est indiquée par le symbole \sim . La syntaxe de la partie supprimée à droite du symbole \sim est la même que celle du modèle, mais cette partie n'est jamais complétée par des sous termes.

5.4.2.3 Modèles, parties à estimer, parties de modèle: syntaxe. Un modèle ou une partie de modèle est une somme de termes, chaque terme étant un produit de facteurs ou pseudofacteurs correspondant à un effet principal ou à une interaction. Noter qu'on peut utiliser indifféremment blanc ou $.$ pour séparer les facteurs dans un terme.

Exemple: $BL + VAR + DOSE + DENS + VAR.DOSE + VAR.DENS + DOSE.DENS$ (9)

Cette dernière expression contient un effet bloc BL, les effets principaux et interactions simples entre les facteurs traitement VAR, DOSE et DENS. Lorsqu'elle figure dans une ligne modèle, on peut la remplacer, en tenant compte du fait que le modèle est complété, par l'expression simplifiée suivante.

Exemple: $BL + VAR.DOSE + VAR.DENS + DOSE.DENS$

L'écriture peut être simplifiée par l'utilisation de parenthèses. Ainsi l'expression (9) s'écrit aussi

$$BL + (VAR + DOSE + DENS) (VAR + DOSE + DENS)$$

Pour retomber sur la 1ère forme, on développe puis on élimine

1. dans chaque terme les facteurs redondants ($VAR.VAR \rightarrow VAR$),
2. puis les termes redondants ($VAR.DOSE + DOSE.VAR \rightarrow VAR.DOSE$).

Lorsqu'on introduit le facteur constant **1** dans un produit, il peut être éliminé lors du développement dans tous les produits le contenant en même temps qu'un autre facteur. Ainsi

$$(1 + VAR + DOSE) DENS \rightarrow DENS + VAR.DENS + DOSE.DENS .$$

Le terme **1.DENS** est remplacé par **DENS**. Ce type d'écriture n'est pas utile dans une ligne modèle puisque les modèles sont automatiquement complétés, mais il peut être utile pour définir une partie à estimer.

5.4.2.4 Utilisation des parties de modèle

On peut réécrire le modèle $BL + (VAR+DOSE+DENS) (VAR+DOSE+DENS)$
 sous la forme simplifiée : $BL + \quad \quad \quad PM \quad \quad \quad \quad \quad \quad PM$

à condition d'introduire dans les parties de modèle la ligne

$$PM: VAR + DOSE + DENS$$

qui définit PM comme une partie de modèle.

Une partie de modèle peut en inclure une autre. Par exemple, les définitions apparaissant à droite et à gauche du tableau ci-dessous sont équivalentes

PM1 : VAR + DOSE	PM1 : VAR + DOSE
PM2 : PM1 + DENS	PM2 : VAR + DOSE + DENS
PM3 : PM1.DENS	PM3 : (VAR + DOSE).DENS

Mais il faut prendre garde à ne pas créer de boucle dans la définition de ces parties de modèle, comme dans l'exemple ci-dessous où la définition de P fait intervenir Q, dont la définition fait intervenir P.

$$\begin{aligned} P &: Q + DOSE \\ Q &: P + DENS \end{aligned}$$

Si par inadvertance on crée une telle boucle, la seule solution est d'interrompre le programme par un \hat{Break} (\hat{Attn}).

5.4.2.5 Suppression de termes Pour écrire un modèle comportant toutes les interactions de deux facteurs à l'exception d'une ou deux, on peut utiliser l'opérateur \sim de "soustraction". La partie à droite de cet opérateur est développée en utilisant la même syntaxe que le modèle et les termes obtenus par ce développement sont supprimés. Cette partie n'est jamais complétée par les sous termes contrairement à la partie figurant à gauche de \sim dans la spécification du modèle.

Ainsi, si PM désigne la même partie de modèle que précédemment, $PM.PM \sim DOSE.DENS$ est équivalent à la somme $VAR.DOSE + VAR.DENS$ qui après complétion par les sous-termes donne $1 + VAR + DOSE + DENS + VAR.DOSE + VAR.DENS$.

5.4.2.6 Les Hiérarchies. Si le niveau d'un facteur A doit rester constant pour chaque combinaison de niveaux de certains autres facteurs, disons B, C, D , on fait figurer la ligne suivante dans les hiérarchies :

A : B C D

On dit alors que A hiérarchise, ou emboîte le facteur produit $B \times C \times D$, ou encore qu'il est marginal à ce facteur.

Exemple. Dans un dispositif comportant des blocs $-BL-$ et des sous-blocs $-SBL-$, le facteur VAR est constant sur chaque bloc et le facteur DOSE constant sur chaque sous-bloc. Cette contrainte est indiquée par:

VAR : BL
DOSE : BL SBL

SBL donne le numéro du sous-bloc dans le bloc.

5.4.2.7 Exemple avec plusieurs modèles et hiérarchies. Dans un plan en bloc où certains facteurs ne peuvent varier intra-bloc, il est souvent indispensable d'introduire plusieurs modèles et parties à estimer associées comme l'illustre l'exemple typique suivant.

	Parties de modèle	Hiérarchies
p1	: A+B+C+D	A : BL
p2	: A+B+C+D+E+F+G+H	B : BL
		C : BL
		D : BL

	Modèles	Parties à estimer associées
[1]	p1 p1	[1] p1
[2]	BL+p2 p2	[2] p2 (E+F+G+H)

La partie à estimer $p2(E + F + G + H)$ du modèle 2 inclue les effets principaux de E, F, G, H et leurs interactions avec chacun des 8 facteurs traitement A, B, C, D, E, F, G, H (soit 26 termes sous forme développée).

Les effets principaux A, B, C, D ne peuvent être estimés dans le modèle incluant l'effet bloc (BL) puisqu'ils ne peuvent varier intra-bloc. La paire *modèle, partie à estimer* 1

permet de s'assurer néanmoins de leur estimabilité dans un modèle inter-bloc comprenant toutes les interactions entre deux de ces facteurs.

