

الجمهورية الجزائرية الديمقراطية الشعبية

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

Université Blida 1

Faculté Des Sciences

Département De maths



Mémoire Présenté Par

Bouchakour Moussa Zineddine et Dini Asma

En vue de l'obtention du Diplôme de Master en Modélisation stochastique

Et statistique

THEME

**EVALUATION DE LA QUALITE DES MODELES POUR
LES PLANS D'EXPERIENCES NUMERIQUES**

MEMBRES DU JURY :

Mme Massoudi nadia Université Blida1 Présidente.

Mme Raber dalila Université Blida 1 Examinatrice.

Mr Elmoossaoui hicham Université Blida1 promoteur.

2019/2020

Dédicaces

A Papa et maman, merci pour votre amour et votre soutien.

Merci d'avoir été à l'école tout au long de ce travail .vous n'avez aidé à aller jusqu'au bout, je vous aime tant

A mes chers frères MOHAMED ELHACEN et AHMED ELHOCINE.

A mes amis.

A tous ceux qui ont contribué à mon bonheur.

ASMA DINI

Dédicaces

A tous ceux qui m'ont enseigné et surtout à. Monsieur **ELMOSSAOUL**.

A mes parents et beaux parents en souvenir de l'intérêt qu'ils ont témoigné

A mes parents.

A mes frères.

A tous mes camarades.

Zineddine

Remerciement

Nous tenons avant toute chose, à remercier DIEU et à lui rendre grâce pour la force et le courage qu'il nous a donné.

Ainsi nous tenons également à exprimer nos vifs remerciements à notre encadreur monsieur Elmoassaoui Hicham.

Pour avoir d'abord proposé ce thème, pour le suivi continué tout long de la réalisation de ce mémoire et qui n'a pas cessé de nous donner ses conseils.

Nous voudrions aussi remercier nos famille, pour le soutien et l'aide inconditionnel qu'elle nous apporte chaque jour.

Nous adressons nos sincères remerciements et gratitude au membre de Jury pour nous avoir consacré son précieux temps.

Nous tenons à remercier vivement toutes les personnes qui nous ont aidés et réaliser ce mémoire, ainsi à tous ceux qui nous ont aidés de près ou de loin à accomplir ce travail.

Merci à tous

Table de matière

Résumé	6
Introduction	9
1. la méthode des plans d'expériences	11
1-1 Historique	11
1-2 A quoi sert un plan d'expérience	11
1-3 Etude d'un phénomène	12
1-4 Terminologie	17
1-5 Matrice d'expérience	20
1-6 Stratégie de la méthodologie de la recherche expérimentale	20
1-7 Les domaines d'application des plans d'expériences	20
1-8 Plans d'expériences orthogonaux	21
1-9 Propriété d'un plan orthogonal	24
1-10 Notion de modélisation mathématique	26
2. plans d'expériences numériques usuels	28
2-1 Plans standards	28
2-2 Plans «marginiaux»	34
2-3 Suite à faible discrédance	42
3. Les surfaces de réponse pour les d'expériences	49
3-1 Régressions linéaires	49

3-2	Modèles additifs	50
3-3	Mars et poly Mars	51
3-4	Réseaux de neurones	53
3-5	krigeage	54
3-6	Variogramme	55
3-7	Critères de qualité des surfaces de réponse	55
4.	Application pour une fonction déterministe à dimension 10	59
4-1	Les données	59
4- 2	Les résultats	60
	Conclusion	62
	Bibliographies	63

TABLE DES FIGURES

1.1 L'environnement du système.	12
1.2 Machine à laver.	14
1.3 Méthode traditionnelle d'expérimentation.	14
1.4 Une expérience à chaque nœud du maillage.	15
1.5 Domaine de variation du facteur.	18
1.6 Définition de l'espace expérimental.	18
1.7 Niveau des facteurs définissant des points expérimentaux dans l'espace expérimental centrée.	19
1.8 Représentation graphique des effets.	22
1.9 Représentations graphique des effets.	24
1.10 Tracé de l'interaction.	25
2.1 Un plan factoriel complet à 5 niveaux.	29
2.2 Un plan de Doehlert à 45 points dans le carré unité avec son simplexe initial: le triangle équilatéral en rouge.	31
2.3 Un échantillonnage par hypercube latin à 5 points en dimension 2.	35
2.4 Un échantillonnage par hypercube latin à 5 points en dimension 2.	36
2.5 Un hypercube latin à 20 points en dimension 2.	36

2.6 Un plan g�n�r� par un tableau orthogonal OA_1 (25,5,5,2) dont les points sont centr�s et projet�s sur le sous-espace (x_1, x_2) .	37
2.7 Un plan g�n�r� par un tableau orthogonal OA_1 (25,5,5,2) dont les points sont randomis�s et projet�s sur le sous-espace (x_1, x_2) .	39
2.8 Une distribution de 49 points issue d'un tableau orthogonal de force 2 lin�aire en dimension 3.	40
2.9 (De gauche � droite). Un hypercube latin, un hypercube latin de Tang et un tableau orthogonal de force 2 � 4 points en dimension 2.	41
2.10 Un hypercube latin g�n�r� par un tableau orthogonal $1OA$ (25, 5, 5,2) dont les points sont randomis�s et projet�s sur le sous-espace (x_1, x_2) .	42
2.11 (De la gauche vers la droite) Les 50, 250 et 500 premiers points d'une suite de Halton en bases 2 et 3.	45
2.12 Des mauvaises projections en 2D. A gauche : les 80 premiers points d'une suite de Halton en bases 11 et 13. A droite : les 100 premiers points d'une suite de Halton en bases 17 et 19.	45
2.13 (De la gauche vers la droite) Les 50, 250 et 500 premiers points d'une suite de Hammersley en base 2.	47
2.14 Les 50, 250 et 500 premiers points d'une suite de Sobol' en dimension 2.	48
4.1 Coupe de la fonction de Jin en fonction de x_1 et x_3 ($x_i = 3, i = 1,3$).	59
4.2 Crit�re de qualit� d'ajustement	60
4.3 Crit�re de validation RMSE	61

LISTE DES TABLEAUX

1.1 Matrice d'expérience.	21
1.2 Des résultats pour deux variables A et B.	22
1.3 Tableau des résultats.	24
2.1 Matrice de Doelhart à 13 expériences sur $[0, 1]^3$	32
4.1 Résultats des critères de qualité d'ajustement (R^2) et de validation (RMSE) pour les suites à faible discrédance	60

ملخص

على الرغم من التقدم في أدوات الكمبيوتر ، فإن الدراسة الرقمية لبعض الظواهر الفيزيائية المعقدة تظل مكلفة للغاية من حيث وقت الحساب. لذلك من الضروري الاستفادة من طرق التخطيط التجريبية. تقدم هذه الأطروحة وتحلل النهجين التجريبيين العدديين الأكثر استخدامًا: الطريقة الكلاسيكية لتصميم التجارب وطريقة kriging وطريقة MARS. وأخيرًا ، نقدم تطبيقًا لمقارنة الأساليب المختلفة.

الكلمات المفتاحية: تصميم التجارب ، سطح الاستجابة ، kriging ، المكعب اللاتيني الفائق ، MARS .

RESUME

Malgré les progrès des outils informatiques, l'étude numérique de certains phénomènes physiques complexes reste très coûteuse en temps de calcul. Il est donc nécessaire de faire appel à des méthodes de planification d'expériences. Ce mémoire présente et analyse les deux approches des expériences numériques les plus utilisées : la méthode classique des plans d'expériences et une méthode par krigeage, la méthode de MARS. Enfin, nous présentons une application pour comparer les différentes approches.

Mots clefs : plan d'expériences, surface de réponse, krigeage, hypercube latin, MARS.

ABSTRACT

Despite advances in computer tools, the digital study of certain complex physical phenomena remains very costly in terms of computation time. It is therefore necessary to make use of experimental planning methods. This thesis presents and analyzes the two most widely used numerical experimental approaches: the classical method of design of experiments and a kriging method and the MARS method. Finally, we present an application to compare the different approaches.

Key words : Experimental designs, response surface, kriging, latin hypercube, MARS.

INTRODUCTION

Les plans d'expériences permettent d'organiser au mieux les essais qui accompagnent une recherche scientifique ou des études industrielles [1]. Ils sont applicables à de nombreuses disciplines et à toutes les industries à partir du moment où l'on recherche le lien qui existe entre une grandeur d'intérêt, Y , et des variables, X_i . Il faut penser aux plans d'expériences si l'on s'intéresse à une fonction du type :

$$Y = f(X_i)$$

Avec les plans d'expériences on obtient le maximum de renseignements avec le minimum d'expériences. Pour cela, il faut suivre des règles mathématiques et adopter une démarche rigoureuse [2]. Il existe de nombreux plans d'expériences numériques et classiques adaptés à tous les cas rencontrés par un expérimentateur.

Dans ce mémoire nous allons présenter et commenter les deux approches des expériences numériques les plus couramment utilisées. La méthode classique des plans d'expériences et une approche plus actuelle par krigeage et méthode de PolyMars. Chacune est articulée autour des deux questions précédentes, à savoir le choix d'un modèle statistique et les plans d'expériences adaptés à ce modèle. L'objectif est de donner un aperçu de ce qui se fait dans le domaine.

La différence fondamentale entre les plans d'expériences numériques et les plans classiques développés en laboratoire (expériences physiques) réside dans le fait qu'en laboratoire, on suppose la présence d'erreurs aléatoires dues aux mesures, aux conditions expérimentales, etc. Ce qui n'est en principe pas le cas avec une expérience numérique (simulation déterministe). Nous n'envisageons pas ici le cas des simulateurs bruités, à savoir des simulations entachées de bruit de simulation (Monte-Carlo ou autre). La présence d'une racine aléatoire dans le simulateur peut aussi induire des différences au niveau des réponses. Nous n'envisageons pas ce cas dans ce travail.

Pour résumer, un plan d'expériences efficace doit être capable d'approcher une assez grande variété de modèles et de fournir une information sur tout le domaine expérimental avec un nombre de simulations le plus faible possible.

A noter que le plan doit convenir à l'analyse souhaitée et que celle-ci devra délivrer autant d'information que possible. Les deux ne peuvent pas être considérés indépendamment. Par exemple, l'utilisation des space filling designs (SFD) paraît être un avantage pour certains métamodèles tels que le krigeage afin d'éviter les problèmes numériques.

Ce méta modèle, nous ne pouvons pas l'utiliser directement sans vérifier qu'il représente correctement la réponse dans le domaine expérimental. Un premier diagnostic consiste à étudier les valeurs du modèle aux points du plan d'expériences : nous parlerons dans ce cas d'ajustement. Pour ce faire, il existe des critères permettant de tester la bonne adéquation entre les observations et les prévisions sur les données utilisées lors de la construction du modèle. C'est par exemple, le rôle du coefficient de détermination R^2 . Une fois le modèle accepté, il est nécessaire de vérifier la capacité du modèle à prédire en n'importe quel point du domaine expérimental la valeur de la réponse étudiée. Il s'agit ici de comparer les prévisions du modèle en des points différents de ceux du plan d'expériences. Pour vérifier que le modèle représente bien le phénomène étudié dans le domaine expérimental, le critère usuel est l'erreur quadratique moyenne ou MSE (Mean Squarred Error). Plus la valeur de la MSE est petite et plus le méta modèle est bon.

Le mémoire est composé de quatre chapitres organisés comme suit :

Le premier chapitre présente la méthodologie de la recherche expérimentale. Le deuxième chapitre présente un état d'art des plans d'expériences couramment utilisés en expérimentation numérique.

Le troisième chapitre est consacré à la présentation des différents métamodèles utilisés pour les plans d'expériences numériques. Ainsi que certains critères pour étudier la qualité de ces plans.

Le quatrième chapitre comporte une application sur une fonction déterministe à dimension 10.

Enfin, une conclusion termine ce travail.

CHAPITRE 1

1 LA METHODE DES PLANS D'EXPERIENCES

1.1 Historique

La méthode des plans d'expériences [1] est à la fois nouvelle et ancienne, les scientifiques n'ont abordé ce sujet que depuis peu d'années. Les premiers qui ont pensé à ce problème sont les agronomes et les statisticiens, les techniques et les notions qu'ils ont développées sont si générales qu'elles peuvent être utilisées dans beaucoup de domaines. Cette méthode est récente puisqu'on peut la faire démarrer avec les travaux de R.A. Fisher [3] (début de vingtième siècle), aussi bizarrée que cela paraisse elle ne porte pas encore de nom. Après 1945, les plans d'expériences suscitent de nombreuses publications et recherches dans le monde anglo-saxon. Des statisticiens comme *Yate*, *Youden*, *Cochran*, *Plakette et Burman*, et bien d'autres, enrichissent et divulguent la méthode de *Box et Hunter*, s'appuyant notamment sur les travaux de *Yate*, développent des méthodes particulières de constructions de plans fractionnaires à deux niveaux. À partir des années 50, en s'attaquant à l'amélioration de la qualité, le Japon imprime un nouveau souffle. *Taguchi* et *Matsuyama* [2] élaborent des tables permettant de construire des plans d'expériences orthogonaux adaptés à la majorité des problèmes industriels. Ces tables ont été publiées en 1959 et 1961 par la suite les méthodes initiées par Taguchi se sont répondues aux états Unis puis en Europe.

1-2- A quoi sert un plan d'expérience ?

Dans toutes les entreprises, et quel que soit le secteur d'activité, le technicien où l'ingénieur est amené à comprendre comment réagit un système en fonction des différents facteurs qui sont susceptibles de le modifier pour vérifier l'évolution du processus, il mesure une réponse et à partir de différents essais, il va tenter d'établir des relations de cause à effet entre la (ou les) réponse (réponses) et les différents facteurs.

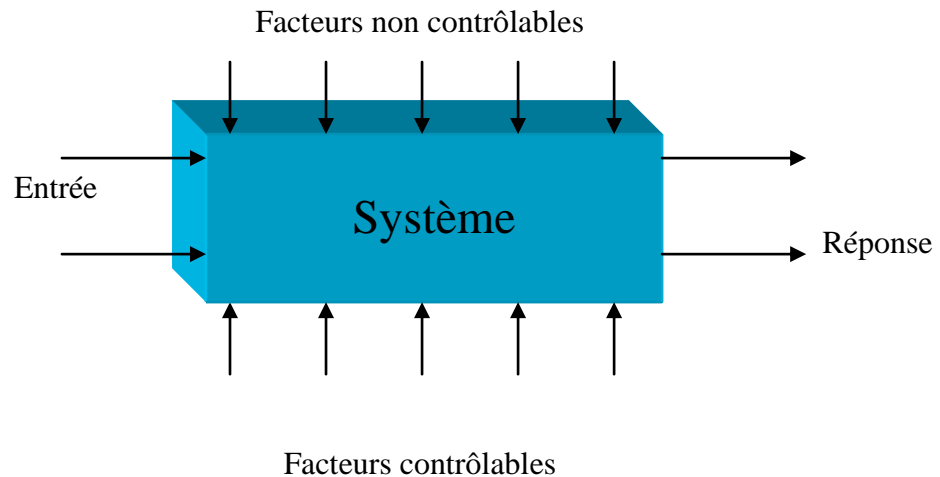


Figure1.1 : L'environnement du système.

Parmi les facteurs on distinguera :

- Les facteurs contrôlables qui dépendent directement du choix du technicien (choix d'un matériau, d'une pression,...).
- Les facteurs non contrôlables qui varient indépendamment des choix du technicien (conditions climatiques, environnement d'utilisation,...)
- Les facteurs d'entrées telles que la matière première dans un processus de fabrication.

Il n'est pas simple de réaliser de telles relations entre réponses et facteurs, surtout s'il existe des interactions entre les facteurs, les plans d'expériences permettent de réaliser ce type de relation, en minimisant le nombre d'expériences tout en maximisant la précision du résultat.

Le domaine de validité est très vaste. Dans le domaine de conception des produits :

- Test de différentes configurations de base.
- Choix de matériaux pour amélioration du fonctionnement.
- Détermination des tolérances de fonctionnement.

On l'applique aussi dans la conception et l'amélioration des processus de production.

- Diminution du pourcentage de défectueux dans une ligne de production.
- Choix des outillages.

