

# BLOCS INCOMPLETS ÉQUILIBRÉS VERSUS PLANS OPTIMAUX

Jérôme PAGÈS, Emmanuel PÉRINEL<sup>1</sup>

## RÉSUMÉ

Dans l'organisation d'une expérimentation en blocs, lorsque le nombre de traitements étudiés excède la taille des blocs, il est classique de recourir à un plan en blocs incomplets équilibrés (B.I.E.). Lorsque ce dernier n'existe pas pour une configuration *nombre de blocs / nombre de traitements / nombre de traitements par bloc* donnée, l'expérimentateur peut préférer se ramener à un B.I.E. dont la configuration est « proche » de celle étudiée, pour bénéficier des propriétés statistiques remarquables de tels plans. Ce faisant, il renonce en général à des données potentiellement disponibles. À l'inverse, l'expérimentateur peut envisager de construire le plan en respectant strictement la configuration initiale afin de prendre en compte la totalité de ces données, quitte à dégrader les propriétés d'équilibre du plan. Ces deux approches sont d'abord comparées à partir de deux exemples où le plan à construire peut être obtenu de façon simple en adaptant un plan en B.I.E. ; les exemples proposés montrent qu'il est préférable, en terme d'efficacité, de conserver la configuration initiale, plutôt que de se ramener à un B.I.E. Pour des configurations plus complexes, les plans peuvent être construits avec profit à partir d'un algorithme de construction de plans optimaux. Les simulations présentées pour un grand nombre de configurations montrent qu'à défaut de présenter systématiquement les équilibres remarquables d'un B.I.E., les plans optimaux ainsi obtenus possèdent une efficacité comparable à ces plans de référence.

*Mots-clés* : blocs incomplets équilibrés, plan optimal, efficacité.

## ABSTRACT

In the organization of an experimental design, if the number of studied treatments is greater than the size of a block, the experimenter classically use a balanced incomplete block design (B.I.B.). When the latter does not exist for the studied configuration *number of blocks / number of treatments / number of treatments per block*, the experimenter may prefer to use a B.I.B. which configuration is close to the studied one, in order to profit from the remarkable properties of such designs. By doing this, he gives up potentially available data. On the other hand, the experimenter may wish to build the design by strictly respecting the studied configuration, in order to take into account of all the potentially available data, even if it means to degrade to balance of the design. These two approaches are first compared on the basis of two examples where the design to be built can be obtained in a simple way by enlarging a B.I.B. The presented examples show that it is preferable, from an efficiency point of view, to respect the initial configuration,

---

1. Laboratoire de mathématiques appliquées, Agrocampus Rennes 65, rue de Saint-Brieuc, 35042 Rennes Cedex, France.

pages@agrocampus-rennes.fr, perinel@agrocampus-rennes.fr

rather than to use a B.I.B. For more complex configurations, the designs can be built using an algorithm for the construction of optimal designs. Simulations obtained for a large number of configurations show that, even if they do not offer the remarkable balance of a B.I.B., optimal designs always have an efficiency comparable to that of these reference designs.

*Keywords* : balanced incomplete block design, optimal design, efficiency.

## 1. Introduction

L'organisation d'une expérimentation en blocs consiste à constituer des ensembles d'unités expérimentales homogènes appelés blocs, puis à leur attribuer au hasard différents traitements en vue de les comparer. Ainsi, en agronomie, si l'on étudie des régimes alimentaires pour animaux (respectivement des variétés), on s'efforcera de les répartir au préalable au sein de groupes d'animaux (respectivement de parcelles de terrain) présentant des caractéristiques aussi similaires que possible. Les dispositifs en blocs sont également largement utilisés en analyse sensorielle où des dégustateurs (jouant le rôle des blocs) évaluent successivement différents produits alimentaires ou cosmétiques, par exemple.

En pratique, on essaie toujours de faire en sorte que l'effet bloc soit non confondu – on dit aussi orthogonal en référence à l'interprétation géométrique du modèle linéaire – avec l'effet traitement. Pour cela, on attribue chacun des  $t$  traitements à chacun des  $b$  blocs (dispositifs en blocs complets). Lorsque le nombre  $k$  de traitements que l'on peut raisonnablement attribuer à un bloc au cours de l'expérimentation est inférieur au nombre de traitements étudiés, l'expérimentateur est alors confronté au problème de l'affectation des traitements aux blocs. Dans ce cas, il a généralement recours aux plans dits en blocs incomplets équilibrés, en abrégé B.I.E. (Bergonzini, 1995 ; Philippeau, 1985).

Ces plans, qui fournissent une affectation des traitements aux blocs, possèdent plusieurs propriétés d'équilibre remarquables ; en particulier, les différents couples de traitements sont présents un même nombre de fois dans un même bloc (pour des plans à une classe associée). Ces équilibres confèrent à ces plans des propriétés statistiques très intéressantes : propriétés d'optimalité, faible confusion entre effets blocs et traitements, formules simplifiées pour le calcul des effets en analyse de la variance. En contrepartie, les B.I.E. n'existent pas pour n'importe quelle configuration  $(b, t, k)$ . Il existe en effet plusieurs conditions nécessaires d'existence de ces dispositifs (voir par exemple Benoist *et al.*, 1994). Citons ici les deux plus connues :

$$bk = tr \quad \text{et} \quad r(k - 1) = \lambda(t - 1)$$

où  $r$  est le nombre de répétitions d'un traitement dans le plan et  $\lambda$  le nombre d'occurrences des différents couples de traitements. Aucune condition suffisante par contre n'est connue. Des tables de plans en B.I.E. peuvent être trouvées dans de nombreux ouvrages ; citons entre autres : Cochran et Cox (1968), Danzart (1994), Mathon et Rosa (1996) ou encore Wakeling *et al.* (2001).