Comment faire figurer tous les niveaux d'un produit de facteurs

Comme cela a été indiqué, on peut contraindre un produit de facteurs à prendre tous ses niveaux sur le plan en le faisant figurer comme terme d'un modèle auquel est associé une partie à estimer vierge.

A titre d'exemple considérons une situation dans laquelle il y a trois facteurs traitement A, B, C à 6, 6 et 2 niveaux. Les 72 traitements du plan factoriel complet doivent être répartis dans un cube comportant 6 positions en abscisse, 3 en ordonnée et 2 hauteurs –facteurs X, Y, Z –. Pour illustrer, on peut imaginer qu'il s'agit de cultiver 6 souches de champignons de Paris (facteur A) avec 6 composts différents (B) à 2 pH (C). Ces champignons sont répartis dans la chambre de culture sur deux plateaux superposés (Z), chacun d'eux contenant 6 lignes (X), 3 colonnes (Y) avec deux bacs à chaque emplacement défini par X, Y, Z .

Compte tenu du système de ventilation et de la position de la porte de la salle de culture, on estime que les trois facteurs X, Y, Z sont susceptibles d'avoir un effet. On désire pouvoir estimer les effets principaux A, B, C et les interactions $A.C, B.C$ dans un modèle comprenant les effets additifs X, Y, Z et toutes les interactions entre facteurs traitement.

Si les facteurs de base utilisés sont X, Y, Z et U , le numéro du bac dans la paire, on est sûr que chacun des 36 triplets de coordonnées X, Y, Z apparaît deux fois. Mais si on utilise A, B, C comme facteurs de base, la paire *modèle-partie à estimer* 1 suivante:

$$\begin{array}{ll} \text{modèle} & \text{partie à estimer} \\ [1] \quad X + Y + Z + A.B.C & [1] \quad A + B + C + A.C + B.C \end{array}$$

peut conduire à des relations de définition où tous les triplets de coordonnées X, Y, Z n'apparaissent pas, telles que celles du tableau 50.

fact. à 2 niv.		
	X_1	Z
A_1	1	1
B_1	1	1
C	1	0
	$X_1 = A_1 B_1 C$	
	$Z = A_1 B_1$	
fact. à 3 niv.		
	X_2	Y
A_2	2	2
B_2	2	2
	$X_2 = A_2^2 B_2^2$	
	$Y = A_2^2 B_2^2$	

TAB. 50 – Solution inadéquate pour une répartition en 3 systèmes de blocs croisés

Pour éviter de telles relations inadéquates, on introduit un modèle comprenant le terme $X.Y.Z$ associé à une partie à estimer vierge :

$$\begin{array}{ccc} \text{modèle} & \text{partie à estimer} & \\ [2] & X.Y.Z & [2] \end{array} \quad (10)$$

Le choix de A, B, C comme facteurs de base est assez naturel. D'une part il n'oblige pas à introduire un pseudofacteur supplémentaire tel que U . D'autre part, il est celui qu'on utilise spontanément lorsqu'on s'interroge, indépendamment de tout programme, sur les constructions possibles. On se demande alors avec quelles interactions entre facteurs traitement confondre chacun des trois systèmes de bloc, et cela amène à définir les pseudofacteurs bloc à partir des traitements et non l'inverse.

Remarques : si X, Y, Z sont marqués comme facteurs blocs dans la table des facteurs définis, le programme détecte les matrices clés pour lesquelles toutes les combinaisons de niveaux des facteurs bloc ne sont pas présentes. Ceci est une raison supplémentaire pour signaler dès le départ les facteurs bloc.

5.4.2.8 Facteurs prédéterminés Les lignes suivantes :

$$\begin{array}{l} Bl_1 : A_1 + B_1 + C_1 \\ Bl_2 : A_2 + 2B_2 \\ E : 2A_2 + D \end{array}$$

dans la fenêtre des facteurs prédéterminés fixent la manière dont les pseudofacteurs Bl_1, Bl_2 déduits du facteur Bl et le facteur E sont calculés à partir des pseudofacteurs A_1, B_1, C_1, A_2, B_2 déduits des facteurs de base A, B, C et du facteur D .

Après avoir vérifié que ces définitions vérifient les contraintes, le programme les prend en compte dans la recherche des autres facteurs.

Il est possible de définir tous les facteurs de cette façon. Cela permet d'étudier les alias dans un plan prédéfini ou de contruire et randomiser ce plan.

5.5 Randomisation

Dans les expériences d'agronomie qui sont à l'origine du développement des plans d'expériences et de la théorie de la randomisation, il est généralement admis que l'observation est la somme d'un effet du traitement et d'un effet non maîtrisé de l'unité expérimentale. Pour éviter que les effets des unités expérimentales ne se superposent de façon systématique aux effets des traitements que l'on compare et ne biaisent les comparaisons, on choisit au hasard l'unité expérimentale affectée à chaque traitement : c'est ce qu'on appelle la *randomisation*.

Lorsque toutes les unités sont équivalentes la randomisation est totalement libre. On l'appelle *randomisation totale*. Elle consiste à déterminer par tirage au hasard sans remise l'unité réelle (parcelle en agronomie, animal en zootechnie, etc . . .) affectée à

chaque unité du plan systématique. Un exemple est donné au tableau 1. Les numéros résultant du tirage sont appelés indices de répétition et figurent dans une colonne repérée par l'en-tête *ind-rep*.

Dans cet exemple le nombre d'unités du plan systématique est égal au nombre maximum d'unités disponibles et la randomisation revient à choisir au hasard une permutation des nombres de 0 à 7, avec une même probabilité de choisir chacune des !8 permutations. Cette situation est fréquente mais il peut aussi arriver que le nombre d'unités disponibles soit plus grand que le nombre d'unités dont on a besoin, situation que l'on prend en compte en introduisant dans la boîte *nombre de niveaux totaux* de l'écran 3 (fig.5) un *nombre de niveaux maximum* égal au nombre total d'unités disponibles.