1.3 Etude d'un phénomène

1.3.1 La démarche expérimentale

Essais la démarche expérimental reste la méthode royale d'acquisition des connaissances dans le domaine industriel de façon naturelle, l'ingénieur fonde sa démarche d'acquisition des connaissances sur des essais en avançant progressivement, puis il

réoriente ses expériences en fonction des résultats obtenus, il cherche en général à augmenter la productivité de ses essais pour cela, il doit :

- Faire un minimum d'essais,
- Obtenir un maximum de précision sur les résultats,
- Etre capable de mettre en évidence des phénomènes secondaires comme des interactions entre les facteurs.

Les difficultés de l'expérimentateur ne s'arrêtent pas à la planification des essais, une fois ceux-ci est réalisés il faut interpréter les résultats.

Une bonne démarche d'acquisition des connaissances devra fournir des résultats aisés à interpréter, ils devront être également faciles à présenter aussi bien à des spécialistes qu'à des non spécialistes.

Tous les industriels connaissent ces interrogations, aussi, il est judicieux de se poser quelques questions :

- Quelle stratégie d'essais adopter pour arriver rapidement aux résultats attendus ?
- Existe-t-il de meilleures stratégies que d'autres ?
- Quel est le nombre minimum d'essais qui doivent être réalisés pour aboutir aux résultats ?
- Peut-on éviter de réaliser des essais inutiles ?
- Comment améliorer la précision de mes résultats ?

La méthode des plans d'expériences apporte une méthodologie capable de répondre à ces questions. Cette méthodologie sans équivalent, permet d'atteindre une meilleure connaissance du système observé par un minimum d'essais et un maximum de précision [2].

1.3.2 Méthode traditionnelle

Traditionnellement, les essais sont effectués de manière séquentielle en faisant varier les variables l'une après l'autre sans planification préalable de l'ensemble des essais à réaliser. Par exemple, supposons que nous ayons comme objectif l'optimisation d'une machine à laver industrielle pour pièces mécaniques[2].

Les paniers des pièces à nettoyer passent dans deux bains, pour chaque bain, il est possible de faire varier le temps de trempage, de mettre des ultrasons, de faire tourner le panier, etc.....

Facteurs :

- ultrasons
- rotation

- agitation
- type de liquide
- période de renouvellement

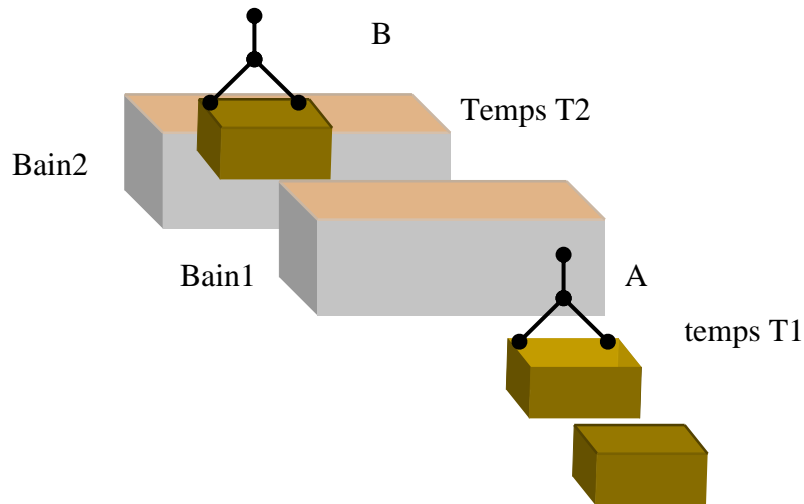


Figure 1.2 : Machine à laver

En simplifiant le problème à l'étude des deux facteurs Temps T_1 (facteur A) et Temps T_2 (facteur B), chacun de ces facteurs peut prendre plusieurs valeurs comprise entre le min (A_{min} , B_{min}) et le max (A_{max} , B_{max}). On souhaite étudier l'influence de chacune de ces grandeurs sur le système. La démarche traditionnelle employée de façon intuitive par les expérimentateurs consiste à étudier séparément les deux temps de trempage A et B . On fixe B à un niveau moyen et on étudie la « réponse » (propriété obtenue) du système lorsque A varie de A_{min} à A_{max} à l'aide de quatre expériences, (figure1. 3).

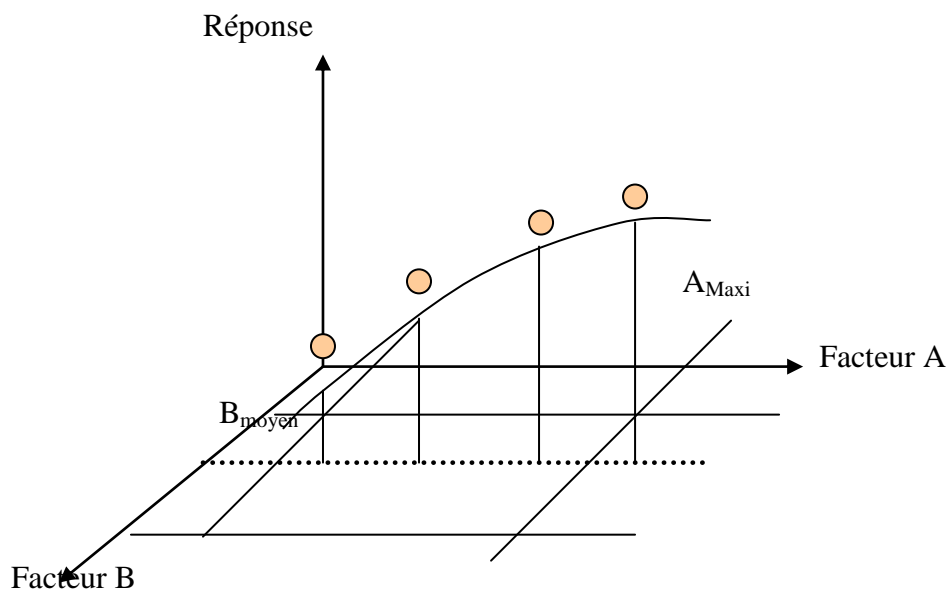


Figure 1.3 : Méthode traditionnelle d'expérimentation.

On réalise la même expérience pour étudier l'influence de B . Au total, nous avons réalisé 8 essais. Le problème est de savoir si nous avons une bonne connaissance du système grâce à ces 8 essais.

Nous savons comment agit B lorsque A est au niveau moyen, mais connaissons-nous l'action de B lorsque A est au niveau A_{max} ? Et au niveau A_{min} ? De même, nous savons comment agit A lorsque B est au niveau moyen, mais connaissons-nous l'action de A lorsque B est au niveau B_{max} ? Et au niveau B_{min} ? Pour être capable de répondre de façon exhaustive à ces questions, il faut faire un maillage du domaine de validité des deux variables A et B et exécuter un test à chacun des nœuds du maillage.

Dans ce cas, il faut réaliser $4 \times 4 = 4^2 = 16$ expériences pour avoir la connaissance du système.

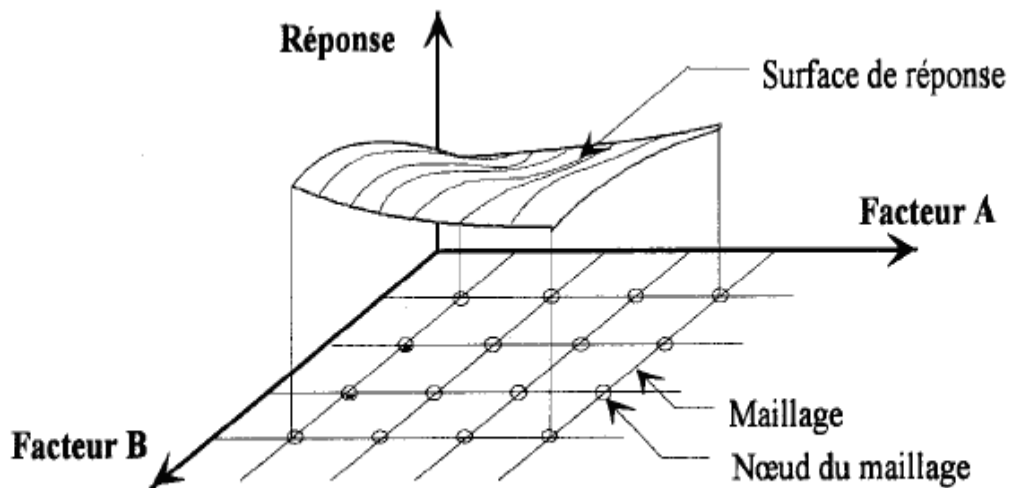


Figure1.4 : une expérience à chaque noeud du maillage.

Mais nous avons étudié uniquement les facteurs A et B , dans notre exemple, il y a encore de nombreux autres facteurs à étudier tels que les ultrasons, la rotation des pièces...etc. Si nous généralisons au cas d'un système répondant à sept facteurs influents et si nous souhaitons réaliser quatre expériences par facteur, nous devons alors réaliser $4^7=16384$ expériences, cela risque de prendre du temps et de coûter assez cher. Même si nous souhaitons réaliser deux points seulement par variable étudiées, il faudra quand même réaliser $2^7=128$ expériences, ce qui est déjà trop important, comme il est impossible de descendre au dessous de 2 points par variable, l'expérimentateur se résout alors souvent à diminuer le nombre de variables étudiées, ce qui nécessairement conduit à laisser un

doute sur les résultats obtenus. Finalement, la solution souvent retenue consiste à procéder par tâtonnements successifs en faisant varier un facteur à la fois, on modifie les ultrasons, puis la rotation, et ainsi de suite jusqu'à obtention d'une configuration satisfaisante, ce qui prend en général beaucoup de temps, d'énergie et donne rarement des résultats satisfaisants. Sans parler du grand nombre d'essais qu'il faut réaliser avec cette méthode, il est difficile d'obtenir l'optimum et impossible d'obtenir les interactions entre les facteurs. On mesure les inconvénients de cette approche quand sont en jeu des risques liés à la sécurité ou des montants financiers importants, c'est pourquoi nous allons examiner la méthode des plans d'expériences.

1.3.3 La méthode des plans d'expériences

Pour donner une définition claire d'un plan d'expériences, il faut auparavant définir les termes : plan et expérience.

- ✓ **Expérience** : c'est une intervention volontaire dans un système en fonction pour observer ou mesurer les effets de cette intervention.
- ✓ **Plans** : il est pris au sens de planification, plan de travail détaillé, organisation, marche à suivre.

La différence capitale avec la méthode traditionnelle tient au fait que l'on fait varier, à la fois, les niveaux de toutes les variables à chaque expérience, mais de manière programmée et raisonnée. Le fait de faire varier toutes les variables à la fois n'est pas un inconvénient mais au contraire offre de nombreux avantages parmi lesquels :

- Diminution du nombre des essais.
- Nombre de facteurs étudiés peuvent être très grand.
- Détection des interactions entre facteurs.
- Modélisation des résultats et l'obtention de l'optimum.

Les plans d'expériences permettent l'étude de nombreux facteurs tout en maintenant le nombre des essais à des valeurs raisonnables. La compréhension de la méthode des plans d'expériences s'appuie sur deux notions essentielles :

- Espace expérimental,
- Modélisation mathématique des grandeurs étudiées.

1.4 Terminologie

1.4.1 Réponses et facteurs

- **Réponses** : on qualifie de réponse la grandeur qui est observée pour chaque expérience réalisée. En d'autres termes c'est le résultat mesuré d'une étude sachant qu'à chaque point du domaine d'étude correspond une réponse. L'ensemble de ces dernières forme la surface de réponse. La valeur d'une réponse ne peut être modifiée que de manière indirecte en faisant varier les facteurs.
- **Facteurs** : un facteur peut être toute variable obligatoirement contrôlable, peuvent influencer sur la réponse observée. La différence fondamentale entre la notion de variable et celle de facteur tient donc dans le fait que tout facteur doit pouvoir être modifié sans difficulté. La valeur donnée à un facteur pour réaliser une expérience est appelée niveau.

Il existe deux types de facteurs, facteurs qualitatifs et quantitatifs :

- ❖ **Un facteur quantitatif** : se traduit par une grandeur numérique mesurable, ce sera une vitesse, une température, une pression, une intensité, une résistance....
- ❖ **Un facteur qualitatif** : ne peut pas donner lieu à quantification directe, on peut seulement identifier ses différents niveaux, ce sera une marque, un procédé, une méthode, un fournisseur, un niveau de qualité. Les facteurs qualitatifs (ou quantitatifs) peuvent être soit des facteurs contrôlés ou non contrôlés. Un facteur contrôlé est un facteur que l'on peut maîtriser, gérer ou modifier. On peut donc en fixant volontairement les valeurs.

Exemple des facteurs contrôlés :

- La température d'un four.
- Le type de fournisseur.

Exemple des facteurs non contrôlés :

- La température ambiante.
- Le taux d'hydrométrie
- L'hétérogénéité d'une éprouvette.

Les facteurs non contrôlés seront, soit des facteurs considérés comme non retenus pour l'étude car non influents et laissés à leurs niveaux habituels, soit des facteurs inconnus ou non discernables que l'on subit au cours de l'expérimentation.

1.4.2 Niveaux

C'est la valeur donnée à un facteur pour réaliser un essai lorsque l'on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes borne inférieure c'est le niveau bas, borne supérieure c'est le niveau haut.

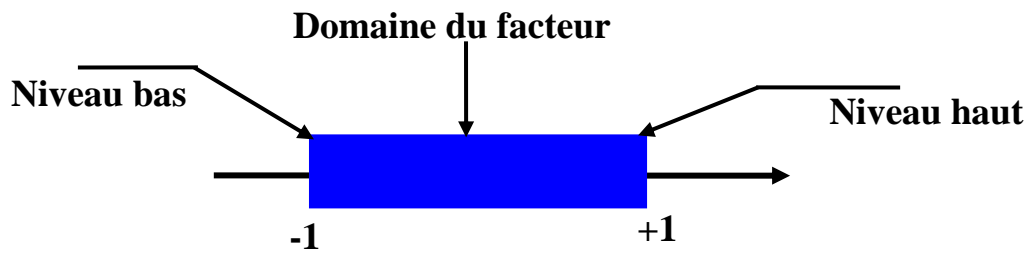


Figure1.5 : Domaine de variation du facteur.

1.4.3 Domaine de variation du facteur

Est l'ensemble de toutes les valeurs, qui peut prendre le facteur entre le niveau bas et le niveau haut.

1.4.4 Espace expérimental

Dans le cas où on a un second facteur, il est représenté lui aussi par un axe et on définit son niveau haut, son niveau bas et son domaine de variation, ce second axe est disposé orthogonalement au premier. On obtient ainsi un repère cartésien qui définit un espace euclidien à deux dimensions, c'est l'espace expérimental.

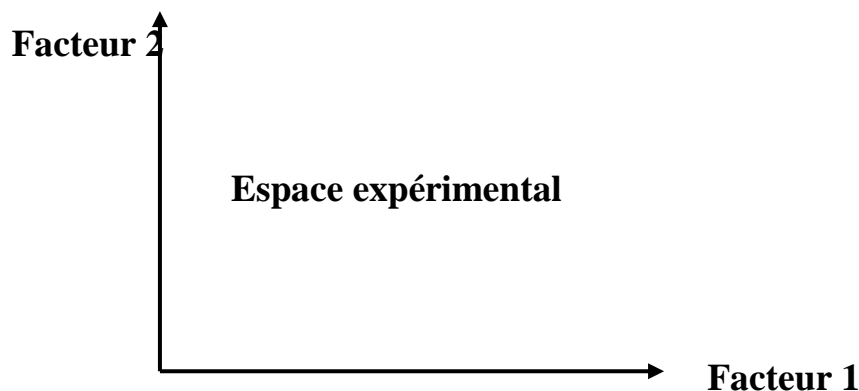


Figure1.6 : Définition de l'espace expérimental pour 2 facteurs.

Le niveau x_1 du facteur 1 et Le niveau x_2 du facteur 2 et peuvent être considérés comme les coordonnées d'un point de l'espace expérimental.