En pratique, le nombre  $t$  de traitements étudiés est généralement fixé et les nombres  $b$  et  $k$  sont en général déterminés par leur maximum :  $b$ , le nombre de blocs disponibles ;  $k$ , le nombre maximum de traitements qu'il est possible ou raisonnable d'affecter à un bloc. Lorsqu'aucun B.I.E. n'existe pour une configuration  $(b, t, k)$  donnée, les deux stratégies suivantes se voient donc proposées à l'utilisateur :

- (1) Modifier les paramètres  $b$ ,  $t$  ou  $k$  de manière à se ramener à un B.I.E., afin de conserver de bonnes propriétés statistiques. Compte tenu des remarques précédentes, ceci conduira l'utilisateur à ne pas utiliser toutes les ressources disponibles (diminuer  $b$ ) et/ou restreindre le nombre de traitements attribués à un bloc (diminuer  $k$ ).
- (2) Conserver la configuration  $(b, t, k)$  de façon à intégrer à l'analyse statistique le plus de données possibles, quitte à « dégrader » les propriétés d'équilibre du plan.

Dans cette seconde approche, pour certaines configurations simples, la solution consistant à « adapter au mieux » un plan en B.I.E. dont les paramètres sont proches de ceux imposés à l'utilisateur peut convenir (par exemple en ajoutant à un B.I.E.  $(b-1, t, k)$  un bloc et en y affectant aléatoirement des traitements). Mais dans la majorité des cas, ces solutions s'avèrent trop complexes, d'où l'idée d'utiliser les possibilités offertes par les algorithmes de construction de plans optimaux. Ces plans sont surtout connus dans le cas de facteurs quantitatifs mais ils s'appliquent aussi au cas de facteurs qualitatifs. Leur intérêt majeur est de respecter parfaitement les contraintes de l'utilisateur. Leur inconvénient majeur est de dépendre du modèle choisi a priori. Or, dans le cas de facteurs qualitatifs, cet inconvénient est atténué : la question, par exemple, de l'introduction ou non de termes quadratiques dans le modèle, si influente sur le plan obtenu, ne se pose pas.

Cet article aborde le problème de l'évaluation d'un plan en blocs incomplets pour des configurations  $(b, t, k)$  quelconques, sous l'angle des stratégies (1) et (2) précédentes. Cette évaluation est réalisée en deux parties. La première partie étudie l'intérêt de conserver l'intégralité des données potentiellement disponibles dans deux situations où le plan en blocs retenu peut être obtenu « simplement ». La seconde partie aborde le problème de façon plus générale, lorsque le plan est construit sous l'approche des plans optimaux. Les plans ainsi obtenus sont ensuite évalués en termes d'efficacité à partir de nombreux exemples.

Dans un premier temps, il est nécessaire d'introduire les critères qui nous permettront d'évaluer la qualité d'un plan.

## 2. Critères d'évaluation d'un plan en blocs incomplets

Nous proposons de construire des plans d'expériences pour des modèles de la forme :

$$y_{ij} = \mu + \beta_i + \tau_j + \epsilon_{ij} \quad (1)$$

où  $y_{ij}$  est la réponse étudiée;  $\mu$  est l'effet moyen;  $\beta_i$  est le coefficient (paramètre) associé au bloc  $i$  ( $i = 1, \dots, b$ );  $\tau_j$  est le coefficient associé au traitement  $j$  ( $j = 1, \dots, t$ );  $\epsilon_{ij}$  est une erreur d'espérance nulle et de variance constante  $\sigma^2$ . Dans la suite,  $\mathbf{X}$  désigne la matrice associée à ce modèle, après prise en compte des contraintes habituelles sur les paramètres ( $\sum_{i=1}^b \beta_i = 0$  et  $\sum_{j=1}^t \tau_j = 0$ ). La matrice  $\mathbf{X}$  est donc de dimensions  $(n, p)$ , où  $n$  est le nombre d'essais et  $p$  est le nombre de paramètres indépendants du modèle, égal à  $(b + t - 1)$ . Les matrices  $(\mathbf{X}'\mathbf{X})$  et  $(\mathbf{X}'\mathbf{X})^{-1}$  sont de dimension  $(p, p)$ .

