

# Notions essentielles pour l'utilisation de la méthodologie des plans d'expériences

Ce chapitre introductif est destiné à expliquer les notions fondamentales nécessaires à la mise en place et les limites d'un plan d'expériences.

Il est essentiel de bien comprendre ces différents points afin de mieux comprendre la construction des plans d'expériences et les différences qui rendront une stratégie expérimentale plus adaptée à un problème donné ou plus judicieuse au regard des objectifs techniques recherchés.

Un vocabulaire est aussi associé à ces notions fondamentales de la méthodologie des plans d'expériences, il faut l'acquérir à l'aide de quelques exemples simples.

## A) Introduction : Dans quelles conditions utiliser cette méthodologie ?

En premier lieu, la méthodologie des plans d'expériences est destinée à l'établissement d'un **modèle empirique purement descriptif** d'un phénomène dépendant de plusieurs variables. Ce modèle peut s'écrire dans les cas les plus complexes :

$$Y = a_0 + \sum_1^i a_i X_i + \sum_1^i a_{ij} X_i X_j + \sum_1^i a_{ijk} X_i X_j X_k \dots + \sum_1^i a_{ii} X_i^2 + \sum_1^i a_{iii} X_i^3$$

Ce modèle décrira le comportement d'un phénomène **Y** en fonction de différents paramètres **X<sub>i</sub>** comporte :

- des termes linéaires ;
- des termes d'interactions faisant intervenir 2 paramètres ou plus ;
- des termes quadratiques, voir cubiques.

Le but des expériences qui seront proposées sera donc d'assurer la détermination la plus précise possible des coefficients **a<sub>i</sub>** de ce modèle **avec un minimum d'expériences réalisées**.

Il est rarement nécessaire d'utiliser des termes cubiques sauf dans quelques cas particuliers.

Cette description d'un phénomène, qui peut paraître simpliste, est cependant largement suffisante si l'objet de cette modélisation est **de pouvoir prédire** le phénomène étudié dans le but, **non pas de le comprendre, mais de l'optimiser**.

Les plans d'expériences sont donc une méthodologie générale pour répondre à une demande précise à un moindre coût économique (notamment en limitant le nombre d'expériences réalisées) tout en assurant certaines propriétés statistiques aux modèles établis (voir chapitre 3).

La démarche mise en œuvre permettra :

- d'explorer un domaine expérimental donné par les valeurs minimales et maximales des différents paramètres ( $X_i$ ) de l'étude de **façon optimale** ;
- de réaliser une étude séquentielle permettant de **limiter le nombre d'expériences** en ajustant les essais réalisés à la complexité du **modèle nécessaire** pour décrire le phénomène étudié ;
- de déterminer **un modèle robuste** donnant une valeur prédite aussi proche que possible de la valeur vraie dans le domaine d'étude choisi ;
- de **détecter facilement** le manque d'adéquation du modèle ;
- de **détecter facilement un point aberrant** dû à une mauvaise mesure ou une mauvaise transcription des données. Et donc de pouvoir l'exclure de l'analyse ;
- de **limiter le nombre d'expériences** mais aussi le **nombre de niveaux nécessaires** pour étudier un paramètre donné ;
- d'estimer de façon précise **l'erreur expérimentale** ;
- de diviser les expériences à réaliser en **blocs** statistiquement équilibrés ;
- d'être insensible à la non normalité éventuelles de certaines réponses.

Certains termes (non normalité, blocs) seront explicités dans cet ouvrage.

## B) Un exemple pour fixer les idées et le vocabulaire

Cet exemple servira de support pour la mise en place et la comparaison de différentes stratégies expérimentales.

Soit un avion en papier construit avec les 4-5 étapes suivantes :

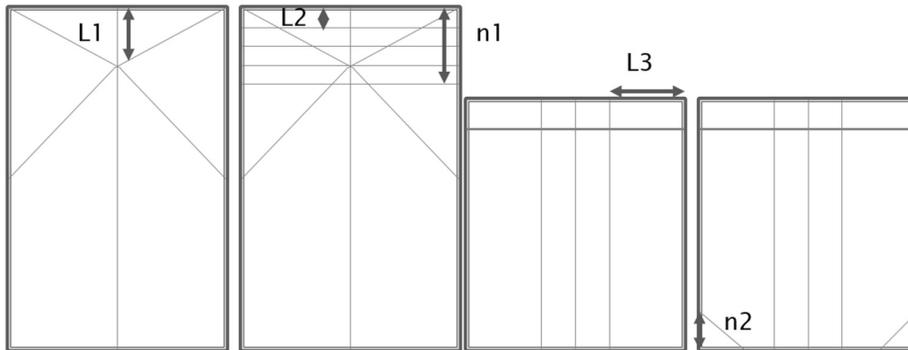


Figure 1 : Construction de l'avion utilisé comme exemple.