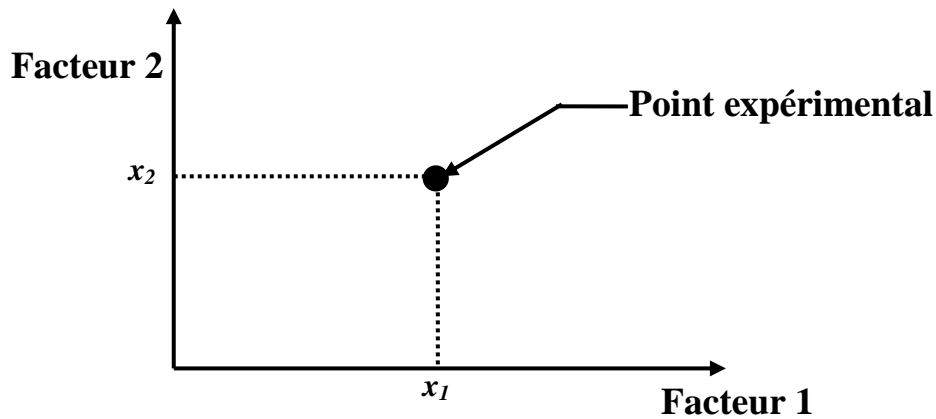


Figure 1.7 : Niveau des facteurs définissant des points expérimentaux dans l'espace expérimental.

1.4.5 Domaine d'étude

Le domaine d'étude est la réunion des domaines de chaque facteur, ce domaine est la partie de l'espace expérimental retenu par l'expérimentateur pour faire ses essais.

1.4.6 Variables centrées réduites

Lorsque l'on attribue la valeur -1 au niveau bas d'un facteur et la valeur +1 au niveau haut, on effectue deux modifications :

- **On change l'unité de mesure :** par exemple, si le niveau bas d'un facteur est 10°C et le niveau haut 30°C, il y a 20°C entre ces deux valeurs, soit 20 fois l'unité courante de température. Entre -1 et +1 il y a deux unités nouvelles, la nouvelle unité vaut donc 10°C. On le donne le nom de **pas**.
- **On déplace l'origine des mesures :** dans l'exemple choisi le milieu de l'intervalle [-1, +1] correspond à une température de 20°C. La nouvelle origine, notée zéro, diffère donc de l'origine exprimée en unités courantes.

Ces deux modifications entraînent l'introduction de nouvelles variables que l'on appelle variables centrées réduites, Centrées pour indiquer le changement d'origine et réduites pour signaler la nouvelle unité. Le passage des variables d'origine A aux variables centrées réduites X, et inversement est donné par la formule suivante :

$$X = \frac{A - A_0}{pas}$$

A_0 : est la valeur centrale en unité courante.

L'intérêt des variables centrées réduites est de pouvoir présenter les plans d'expériences de la même manière quels que soient les domaines d'étude retenus et quels que soient les facteurs.

1.5 Matrice d'expérience

Une matrice d'expériences est un objet mathématique qui représente l'ensemble des expériences à réaliser. Elle est toujours écrite sous forme codée. Elle est constituée de N lignes correspondants au nombre d'expériences et K colonnes correspondants au nombre de variables.

1.6 Stratégie de la méthodologie de la recherche expérimentale

Selon l'objectif tracé, le choix de la stratégie adaptée doit comporter quelques informations qui sont censés répondre à l'objectif, on peut distinguer quatre types :

1.6.1 Recherche exploratoire

Il s'agit de cas où nous ne connaissons presque rien sur le domaine. La méthode des plans d'expériences offre des techniques permettant de cadrer la démarche pour éviter de travailler de façon anarchique.

1.6.2 Criblages de facteurs

Les plans de criblages ont l'objectif de détecter les facteurs les plus influents sur une réponse donnée avec un minimum d'essais.

1.6.3 Etudes quantitatives des facteurs

Contrairement à la stratégie classique qui consiste à faire varier un facteur à la fois tout en gardant les autres constants et où le nombre d'essais est élevé, la méthodologie de la recherche expérimentale permet d'avoir les informations désirées avec un minimum d'essais. Elle fait aussi apparaître les interactions entre les facteurs qui sont complètement ignorées dans la stratégie classique.

1.6.4 Études quantitatives des réponses

Cette stratégie correspond à la modélisation du phénomène sous la forme d'une relation empirique qui permet de déterminer la valeur de la réponse en tout point de domaine expérimental sans qu'on soit obligé de faire l'expérience. Ainsi l'expérimentateur peut accéder à la zone la plus intéressante du domaine.

1.7 Les domaines d'applications des plans d'expériences

Les plans d'expériences sont habituellement appliqués dans la recherche et le développement de nouveaux produits ou procédés. Les secteurs d'activité concernés peuvent être aussi différents que l'agronomie, la chimie, l'électronique, la mécanique ou le marketing, ce qui est nouveau, c'est leur application à la qualité.

1.8 Plans d'expériences orthogonaux

Supposons que nous voulions mettre en évidence les effets de la vitesse d'un véhicule et de la pression. Un plan d'expériences est concrètement la liste des essais que l'on a choisis de gonflage, sur la durée de vie d'un pneumatique d'un type d'une marque donné. Prenons par exemple les deux niveaux des facteurs V et P suivants :

$$V_1 = 120 \text{ km/h} \quad P_1 = 1.8 \text{ bars}$$

$$V_2 = 140 \text{ km/h} \quad P_2 = 2 \text{ bars}$$

Le plan d'expériences que nous proposons est défini par les 4 essais suivants : V_1P_1 , V_2P_1 , V_1P_2 , V_2P_2 .

Les niveaux des facteurs sont représentés par des chiffres (1, 2). Pour notre exemple nous obtenons la matrice d'expériences suivante :

Facteur N° essais	V	P	Y
1	1	1	Y_1
2	2	1	Y_2
3	1	2	Y_3
4	2	2	Y_4

Tableau 1.1 matrice d'expérience

Ou les Y_i sont des durées de vie du pneumatique mesurées lors des 4 expériences. Ce plan à quatre expériences permet d'estimer les effets des deux facteurs V et P avec une bonne précision car les essais sont, dans certain sens, doublés. En effet, nous disposons de deux valeurs de la réponse Y pour le niveau V_1 ainsi que deux valeurs pour le niveau V_2 .

Représentons graphiquement l'effet de facteur V et celui de P sur la durée de vie d'un pneumatique :

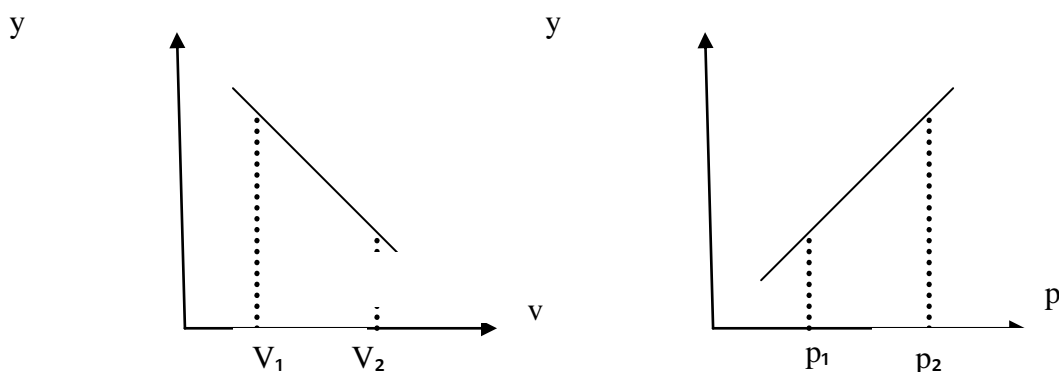


Figure 1.8 : Représentation graphique des effets.

Les effets de chaque facteur sont estimés par différence entre moyennes. Ici, l'effet du facteur vitesse V est estimé par la différence des moyennes des résultats pour $V=V_1$ et $V=V_2$

$$\text{On a effet total de } V = \frac{Y_2 + Y_4}{2} - \frac{Y_1 + Y_3}{2}$$

De même, l'effet de P est calculé de la façon suivante :

$$\text{effet de } P = \frac{Y_3 + Y_4}{2} - \frac{Y_1 + Y_2}{2}$$

Lorsque V est au niveau 2 (essais $n^\circ 2, 4$), P est une fois au niveau 1 ($n^\circ 2$) et une fois au niveau 2 ($n^\circ 4$). Lorsque V est au niveau 1 (essais $n^\circ 1, 3$), P est une fois au niveau 1 ($n^\circ 1$) et une fois au niveau 2 ($n^\circ 3$).

L'effet du facteur vitesse, tiré des moyennes, n'est donc pas biaisé par P puis que les mêmes niveaux de P sont présentés pour chaque niveau de V .

La propriété d'un plan qui permet que chaque effet soit estimé de sorte que les autres n'apportent aucun biais à son estimation, est l'orthogonalité. Prenons par exemple le plan complet suivant issu d'une table standard classique :

N°	A	B	C
1	1	1	$Y_1 = 100$
2	1	2	$Y_2 = 50$
3	2	1	$Y_3 = 180$
4	2	2	$Y_4 = 200$

Tableau 1.2 des résultats pour deux variables A et B

L'exemple est un cas simple qui permet de comprendre le principe des plans d'expériences. Il s'agit d'un plan composé de 4 lignes, soit 4 essais ou 4 expériences à réaliser. Les 4 combinaisons possibles sont : A_1B_1 , A_1B_2 , A_2B_1 , A_2B_2 , pour chaque combinaison, l'essai réalisé donne un résultat Y .

L'effet moyen est :

$$I = (Y_1 + Y_2 + Y_3 + Y_4)/4$$

Moyenne générale : $I = (100 + 50 + 180 + 200)/4 = 132.5$

L'effet du facteur A est obtenu par la formule :

$$\begin{aligned} \text{effet de } A &= \frac{Y_3 + Y_4}{2} - \frac{Y_1 + Y_2}{2} \\ &= \text{niveau plus haut} - \text{niveau plus bas} \end{aligned}$$

L'effet du facteur B est obtenu par la formule :

$$\begin{aligned} \text{effet de } B &= \frac{Y_2 + Y_4}{2} - \frac{Y_1 + Y_3}{2} \\ &= \text{niveau plus haut} - \text{niveau plus bas} \end{aligned}$$

$$\text{effet de } B = (50 + 200)/2 - (100 + 180)/2 = 125 - 140 = -15$$

$$\text{effet de } A = (180 + 200)/2 - (100 + 50)/2 = 190 - 75 = 115$$

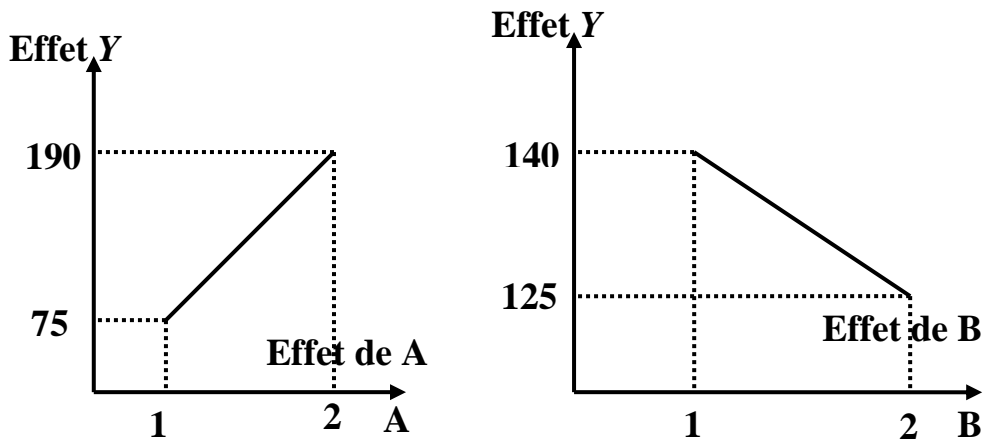


Figure 1.9 : Représentations graphique des effets.

1.9 Propriété d'un plan orthogonal

Lorsqu'un plan est orthogonal, l'effet de A n'est pas biaisé par B si, l'on reprend le plan à 4 essais et 2 facteurs A et B , la solution peut se résumer de la manière suivante :

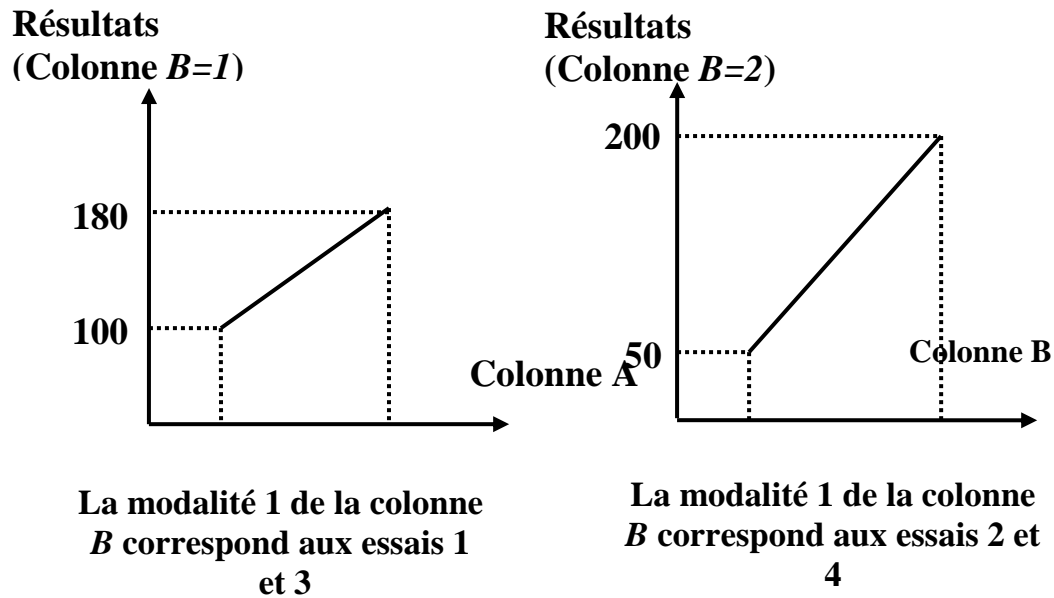
- Lorsque A est au niveau 1, B est au niveau 1,
- Lorsque A est au niveau 1, B est au niveau 2,
- Lorsque A est au niveau 2, B est au niveau 1,
- Lorsque A est au niveau 2, B est au niveau 2.

1.9.1 Calcul des interactions

L'interaction AB est étudiée, la modalité 1 de la colonne B correspond aux essais 1 et 3, la modalité 2 de la colonne B correspond aux essais 2 et 4.

N°	A	B	<i>Résultat</i>
1	1	1	100
2	1	2	50
3	2	1	180
4	2	2	200

Tableau 1.3 : Tableau des résultats.



1.9.2 Visualisation par superposition des deux graphiques

L'interaction est significative si les droites se coupent franchement, dans le cas présent, l'interaction semble peu significative. Les droites se coupent mais ne sont pas perpendiculaires.

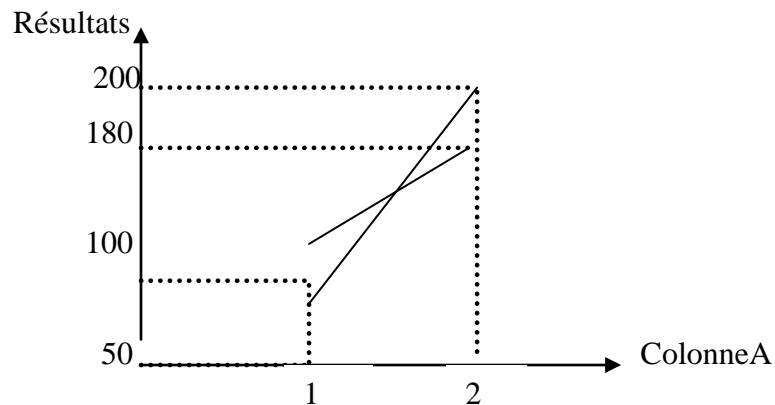


Figure 1.10 : Tracé de l'interaction.