**Qualité d'un plan.** Dans une expérimentation en blocs, la connaissance des différences existant entre les traitements revêt une importance particulière. Il est donc important de fournir des plans tels que les coefficients des traitements dans le modèle seront estimés avec une grande précision. La précision d'un coefficient traitement  $\tau_j$  est mesurée classiquement par la variance de l'estimateur de ce coefficient, que l'on notera  $V(\hat{\tau}_j)$  :

- dans le cas général, cette variance est obtenue à partir de la matrice variance du vecteur des paramètres qui est donnée par la formule classique :

$$\sigma^2(\mathbf{X}'\mathbf{X})^{-1} \tag{2}$$

- dans le cas d'un plan complet comportant  $(bt)$  observations, la variance  $V(\hat{\tau}_j)$  s'exprime simplement par :

$$\frac{(t-1)}{bt} \sigma^2 \tag{3}$$

- dans le cas d'un plan en B.I.E. comportant  $(bk)$  observations, la variance  $V(\hat{\tau}_j)$  est calculée par la formule (Philippeau, 1985, page 124) :

$$\frac{(t-1)^2}{bt(k-1)} \sigma^2 \tag{4}$$

Cependant, lorsque l'on souhaite comparer la précision de deux coefficients, la variance a peu de sens si les traitements n'ont pas le même nombre de répétitions. On définit pour cela une variance normalisée, en multipliant la variance par le nombre  $n_j$  de répétitions associé au traitement en question (Kobilinsky, 1997, page 173) :

$$V_u(\hat{\tau}_j) = n_j V(\hat{\tau}_j) \tag{5}$$

$V_u(\hat{\tau}_j)$  est appelée *variance par unité* du  $j^{\text{e}}$  coefficient traitement.

**Comparaison de deux plans.** La variance par unité définie ci-dessus permet de définir de façon immédiate une mesure de comparaison de deux plans pour un traitement donné (sous réserve bien entendu que la variance  $\sigma^2$  reste la même). On calcule pour cela le rapport entre les variances par unité du traitement en question dans les deux plans. Lorsque l'on compare deux plans,

on est souvent en présence d'un plan à évaluer et d'un plan de référence. La mesure suivante, appelée *V*-efficacité, permet d'évaluer la précision relative d'un coefficient traitement  $\tau_j$  du plan à évaluer par rapport au plan de référence :

$$V\text{-eff}(\hat{\tau}_j) = \left[ \frac{V_u^*(\hat{\tau}_j)}{V_u(\hat{\tau}_j)} \right] \quad (6)$$

où  $V_u(\hat{\tau}_j)$  est la variance par unité du  $j^e$  traitement pour le plan à évaluer et  $V_u^*(\hat{\tau}_j)$  la variance par unité de ce même coefficient dans le plan de référence. Dans ce travail, le plan de référence associé au plan à évaluer sera dans tous les cas de figure le plan complet qui lui correspond (incluant le même nombre  $bk$  d'essais).

Cette mesure suffit à comparer deux plans si les traitements sont équirépétés dans chacun des plans; les mesures  $V\text{-eff}(\hat{\tau}_j)$  sont alors égales pour tous les traitements. Ainsi, l'efficacité d'un B.I.E., par exemple, est définie par le rapport inverse entre la variance d'un coefficient traitement quelconque du B.I.E. et la variance d'un coefficient associé au plan complet à  $bk$  observations qui lui correspond, soit :

$$\frac{\frac{(t-1)}{bk} \sigma^2}{\frac{(t-1)^2}{bt(k-1)} \sigma^2} = \frac{t(k-1)}{(t-1)k} \quad (7)$$

Dans le cas où les traitements n'ont pas le même nombre de répétitions, une mesure globale d'efficacité, notée *V*-eff, peut être obtenue en réalisant la moyenne des efficacités individuelles  $V\text{-eff}(\hat{\tau}_j)$  sur l'ensemble des traitements (Jones et Kenwards, 1989) :

$$V\text{-eff} = \frac{1}{t} \sum_{j=1}^t V\text{-eff}(\hat{\tau}_j) \quad (8)$$

### 3. Faut-il utiliser un B.I.E. lorsque cela implique de limiter le nombre de mesures ?

#### 3.1. Exemple où recourir à un B.I.E. implique de réduire le nombre de blocs : $b = 9$ , $t = 7$ , $k = 3$

Supposons qu'un expérimentateur dispose de 9 blocs. Comme il existe un B.I.E. (7, 7, 3), il peut être tenté de n'utiliser que 7 blocs afin de s'y ramener. Dans ce cas particulier, s'il souhaite utiliser les 9 blocs, un plan satisfaisant est simple à construire : au B.I.E. (7, 7, 3) on ajoute deux blocs regroupant à eux deux, 6 des 7 traitements (cf. figure 1).

La comparaison entre ces deux plans se fait d'abord au travers des variances des estimateurs des coefficients traitements associées à ces deux plans (tableau 1).

BLOCS INCOMPLETS ÉQUILIBRÉS *VERSUS* PLANS OPTIMAUX

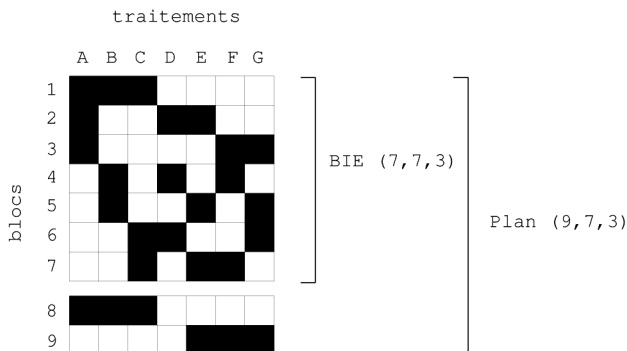


FIG 1. — Deux plans confrontés : (7, 7, 3) et (9, 7, 3). À l'intersection de la ligne  $i$  et de la colonne  $j$ , la case est noircie si le le traitement  $j$  est présent dans le bloc  $i$ .