Lorsque les unités expérimentales sont structurées en blocs, la randomisation doit respecter cette structure. Cette structure en bloc est précisée par les facteurs blocs. Mais les niveaux donnés à ces derniers dans le plan systématique sont sans rapport avec les labels ou numéros des véritables unités expérimentales. La randomisation décrite ci-après aboutit précisément à remplacer ces niveaux du plan systématique par les niveaux repérant les véritables unités expérimentales. Elle s'effectue de telle façon que toutes les unités ayant même niveau d'un facteur bloc dans le plan systématique ont même niveau dans le plan randomisé résultant de ce remplacement. La randomisation du plan ROBOT1 au § 3.1.4 donne un exemple (tableaux 10 et 11).

La spécification des facteurs blocs se fait par l'intermédiaire du modèle de randomisation. Ces facteurs doivent pouvoir s'exprimer à partir d'un ensemble de pseudofacteurs de base, appelés *pseudofacteurs blocs*, qui sont tels que le nombre d'unités par combinaison de niveaux de ces pseudofacteurs est constant. Si ce nombre d'unités est $k > 1$, le programme introduit un pseudofacteur bloc supplémentaire à k niveaux, l'indice de répétition noté *ind-rep* en abrégé, qui précise le numéro d'unité.

La figure 5 montre comment les pseudofacteurs et facteurs blocs peuvent être indiqués au moment de la randomisation en cochant la case ad hoc de la colonne *bl*, dans la boîte *sélection des blocs* à gauche de l'écran. Un choix standard est proposé si l'utilisateur a antérieurement précisé quels sont les facteurs blocs dans l'écran 2. La seconde boîte introduite après validation de la boîte de gauche lors de la randomisation permet, lorsqu'il y a d'avantage de blocs disponibles que de blocs effectivement expérimentés de le spécifier en modifiant les *nombres de niveaux maximum* qui donnent les nombres de blocs disponibles au sein desquels s'opère le choix.

Le modèle de la randomisation introduit ensuite donne la liste des facteurs bloc exprimés comme produit de pseudofacteurs et séparés par des +. Le germe introduit enfin initie le tirage au hasard. La randomisation effectuée est entièrement déterminée par le germe, ce qui permet de la refaire à l'identique: il suffit pour cela de garder le même germe.

Noter que si le fichier randomisé est perdu, on ne peut le recréer à l'identique qu'à condition de connaître le germe correspondant. En fait, le dernier germe utilisé est stocké dans le fichier .PS contenant le plan systématique et est reproposé en cas de nouvelle randomisation. Si il existe un fichier de même nom suffixé par REG, le germe y est aussi stocké, ce qui permet de réobtenir la dernière randomisation également si le fichier PS a

été perdu.

Dans l'exemple de la figure 5, les pseudofacteurs blocs sont la plaque *pl* et la colonne *col*. On utilise pour l'expérimentation 6 plaques et pour chaque plaque les 4 colonnes. Il y a donc égalité, dans la seconde boîte de l'écran 3, entre les *nombres de niveaux* et les *nombres de niveaux maximum*. Il en serait différemment si seulement 3 des colonnes étaient utilisées sur chaque plaque. Pour choisir ces trois colonnes au hasard parmi les 4 de chaque plaque, il faudrait alors remplacer le nombre de niveaux maximum 3 donné en standard pour le pseudofacteur *col* par un 4.

Le facteur plaque coïncide avec le pseudofacteur *pl*. Par contre, le pseudofacteur *col* n'a dans cet exemple pas de signification s'il est séparé du facteur *pl*. Les colonnes sont considérées là comme une subdivision des plaques. Il y a 6×4 colonnes au total. Chaque colonne est spécifiée par le numéro de la plaque à laquelle elle appartient *pl* et son numéro *col* dans cette plaque. Le modèle de la randomisation comprend donc deux termes *pl* et *pl.col*, le second correspondant au facteur colonne. La randomisation est décrite pour cet exemple particulier dans le paragraphe 3.1.4.

Un autre exemple est donné en bas à droite de l'écran 3 (fig. 5), où apparaissent quatre facteurs blocs : des lignes (LIG) croisées avec des colonnes (COL), des minilignes emboîtées dans les lignes (MINILIG) et des microcolonnes (MICROCOL) subdivisant les cellules formées par intersection des lignes et des colonnes (figure 7). Les unités sont dans cet exemple définies par les quadruplets de niveaux des 4 pseudofacteurs LIG, COL, MINILIG, MICROCOL. Les facteurs ligne, colonne coïncident respectivement avec les pseudofacteurs LIG, COL. Les facteurs miniligne et microcolonne sont définis respectivement par les produits MINILIG.LIG et MICROCOL.LIG.COL.

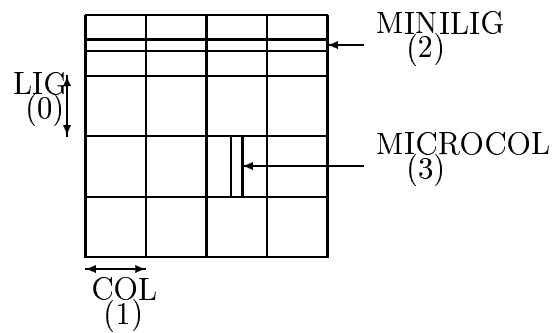
Dans le cas général, le modèle est donc formé par la liste des facteurs blocs qui sont définis comme produits de pseudofacteurs blocs.

Les pseudofacteurs blocs n'apparaissant pas dans le modèle introduit par l'utilisateur sont automatiquement rajoutés. Cela permet de ne pas introduire de modèle de randomisation si on a pas à introduire de produits dans le modèle de la randomisation, ce qui se produit lorsque les facteurs blocs sont tous croisés (pas de hiérarchies) et en particulier lorsqu'il n'y a qu'un seul facteur bloc.