1.9.3. Interprétation graphique des résultats

Si un facteur est significatif, il faut évidemment en tenir compte et choisir la modalité qui donne le meilleur résultat, dans l'autre cas si le facteur n'est pas significatif, cela signifie que sa valeur n'a pas beaucoup d'importance et qu'en conséquence, sa tolérance peut être large. Le raisonnement est le même pour les interactions. Cette propriété d'orthogonalité est indispensable pour obtenir une estimation de l'effet propre à chaque facteur, l'astuce consiste à tester l'effet d'un facteur gratuitement pendant les répétitions nécessaires à l'estimation de l'effet d'un autre facteur, ce qui évite de doubler chaque

expérience. Alors l'orthogonalité est la propriété qui fait que pour un niveau fixé de tout facteur du plan, chaque autre facteur apparaît un nombre égale de fois à chacun de ses niveaux possibles.

1.10 Notion de modélisation mathématique

Le modèle est une relation entre les facteurs x_1, x_2, \dots, x_k et la réponse que l'on désire étudier $Y = f(x_1, x_2, \dots, x_k)$ correspond à une surface de réponse. En utilisant certaines classes de fonction usuelle, cette fonction prendra la forme d'un polynôme de degré plus ou moins élevé :

$$Y = a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k$$

Y : est la grandeur à laquelle s'intéresse l'expérimentateur.

x_i : le niveau attribué au facteur i

a_0, a_i, \dots Sont les coefficients du modèle, ils ne sont pas connus et doivent être calculés à partir des résultats des expériences.

1.11 Système d'équation :

Chaque point expérimental permet d'obtenir une valeur de la réponse. Cette réponse est modélisée par un polynôme dont les coefficients sont les inconnues qu'il faut déterminer.

A la fin du plan d'expériences. On a un système de n équations (s'il y a n essais) à p inconnues (s'il y a p coefficients dans le modèle choisi a priori). Ce système s'écrit d'une manière simple en notation matricielle :

$$Y = Xa + e \quad \dots \dots \dots (1.1)$$

Y est le vecteur des réponses

X la matrice de modèle, qui dépend des points expérimentaux choisis pour exécuter le plan et le modèle postulé.

a : est le vecteur de coefficients.

e : est le vecteur des écarts types.

Ce système possède un nombre d'équations inférieur au nombre d'inconnues.

Il y a n équations et $p+n$ inconnues. Pour le résoudre, on utilise une méthode de régression basée sur le critère des moindres carrée. On obtient ainsi les estimations des coefficients que l'on note

$$\hat{a} = (X^t X)^{-1} X^t y \quad \dots \dots \dots (1.2)$$

Avec :

X^t : est la matrice transposée de x .

$(X^t X)$: la matrice d'information.

$(X^t X)^{-1}$: la matrice de dispersion.

CHAPITRE 2

PLANS D'EXPERIENCES NUMERIQUES USUELS

Comme nous avons pu l'indiquer en introduction, les objectifs attendus des plans d'expériences pour l'exploitation des codes de calcul en phase initiale sont principalement le remplissage de l'espace et le non redondance des points projetés sur les sous-espaces. Nous garderons à l'esprit également que les méthodes de génération des plans doivent être peu coûteuses en temps d'exécution et doivent permettre d'obtenir des plans efficaces avec un faible nombre d'expériences.

Ce chapitre a pour but de fournir un état de l'art critique et le plus concis qui soit des plans d'expériences paraissant les plus adaptés en phase d'exploration vis-à-vis essentiellement des propriétés énoncées ci-dessus.

2.1. Plans standard

Nous avons choisi de discuter des plans standards dans ce mémoire car ils ont été développés à l'origine pour des applications de surface de réponse. Parmi ces plans, les plus courants sont les plans factoriels, les plans de Box-Behnken et composites qui sont faciles à générer.

2.1.1. Plans factoriels complets

La méthode la plus simple pour répondre à la problématique du remplissage de l'espace consiste à choisir les points sur une grille régulière dans le domaine expérimental. Pour obtenir une grille régulière à k niveaux, il suffit de choisir k valeurs réparties à différents intervalles réguliers pour chaque facteur. Par exemple, dans le carré unité $[0,1] \times [0,1]$ et en choisissant 5 niveaux, on peut considérer la grille de points (voir figure (2-1)) suivante $\{0, 0.25, 0.5, 0.75, 1\} \times \{0, 0.25, 0.5, 0.75, 1\}$.

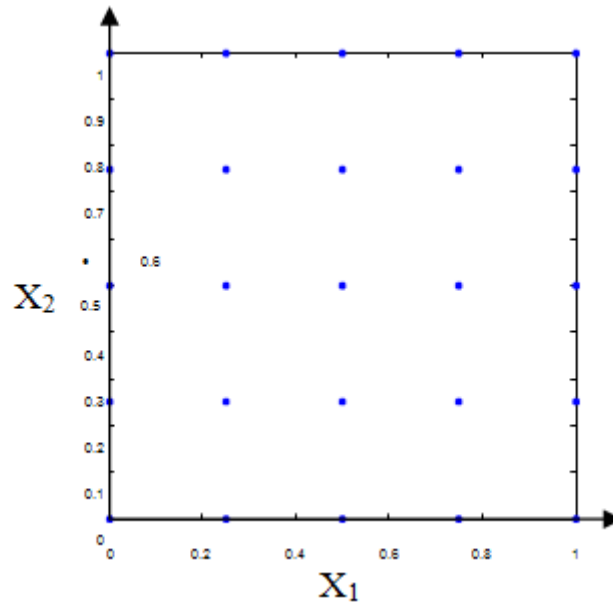


Figure (2.1). Un plan factoriel complet à 5 niveaux.

Il est évident que plus le nombre de niveaux ne sera important et plus le remplissage de l'espace ne sera satisfaisant mais cela entraîne une « explosion » du nombre de simulations. Il est donc essentiel de trouver un bon compromis et de choisir les niveaux paraissant les plus adaptés au problème [5].

L'utilisation de ce type reste intéressante si la dimension du problème est petite, dès que la dimension est plus élevée, le nombre de simulations $d = nk$ augmente considérablement, et l'idée d'utiliser une grille devient alors inenvisageable. De plus, seulement k valeurs différentes apparaissent pour chaque dimension. Dans le cas où la réponse dépend très fortement de seulement quelques variables (par exemple, une ou deux en dimension 5), le plan factoriel conduit à beaucoup de réplifications inutiles. Ce type de plan devient donc inutilisable lorsque la dimension augmente.

2.1.2. Plans factoriels fractionnaires

Le contexte dans lequel nous sommes ne nous autorisant pas à effectuer un nombre importants de simulations, les plans factoriels complets sont inadapés. Pourtant le principe sur lequel ils reposent reste intéressant. C'est pourquoi, les plans factoriels fractionnaires peuvent constituer une bonne alternative. Le principe consiste à prendre des sous-ensembles de plans factoriels complets permet de réduire considérablement le coût en nombre de simulations. Pour plus de détails [5].

2.1.3. Plans composites

Une matrice d'expériences composite est une combinaison :

- D'une matrice d'expériences factorielle à 2 niveaux, complète (2^p) ou fractionnaire (2^{p-f}), dont les points sont pris parmi les sommets de l'hyper cube ($[-1,1]$ par exemple).
- D'une matrice d'expériences axiale composée de points disposés symétriquement sur chaque axe, à une distance α du centre du domaine d'étude.
- D'un point central qui permet d'obtenir une information sur la variabilité du phénomène ou tester la validité du modèle.

Ces plans sont souvent utilisés en expérimentation classique pour approcher des surfaces de réponse de degré 2 [6]. Il est possible de générer différents plans de ce type en faisant varier la distance entre le point central et les points aux bords du domaine. On peut citer par exemple, les plans composites centrés ordinaires, les plans composites à faces centrées, les plans composites centrés inscrits... etc.

Le nombre d'expériences des plans composites augmente rapidement avec le nombre de variables étudiées et cette augmentation de points est surtout due à la matrice factorielle. Ces plans ne répondent pas convenablement aux objectifs de remplissage de l'espace. En effet, ces plans ne testent que 3 ou 5 niveaux (selon la valeur de α) par paramètres et cela avec n'importe quelle taille du plan.

2.1.4. Plans de Box-Behnken

Un plan de Box-Behnken est un plan dont les variables ne prennent que 3

niveaux (+1, -1, 0), si l'on considère pour le domaine expérimental l'hyper cube $[-1,1]^p$, ces Plans d'expériences sont composés d'une matrice factorielle à 2 niveaux et de blocs incomplets équilibrés en suivant un arrangement particulier. Ses matrices sont ensuite complétées d'un point au centre du domaine. Ces plans sont une alternative aux plans composites parce qu'ils requièrent seulement 3 niveaux pour chaque facteur [6] et permettent aussi de modéliser une surface de réponse de degré 2.

En termes de nombre de points, un plan de Box-Behnken est comparable à un Plan composite en dimension 3 et 4. A noter qu'il n'existe pas de Box-Behnken en dimension 2. Ces plans

ayant leurs points sur les axes factoriels du domaine, ils ne remplissent pas correctement l'espace.

2.1.5. Plans de Doehlert

Les plans de Doehlert [7] sont une partie intégrante de la famille des réseaux. La méthode de génération est itérative et consiste à considérer un simplexe initial dans le domaine d'exploration et à lui appliquer des isométries (translations et rotations) à partir d'un de ses sommets. En itérant ce procédé, on obtient alors une distribution de points particulière (Figure (2-2)). Pratiquement, cela revient, Pour chaque variable sur $[-1,1]$, à soustraire successivement les coordonnées de chaque point du simplexe initial aux autres.

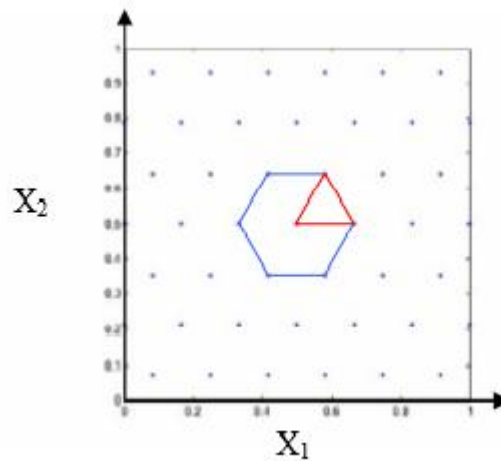


Figure (2.2). Un plan de Doehlert à 45 points dans le carré unité avec son simplexe initial: le triangle équilatéral en rouge

Par exemple pour $p = 3$, les coordonnées du simplexe initial sont, sur $[0,1]$, données par les 4 premières expériences du tableau ci-dessous la matrice à 13 expériences de doehlert est présentée dans le Tableau 2.1

N°	X ₁	x ₂	x ₃
1	0.5	0.5	0.5
2	1	0.5	0.5
3	0.75	0.933	0.5
4	0.75	0.6445	0.9085
5	0	0.5	0.5
6	0.25	0.067	0.5
7	0.75	0.067	0.5
8	0.25	0.9330	0.5
9	0.25	0.3555	0.0915
10	0.75	0.3555	0.0915
11	0.5	0.7885	0.0915
12	0.25	0.6445	0.9085
13	0.5	0.2115	0.9085

Tableau (2. 1). Matrice de Doelhert à 13 expériences sur $[0, 1]^3$

Ces plans permettent avant tout de générer des points uniformément dans l'espace ; ce qui est particulièrement utile lorsque l'on veut explorer la totalité du domaine

(Bornes et intérieur) sans proposer de modèle a priori représentant la réponse. La densité de points sera d'ailleurs déterminée par la taille que l'on accorde au simplexe initial. Ces plans ont aussi de nombreux autres avantages.[5]

Une de leurs caractéristiques importantes est de permettre une démarche séquentielle dans l'étude d'une surface de réponse. Par exemple, si la zone expérimentale la plus intéressante ne se trouve pas dans le domaine préalablement choisi mais plutôt située au voisinage de celui-ci ; il est possible de construire autour de l'un des points du plan initial, un nouvel ensemble d'expériences qui comprend une partie des simulations déjà effectuées sans pour autant détruire l'uniformité préexistante. A noter que cette propriété est par exemple non valable concernant les plans factoriels.

Ces plans permettent aussi de simplifier l'étude à une zone restreinte en changeant les dimensions du simplexe initial.

Le nombre minimal de points distincts, n , d'un plan de Doehlert pour un nombre de facteurs p donné est $n = d^2 + d + 1$, ce qui en grande dimension paraît être un avantage. Par exemple si $d = 20$, n sera égal à 421.[4]

2.1.6 Pourquoi utilise-t-on encore des plans d'expériences classiques en simulation?

Les plans d'expériences classiques ne sont pas remis en cause. Si on les utilise dans le domaine où ils sont applicables, ils sont de très bonne qualité. Mais si les hypothèses initiales ne sont pas vérifiées, comme cela semble être souvent le cas en simulation, leur utilisation n'est pas appropriée. Cependant, malgré la complexité des modèles numériques, un simple modèle polynomial peut être suffisant pour approcher la réponse du simulateur et ce pour diverses raisons.

- La réponse du simulateur peut-être lisse même si le modèle mathématique décrivant le phénomène simulé est très complexe. Par exemple l'industrie pétrolière utilise des simulateurs d'écoulement pour prédire des productions cumulées d'huile. On constate que ce type de réponse est suffisamment régulier pour être résumée par une surface de réponse polynomiale on trouve dans Dejean et Blanc [8] .

- Dans de nombreuses applications, le domaine d'étude peut être plus petit que le domaine de variation des variables d'entrée. Sur un domaine restreint il est alors possible de représenter la réponse du simulateur par un simple modèle polynomial même si globalement la réponse a un comportement plus complexe.

Enfin une dernière raison pour laquelle la méthode des plans d'expériences classique est souvent utilisée vient du fait qu'elle est bien connue des ingénieurs et l'interprétation physique du modèle polynomial est simple, de plus elle est disponible en logiciel.

Remarque

Les plans étudiés auparavant sont généralement de bonne qualité concernant le remplissage de l'espace. Cependant, la plupart de ces plans sont soit coûteuses en nombre d'expériences ou et soit une cause de perte d'information dans le contexte des plans

expériences numériques. C'est essentiellement pour ces raisons que nous avons choisi d'étudier qu'une infime partie de tous les plans d'expériences classiques connus. Cela ne

signifie pas qu'il faut les écarter pour autant des plans d'expériences numériques, le lecteur souhaitant en savoir davantage sur ces plans peut se référer aux références [5]

2.2. Plans « marginaux »

Dans cette partie la notion de marges se réfère à des sous espace factoriels ,par exemple ,les marges de dimension 1 correspondent aux axes factoriels .parmi ces plans on trouve :

2.2.1. Hyper cubes latins

La méthode d'échantillonnage par hyper cube latin a été introduite pour évaluer numériquement les intégrales multiples. Elle permet d'assurer le non redondance de l'information au travers d'une bonne répartition des projections sur les axes factoriels. En pratique, les hyper cubes latins sont très utilisés en planification d'expériences numériques notamment pour leur simplicité d'usage et de construction.[12]

Chaque axe $[0,1]$ du cube unité est découpé en q segments de même longueur

de la façon suivante $\left\{ \left[0, \frac{1}{n}\right], \left[\frac{1}{n}, \frac{2}{n}\right], \dots \dots \dots, \left[\frac{n-1}{n}, 1\right] \right\}$. En faisant le produit cartésien de ces intervalles, on obtient un maillage de q^p cellules de même taille. Ensuite, on choisit q cellules parmi les q^d possibles de telle manière que les cellules des marges de dimension 1 soient représentées une fois chacune. Enfin, on tire un point au hasard dans chaque cellule présélectionnée.

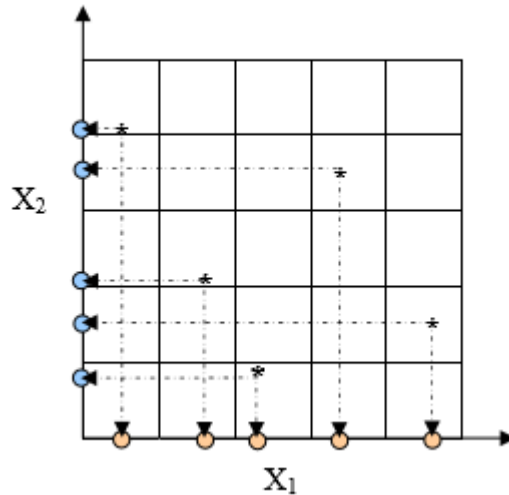


Figure (2.3).Un échantillonnage par hypercube latin à 5 points en dimension 2.