TABLEAU 1. — Comparaison entre les deux plans de la figure 1. Les variances des effets traitement sont calculées à l'aide de la formule (4) pour le plan (7, 7, 3) et de la formule (2) pour le plan (9, 7, 3).

Configuration $(b, t, k)$	Variances des effets des traitements						
	A	B	C	D	E	F	G
7-7-3 (B.I.E.)	0.367	0.367	0.367	0.367	0.367	0.367	0.367
9-7-3	0.281	0.281	0.281	0.367	0.281	0.281	0.281

On remarque que, en ajoutant deux blocs, les variances diminuent globalement : ceci est dû à l'augmentation importante des données (27 au lieu de 21) : 0.367 pour les variances du plan (7, 7, 3); 0.281 pour les variances du plan (9, 7, 3) (excepté pour le traitement  $D$  dont la variance n'a pas changé en raison d'un nombre de mesures inchangé). Le tableau 2 rassemble les différents éléments permettant une comparaison en terme d'efficacité.

TABLEAU 2. — Efficacité du plan (9, 7, 3) comparée à celle du B.I.E. (7, 7, 3).

Configuration $(b, t, k)$	Traitement $D$		Autres traitements		Efficacité globale
	$V_u(\hat{\tau}_j)$	$V\text{-eff}(\hat{\tau}_j)$	$V_u(\hat{\tau}_j)$	$V\text{-eff}(\hat{\tau}_j)$	$V\text{-eff}$
7-7-3 (B.I.E.)	$0.367 \times 3 = 1.102$	0.777	$0.367 \times 3 = 1.102$	0.777	0.777
9-7-3	$0.367 \times 3 = 1.102$	0.779	$0.281 \times 4 = 1.124$	0.769	0.770

Les efficacités globales données dans ce tableau indiquent que l'ajout de deux nouveaux blocs n'a pas trop endommagé la structure du plan (0.770 contre 0.777 pour le B.I.E.). En effet, les variances (exceptée celle du traitement  $D$ , inchangée) ont diminué dans une proportion ( $\frac{367}{281} = 1.31$ ) à peine inférieure à l'inverse de l'augmentation du nombre de mesures par traitement ( $\frac{4}{3} = 1.33$ ).

*Conclusion.* Dans un tel cas où les 9 blocs sont disponibles, il eut été dommage de ne pas les utiliser tous.

**3.2. Exemple où recourir à un B.I.E. implique de diminuer le nombre de traitements par bloc :  $b = 7, t = 7, k = 5$**

Dans ce cas, l'expérimentateur a la possibilité d'affecter 5 traitements par bloc. Comme il existe un B.I.E.  $(7, 7, 4)$ , il peut être tenté de limiter à 4 le nombre de traitements par bloc. S'il souhaite conserver  $k = 5$ , il se heurte au problème de construction du plan. Celui-ci peut toutefois être construit relativement facilement sous la contrainte du double équilibre entre les blocs d'une part et entre les traitements d'autre part. Ces deux plans sont présentés dans la figure 2.

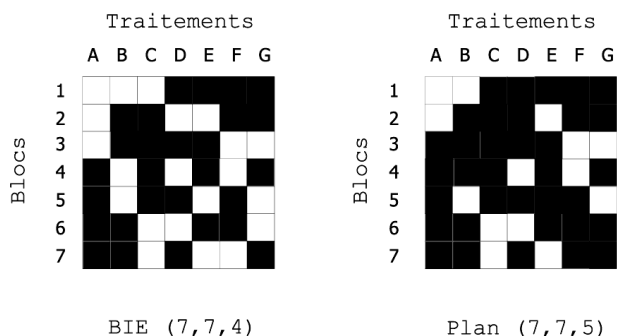


FIG 2. — Plan  $(7, 7, 5)$  construit en complétant le plan  $(7, 7, 4)$ . Remarque : le B.I.E.  $(7, 7, 4)$  est le complément (défini par les cases blanches) du B.I.E.  $(7, 7, 3)$  de la figure 1.

La comparaison entre ce plan  $(7, 7, 5)$  et le B.I.E.  $(7, 7, 4)$  tient pour l'essentiel dans le tableau 3.

TABLEAU 3. — Efficacité comparée des deux plans de la figure 2. La variance de chaque coefficient traitement est constante dans les deux plans.

Configuration $(b, t, k)$	Répétitions $n_j$	Variance $V(\hat{\tau}_j)$	Var. par unité $V_u(\hat{\tau}_j)$	V-eff
7-7-4 (B.I.E.)	4	0.245	0.980	0.875
7-7-5	5	0.184	0.921	0.931

En passant de 4 à 5 traitements par bloc (et donc de  $n = 28$  à  $n = 35$  observations), l'efficacité du plan (mesurée par  $V$ -eff) a augmenté, résultat à relier avec le fait que l'on se rapproche du plan complet. En conclusion, dans ce cas, il serait dommage de restreindre le nombre de traitements par bloc.