La construction de cet avion dépend de nombreux paramètres que nous listerons plus loin, mais pour peu que l'on limite l'étude aux seuls paramètres géométriques présentés ici (L1, L2, n1, L3, n2) on peut choisir d'étudier le temps de vol de l'avion en fonction de ces 5 paramètres. Comment procéder ?

### B-1) Méthode OVAT (one variable at time)

La réalisation d'un avion dont le temps de vol sera celui recherché (ou défini par l'auteur de l'étude) va nécessiter la description physique des différents phénomènes physiques de portance et de trainée en fonction des différents paramètres choisis. Cela va donc nécessiter, pour chaque paramètre, une étude afin de déterminer dans les équations correspondant aux modèles physiques toutes les valeurs nécessaires.

Cette étude, destinée à la détermination d'un modèle de compréhension permettra d'établir des lois physiques en testant chaque paramètre dans une grande gamme de valeurs différentes.

Les expériences servant à la détermination de ces modèles physiques seront donc conduites en fixant certains paramètres pendant que d'autres seront modifiés et sont destinées à comprendre les lois physiques gérant le vol de cet avion.

**Ce type d'étude n'est pas l'objet de la méthodologie des plans d'expériences.**

### B-2) Optimisation par la méthode OVAT

L'expérimentateur peut choisir de chercher à optimiser le temps de vol de son avion **sans passer** par la détermination des modèles physiques de portance et de trainée.

Il cherche donc, à être en mesure de déterminer l'avion ayant le **meilleur** temps de vol en réalisant des avions différents dont il va mesurer les temps de vols. Son objectif est donc de trouver, par une série d'expériences, quel avion aura le « meilleur » temps de vol.

Sans connaissances particulières en plan d'expériences, il va, comme c'est souvent le cas, fixer certains paramètres et n'en faire varier qu'un à la fois.

Si on limite la représentation à 2 paramètres il réalisera donc les expériences suivantes :

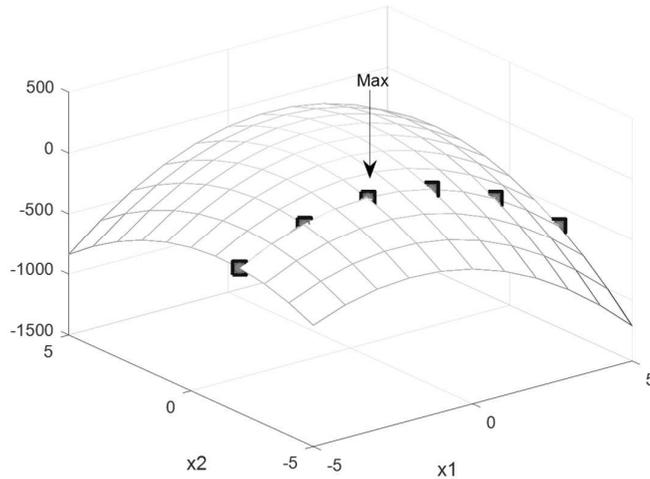


Figure 2 : Optimisation OVAT du 1<sup>er</sup> paramètre

Une fois obtenue la « meilleure » valeur pour le 1<sup>er</sup> paramètre, il fera varier un second paramètre en fixant le premier à la valeur « optimale » déterminée.