Définition:

Un hypercube latin à n points sur $[0,1]^d$. est défini par l'ensemble de points X^i tel que

$$X_j^i = \frac{\pi_j(i) + U_j^{(i)}}{n} ; \quad 1 \leq i \leq n ; \quad 1 \leq j \leq d$$

Où π_j est une permutation de $1, \dots, n$ et $U_j^{(i)} \sim U[0,1]$ est une valeur aléatoire de distribution uniforme sur $[0,1]$. Ainsi, $(\pi_1(i), \dots, \pi_d(i))$ représente la cellule dans laquelle se trouve le point X^i , et $(U_1^{(i)}, \dots, U_d^{(i)})$ désigne à quel endroit le point se trouve dans la cellule. L'hypercube latin obtenu correspond donc à la matrice de n lignes et d colonnes à coefficients X_j^i .

Remarques :

- Nous pouvons choisir de placer le point au centre des cellules pour ainsi supprimer l'aléa de la planification.
- Un hyper cube latin, défini par la matrice π , est un plan très facile à construire puisque chaque colonne est une permutation de 1 à n .

Les points d'un hyper cube latin ont la propriété intéressante d'être uniformément distribués sur les axes factoriels (voir Figure (2-3)). Cependant, cette propriété n'assure pas l'uniformité des points sur tout le domaine expérimental. Pour q fixé, il existe $q!$ permutations possibles pour chacune des p colonnes, et donc $(q)^{p-1}$ hypercubes latins éventuels qui n'assurent pas tous l'uniformité des points dans l'espace. C'est le cas par exemple de l'hypercube latin de

la Figure (2-4) où les points sont placés sur l'une des diagonales du domaine expérimental. Si le processus réel est une fonction de $(x_2 - x_1)$, alors l'information donnée par ce plan d'expériences se réduit à 1 point au lieu de 5.

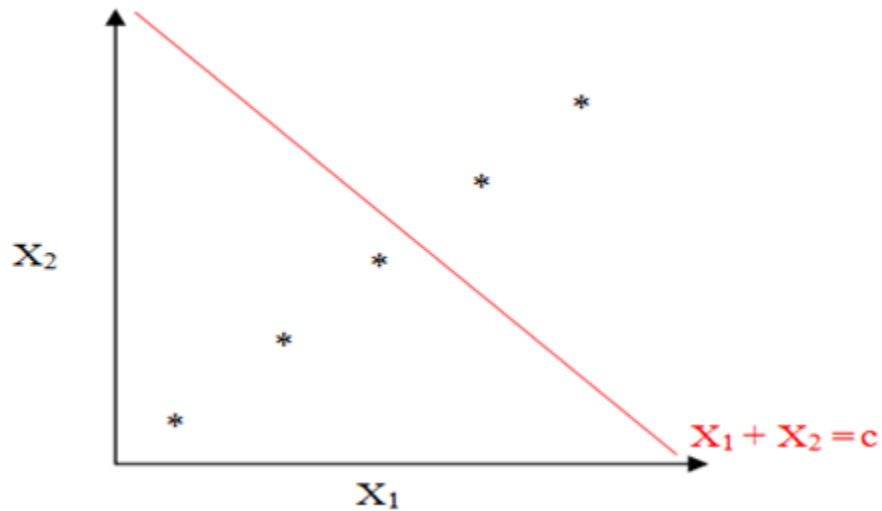


Figure (2.4) un échantillonnage par hypercube latin à 5 points en dimension 2

De façon générale, le contexte dans lequel nous sommes ne nous autorisant pas à prendre un grand nombre de points, nous aurons un nombre de mailles réduit. Il sera donc fréquent d'obtenir des plans où l'on peut observer des zones lacunaires et des agglomérats de points comme on peut l'observer sur la Figure (2-5).

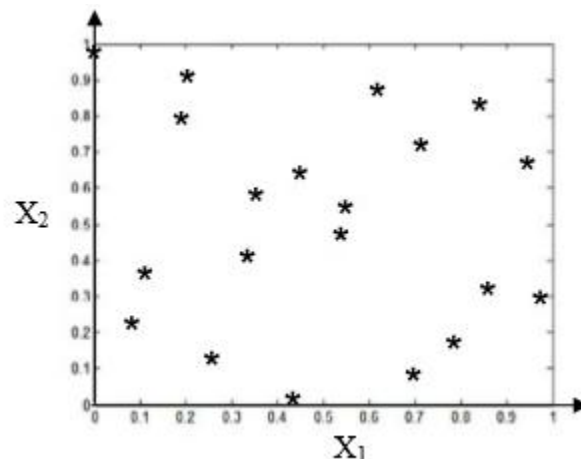


Figure (2.5). Un hypercube latin à 20 points en dimension 2

2.2.2. Tableaux orthogonaux

Définition:

Un tableau orthogonal de force t à q symboles est une matrice de n lignes et $P(> t)$ colonnes avec ses éléments prenant q valeurs possibles, telle que toutes les sous matrices de taille $n \times t$ aient q^t lignes apparaissant le même nombre λ fois. Il est évident que

$$n = \lambda q^t. \text{ On note un tel tableau orthogonal } OA_\lambda(n, p, q, t).$$

Géométriquement, cela revient à découper les axes du cube unité en q segments égaux afin d'obtenir q^p cellules de même taille. On choisit n cellules de telle façon qu'elles forment un tableau orthogonal défini comme ci-dessus. Cette définition assure que chaque t colonnes de la matrice du plan, soit chaque t -uplet, apparaît exactement λ -fois.

Remarque. Un tableau orthogonal de force 1 est un hyper cube latin.

Comme pour les hyper cubes latins, on peut choisir le point aléatoirement dans la cellule mais aussi au centre de la cellule. Dans ce dernier cas, toutes les projections sur les sous-espaces de dimensions t conduisent alors à une grille régulière comme on peut le constater sur la Figure (2-6).

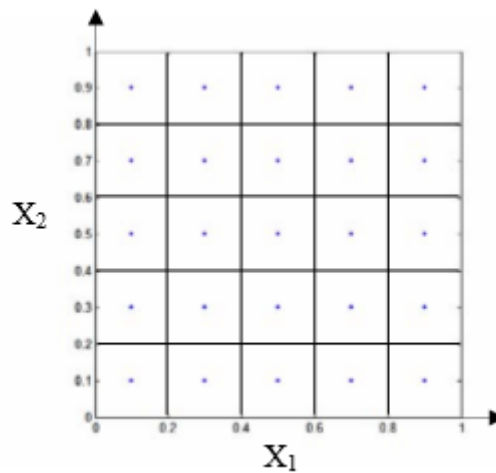


Figure (2-6). Un plan généré par un tableau orthogonal $OA_1(25,5,5,2)$ dont les points sont centrés et projetés sur le sous-espace (x_1, x_2) .

Définition

Un échantillonnage par tableau orthogonal à n points sur $[0,1]^p$ est un ensemble de points X^i tel que

$$X_j^i = \frac{\pi_j(A_j^{(i)}) + U_j^{(i)}}{q}, 1 \leq i \leq n, 1 \leq j \leq d.$$

où π_j est une permutation de $0, \dots, q-1$, $(A_j^{(i)})$ sont les éléments du tableau orthogonal et $U_j^{(i)} \sim U[0,1]$ est une valeur aléatoire de distribution uniforme sur $[0,1]$.

Ainsi, $(\pi_1(A_1^{(i)}), \dots, \pi_d(A_d^{(i)}))$ représente la cellule dans laquelle se trouve le point X^i , et $(U_1^{(i)}, \dots, U_d^{(i)})$ désigne à quel endroit le point se trouve dans la cellule. Le tableau orthogonal correspond à la matrice de n lignes et p colonnes à coefficients X_j^i .

Concernant la génération des plans, on a une propriété analogue à celle des hyper cubes Latins : [14]

Propriété

Si on effectue une permutation des symboles sur chaque colonne d'un tableau orthogonal de force t , alors on obtient toujours un tableau orthogonal de force t .

Discussion. La Figure (2-7) montre les points générés par un tableau orthogonal de type $OA_1(25, 5, 5, 2)$ projetés sur le sous-espace (x_1, x_2) . Nous pouvons observer qu'il y a bien un unique point par cellule et donc que les points projetés sur le sous-espace (x_1, x_2) sont répartis uniformément.

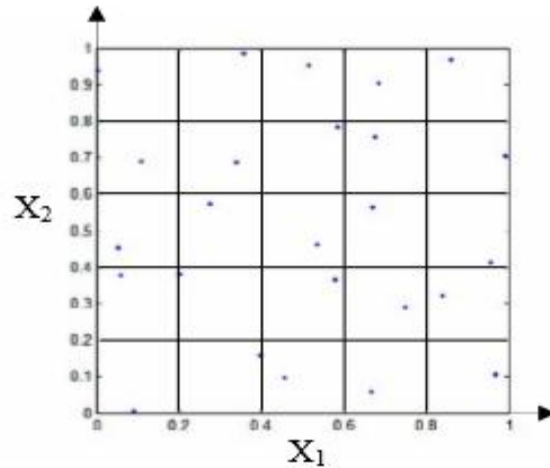


Figure (2.7) Un plan généré par un tableau orthogonal $OA_1(25,5,5,2)$ dont les points sont randomisés et projetés sur le sous-espace (x_1, x_2)

2.2.2.1. Cas particulier des tableaux orthogonaux linéaires

Les tableaux orthogonaux linéaires ont été considérés car ils sont plus faciles à mettre en œuvre que dans le cas général.

Définition:

Un tableau orthogonal linéaire est un tableau orthogonal tel que :

- Le nombre de symboles q est un nombre premier
- Les lignes sont toutes distinctes et constituent un sous-espace vectoriel de $(Z/qZ)^d$ dans ce cas, on le note de $OA((Z/qZ)^d, t)$.

$Z/qZ = \{0, \dots, q-1\}$ (Ou bien encore noté Z^q) est un corps sur Z car on suppose q

Premier.

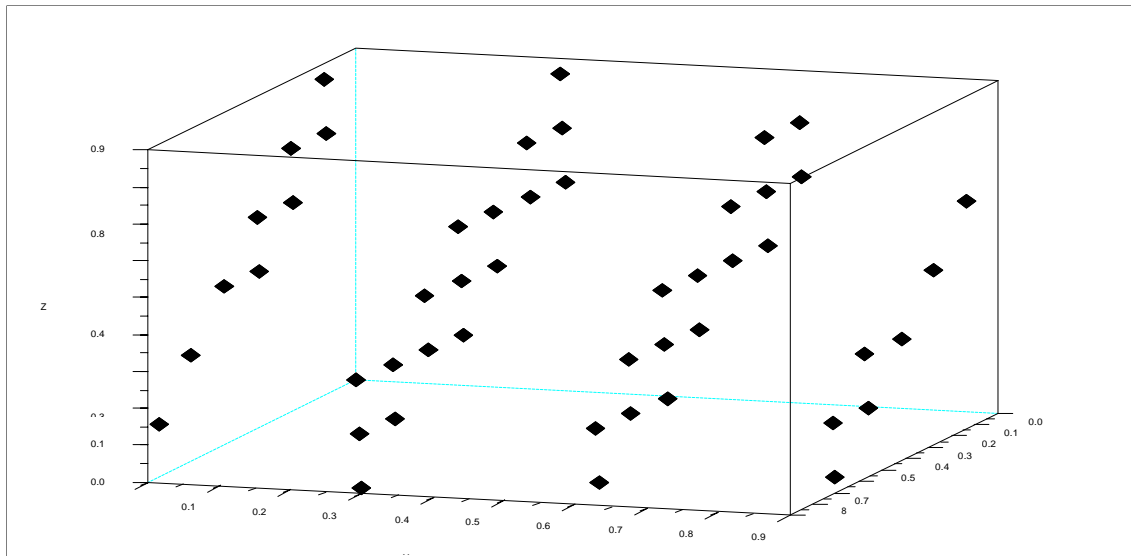


Figure (2.8). Une distribution de 49 points issue d'un tableau orthogonal de force 2 linéaire en dimension 3

Remarque. Un tableau orthogonal linéaire de force t , n'est autre qu'un tableau orthogonal de force t avec une structure de sous-espace vectoriel. Notamment, un tableau orthogonal linéaire de force 1 est toujours un hyper cube latin.[5]

2.2.3. Hyper cubes latins basés sur les tableaux orthogonaux de force 2

Les tableaux orthogonaux sont couramment employés pour la planification d'expériences et leur succès est dû à leurs bonnes propriétés d'uniformité. Concernant les tableaux orthogonaux de force $t > 1$, cette uniformité n'est garantie que sur les sous-espaces de dimension t . Ainsi, les tableaux orthogonaux présentent des répétitions sur les axes factoriels.

Il semble alors préférable d'utiliser des hyper cubes latins afin de garantir une bonne représentation des axes factoriels. Cependant, ces plans ne garantissent pas, en général, une distribution uniforme sur les sous-espaces de dimension $t > 1$. Pour résumer, aucune des 2 méthodes n'est complètement satisfaisante.[5]

L'idée proposée par Tang [13] pour remédier à ces problèmes est de combiner les propriétés d'orthogonalité d'un tableau orthogonal avec les bonnes propriétés de projection des hyper cubes latins. Tang fournit un algorithme pour générer ces plans à partir de tableaux orthogonaux de force 2, ce qui permet de les faire hériter des bonnes propriétés d'uniformité

Sur les marges de dimension 1.

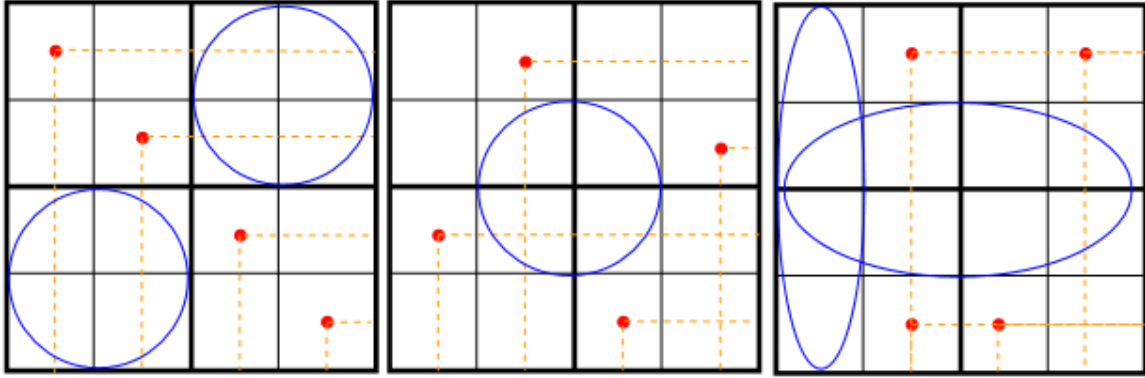


Figure (2.9). (De gauche à droite). Un hypercube latin, un hypercube latin de Tang et d'un tableau orthogonal de force 2 à 4 points en dimension 2.

2.2.3. 1 Proposition

Soit A un tableau orthogonal de type $OA(n, d, q, 2)$ Pour chaque colonne de A , on remplace chaque élément par une permutation de l'ensemble à q éléments suivant :

$[kq + 1, kq + 2, \dots, (k + 1)q, \forall k = 0, \dots, q - 1]$ Pour obtenir un hyper cube latin.

Par exemple, si $A = {}^t \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$ et que l'on considère les

permutations suivantes

$0 \rightarrow 0$ puis $0 \rightarrow 1$ et $1 \rightarrow 3$ puis $1 \rightarrow 2$ pour chaque dimension (colonne de la matrice A),

Alors on obtient l'hyper cube latin ${}^t \begin{pmatrix} 0 & 1 & 3 & 2 \\ 0 & 3 & 1 & 2 \end{pmatrix}$

réciroquement, à partir de l'hyper cube latin, il est possible de générer le tableau Orthogonal de type $OA1(n, d, q, 2)$ où les coefficients sont définis par $[q_{ij}/q], i = 1, \dots, q^2$. Avec [...] qui désigne la partie entière inférieure.