## 4. La solution des plans optimaux

### 4.1. Principe

Dans la majorité des situations, l'approche précédente consistant à adapter au mieux un B.I.E. n'est pas envisageable (plans volumineux, équilibres du plan difficiles à gérer «à la main»). On peut dans ce cas construire un plan respectant strictement la combinaison  $(b, t, k)$  proposée en utilisant un algorithme de construction de plans optimaux du type de ceux décrits dans (Fedorov, 1972), (Atkinson *et al.*, 1992) ou (Benoist *et al.*, 1994). Le principe de ces algorithmes est le suivant :

1. Définir l'ensemble des observations candidates ; dans notre cas, cet ensemble est défini par l'ensemble des  $b \times t$  couples (bloc, traitement) du plan complet.
2. Générer un plan initial par tirage au hasard de  $b \times k$  observations dans cet ensemble candidat.
3. Améliorer ce plan pas à pas ; à chaque pas, une observation du plan construit est remplacée par une observation candidate de façon à accroître le plus possible  $\det(\mathbf{X}'\mathbf{X})$  (ou, de manière équivalente, à réduire le plus possible  $\det(\mathbf{X}'\mathbf{X})^{-1}$ ).

*D'où le nom d'algorithme d'échanges.*

Le plan obtenu par un tel algorithme est dit  $D$ - optimal (en référence la valeur du déterminant qu'il optimise). De la même façon, on peut construire un plan  $A$ - optimal en substituant au critère du déterminant celui défini par la trace de la matrice  $(\mathbf{X}'\mathbf{X})^{-1}$ . Naturellement, ce type d'algorithme ne garantit pas l'obtention d'un optimum global. En pratique, on les exécute plusieurs fois et l'on conserve la meilleure solution. La pratique en maints domaines montre que l'on obtient ainsi des plans au moins acceptables à défaut d'être optimaux.

Dans l'algorithme d'échanges usuel, on introduit le nombre de modalités de chacun des facteurs (ici  $b$  et  $t$ ) et le nombre d'essais (ici  $b \times k$ ) du plan à construire. Introduire la contrainte de l'équilibre des blocs (même nombre de mesures par bloc) nécessite une modification de l'algorithme. Afin d'obtenir des résultats plus généraux, nous n'avons pas modifié l'algorithme à ce niveau, nous limitant à vérifier en final si la solution proposée satisfait cette contrainte. Ainsi, les plans décrits ci-après peuvent être obtenus avec n'importe quel logiciel de construction de plans optimaux : par exemple, la procédure OPTX du logiciel SAS, le logiciel NEMROD (NEMROD, 1997) ou encore l'algorithme OPTDSGN développé par Tjur (Tjur, 1989). Pour les besoins de cet article – et en particulier des tableaux de synthèse des différentes simulations réalisées – nous avons développé un programme spécifique en suivant l'algorithme de base décrit dans (Benoist *et al.*, 1994). Une version de cet algorithme peut être obtenue sur simple demande auprès des auteurs.



## 4.2. Exemples de plans construits par algorithme dans le cas où l'optimum est connu

Nous avons soumis l'algorithme d'échanges décrit ci-dessus à différentes configurations  $(b, t, k)$  pour lesquelles il existe un B.I.E. Pour chacune de ces configurations, 1000 plans sont construits sur la base d'autant d'initialisations aléatoires. On calcule alors le nombre de plans équilibrés, le nombre de fois où l'optimum est trouvé par l'algorithme et une mesure d'écart à l'optimum en formant le rapport entre l'efficacité moyenne des 100 meilleurs plans construits et l'efficacité (connue) du B.I.E. Les résultats sont synthétisés dans le tableau 4.

TABLEAU 4. — Constructions algorithmiques de 7 plans dans le cas où l'optimum (un B.I.E.), est connu. Résultats obtenus pour 1000 initialisations aléatoires.

PROPRIÉTÉS	Configuration $(b, t, k)$ et nombre $n$ d'essais						
	(7, 7, 3) $n = 21$	(11, 11, 5) $n = 55$	(15, 6, 4) $n = 60$	(16, 16, 6) $n = 96$	(42, 7, 4) $n = 168$	(64, 16, 6) $n = 384$	(90, 10, 5) $n = 450$
Plans équilibrés <sup>1</sup>	100%	100%	100%	100%	100%	100%	100%
Optimum <sup>2</sup>	20%	0%	15%	0%	8%	0%	0%
$\frac{V - \text{eff}}{V - \text{eff (b.i.e.)}}$ <sup>3</sup>	1	0.993	1	0.993	0.999	0.998	0.999

- (1) pourcentage de plans construits où chaque bloc contient le même nombre de traitements et où chaque traitement est présent un même nombre de fois dans le plan.
- (2) pourcentage de B.I.E. obtenus pour les 1000 plans construits.
- (3) ratio d'efficacité calculé pour les 100 meilleurs plans.