Lorsque le nombre d'unités par combinaison de niveaux des pseudofacteurs blocs est strictement supérieur à 1 ($k \geq 1$), le programme rajoute automatiquement au modèle un facteur unité, produit de *ind-rep* et de tous les autres pseudofacteurs.

Du modèle, complété le cas échéant par tous les pseudofacteurs bloc, on déduit immédiatement un *ordre partiel* entre les pseudofacteurs bloc. On a par définition $A \leq B$ si tout facteur du modèle contenant A contient aussi B . En particulier, l'indice de répétition quand il est introduit est inférieur à tous les autres pseudofacteurs. Cet ordre est souvent représenté sous forme d'un diagramme, dit diagramme de Hasse, tel que celui qui apparaît dans la figure 7.

★ En fait il s'agit plus précisément d'un préordre que d'un ordre car deux pseudofacteurs distincts A_1 et A_2 peuvent être systématiquement associés et on a alors à la fois $A_1 \leq A_2$ et $A_2 \leq A_1$ sans avoir l'égalité. On peut cependant ignorer cette situation que le programme traite simplement en remplaçant



Modèle de la randomisation
 LIG+COL+MINILIG.LIG+MICROCOL.LIG.COL

Ordre partiel entre pseudofacteurs
 MINILIG < LIG {(2) < (0)}
 MICROCOL < LIG {(3) < (0)}
 MICROCOL < COL {(3) < (1)}

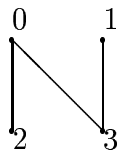


FIG. 7 – Exemple de structure en bloc (extrait de [3])

chaque ensemble de pseudofacteurs constamment associés par leur pseudofacteur produit.

Le détail de la randomisation peut être obtenu en activant l'option *impressions intermédiaires* décrite au § 5.9, qui s'obtient à partir du choix *Initialisation* du le menu général (tab. 48). Un exemple de sortie ainsi obtenue est donné au tableau 14.

La structure du système de blocs pour ce dernier exemple, schématisée sur la figure 6, peut être décrite par le modèle $col1 + col1.col2 + lig1 + lig1.lig2$ auquel correspond l'ordre $col2 < col1, lig2 < lig1$. Le programme de randomisation renumérote les pseudofacteurs de façon compatible avec cet ordre. Plus exactement, dans le cas où certains facteurs apparaissent systématiquement associés dans les termes du modèle, il détermine d'abord les *classes* de pseudofacteurs associés, et renumérote ces classes. Chaque classe est ici réduite à un facteur et la numérotation interne utilisée est 0 pour $col1$, 1 pour $lig1$, 2 pour $col2$, 3 pour $lig2$. Noter que si une classe comprenait plus d'un pseudofacteur, le programme remplacerait les facteurs de cette classe par leur produit. On peut donc toujours considérer dans ce qui suit qu'il y a un unique pseudofacteur par classe.

Pour chaque pseudofacteur j est déterminée la liste $]j) = \{i/i > j\}$ des pseudofacteurs qui lui sont strictement supérieurs. Dans l'exemple, les listes $]0)$ et $]1)$ associées aux pseudofacteurs 0 ($col1$) et 1 ($lig1$) sont vides. Les listes $]2)$ et $]3)$ associées au pseudofacteurs 2 ($col2$) et 3 ($lig2$) contiennent respectivement les facteurs 0 ($col1$) et 1 ($lig1$).

Pour chaque pseudofacteur bloc j , les niveaux réels sont déterminés par des tirages au hasard sans remise effectués indépendamment pour chacune des combinaisons de niveaux des pseudofacteurs supérieurs, c'est à dire des pseudofacteurs de $]j)$. Ainsi pour le pseudofacteur colonne $col2$, un tirage séparé est effectué pour chacune des deux macrocolonnes $col1 = 0$ et $col1 = 1$.

La détermination du niveau réel des pseudofacteurs bloc pour une unité du plan systématique utilise le tirage ad hoc. Celui-ci dépend, pour chaque pseudofacteur, des niveaux des pseudofacteurs supérieurs. Par exemple pour l'unité du plan systématique définie par $col1 = 1, lig1 = 1, col2 = 0, lig2 = 0$, la détermination de $col2$ dans le plan randomisé utilise la permutation de la macrocolonne 1 puisque $col1 = 1$. L'information du tableau 14 permet de trouver que les niveaux des quatres pseudofacteurs dans le plan randomisé sont $col1 = 0, lig1 = 0, col2 = 1, lig2 = 0$.

La théorie de ce type de randomisation est décrite dans [3]. Une description simplifiée figure dans [2]). Des résultats figurant dans [3], on déduit aisément qu'une telle randomisation induit une structure de covariance des effets aléatoires des unités identique à celle d'un modèle à effets aléatoires classique contenant un effet bloc aléatoire pour chaque terme *ancestral* (cf [14]). Est ancestral tout produit de pseudofacteurs qui, quand il contient un pseudofacteur A contient aussi tous les pseudofacteurs supérieurs, i.e. tous les B tel que $A \leq B$. Dans le modèle classique en question, les effets sont tous non corrélés, leur variance est constante pour chaque effet.

Pour illustrer ce dernier point, considérons de nouveau les différents modèles de randomisation déjà considéré dans ce paragraphe.

- Modèle $pl + pl.col$ de la figure 5. Chaque colonne de chaque plaque contient 2 unités et le modèle est donc automatiquement complété en interne par le terme $pl.col.ind-$

5.6.1 Création de nouveaux facteurs (fig. 4)

Ce choix permet de créer des facteurs produits et de coder leurs niveaux en en agréant certains si besoin est. En particulier il permet de recoder un facteur en agréant éventuellement certains niveaux.