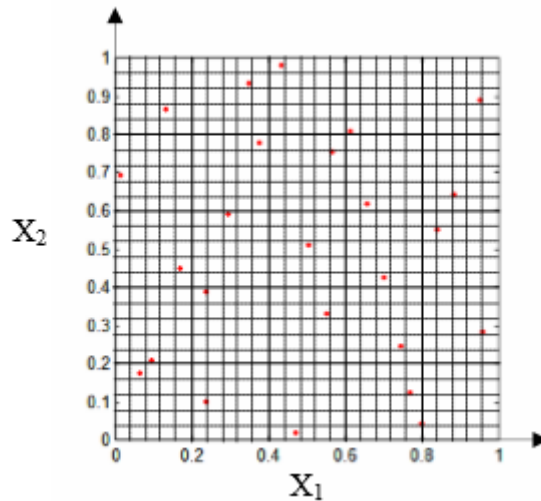


Figure (2.10). Un hypercube latin g n r  par un tableau orthogonal OA1 (25, 5, 5,2) dont les points sont randomis s et projet s sur le sous-espace (x_1, x_2) .

Cette technique est une m thode d' chantillonnage   trois degr s.

G om triquement, on effectue un premier  chantillonnage du cube unit  en q^p cellules. Parmi lesquelles, on en choisit $n = q^2$ de telle mani re qu'elles forment un tableau orthogonal de force 2. On choisit une petite cellule   l'int rieur des n cellules pr s lectionn es de fa on   ce qu'elles constituent un hyper cube latin. A ce stade, nous avons donc un hyper cube latin bas  sur un tableau orthogonal. Il nous suffit alors de choisir un point au hasard dans chacune des petites cellules (voir Figure (2.10)).

2.3. Suites   faible discr pance

Ces suites de points sont initialement utilis es pour remplacer les suites al atoires dans la m thode de Monte-Carlo, d'o  la d nomination de m thodes de quasi Monte-Carlo. la plupart de ces suites   faible discr pance sont g n r es par des algorithmes d terministes afin d'obtenir des points r partis le plus uniform ment possible dans le domaine exp rimental.[15]

La discr pance mesure l' cart entre une distribution de points uniforme et une distribution de points donn e ; en d'autres termes elle permet de mesurer l'irr gularit  de la distribution. Si on consid re la fonction de r partition empirique \hat{F}_n des points x^0, \dots, x^{n-1} on peut d finir la discr pance par :

$$D_n(X) = \max_{x \in [0,1]} |\hat{F}_n(x) - F_U(x)|$$

Avec F_U la fonction de r partition de la loi uniforme sur $[0,1]$.

Définition

Soit X un espace compact et μ une probabilité régulière sur la tribu borélienne de X . Une suite $(x_n)_{n \in \mathbb{N}}$ de points de X est uniformément distribuée si pour tout $f \in C(X)$, nous avons

$$\frac{1}{n} \sum_{k=1}^n f(x_k) \xrightarrow{n \rightarrow \infty} \int_X f d\mu$$

Remarque :

C'est la loi forte des grands nombres qui assure presque sûrement cette convergence. Une propriété importante, ici est la suivante :[5]

Propriété:[15]

Une suite (x_n) est uniformément distribuée si $\lim_{n \rightarrow +\infty} D_n(x) = 0$.

il existe de nombreux résultats de majoration de la discrétance. Commençons par énoncer le résultat le plus fameux, connu sous le nom d'inégalité de Koksma-Hlawka.

Théorème:(Hlawka, 1961)

Si f est une fonction à variation $v(f)$ bornée au sens de Hardy et Krause, alors pour toute suite de points x^1, \dots, x^n sur $[0,1]^p$, nous avons :

$$\left| \frac{1}{n} \sum_{i=1}^n f(x^i) - \int_{[0,1]^p} f(t) dt \right| \leq v(f) D_n(X). \quad (2.1)$$

Ainsi, l'erreur d'approximation est, dans le pire des cas, égale au produit de la variation et de la variation et de la discrétance $D_n(X)$.

Il est possible de construire des suites pour lesquelles la discrétance est plus faible que celle d'une suite aléatoire égale à $\frac{1}{\sqrt{n}}$. On les appelle des suites à faible discrétance. Ces suites ont la particularité de remplir le cube unité uniformément et de manière extrêmement régulière.

L'idée naturelle venant à l'esprit pour avoir les points les plus uniformément répartis est de considérer une grille régulière. Cependant, il est possible de montrer que la discrétance d'une telle distribution est de l'ordre de $\frac{1}{n}$ ce qui est en fait un bien médiocre résultat.

Une notion importante sur laquelle repose la construction de la plupart de ces suites est la définition de la fonction radicale inverse en base b donnée ci-dessous. .

Définition:

Soit b un entier ≥ 2 , la fonction radicale inverse en base b est donnée par

$$\varphi_b(i) = \frac{p_0}{b} + \frac{p_1}{b^2} + \dots + \frac{p_m}{b^{m+1}}$$

Avec $i = p_0 + p_1b + \dots + p_mb^m$ et $m = \begin{cases} 1 + [\log_b(i)] & \text{si } i \neq 0 \\ 1 & \text{sinon} \end{cases}$

La suite $C_b = \{x^0, x^1, \dots, x^{n-1}\}$ avec $x^i = \varphi_b(i)$ est appelée suite de Van Der Corput en base b .

2.3.1. Suites de Halton

Les suites de Halton sont la version en dimension $d \geq 1$ des suites de Van Der Corput qui quant à elles constituent la version unidimensionnelle de ces suites. L'idée pour générer les suites de Halton consiste à utiliser une base différente pour chaque dimension.

Définition :

Une suite de Halton $H_{b_1; \dots; b_d} = \{x^0; x^1; \dots; x^{n-1}\}$ en bases $b_1; \dots; b_p$ est définie par :

$$x^i = (\varphi_{b_1}(i); \dots; \varphi_{b_p}(i)) \in [0,1]^p$$

Où $b_1; \dots; b_p$ sont des entiers positifs premiers entre eux deux à deux.

Afin de minimiser au mieux la discrédance, il convient de prendre pour les bases les p premiers nombres premiers. En effet, cela permet de minimiser la constante intervenant dans le terme dominant de la majoration de la discrédance d'une telle suite [5] .

Ces suites ont l'avantage d'être faciles à implémenter et d'avoir un temps d'exécution très faible. Pour passer de $x^i = \varphi_b(i)$ à $x^{i+1} = \varphi_b(i + 1)$ il suffit d'effectuer une addition, en base b , à droite de $\frac{1}{b}$.

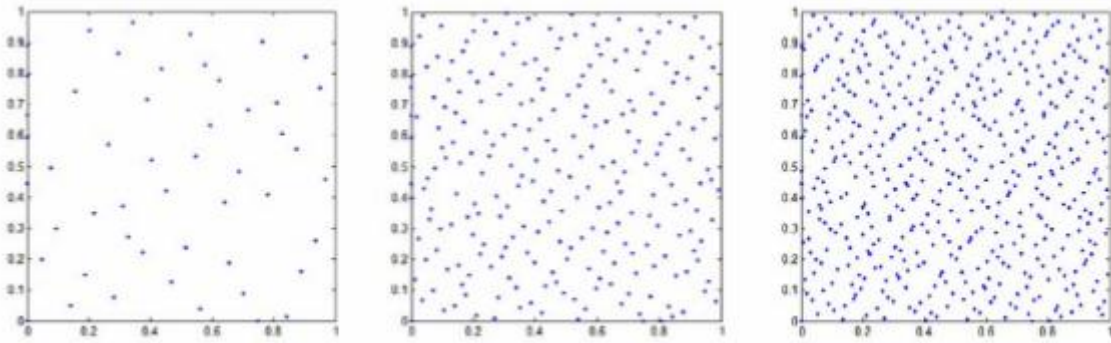
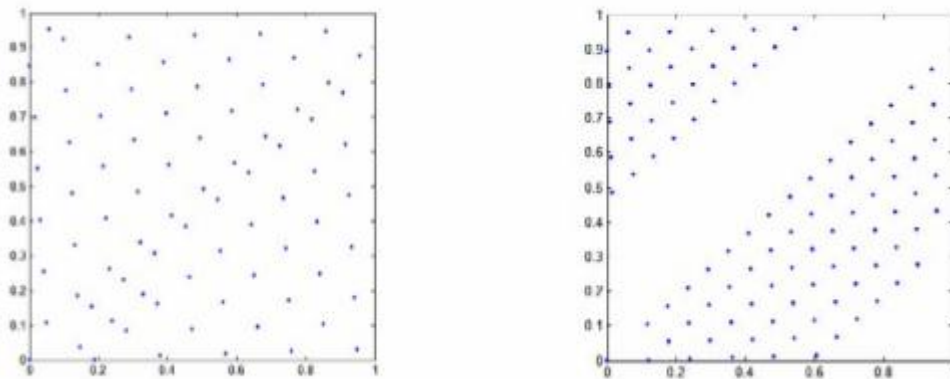


Figure (2.11). (De la gauche vers la droite) Les 50, 250 et 500 premiers points d'une suite de Halton en bases 2 et 3.

Malgré l'uniformité de la distribution dans le cube unité (Figure (2.11)), ces suites présentent des pathologies en dimension élevée. En effet, lorsque l'on projette les points d'une telle suite dans le carré unité, l'allure de la structure est parfois décevante (Figure (2.12)). Nous pouvons observer des diagonales décalées à chaque itération, qui laissent place à de nombreuses zones vides dans le carré unité. Ce phénomène résulte de la régularité des sous-suites de longueur b des suites de Van der Corput. A noter cependant que cette pathologie ne se manifeste pas pour n'importe quelle paire de bases. En général, il est plus fréquent d'observer ce phénomène en dimension élevée et lorsque la différence entre les deux bases est très petite par rapport aux bases utilisées.[5]



Figure(2.12). Des mauvaises projections en 2D. A gauche : les 80 premiers points d'une suite de Halton en bases 11 et 13. A droite : les 100 premiers points d'une suite de Halton en bases 17 et 19.

Une solution pour corriger cette pathologie est de casser les corrélations entre les suites de Van Der Corput des différentes dimensions. Braaten et Weller (1979) ont amélioré l'uniformité des points d'une suite de Halton en appliquant sur chaque nombre premier b_j

une permutation π_{b_j} sur $\{0; \dots \dots \dots; b_j - 1\}$ tel que $\pi_{b_j}(0) = 0$. Ainsi, à partir du nouveau développement de i en base b_j :

$$\varphi_{b_j}(i) = \frac{\pi_{b_j}(p_0)}{b_j} + \frac{\pi_{b_j}(p_1)}{b_j^2} + \dots \dots \dots + \frac{\pi_{b_j}(p_m)}{b_j^{m+1}}$$

On obtient une suite de Halton $Hg_{b_1; \dots \dots \dots; b_d} = \{x^0; x^1; \dots \dots \dots; x^{n-1}\}$ avec $x^i = (\varphi_{b_1}(i); \dots \dots \dots; \varphi_{b_d}(i))$ dite généralisée. La démarche est aussi dénommée comme une méthode de scrambling.[17]

2.3.2. Suites de Hammersley

Une suite de Hammersley en dimension d est construite à partir d'un terme dépendant du nombre de points et d'une suite de Halton en dimension $p-1$.

Définition :

Une suite de Hammersley $H_{b_1; \dots \dots \dots; b_{p-1}}^n = \{x^0; \dots \dots \dots; x^{n-1}\}$ en bases $b_1; \dots \dots \dots; b_{p-1}$ est définie par :

$$x^i = \left(\frac{i}{n}, \varphi_{b_1}(i), \dots \dots \dots; \varphi_{b_{p-1}}(i)\right) \in I^p$$

Où $b_1; \dots \dots \dots; b_{p-1}$ sont des entiers positifs premiers entre eux deux à deux.

Nous prendrons également le soin de prendre les $p - 1$ premiers nombre afin de réduire au mieux la discrédance de $H_{b_1; \dots \dots \dots; b_{p-1}}^n$.

Ces suites étant construites à partir de suites de Halton, elles présentent le même phénomène de diagonales successives. De plus, il est impossible de rajouter des points supplémentaires à ces suites sans pour autant perturber la discrédance. Dans le cas où le nombre de points à générer n 'est pas connu à l'avance, il est déconseillé d'utiliser une suite de Hammersley. Aussi, ces suites ont perdu l'aspect itératif des suites de Halton qui permettait de rajouter des points facilement. [20]

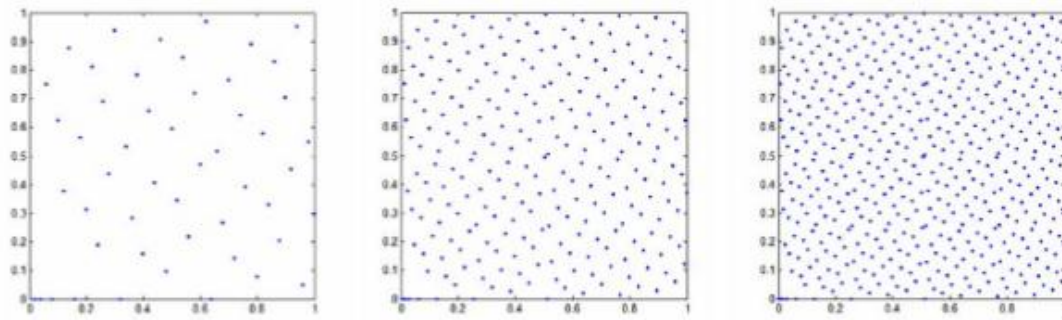


Figure (2.13). (De la gauche vers la droite) Les 50, 250 et 500 premiers points d'une suite de Hammersley en base 2.

2.3.3. Suites de Sobol'

Elles sont définies à partir de polynômes primitifs sur $Z_2 = \{0,1\}$. On rappelle à cet effet la définition d'un polynôme primitif.

Définition :

Un polynôme $p(t)$ de degré s de la forme $t^s + \mu_1 t^{s-1} + \dots + \mu_{s-1} t + \mu_s$ est primitif sur le corps Z_2 s'il est irréductible sur Z_2 et si le plus petit entier i pour lequel il divise $t^i - 1$ (ou $t^i + 1$) est égal à $2^s - 1$.

La notion d'irréductibilité consiste à vérifier si le polynôme $p(t)$ n'est divisible par aucun autre polynôme de degré inférieur. Le plus petit entier positif i tel que $p(t)$ divise $t^i - 1$ est aussi appelé l'ordre du polynôme.

Un polynôme primitif de degré s doit posséder les monômes 1 et t^s , ainsi qu'un nombre impair de termes.

Définition :

Une suite de Sobol' $S = \{x^0; x^1; \dots; x^{n-1}\}$ en dimension 1 est définie comme suit :

$$x^i = \frac{1}{2^m} \left(\bigoplus_{k=1}^m p_k l_k \right)$$

Où $(p_1; \dots; p_m)$ est la représentation binaire de i ; $m = \begin{cases} 1 & \text{si } i = 0 \\ 1 + \lfloor \log_2(i) \rfloor & \text{sinon} \end{cases}$ et où

le symbole \bigoplus désigne simplement l'addition dans Z_2 .

Les l_k avec $k > s$ sont obtenus en utilisant la relation de récurrence suivante

$$l_k = 2\mu_1 l_{k-1} \oplus 2^2 \mu_2 l_{k-2} \oplus \dots \oplus 2^{s-1} \mu_{s-1} l_{i-s+1} \oplus (2^s l_{i-s} \oplus l_{i-s})$$

Où les μ_k sont les coefficients d'un polynôme $t^s + \mu_1 t^{s-1} + \dots + \mu_{s-1} t + \mu_s$ primitif de degré s sur Z_2 et les nombres $l_1; \dots; l_s$ doivent être des entiers impairs tels que $1 < l_k < 2^k$ pour $k = 1, \dots, s$.