Exemple : pour les 1000 plans construits dans la configuration  $(15, 6, 4)$  : tous sont équilibrés ; 150 sont des B.I.E. ; le ratio d'efficacité est donc égal à 1.

*Conclusion.* Au vu des simulations réalisées, quelle que soit la configuration  $(b, t, k)$  proposée, l'algorithme fournit toujours un plan équilibré. L'optimum est assez souvent trouvé pour des plans de petite taille. Lorsqu'il n'est pas obtenu (comme pour la configuration  $(90, 10, 5)$ ), les indicateurs d'efficacité montrent que les 100 meilleurs plans construits ont des propriétés toujours comparables à celles de l'optimum (efficacité voisine de 1).

**Autres exemples.** L'algorithme a également été appliqué aux plans proposés dans les sections 3.1 et 3.2. Il a conduit aux résultats suivants :

- **Plan (9,7,3).** Le meilleur plan fourni par l'algorithme a une  $V$ -efficacité (0.772) à peine supérieure à celle associée au plan proposé dans la figure 1 (0.770). Sans être optimale, l'approche précédemment proposée conduisait donc bien à un plan très satisfaisant.
- **Plan (7,7,5).** Dans ce cas, l'algorithme a fourni un plan identique à celui de la figure 2. Ceci renforce l'approche intuitive adoptée plus haut.

### 4.3. Cas où l'optimum n'est pas connu

Dans une situation de blocs incomplets, l'utilisation d'un dispositif en B.I.E. assure d'obtenir une efficacité optimale (voir équation (7)). Le graphique 3 montre ainsi l'efficacité de différents plans en B.I.E. en fonction du nombre  $t$  de traitements et du nombre  $k$  de traitements par bloc (pour les configurations les plus courantes vérifiant  $k < 10$ ).

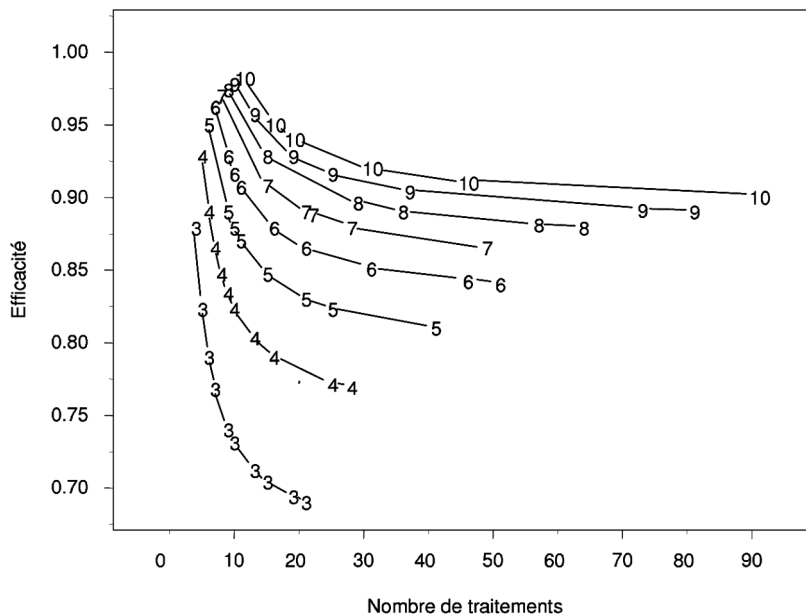


FIG 3. — Efficacité de quelques plans en B.I.E. en fonction du nombre  $t$  de traitements (en abscisse) et du nombre  $k$  de traitements par bloc (valeur affectée à un point du graphe). Chaque courbe décrit donc l'efficacité calculée par la formule (7) à  $k$  fixé.

Lorsqu'il n'existe pas de B.I.E. pour la configuration  $(b, t, k)$  étudiée, l'efficacité maximale que l'on peut obtenir dépend de plusieurs contraintes existant sur les paramètres  $b, t$  et  $k$ . (Williams *et al.*, 1977) proposent ainsi différentes bornes supérieures selon que  $t \leq b$ ,  $t > b$  ou que le plan est décompensable (*resolvable design*) ou non. Des valeurs pour ces bornes sont également fournies par le logiciel OPTDSGN (Tjur, 1989) lors de la construction d'un plan en blocs incomplets optimal. En règle générale, les valeurs de ces bornes sont très proches (et inférieures) à celles données par la formule (7). Par exemple, pour une configuration  $(b = 24, t = 12, k = 5)$ , la plus grande borne supérieure connue (fournie par l'algorithme OPTDSGN développé par Tjur) est de 87.16% et la formule (7) conduit à une efficacité égale à  $\frac{(5-1)12}{5(12-1)} = 87.27\%$ .