La syntaxe pour créer un facteur produit est simple. On indique sur une ligne de la fenêtre *Définition de nouveaux facteurs* le nom du nouveau facteur, suivi du signe (:) et de la liste d'anciens facteurs à partir duquel il est formé. Pour faciliter la tâche de l'utilisateur, la liste des anciens facteurs figure dans une autre fenêtre.

Par exemple, si les facteurs déjà existant sont

Prod Temp pH Bloc

on peut changer les niveaux du facteur *Prod* et définir un facteur produit *Temp×pH* en tapant dans la fenêtre de définition :

Tp : Temp pH
Prod : Prod

On clique sur le bouton “niveaux” pour redéfinir les niveaux du facteur sur lequel est positionné le curseur. Si on omet de le faire, le programme utilise une numérotation séquentielle 0, 1, 2, ...

5.7 Etude des alias

Une fois obtenues les matrices clés répondant aux spécifications, l'option *étude des alias* du menu général (tableau 48), permet de trouver les effets confondus pour tout ou partie des solutions. Cette option fait apparaître l'écran 6 de la figure 8, où est défini le modèle utilisé. Le modèle proposé est celui qui a été antérieurement introduit pour la recherche du plan (ou les modèles). Il est possible de le modifier ainsi que les parties de modèles dans les deux boîtes du haut de l'écran 6.

L'étude des alias différencie les facteurs traitement et les facteurs blocs, marqués par une flèche au bas de l'écran 6 de la figure 8. Il est donc important, pour avoir des sorties adaptées au problème posé, d'avoir préalablement indiqué quels sont les facteurs blocs dans les boîtes figurant à gauche de l'écran 2 (figure 3).

Les exemples de sorties de l'étude des alias abondent dans cette notice. On se reportera en particulier au § 2.2.2, 2.3.2. L'exemple du § 3.3.2 illustre comment on utilise la connaissance des alias pour choisir une bonne solution.

5.8 Contenu des fichiers .REG du répertoire actif

Cette option permet de connaître le contenu de tous les fichiers .REG du répertoire actif qui est celui qui apparaît en haut de l'écran 5 d'initialisation (figure 9).

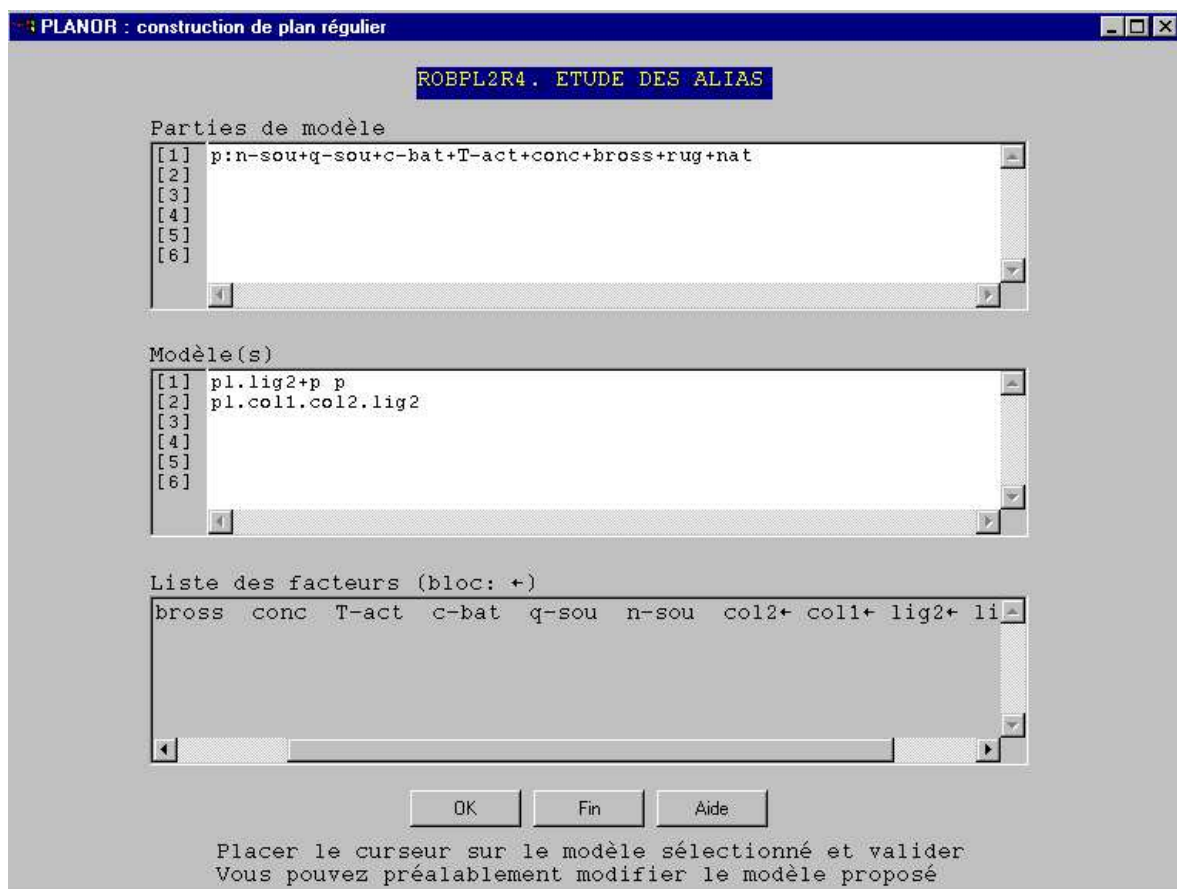


FIG. 8 – Sélection du modèle pour l'étude des alias. Ecran 6

5.9 Initialisation

Cette option du menu général fait apparaître l'écran 5 figurant dans la figure 9, dont les champs sont décrits ci-après.



FIG. 9 – Initialisation, écran 5

Librairie des fichiers créés : chemin d'accès au répertoire (directory) qui contient tous les fichiers lus ou créés par le programme. Il peut être librement modifié avant chaque nouvelle étude.