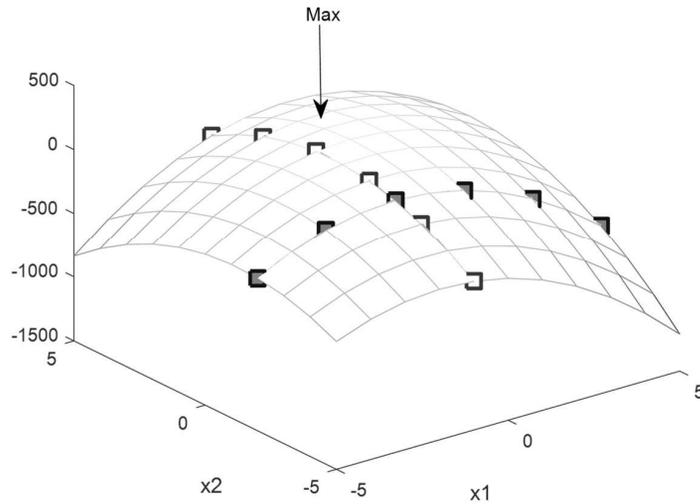


Figure 3 : Optimisation OVAT du 2<sup>ème</sup> paramètre

S'il s'arrête à ce stade, il pensera à tort, pour ces deux paramètres, avoir trouvé un optimum alors que la représentation de la nappe montre bien qu'il n'est pas à l'optimum (en 0,0).

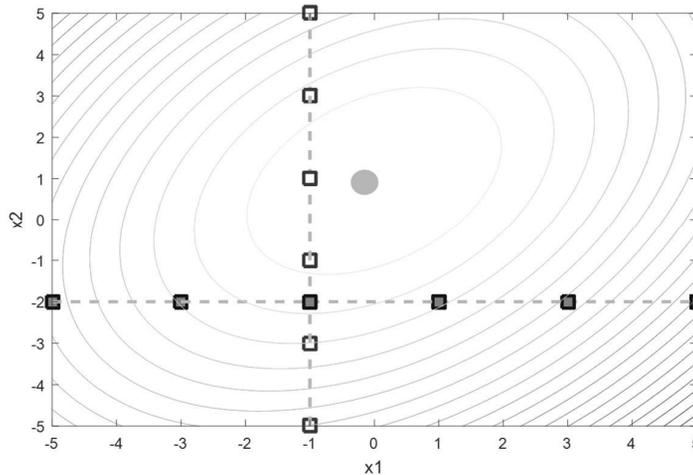


Figure 4 : Optimum déterminé par OVAT

Cette démarche, où l'expérimentateur ne se base que sur le résultat d'expériences pour déterminer un optimum, est très courante et montre très vite (comme ici à seulement 2 dimensions) ses limites. Pour obtenir l'optimum du problème, on voit bien que l'expérimentateur a deux solutions :

- répéter sa démarche (fixer à nouveau le 1<sup>er</sup> paramètre, puis le second, etc.) et donc multiplier le nombre d'expériences réalisées sans connaître a priori le nombre d'expériences qui devront être réalisées en tout. C'est l'objet d'une méthode décrite plus loin qui ne conduit pas à l'établissement d'un modèle ;
- essayer de déterminer l'équation de la nappe décrivant le phénomène étudié et en trouver le maximum par un calcul mathématique avec un nombre limité d'expériences.

**C'est cette seconde solution qui relève de la méthodologie des plans d'expériences.**

### **B-3) Optimisation expérimentale par la méthodologie des plans d'expériences**

Pour cet exemple (en se limitant à 2 paramètres), la méthodologie serait la suivante :

- postuler un modèle linéaire sans interaction  $Y = a_0 + a_1X_1 + a_2 X_2$  ;
- construire une liste d'expériences permettant de déterminer à moindre coût les coefficients  $a_0$ ,  $a_1$ ,  $a_2$  ;
- calculer ces coefficients ;
- vérifier si ce modèle est suffisant pour décrire le phénomène (on considère ici que **non**) ;
- postuler un modèle linéaire avec interaction  $Y = a_0 + a_1X_1 + a_2 X_2 + a_{12}X_1X_2$  ;
- construire une liste d'expériences permettant de déterminer à moindre coût les coefficients  $a_0$ ,  $a_1$ ,  $a_2$  et  $a_{12}$  ;
- vérifier si ce modèle est suffisant (on considère ici que **non**) ;
- postuler un modèle quadratique  $Y = a_0 + a_1X_1 + a_2 X_2 + a_{12}X_1X_2 + a_{11}X_1^2 + a_{22}X_2^2$  ;
- construire une liste d'expériences permettant de déterminer à moindre coût les coefficients  $a_0$ ,  $a_1$ ,  $a_2$ ,  $a_{12}$ ,  $a_{11}$  et  $a_{22}$  ;
- vérifier si ce modèle est suffisant (on considère ici que **oui**).

On peut avoir l'impression, avec cette étude séquentielle, que le nombre d'expériences à réaliser est aussi important qu'avec la méthode OVAT mais il n'en est rien, seuls l'analyse et la réflexion ont augmenté.