L'évaluation d'un plan construit par algorithme pour une configuration  $(b, t, k)$  quelconque peut donc être réalisée en considérant l'écart entre le niveau

d'efficacité obtenu par l'algorithme et l'efficacité maximale (approximée par (7)) que l'on peut attendre pour la configuration en question. Pour ce faire, nous avons procédé comme suit :

1. Plusieurs configurations  $(b, t, k)$  ont été soumises à un algorithme de construction de plan en blocs incomplets optimal. Nous nous sommes limités ici à des situations où le nombre d'observations par traitement ( $r$ ) est supérieur à 8 (à l'exception de quelques situations discutées ci-dessous) et où l'équilibre des traitements est possible (c'est-à-dire avec  $r = \frac{bk}{t}$  entier). Les résultats numériques obtenus sont présentés dans le tableau 5 ;
2. Les efficacités des plans construits sont ensuite portées sur un graphique tel celui de la figure 3, permettant d'apprécier visuellement le niveau d'efficacité fourni par de tels algorithmes. Les résultats graphiques sont présentés dans la figure 4.

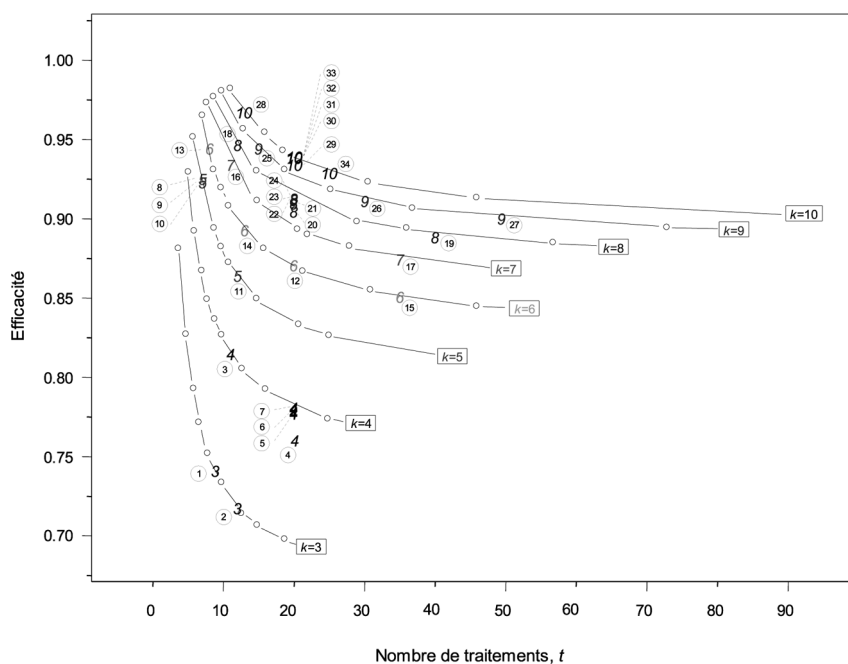


FIG 4. — Efficacité des plans optimaux obtenus par algorithme en comparaison de l'efficacité des B.I.E. Comme pour les B.I.E. de la figure 3, les 34 configurations du tableau 5 sont repérées par leur nombre  $k$  de traitements par bloc. Certaines configurations étant quasi superposées, on leur a également associé leur numéro *conf.* défini dans le tableau 5.

L'étude de la figure 4 conduit aux remarques suivantes :

- Pour chaque configuration proposée, le niveau d'efficacité obtenu n'est jamais très éloigné de sa courbe de référence (i.e. celle associée à un même

BLOCS INCOMPLETS ÉQUILIBRÉS *VERSUS* PLANS OPTIMAUX

TABLEAU 5. — Efficacité de plans construits par algorithme pour 34 configurations  $(b, t, k)$  ( $r$  est le nombre d'observations par traitement). *conf.* est le numéro de la configuration testée;  $V$ -eff est l'efficacité du plan construit par algorithme (reportée sur le graphe de la figure 4);  $V$ -eff<sub>BIE</sub> est l'efficacité de référence fournie par la formule (7); Eff<sub>rel</sub> est l'efficacité relative obtenue par le rapport des deux efficacités précédentes.

<i>conf.</i>	$b$	$t$	$k$	$r$	$V$ -eff	$V$ -eff <sub>BIE</sub>	Eff <sub>rel</sub>
1	12	9	3	4	0.740	0.750	0.9867
2	24	12	3	6	0.721	0.727	0.9914
3	22	11	4	8	0.822	0.825	0.9964
4	20	20	4	4	0.765	0.789	0.9690
5	40	20	4	8	0.784	0.789	0.9931
6	60	20	4	12	0.787	0.789	0.9969
7	70	20	4	14	0.788	0.789	0.9981
8	7	7	5	5	0.931	0.933	0.9975
9	14	7	5	10	0.933	0.933	0.9996
10	28	7	5	20	0.933	0.933	0.9996
11	24	12	5	10	0.870	0.873	0.9969
12	40	20	6	12	0.876	0.877	0.9986
13	12	8	6	9	0.952	0.952	0.9996
14	13	13	6	6	0.900	0.903	0.9969
15	35	35	6	6	0.851	0.858	0.9920
16	11	11	7	7	0.942	0.943	0.9991
17	35	35	7	7	0.878	0.882	0.9951