Nom de l'expérience. Ce nom est utilisé, avec un suffixe ad-hoc, pour tous les fichiers spécifiques créés dans l'étude (fichiers REG, PS, PR, HIS, HIR). Il est aussi proposé avec le suffixe OUT comme nom standard pour les fichiers de résultats. Ce nom est mis à jour après chaque utilisation du programme et fait donc toujours référence au dernier plan mis en place.

Largeur de page à l'impression. Donne la largeur maximum des lignes dans les fichiers de résultats. Les lignes dépassant cette largeur sont coupées. Le symbole > est utilisé pour indiquer qu'un ligne est la continuation de la précédente. Les matrices sont découpées en blocs pour faciliter la lecture.

Largeur 0: les lignes ne sont pas coupées et peuvent avoir une largeur quelconque. L'éditeur intégré, de même que la plupart des éditeurs classiques, permet sans difficulté de lire la partie de texte dépassant la largeur de l'écran, mais les fichiers de résultats ne peuvent être imprimés sans modification préalable dans le cas où le fichier de résultat contient des matrices trop larges.

Impressions intermédiaires. Ces impressions n'intéresseront que des spécialistes désireux de mieux comprendre l'algorithme de recherche de matrice clé décrit dans [17].

Caractères accentués conservés en sortie. Certains caractères (é, à, ≤, ... ,) peuvent poser des problèmes à l'impression. Si la case n'est pas cochée, ces caractères sont remplacés par des caractères standards (e, a, < ou =, ...).

N0 de référence. Chaque utilisateur se voit attribuer un numéro personnel de référence à 12 chiffres qu'il doit introduire lors de la première utilisation sur un microordinateur du logiciel PLANOR (ou d'ANALYS).

5.10 Contenu des fichiers REG, PS, PR, HIS, HIR

5.10.1 Fichiers .REG

Les enregistrements 1 et 2 contiennent les informations de l'écran 1, les enregistrements 3, 4, 5 les informations introduites dans l'écran 2 et ses écrans auxiliaires, à part *modrand* et RLRAND qui proviennent de l'écran 3. De ces informations, on déduit les variables contenues dans les enregistrements 6 et 7. Enfin l'enregistrement 8 contient pour l'essentiel les matrices clés résultant de la recherche.

On donne dans ce qui suit la liste des variables contenues dans chaque enregistrement, puis le contenu de ces variables. Certaines de ces variables sont ce qu'on appelle en APL des tableaux généralisés (structures dans d'autres langages) qui peuvent contenir par exemple plusieurs matrices de tailles différentes.

Dans les noms des variables, le suffixe *f* fait référence à un facteur, le suffixe *p* à un pseudofacteur, le suffixe *u* à un facteur ou pseudofacteur de base (*u* pour unité), le suffixe *t* à un facteur ou pseudofacteur défini (*t* pour traitement)

L'ordre de rangement des facteurs et pseudofacteurs dans le fichier REG est l'ordre inverse de l'ordre d'introduction, modifié pour tenir compte des hiérarchies et facteurs prédéfinis. L'algorithme de recherche backtrack procède en effet dans l'ordre décroissant des numéros de facteurs ou pseudofacteurs, ce qui oblige à classer à l'intérieur du programme tout facteur ou pseudofacteur à définir avant ceux qui servent à le définir.

Rappelons que les facteurs initialement introduits sont d'abord décomposés en pseudofacteurs ayant des nombres de niveaux puissances d'un premier. Pour ceux de ces facteurs, de libellés LIBft, qui figurent dans la partie droite de l'écran 1 (i.e. dans les modèles, termes à estimer, hiérarchies, facteurs prédéfinis), les premiers et exposants associés dans ces décompositions sont données par les variables Ppt et Ept. Pour la recherche des matrices clés, il convient de réordonner les pseudofacteurs en fonction du premier divisant leur nombre de niveaux, ce qui conduit aux listes LIBmug, LIBnug. Pour repasser ultérieurement des pseudofacteurs de LIBnug aux facteurs figurant dans LIBft, on utilise les numéros d'ordre contenus dans N0nug.

Variables du fichier REG

Enregistrement 1 : NBUNIT CHOIXfu CHOIXdf
Enregistrement 2 : TMAX NBSOL RLINK comEXP FACINEL
Enregistrement 3 : LIBfu NIVfu BLOCfu LIBNfu NIVpsu
Enregistrement 4 : LIBf NIVf BLOCf LIBNf NIVps modrand RLRAND
Enregistrement 5 : pmod mod esta hieralpha pd
Enregistrement 6 : BLOCft LIBft NIVft Npft LIBpt Ppt Ept
Enregistrement 7 : BLOCnug N0nug LIBmug LIBnug Pg mug nug
Enregistrement 8 : INDPT RLINKS fUg fUss

Enr. 1 : haut de l'écran 1

NBUNIT : nb. d'unités
CHOIXfu : option utilisée pour définir les pseudofacteurs de base
1 : choix standard (2_1, 2_2, 3_1, etc ...)
2 : utilisateur
CHOIXdf : type de décomposition des facteurs en pseudofacteurs
1 : maximal, exemple 36 $\rightarrow 2 \times 2 \times 3 \times 3$
2 : minimal, exemple 36 $\rightarrow 4 \times 9$
3 : au choix, exemple 36 $\rightarrow 4 \times 3 \times 3$

Enr. 2 : bas de l'écran 1

TMAX : temps maximum en mn pour la recherche Backtrack
NBSOL : nb. de solutions recherchées (999 = toutes les solutions)
RLINK : germe définissant l'ordre aléatoire d'exploration des différentes possibilités pour les facteurs à définir
comEXP : commentaire
FACINEL : inclusion facteurs dans ens. inéligible (1 oui, 2 non)

Enr. 3 : écran 2, facteurs de base.