La figure suivante montre qu'il suffit de 9 expériences (choisies par la méthodologie des plans d'expériences) pour déterminer le modèle quadratique (à 2 paramètres). Cette construction d'expériences obéit à des critères précis qui ne relèvent en rien du hasard. C'est l'objet de cet ouvrage que d'apprendre à construire et surtout choisir ces plans.

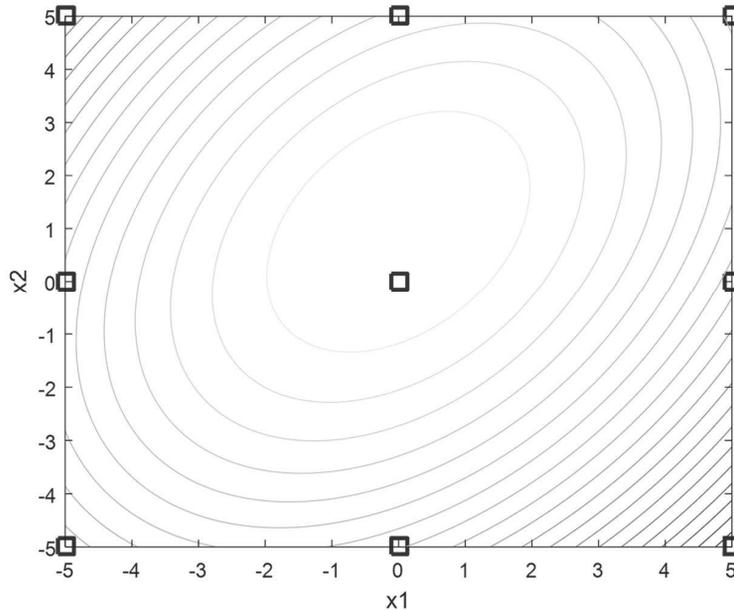


Figure 5 : Points expérimentaux construits dans un plan  $3^2$

Avec le résultat de ces expériences, une simple régression permettra de déterminer les coefficients du modèle quadratique, et de simples dérivées partielles donneront l'emplacement **précis (à l'erreur expérimentale près)** de l'optimum même si l'expérience correspondant au maximum ne fait pas partie des expériences réalisées.

#### B-4) Utilisation

Quel que soit le secteur industriel, cette méthodologie fait partie des outils de base de l'ingénieur « procédé », de l'ingénieur « qualité », de l'ingénieur « industrialisation » ou de l'ingénieur « développement ».

De nombreux managers sont aussi formés à cet outil (ou à cette palette d'outils) et bien évidemment, il est plus facile d'expliquer la mise en place et la démarche utilisées, si techniciens, agents de maîtrise, ouvriers sont aussi formés aux bases de cette méthodologie.

Pour éviter d'en rester au seul exemple de l'avion en papier, des études de cas seront proposées en fin d'ouvrage.

### C) Vocabulaire de base

#### C-1) La réponse Y

Le phénomène étudié noté **Y** est dénommé une réponse et correspond à la caractéristique qui sera mesurée à chaque expérience. En général, lors de la réalisation d'un plan d'expériences, on étudie simultanément plusieurs caractéristiques (ou réponses).

On peut, par exemple, chercher à obtenir un avion possédant un grand temps de vol et une grande distance de vol mais aussi étudier la dispersion du temps de vol d'un avion. Il faut bien évidemment dans ce cas que l'avion vole plusieurs fois.

L'étude de plusieurs réponses simultanées est conseillée même quand elle n'est pas nécessaire ou imposée par l'étude car elle ne conduit pas à faire plus d'expérience mais à augmenter le nombre de caractérisations d'une expérience.

Pour chaque réponse on déterminera un modèle :

$$Y = a_0 + \sum_1^i a_i X_i + \sum_1^i a_{ij} X_i X_j + \sum_1^i a_{ij} X_i X_j X_k \dots + \sum_1^i a_{ii} X_i^2 + \sum_1^i a_{iii} X_i^3$$

Mais cette seule écriture ne suffit pas à décrire toutes les réponses étudiées. Il est parfois nécessaire de transformer une réponse. Ces transformations (aussi utilisées en statistiques appliquées) répondent à deux nécessités :

- la dilatation ou la contraction de l'échelle de la valeur Y ;
- l'amélioration de l'indépendance de la moyenne des expériences vis-à-vis de la variance.

Cette deuxième raison est explicitée un peu plus loin.