Afin de construire une suite de Sobol' en dimension p , il suffit de choisir p polynômes primitifs distincts.

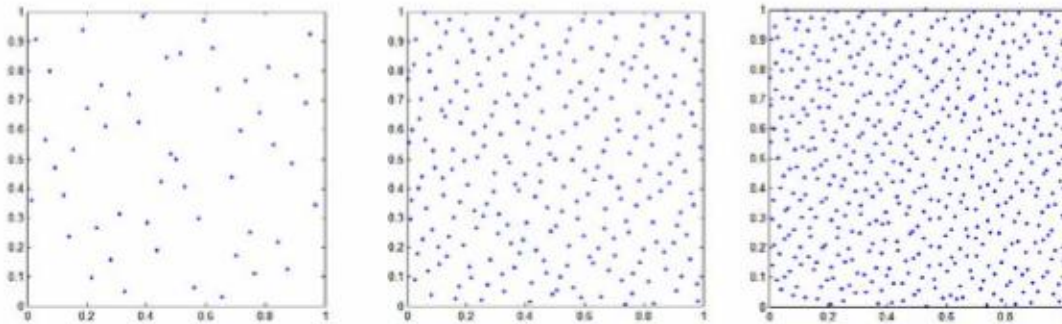


Figure (2.14). Les 50, 250 et 500 premiers points d'une suite de Sobol' en dimension 2.

Ces suites ont l'avantage d'être rapides à construire car la nature binaire des ordinateurs permet de réduire le temps de calcul. Elles ont aussi l'avantage de préserver, généralement, l'uniformité de la distribution lorsque la dimension augmente.[14]

CHAPITRE 3

LES SURFACES DE REPONSE POUR LES PLANS D'EXPERIENCES

Dans ce chapitre, nous présentons l'écriture pour cela, différents type de métamodèle seront représentés.

3.1 Régression linéaire

Le modèle est une relation entre les facteurs x_1, x_2, \dots, x_k et la réponse que l'on désire étudier. Considérons un phénomène aléatoire dépendant de k variables et supposons que l'on cherche à modéliser au mieux ce phénomène. La démarche statistique consiste alors à effectuer n expériences, judicieusement choisies dans le cas des plans d'expériences. Chacune est représentée par un point x de \mathbb{R}^k (ceci est possible si les variables étudiées sont quantitatives, pour le cas qualitatif on utilise un sous-ensemble de \mathbb{N}^k)

En désignant par $Y(x)$ la réponse mesurée en x on suppose classiquement que cette réponse résulte de la somme de la loi de réponse f en x (i.e. la réponse réelle recherchée) et du résidu ε en x (i.e. l'erreur commise). Donc :

$$Y(x) = f(x) + \varepsilon(x)$$

Le résidu peut rendre compte de bon nombre de causes telles que des erreurs dues à l'expérimentateur, un mauvais modèle postulé, l'oubli de certaines variables, ...etc. Nous supposons généralement que les résidus sont des variables aléatoires réelles vérifiant les trois hypothèses suivantes:

$$\begin{cases} \forall x, & E(\varepsilon(x)) = 0, \\ \forall x \neq x', & Cov(\varepsilon(x), \varepsilon(x')) = 0 \\ \forall x, & Var(\varepsilon(x)) = \sigma^2 \end{cases}$$

Nous considérons dans cette partie, un modèle statistique dépendant de k variables avec f fonction linéaire par rapport à p paramètres inconnus. Mathématiquement, un modèle est linéaire par rapport aux paramètres $\beta_{i(i=1,\dots,p)}$ si et seulement si chacune des dérivées partielles $\partial f(x)/\partial \beta_i$ ne dépendent plus de β_i . Etant donné un phénomène aléatoire à

expliquer, il n'est généralement pas simple de proposer un modèle adéquat. La fonction f est souvent trop complexe. C'est pourquoi il est courant de l'approcher à l'aide d'une classe de fonctions usuelles (développement de Taylor, séries de Fourier...etc). Si n expériences ont été réalisées, repérées par les points x_i ($i = 1, \dots, n$) de \mathbb{R}^k , nous avons donc :

$$\forall i = 1, \dots, n, \quad Y(x_i) = f(x_i) + \varepsilon(x_i)$$

Puisque f est une fonction linéaire par rapport aux paramètres inconnus, Nous pouvons donc aussi écrire ce modèle matriciellement sous la forme suivante :

$$Y = X\beta + \varepsilon$$

Avec :

- $Y \in \mathbb{R}^n$: Vecteur de réponses observées,
- $X(n, p)$: Matrice de calcul, qui dépend des points expérimentaux choisis pour exécuter le plan et du modèle postulé,
- $\beta \in \mathbb{R}^p$: Vecteur des coefficients inconnus,
- $\varepsilon \in \mathbb{R}^n$: Vecteur des résidus. Les hypothèses mentionnées au paragraphe précédent se traduisent simplement par : $E(\varepsilon) = 0$ et $Var(\varepsilon) = \sigma^2 I_n$.

En conséquence, $X\beta$ est donc la réponse moyenne donnée (prédite) par ce modèle.

3.2 Modèles additifs

Les modèles additifs sont un cas particulier des modèles additifs généralisés (GAM) développés par Hastie et Tibshirani [16]. Ce modèle présente l'avantage d'accorder à la surface plus de flexibilité qu'un modèle linéaire. La structure additive offre un réel outil de visualisation qui permet de représenter l'effet de chaque variable sur la réponse. Cette méthode est aussi particulièrement adaptée aux cas de grande dimension. L'hypothèse d'additivité sur laquelle est basé ce modèle suppose que la réponse s'écrit comme une somme de fonctions en chacune des variables explicatives. En comparaison avec la régression linéaire, chaque terme linéaire est remplacé par une fonction plus générale et la réponse Y se décompose sous la forme suivante :

$$Y = \alpha + f_1(x_1) + \dots + f_p(x_p) + \varepsilon$$

Où les f_j ($j = 1, \dots, p$) sont des fonctions de la variable associée x_j à estimer et ε un bruit blanc indépendant des variables. L'unicité de décomposition est assurée en supposant

que l'espérance mathématique $E(f_j)$ égale à 0. La constante α représente alors le niveau moyen des données.

Cette décomposition présente l'avantage d'ajuster des modèles non paramétriques simples et faciles à interpréter. Si le problème contient des interactions importantes, le modèle additif peut conduire à des résultats erronés. La généralisation du modèle ci-dessus peut se faire prenant des termes dépendant de plusieurs prédicteurs, souvent limité au nombre de 2.

Les modèles additifs ont l'avantage d'être un outil de visualisation des variations de la réponse dans chaque direction principale, ce qui peut amener à simplifier le modèle. L'étude des données par cette méthode fournit une information a priori qui pourra par la suite être utilisée pour régler les paramètres lors de l'ajustement d'un autre modèle. Citons par exemple l'étude de la nuée variographique empirique pour le krigeage. [5]

3.3 MARS et PolyMARS (Multivariate Additive Régression Spline)

Ces méthodes sont des modèles polynomiaux par morceaux qui sont particulièrement bien adaptés en grande dimension. La méthode MARS proposée par Friedman [17] utilise des polynômes par morceaux de degré 1. c'est à dire des splines linéaires pour estimer les fonctions de régression. Une extension de MARS proposée par Kooperberg [18] permet de construire des modèles PolyMARS (Polychotomous regression based on MARS). Ces modèles peuvent s'appliquer dans le cas des réponses multiples et paraissent être davantage adaptée à la grande dimension que la méthode MARS.

La principale similitude entre ces deux méthodes est la construction du modèle ajusté qui se déroule en trois phases :

- phase d'addition de fonctions de base dans le modèle.
- phase de suppression.
- Et sélection du modèle final selon un critère de validation croisée généralisée (GCV).

3.3.1 Méthode MARS

C'est une généralisation de la régression linéaire stepwise où les fonctions de base sont des splines linéaires. Le modèle s'écrit sous la forme suivante : $Y = \beta_0 + \sum_{j=1}^J \beta_j h_r(x)$

Où les paramètres β_j sont estimés par moindres carrés et les fonctions de base h_r sont des splines linéaires avec un nœud à la valeur t c'est à dire Des fonctions de la forme :

$$(x - t)_+ = \begin{cases} x - t & \text{si } x > t \\ 0 & \text{sinon} \end{cases} \text{ Ou } (t - x)_+ = \begin{cases} t - x & \text{si } x < t \\ 0 & \text{sinon} \end{cases}$$

Les fonctions de base sont construites en définissant un nœud pour chaque valeur x_j^i du plan d'expériences. Notons que seuls les nœuds « lointains » du nœud déjà choisi peuvent être candidats pour un nouveau nœud en cette variable.

La méthode MARS peut donc être vue comme une régression linéaire avec des fonctions de base construites à partir des observations au lieu d'être imposées.

3.3.2 Méthode PolyMARS

Cette méthode permet de prendre en compte plusieurs réponses simultanément. Nous considérons ici uniquement le cas d'une seule réponse. La réponse Y est approchée par des fonctions qui dépendent au plus de deux prédicteurs x_j . Similairement à MARS, les fonctions de base utilisées sont des splines linéaires et peuvent être ainsi de la forme 1

$$x_j ; (x_j - t)_+ ; x_j x_k ; (x_j - t)_+ x_k \text{ ou } (x_j - t)_+ (x_k - s)_+ .$$

3.3.3 Différences entre MARS et PolyMARS [5]

La principale différence se situe au niveau des possibilités pour les fonctions de base. Les nombreux choix concernant les fonctions de base pour la méthode PolyMARS peuvent parfois induire un temps de calcul plus important.

MARS explore un espace de fonctions alors que PolyMARS utilise la notion d'espace autorisé c'est-à-dire que la construction de l'espace satisfait les contraintes suivantes :

- ajout des effets simples x_j .
- ajout d'une spline linéaire $(x_j - t)_+$ si le prédicteur x_j est inclus dans le modèle.
- Ajout d'un terme d'interaction $x_j x_k$ si les deux prédicteurs sont inclus dans le modèle individuellement.
- ajout de $(x_j - t)_+ x_k$ si les deux termes sont inclus dans le modèle.

- ajout de $(x_j - t)_+(x_k - s)_+$ si $(x_j - t)_+ x_k$ et $(x_k - s)_+ x_j$ sont dans le modèle.

Le respect de cet ordre permet d'obtenir des modèles plus simples et donc plus faciles à interpréter. D'autres raisons telles que la réduction de la variance résiduelle et des arguments liés au taux de convergence permettent de justifier cette approche. Cependant, les contraintes de construction peuvent parfois faire obstacle à la bonne représentation de la réponse.

Lors de la phase de suppression, cet ordre doit aussi être respecté. Par exemple, il faut que le terme d'interaction ait été préalablement supprimé avant que les effets principaux le soit.

La méthode PolyMARS permet de diminuer le nombre de nœuds et donc de gagner en efficacité par rapport à la méthode MARS qui, quand à elle, teste toutes les valeurs observées. D'autre part, les procédures d'addition et de suppression dans PolyMARS sont basées sur des tests statistiques et non sur un critère GCV.

3.4 Réseaux de neurones

Un neurone est une fonction non linéaire, paramétrée, bornée, en général, une fonction sigmoïde :

$$a\left(\sum_{j=1}^{p+1} \theta_j x_j\right)$$

Où $x_j ; j = 1, \dots, p$ sont les variables ; $x_{p+1} = 1$; et $\theta_j ; j = 1; \dots, p + 1$. sont les paramètres de la fonction. La fonction sigmoïde est appelée fonction d'activation. On trouve par exemple ; $a(x) = x$ ou bien encore $a(x) = \frac{1}{1+e^{-x}}$.

Un réseau de neurones est une combinaison de neurones :

$$f(x; \theta) = \theta + \left(\sum_{k=1}^{N_c} \theta_k\right) a\left(\sum_{j=1}^{d+1} \theta_j x_j\right)$$

Où N_c est le nombre de neurones dans la couche caché et $\{\theta; \theta_k; \theta_{jk}\}, k = 1, \dots, N_c$ et $j = 1, \dots, p + 1$ sont les paramètres inconnues du réseau, encore appelés poids.

Les réseaux de neurones déterminent la nature de la relation entre les données à travers les interconnexions entre les neurones. Le réseau est nourri par les valeurs des paramètres en entrée

et dans le cas d'un réseau multicouche, les résultats en sortie sont obtenus par propagation des valeurs suivant la direction des liens et en classant les nœuds suivant la topologie du réseau (les nœuds de la première couche sont activés avant les nœuds de la deuxième, etc.). [5]

3.5 Krigeage

Cette méthode a été développée dans le domaine de la géostatistique par Matheron dans les années 60 [21] et a été utilisée en planification d'expériences numériques par Sacks [22]. Le krigeage présente l'avantage d'être une méthode interpolante aboutissant à la construction d'un modèle probabiliste. Cette méthode va donc dépendre fortement des points d'expériences proposés. Le krigeage consiste à construire un interpolateur qui est une fonction linéaire des observations tenant compte de la structure spatiale des données.

Le modèle d'interpolation dans sa forme générale est

$$Y(x) = \sum_{j=0}^p \beta_j f_j(x) + Z(x)$$

Où $f_1(x); \dots; f_p(x)$ sont des fonctions de base connues, $\beta = (\beta_1; \dots; \beta_p)^T$

Un vecteur de coefficients estimés et Z un processus gaussien stationnaire centré et de forme de covariance factorisée gaussienne.

Ce type de modèle est référé sous le nom de krigeage universel. Il est aussi possible de considérer un krigeage ordinaire de la forme $Y(x) = \mu + Z(x)$ où μ est une constante à estimer. Cette forme de krigeage est généralement la plus utilisée en pratique. Le krigeage est une méthode intermédiaire entre des méthodes rigides telles la régression où les fonctions utilisées sont imposées en amont et des méthodes souples telles les modèles additifs ou PolyMARS où l'estimation repose sur la construction de splines dans chaque direction. Concernant le krigeage, la régularité de la fonction est choisie au travers une fonction de corrélation et de paramètres déterminés par les observations. Un autre avantage du krigeage est de pouvoir associer une variance de prévision à la surface. Ainsi, il est intéressant de tenir compte de cette information en jouant différemment sur les zones connues avec précision et celles où on ne dispose que de très peu d'observations.

Les principaux inconvénients du krigeage se situent au niveau numérique. En effet, l'inversion de la matrice de corrélation est difficile à cause du mauvais conditionnement. Ce

qui se justifie, en grande dimension, par le nombre d'observations souvent insuffisant pour permettre l'identification des paramètres de portée. Ainsi, la fonction de vraisemblance est souvent plate au voisinage de l'optimum. Pire encore, le profil de vraisemblance peut présenter des optima locaux. Il convient alors de choisir des algorithmes d'optimisation adaptés. Contrairement aux autres méthodes présentées ci-dessus, l'interprétation de l'estimateur du krigeage est difficile. [5]

3.6 Variogramme [5]

En géostatistique, la fonction la plus utilisée pour décrire la continuité d'un processus est le variogramme. Et cela surtout parce qu'elle est plus simple à estimer que la covariance (qui demande l'estimation préalable de l'espérance mathématique), mais aussi parce qu'elle permet d'accommoder des situations où la variance du processus Z à modéliser est inconnue.

Le variogramme théorique est défini par :

$$\gamma(h) = \frac{1}{2} \text{var}[Z(x) - Z(x + h)] = \frac{1}{2} E[(Z(x) - Z(x + h))^2]$$

Sous l'hypothèse d'un processus est stationnaire.

Cette fonction, croissante en fonction de h , synthétise beaucoup d'information concernant le comportement conjoint des variables aléatoires. Si la valeur du variogramme converge vers une valeur limite, appelée palier, cela signifie qu'il y a une distance h à partir de laquelle $Z(x)$ et $Z(x + h)$ sont non corrélés on définit alors :

- portée α : il s'agit de la distance à partir de laquelle deux observations ne sont plus corrélées.
- Palier: c'est la variance de la variable aléatoire $Z(x)$.
- effet de pépite : il s'agit de variations à courte échelle, erreur de localisation, d'analyse et de précision analytique. La pépite traduit la discontinuité du processus à très courtes distances.