LIBfu : libellés des facteurs de base (ordre inverse de l'ordre d'introduction)
NIVfu : nombres de niveaux correspondants
BLOCfu : indicateur facteurs blocs
LIBNfu : libellés des niveaux
NIVpsu : décompositions des niveaux introduites par l'utilisateur lorsque CHOIXdf=3.

Enr. 4 : écran 2, fact. définis + écran 3, modrand.

LIBf : libellés des facteurs définis
NIVf : nb. de niveaux correspondants
BLOCf : indicateur facteurs blocs
LIBNf : libellés des niveaux
NIVps : décompositions des niveaux introduites par l'utilisateur lorsque CHOIXdf=3.
modrand : modèle de la randomisation précisant la structure des systèmes de blocs (introduit dans l'écran 3).

Enr. 5: écran 2, partie droite

pmod : parties de modèle
mod : modèle(s)
esta : partie(s) à estimer associée(s)
hialpha : hiérarchies
pd : facteurs prédéfinis

Enr. 6: facteurs apparaissant en colonne des matrices clés

BLOCft : indicateur des facteurs blocs dans LIBft
LIBft : sous liste des facteurs de base ou à définir figurant dans le ou les modèles, les termes à estimer, les hiérarchies ou les facteurs prédéfinis. Les pseudofacteurs associés sont ceux qui apparaissent en colonne des matrices clés.
NIVft : nombre de niveaux associés
Npft : nombre de pseudofacteurs par facteur de LIBft
LIBpt : libellés des pseudofacteurs associés aux éléments de LIBft
Ppt, Ept : premiers, exposants associés dans la décomposition en pseudofacteurs de chaque facteur figurant dans LIBft.

Enr. 7: classement par premier.

BLOCnug : indicateur des pseudofacteurs bloc dans LIBnug
N0nug : numéros des pseudofacteurs de LIBnug dans LIBft.
LIBmug : listes par premier des pseudofacteurs unités
LIBnug : listes par premier des pseudofacteurs traitement
Pg : premiers distincts divisant l'un des nombres de niveaux
mug : exposants des premiers dans les nb. de niv. des pseudofacteurs unités
nug : exposants des premiers dans les nb. de niv. des pseudofacteurs à définir.

Enr. 8: résultats de la recherche.

INDPT : 1 si on peut choisir indépendamment les parties de matrice clé associées aux différents premiers, 0 sinon
RLINKS : germe(s) utilisé pour trouver la (les) solution(s)
fUg : parties fixes des matrices clés associés aux différents premiers
fUss : parties non fixes des matrices clés pour les différentes solutions.
En règle générale, fUss contient un item par solution, item formé des matrices clés associées aux différents premiers.
Cependant si INDPT=1, que la recherche est exhaustive (NBSOL=999), et qu'il y a plusieurs premiers en jeu, fUss contient un item par premier, item formé de l'ensemble des solutions obtenues pour ce premier. On peut dans ce cas choisir librement pour chaque premier une des solutions trouvées.

5.10.2 Fichiers PS et PR

Les enregistrements 1, 2, 3 des fichiers PS (Plan Systématique) et PR (Plan Randomisé) ont le même contenu, décrit ci-après. Noter que dans la variable PLAN contenant le plan proprement dit, les niveaux sont donnés par leur numéro compté à partir de 0. Les libellés correspondants figurent dans la variable LIBNf.

Les fichiers PS contiennent deux enregistrements supplémentaires. L'enregistrement 4 contient des informations utiles pour randomiser le plan, en particulier, si le plan a déjà été randomisé, le modèle de randomisation et le germe antérieurement introduit. L'enregistrement 5 donne les informations permettant de savoir quelle matrice clé a servi à construire le plan dans le cas où le fichier a été obtenu directement à partir d'un fichier REG dans lequel sont stockées plusieurs solutions.

Contenu des 3 premiers enregistrements des fichiers PS et PR

Enregistrement 1 : LIBf
Enregistrement 2 : PLAN
Enregistrement 3 : NIVf BLOCf LIBNf

Variables des 3 premiers enregistrements des fichiers PS et PR

LIBf : libellés des facteurs et pseudofacteurs associés aux colonnes du Plan
PLAN : plan d'expérience. Les niveaux sont repérés par leur numéro comptés à partir de 0, i.e. 0, 1, ...
NIVf : nombre de niveaux des facteurs et pseudofacteurs.
BLOCf : indicatrice des facteurs bloc.
LIBNf : listes de libellés de niveaux.

Contenu des enregistrements complémentaires 4 et 5 des fichiers PS

Enregistrement 4 : PSEUDO NDELTA modrand RLRAND
Enregistrement 5 : Pg INDPT NBSOL NUM

Détail de l'enregistrement 4 d'un fichier PS

PSEUDO : liste des pseudofacteurs blocs utilisés pour la randomisation.
NDELTA : donne pour chaque pseudofacteur bloc le nombre total de niveaux parmi lesquels sont tirés au hasard les niveaux effectivement utilisés
modrand : modèle décrivant la structure en bloc
RLRAND : germe utilisé pour la randomisation

Détail de l'enregistrement 5 d'un fichier PS

Pg : liste des premiers
INDPT : indique si les matrices clés associées aux différents premiers peuvent être choisies indépendamment (INDPT=1) ou non (INDPT=0).
NBSOL : nombre de solutions requises (recherche exhaustive si NBSOL=999).
NUM : numéro de la solution ou numéros pour les différents premiers dans le cas NBSOL=999, INDPT=1.

5.10.3 Fichiers HIS et HIR

Rappelons qu'à chaque fichier PS ou PR est associé un fichier suffixé par HIS ou HIR dans lequel sont enregistrées les modifications effectuées par l'option *Recodage, sélection des facteurs, tri, ...* du menu général (tableau 48). Chaque enregistrement de ce fichier donne le code de l'opération effectuée, puis les indications qui ont permis de l'effectuer.