<i>conf.</i>	$b$	$t$	$k$	$r$	$V$ -eff	$V$ -eff <sub>BIE</sub>	Eff <sub>rel</sub>
18	15	12	8	10	0.954	0.955	0.9994
19	40	40	8	8	0.892	0.897	0.9939
20	10	20	8	4	0.912	0.921	0.9902
21	20	20	8	8	0.920	0.921	0.9989
22	25	20	8	10	0.920	0.921	0.9989
23	30	20	8	12	0.920	0.921	0.9989
24	35	20	8	14	0.921	0.921	0.9999
25	30	15	9	18	0.952	0.952	0.9996
26	60	30	9	18	0.919	0.920	0.9994
27	50	50	9	9	0.940	0.907	0.9967
28	13	13	10	10	0.975	0.975	1.0000
29	8	20	10	4	0.942	0.947	0.9943
30	16	20	10	8	0.946	0.947	0.9986
31	20	20	10	10	0.947	0.947	0.9996
32	24	20	10	12	0.947	0.947	0.9996
33	28	20	10	14	0.947	0.947	0.9996
34	25	25	10	10	0.937	0.938	0.9995

nombre de traitements par bloc). Cela signifie que les propriétés statistiques du plan obtenu sont comparables à celles d'un B.I.E.

- Pour des nombres  $t$  et  $k$  fixés, le niveau d'efficacité obtenu dépend du nombre  $r$  de répétitions de chaque traitement. Par exemple, avec  $k = 4$  et  $t = 20$ , on observe que le niveau d'efficacité se rapproche de son niveau optimal lorsque le nombre de blocs passe de 20 à 70, le nombre de répétitions passant de 4 à 14.
- Pour un même nombre de répétitions par traitement, le niveau d'efficacité obtenu est d'autant plus proche de sa borne supérieure que le rapport  $\frac{k}{t}$  est élevé. Ainsi, pour  $t = 20$  traitements et  $r = 4$  répétitions par traitement, l'efficacité atteint 99.43% de sa borne supérieure si  $k = 10$ , contre 96.90% pour  $k = 4$ .

En règle générale, on observe donc qu'à partir de  $r = 8$  observations par traitement (la quasi-majorité de nos configurations testées), l'efficacité d'un plan optimal construit par algorithme est toujours très satisfaisante.

## 5. Conclusion

Les différents exemples étudiés dans cette section montrent que l'utilisation d'algorithmes de construction de plans optimaux est une alternative très intéressante aux plans en B.I.E., car :

- dans le cas où le plan en B.I.E. existe, lorsque l'algorithme ne parvient pas à retrouver le B.I.E. (cas des plans de grande taille), les plans construits sont toujours acceptables du point de vue de leur efficacité;
- ils permettent de prendre en compte l'ensemble des observations potentiellement disponibles (c'est-à-dire de ne pas écarter des observations en vue de se ramener à un B.I.E.) : les plans ainsi obtenus possèdent une efficacité comparable (voire supérieure) à ces plans de référence;
- ils satisfont une contrainte particulièrement importante en pratique : l'équilibre du plan pour les blocs et les traitements (chaque bloc contient le même nombre de traitements et chaque traitement est présent un même nombre de fois dans le plan) lorsque celui-ci est possible.

## Références

- ATKINSON A.C. and A.N. DONEV (1992), Optimum experimental designs, *Oxford Statistical Science*, Series 8, Oxford University Press.
- BENOIST D., Y. TOURBIER and S. GERMAIN-TOURBIER (1994), Plans d'expériences : construction et analyse, Paris, *Lavoisier*.
- BERGONZINI J.C. (1995), Analyse et planification expériences. Les dispositifs en blocs, Ed. *Masson*, Paris.
- COCHRAN W.G. and G.M. COX (1968), Experimental designs, 2nd Edition. New York, *Wiley*
- DANZART M. (1994), Les plans d'expériences en analyse sensorielle, Multigraphié ENSIAA. MASSY.

- FEDOROV V.V. (1972), Theory of optimal experiments, *Academic Press*, New York.
- JONES B. and M.G. KENWARD (1989), Design and Analysis of Cross-Over Trials, *Chapman and Hall / CRC*.
- KOBILINSKY A. (1997), Les plans factoriels, in : *Plans d'expériences. Application à l'entreprise, Technip*, Dreesbeke, J.J., Fine J. and G. Saporta Eds., Paris.
- MATHON R. and A. ROSA (1996), The CRC handbook of combinatorial designs, CRC Press Series on Discrete Mathematics and its Applications, Boca Raton.
- NEMROD (1997), Logiciel de construction de plans d'expériences, Version 3.1., diffusé par : LPRAI – 17, rue des Ecoles, F-13004 Marseille.
- PHILIPPEAU G. (1985). Théorie des plans d'expérience : application à l'agronomie. 2e édition. Paris : Service des études statistiques de l'I.T.C.F., 205 p.
- SAS Version 8, 1989-1996, SAS Institute Inc., Cary, NC, USA.
- TJUR T. (1989), OPTDSGN Software, Optimization of block designs, Institute of mathematical statistics, Copenhagen.
- WAKELING I.N., HASTED A. and D. BUCK (2001), Cyclic presentation order designs for consumer research, *Food Quality and Preference*, 12, 39-46.
- WILLIAMS E.R. and H.D. PATTERSON (1977), Upper bounds for efficiency factors in block designs, *Australian Journal of scientific research*, Series A, **19(3)**, 194-201.