3.7 Critères de qualité des surfaces de réponse

Cette partie traite des différents critères permettant de mesurer la qualité de la surface ajustée tant en ajustement du modèle qu'en prédiction.

Après avoir construit un modèle, la question de la validité du modèle obtenu se pose. Un premier diagnostic consiste à étudier les valeurs du modèle aux points du plan d'expériences : nous parlerons dans ce cas d'**ajustement**. Pour se faire il existe des critères permettant de tester la bonne adéquation entre les observations et les prévisions sur les données utilisées pour la construction du modèle. C'est notamment le rôle des coefficients de détermination R^2 et R^2 ajusté, de l'étude des résidus, etc.

Une fois le modèle accepté, il est nécessaire de quantifier la qualité de la surface ajustée en termes de **prédiction**. Il s'agit ici de comparer les prévisions du modèle en des points différents de ceux du plan d'expériences. Dans le cas où le nombre de simulations réalisables n'est pas contraint, il est alors possible de définir un ensemble de points tests sur lequel évaluer des critères d'erreur en prédiction (MAE, RMSE, etc.). Dans le cas contraire, il est possible d'utiliser les techniques de validation croisée.

3.7.1 Ajustement

Le but ici est de proposer des indicateurs permettant de vérifier la bonne adéquation entre les données simulées et leur approximation calculée à partir du métamodèle. A noter que ce paragraphe ne concerne pas les méthodes d'interpolation où le métamodèle passe par les points du plan d'expériences.

3.7.2 Coefficients de détermination

Le coefficient de détermination est le critère généralement utilisé en régression linéaire pour tester l'ajustement du modèle. Ce coefficient est défini par :

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Où \bar{y} est une estimation de la moyenne de la réponse et n est le nombre de points du plan d'expériences.

Ce critère permet de mesurer le pourcentage de la variabilité totale de la réponse expliquée par le modèle. Le coefficient ne doit pas être utilisé pour comparer des modèles entre eux dans la mesure où il dépend fortement du modèle. En effet, le R^2 augmente lorsque le nombre de termes grandit même si tous les prédicteurs ne sont pas significatifs. Pour pallier ce problème, on utilise plutôt le coefficient de détermination ajusté défini par :

$$R_{adj}^2 = 1 - \frac{\frac{1}{n-(p+1)} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2}$$

Où p désigne le nombre de termes dans le modèle (constante non incluse). Si on exprime R_{adj}^2 en fonction de R^2 , on constate que le R_{adj}^2 est toujours plus faible que le R^2 et que la différence entre ces deux coefficients est d'autant plus grande que le nombre de prédicteurs est élevé. Le R_{adj}^2 est donc un compromis entre un modèle qui représente fidèlement la variabilité de la réponse et un modèle qui n'est pas trop riche en prédicteurs.

Un coefficient de détermination R^2 évalué sur un ensemble test (aussi appelé R^2 externe) donne une indication sur les qualités d'un métamodèle. La valeur du R^2 en prédiction peut être négative ce qui signifie que le modèle crée de la variabilité par rapport à un modèle constant.

3.7.3 Erreur quadratique moyenne : *MSE*

La *MSE* (Mean Square Error) est définie par :

$$MSE = \frac{1}{m} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Où n représente le nombre d'essais, Ce critère mesure l'erreur quadratique moyenne commise sur les données tests. Concrètement, une faible valeur de la *MSE* signifie que les valeurs prédites sont proches des valeurs réelles.

3.7.4 Erreur absolue moyenne : *MAE*

Le critère *MAE* (Mean Absolute Error) est défini par :

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|$$

Ce critère est une mesure similaire à la *RMSE*. Il est néanmoins plus robuste car moins sensible aux valeurs extrêmes que la *MSE*.

Toutes les mesures de distance (*MSE* et *MAE*) sont équivalentes et permettent de quantifier la proximité de la solution approchée par rapport aux données simulées. Une petite valeur de ces critères signifie que le modèle estimé est proche de la réalité.

3.7.5 Résidu maximal: RM

Le résidu maximal (correspondant à la norme L_∞) permet de vérifier si toutes les observations ont été correctement prédites. Il est défini par :

$$RM = \max_{1 \leq i \leq m} |y_i - \hat{y}_i|$$

On peut également introduire le critère RMA (Relative Maximum Absolute Error) comme étant le résidu maximal standardisé :

$$RMA = \frac{\max_{1 \leq i \leq m} |y_i - \hat{y}_i|}{\sigma_y}$$

Une grande valeur de ces critères indique une erreur importante d'approximation en un point du plan d'expériences même si la précision indiquée par le coefficient R^2 est satisfaisante. A noter que dans le cas de données normalisées, ces deux critères sont égaux.

CHAPITRE 4

APPLICATION POUR UNE FONCTION DETERMINISTE A DIMENSION 10

Dans ce chapitre, nous représentons une étude de comparaison entre certains nombre de plans d'expériences usuels sur une certaine fonction déterministe de 10 variables. Cette étude a été présentée par la première fois dans le travail de Jessica franc (voir [5])

4-1 Les données

La fonction déterministe de 10 facteurs définie sur $[-10,10]^{10}$ est donnée par :

$$f(x_1; \dots \dots \dots; x_{10}) = \sum_{j=1}^{10} \exp x_j \left\{ c_j - x_j - \ln \left[\sum_{k=1}^{10} \exp(x_k) \right] \right\}$$

Avec

$(c_1; \dots; c_{10}) =$

$(-6.089, -17.164, -34.054, -5.914, -24.721, -14.986, -24.100, -10.708, -26.662, -22,179)$

La figure suivante représente la coupe²⁷ à $x_i = 3$ pour $i = 1,3$.

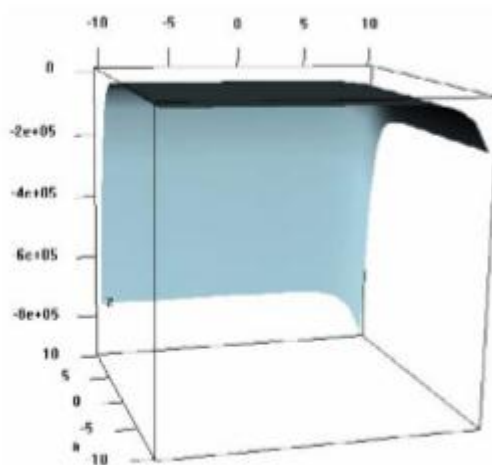


Figure (4.1) Coupe de la fonction de Jin en fonction de x_1 et x_3 ($x_i = 3, i = 1,3$)

Les plans utilisés dans la comparaison sont :

- Plans aléatoires (Aléatoire).
- Plans minimax (Minimax).
- Plans à entropie maximale (Dmax).

Les plans testés ici introduisant tous de l'aléa stochastique, l'étude a été conduite sur 50 plans de chaque type afin de donner un sens aux résultats.

Les différents plans sont comparés à l'aide de critères R^2 et $RMSE$ (voir chapitre 3) pour valider le modèle en ajustement et en prédiction. Nous avons choisi un modèle de régression de type polynomiale de degré 2. Le choix de ce type de modèle réside dans le fait qu'il est rapide à le mettre en œuvre en grande dimension.

4.2. Les résultats de comparaison et discussions

Dans cette partie, nous comparons les plans d'expériences entre eux à l'aide du critère R^2 pour mesurer la qualité de l'ajustement et du critère de $RMSE$ pour mesurer la qualité de prédiction. Les résultats concernant les plans introduisant de l'aléa stochastique seront présentés sous forme de box plot, et ceux des plans déterministes sous forme de tableau.

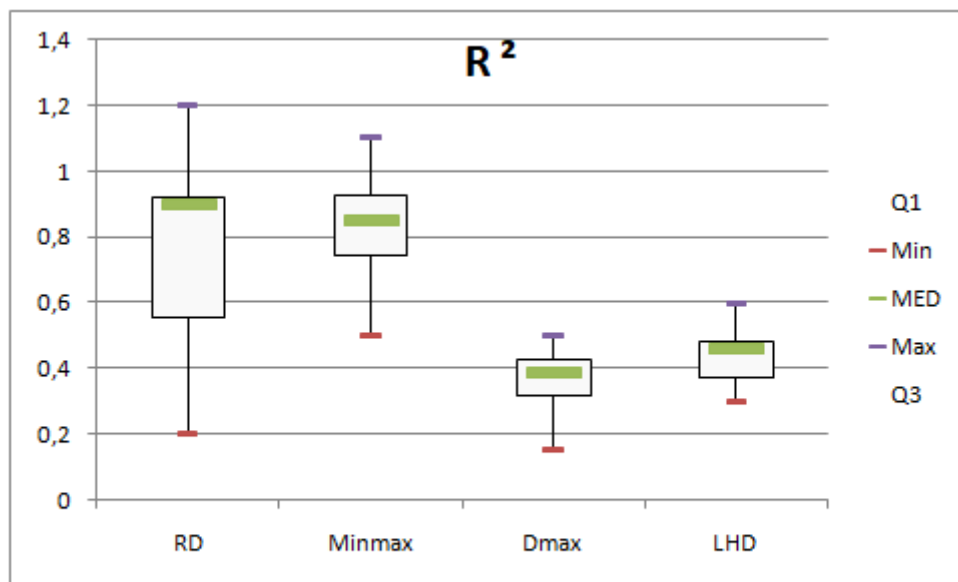


Figure 4.2 Critère de qualité d'ajustement

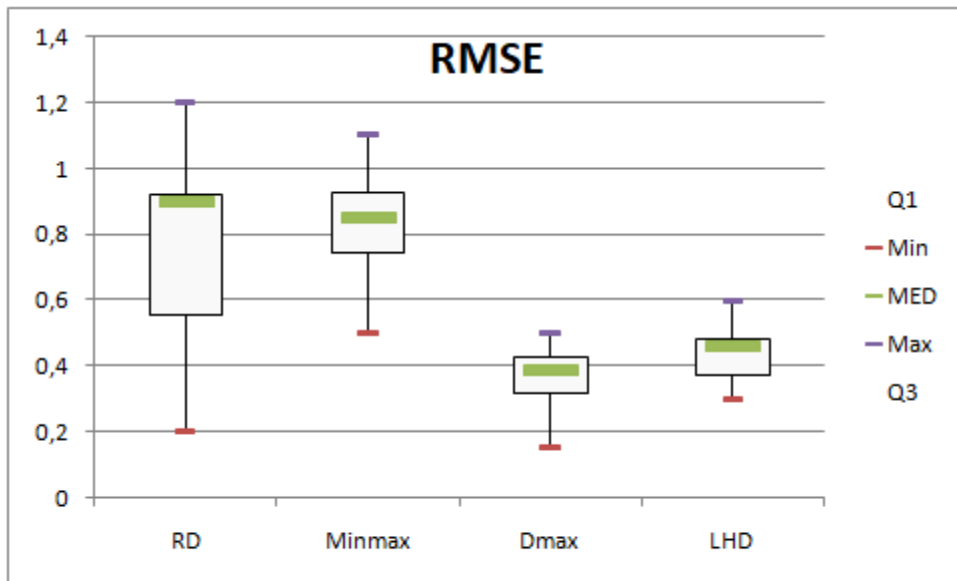


Figure 4.3 Critère de validation RMSE

En analysant les résultats montrés par la figure 4.2, nous remarquons que les plans minimax sont les plus mauvaise que ce soit prédiction ou en ajustement.

Les résultats concernant les suites à faible discrédance sont donnés dans le tableau ci-dessous :

plans \ critères	R^2	RMSE
Halton	0.93	0.49
Faure	0.88	0.63
Sobol	0.88	0.56

Tableau (4.1) Résultats des critères de qualité d'ajustement (R^2) et de validation (RMSE) pour les suites à faible discrédance

La suite de Faure et Sobol paraît être ici de mauvaise qualité, et toutes les suites sont mauvaises vis à vis du critère d'ajustement.

CONCLUSION

La Méthodologie de la Recherche Expérimentale, ou méthode de plans d'expériences, est une discipline mathématique faisant partie de la statistique inférentielle. La théorie de la méthode des plans d'expériences reste toujours en développement et elle a pris une importance si grande, grâce à l'existence de logiciels spécifiques.

La méthode par krigeage semble être l'alternative aux plans d'expériences la plus communément utilisée, notamment parce qu'elle propose un modèle statistique plus flexible elle n'est cependant pas la seule et autres types de modélisation peuvent être envisagés. Par exemple les réseaux de neurones sont connus pour leur capacité à modéliser des surfaces irrégulières.

Dans ce travail, nous nous sommes intéressés en premier lieu à présenter un état d'art sur les plans d'expériences numériques. En second lieu, une présentation a été réalisée concernant les différents méta-modèles qu'on peut utiliser dans les plans d'expériences numériques, et une application a été effectuée pour réaliser une comparaison entre certains nombre de plans numériques afin d'évaluer leur qualité de prédiction et d'ajustement.

BIBLIOGRAPHIE

1. GOUPY Jacques «Introduction aux Plans d'expériences». Dunod. Paris. 303 pages. (2001).
2. BOX George. E. P, HUNTER William G, HUNTER J. Stuart «Statistics for Expérimentateur» deuxième édition. John Wiley and Sons. New-York. 633 pages. (2005).
3. Fisher, R.A., "The Design of Experiments", 2nd edn, Oliver and Boyd, Edinburgh (1935).
4. Goupy, J., "La méthode des plans d'expériences", DUNOD, Paris, (1988), 303 pages.
5. Franco, J, « planification d'expériences Numériques en phase Exploratoire pour des codes de calculs Simulant des phénomènes complexes », Doctorat thèses, l'Ecole Nationale Supérieure des Mines de saint Etienne, (2008)
6. Box G.E.P., Behnken D.W. (1960). Some New Three-Level Designs for the Study of Quantitative Variables, *Technometrics*, 2, 455-475.
7. Doehlert D.H. (1970). Uniform shell designs, *Applied Statistics*, 19, 231-239.
8. Dejean J.-P., Blanc G. (1999). Managing Uncertainties on Production Prediction Using Integrated Statistical Methods. Paper SPE 56696 presented at the 1999 SPE Annual Technical Conference and Exhibition, Houston, U.S.A., October 3-6.
9. Jourdan A., Zabalza-Mezghani I. (2004). Response surface designs for scenario management and uncertainty quantification in reservoir production. *Mathematical Geology* 36 (8), 965-985.
10. Mc Kay M.D., Conover W.J. Beckman R.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, 21, 239-245.
11. Tarry G. (1901). Le problème de 36 officiers, *Compte-rendu de l'Assoc. Français Avanc. Sci. Naturel* 2, 170-203.
12. Jourdan A. (2000). Analyse statistique et échantillonnage d'expériences simulées, Thèse de l'Université de Pau et des Pays de l'Adour.
13. Tang B. (1993). Orthogonal array-based latin hypercubes. *Journal of the American Statistical Association*, 88, 1392-1397.

14. Thiémarc E. (2000). Sur le calcul et la majoration de la discrédance à l'origine. Thèse No 2259, Département de mathématiques, école polytechnique fédérale de Lausanne.
15. Niederreiter H. (1987). Point sets and sequences with small discrepancy. *Monath. Math.* 104, 273-337.
16. Hastie T., Tibshirani R., Friedman J. (2001). *The Elements of Statistical Learning*, Springer.
17. Friedman J. Multivariate Adaptive Regression Splines (invited paper). (1991) *Annals of Statistics*, 10 1-141.
18. Kooperberg C., Bose S., Stone C.J. (1997). Polychotomous Regression, *Journal of the American Statistical Association*, 92, 117-127.
19. Stone C.J., Hansen M.H., Kooperberg C., Truong Y.K. (1997). Polynomial splines and their tensor products in extended linear modeling. *Ann. Statist.* 25, 1371-1470.
20. Thiria S. Lechevallier Y., Gascuel O., Canu S. (1997). *Statistique et méthodes neuronales*, éditions Dunod : Discrimination, Prévision, Représentation structurée, Validation, Validité.
21. Matheron G. (1963). Principles of geostatistics, *Econm. Geol.* 58, 1246-1266.
22. Sacks J., Welch W.J, Mitchell T.J., Wynn H.P (1989). Design and Analysis of Computer Experiments, *Statistical Science*, 4, 409-435.