Exemple:

LEC	TEST1.PS	Lecture du fichier TEST1.PS
NEW	$ab:a b$	Définition d'un nouveau facteur ab comme produit des facteurs a et b .
SEL	c	Sélection de facteurs: élimination du facteur c .
TRI	$d e$	Tri des unités sur d , puis sur c pour d constant.
REP	5 6 2 0	l'unité 5 est dupliquée, l'unité 6 supprimée.
FUS	TEST2.PS	on ajoute à la suite le fichier TEST2.PS en ne retenant que les facteurs communs.
SAV	TEST1A.PS	fichier modifié sauvé sous le nom TEST1A.PS.

Références

- [1] Agrawal V., Dey A. (1983). Orthogonal Resolution-IV Designs for Some Asymmetrical Factorials. *Technometrics*, **25**, n° 2, 197-199.
- [2] Bailey R.A. (1981). A Unified Approach to Design of Experiments. *J. Roy. Statist. Soc. A*, **144**, Part 2, 214-223.
- [3] Bailey R.A., Praeger C.E., Rowley C.A., Speed T.P. (1983). Generalized wreath products of permutation groups. *Proc. London Math. Soc.*, (3), **47**, 69-82.
- [4] Bailey R.A. (1985). Factorial Design and Abelian groups. *Linear Algebra Appl.*, **70**, 349-368.
- [5] Brouillaud-Delattre A., Kobilinsky A., Cerf O. (1994). Méthode de mesure de l'efficacité des procédés de nettoyage et de désinfection des surfaces ouvertes. *Lait*, **74**, 79-88.
- [6] Cliquet S., Durier C., Kobilinsky A. (1994). Principe of a fractional factorial design for qualitative and quantitative factors: application to the production of *Bradyrhizobium japonicum* in culture media. *Agronomie*, **14**, 569-587. Traduction en Anglais du texte référencé ci-dessous.
- [7] Cliquet S., Durier C., Kobilinsky A. (1994). Principe d'un plan factoriel fractionné intégrant des facteurs qualitatifs et quantitatifs; application à la production d'un milieu de culture pour *Bradyrhizobium japonicum*. Document interne. Laboratoire de Biométrie. INRA Versailles.
- [8] Edmondson R.N. (1991). Agricultural response surface experiments based on four-level factorial designs. *Biometrics*, **47**, 1435-1448.
- [9] Edmondson R.N. (1993). Systematic row-and-column designs balanced for low order polynomial interactions between rows and columns. *J. R. Statist. Soc. B*, **55**, 707-723.
- [10] El Mossadeq A. (1993). Construction des fractions régulières de plans d'expériences factoriels. Thèse d'Etat. Université de Pau et des pays de l'Adour.
- [11] Franklin M.F. and Bailey R.A. (1977). Selection of Defining Contrasts and Confounded Effects in Two-level Experiments. *Appl. Statist.*, **26**, No. 3, 321-326.
- [12] Franklin M.F. (1985). Selecting Defining Contrasts and Confounded Effects in p^{n-m} Factorial Experiments. *Technometrics*, **27**, No. 2, 165-172.
- [13] Kobilinsky A. (1985) Orthogonal factorial designs for quantitative factors. *Statist. Decisions, Suppl.*, **2**, 275-285.
- [14] Kobilinsky A. (1989). Randomization of a cartesian block structure. Document interne.
- [15] Kobilinsky A. (1990). Complex linear models and cyclic designs. *Linear Algebra and Applications*, **127**, 227-282.
- [16] Kobilinsky A. (1997). Les plans factoriels. Chap 3, p69–209. In : *Plans d'expériences : applications à l'entreprise*, Eds: Driesbeke, Fine, Saporta. Technip, Paris. 509p.
- [17] Kobilinsky A. (1994). Automatic generation of asymmetrical regular designs. Document interne.
- [18] Kobilinsky A., Fliss M., 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.

- [19] Kobilinsky A. and Monod H. (1991). Experimental design generated by group morphisms: an introduction. *Scand. J. Statist.*, **18**, 119-134.
- [20] Kobilinsky A. and Monod H. (1995). Juxtaposition of regular factorial designs and the complex linear model. *Scand. J. Statist* **22**, n° 2, 223-254.
- [21] Margolin B.H. (1969). Resolution IV Fractional Factorial Designs. *J. Roy. Statist.Soc. Ser. B*, **31**, 514-523.
- [22] Patterson H.D. (1976). Generation of factorial designs. *J. Roy. Statist.Soc. Ser. B*, **38**, 175-179.
- [23] Patterson H.D. and Bailey R.A. (1978). Design keys for factorial experiments. *Appl. Statist.*, **27**, 335-343.
- [24] Voss D.T. (1988). Single-generator generalized cyclic factorial designs as pseudo-factor designs. *Ann. Statist.*, **16**, 1723-1726.
- [25] Voss D.T. (1993). Comparing the classes of single replicate generalized cyclic factorial designs with and without pseudofactors. *J. Statist. Plann. Inference* **35**, 113-120.

Liste des fichiers .REG correspondant aux exemples de la notice

ex333.reg	§ 2.2.1
ex664.reg	§ 2.3.1
robot1a.reg	§ 3.1.2
robot1b.reg	§ 3.1.2
robpl1r4.reg	§ 3.2.1
robpl1r5.reg	§ 3.2.2
essai.reg	§ 3.3
robpl2r5.reg	§ 3.3.1
robpl2r4.reg	§ 3.3.2
rop2f4r5.reg	§ 3.3.3
essai1.reg	§ 3.3.4
rop2f4r4.reg	§ 3.3.4
hygien1.reg	§ 4.1.4
hygien2.reg	§ 4.1.4
hygien3.reg	§ 4.1.4
hygien4.reg	§ 4.1.4
quant1.reg	§ 4.2.2
quant2.reg	§ 4.2.3.1
rhizo6.reg	§ 4.2.3.2