



République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur
et de la Recherche Scientifique
Université de Tissemsilt
Faculté des Sciences et de la Technologie
Département des Sciences et de la Technologie



Polycopié de cours

Plans d'expériences

Filière : Génie Mécanique

Spécialité : Fabrication Mécanique et Productique

Préparé par : Dr. HAMDI Amine

SOMMAIRE

SOMMAIRE.....	2
AVANT-PROPOS.....	5

CHAPITRE 1 : HISTORIQUE ET OBJECTIFS

1.1- Introduction	6
1.2- Qu'est-ce qu'une expérience	6
1.3- Planification des expériences	7
1.3.1- Méthode uni-factorielle	7
1.3.2- Méthode multifactorielle	8
1.4- Liste de contrôle	9
1.4.1- Définir les objectifs de l'expérience	9
1.4.2- Identifier toutes les sources de variation	10
1.4.3- Choisir une règle pour attribuer les unités expérimentales	11
1.4.4- Spécifier les mesures à effectuer	11
1.4.5- Effectuer une expérience pilote	12
1.4.6- Spécifier le modèle	12
1.4.7- Détailler l'analyse	13
1.4.8- Calculer le nombre d'observations nécessaires	13
1.4.9- Examiner les décisions précédentes	14
1.5- Aperçu historique	14
1.6- Objectifs	15
1.7- Terminologie	15
1.7.1- Facteurs	15
1.7.2- Niveau des facteurs	16
1.7.3- Effet	16
1.7.4- Interaction	16
1.7.5- Espace expérimental	17
1.7.6- Domaine d'étude	17
1.7.7- Réponses	17

CHAPITRE 2 : QUELQUES NOTIONS DE STATISTIQUES

2.1- Introduction.....	19
2.2- Terminologie	19
2.3- Mesures numériques	20
2.4- Diagrammes de distribution	21
2.5- Régression linéaire	23
2.5.1- Régression linéaire simple	24
2.5.2- Méthode des moindres carrés	24
2.5.3- Coefficient de corrélation	33
2.5.4- Coefficient de détermination	35
2.5.5- Régression linéaire multiple	37
2.5.5.1- Intervalle de confiance	38
2.5.5.2- Niveau de confiance	39

CHAPITRE 3 : PLANS OPTIMAUX

3.1- Introduction	41
3.2- Concepts des plans expérimentaux	41
3.2.1- Réplication	42
3.2.2- Blocage	42
3.2.3- Randomisation	43
3.3- Plan optimal de premier ordre	43
3.3.1- Plan optimal à deux niveaux de type 2^k	44
3.3.2- Plan optimal simplifié de type 2^{k-p}	59
3.4- Plan optimal de deuxième ordre	64

CHAPITRE 4 : OPTIMISATION DES PHENOMENES

4.1- Introduction	65
4.2- Approche expérimentale classique de l'ingénierie	65
4.3- Approche de modélisation numérique	66
4.4- Modélisation des relations entre les paramètres d'entrée-sortie	67
4.4.1- Méthode Taguchi	68
4.4.2- Méthodologie de surface de réponse	69
4.4.2.1- Plan factoriel complet	71
4.4.2.2- Plan factoriel fractionnaire.....	72
4.4.2.3- Conception composite centrale.....	73
4.4.2.4- Plan de Plackett-Burman.....	74
4.4.2.5- Plan Box-Behnken	75
4.4.3- Réseau de neurones artificiel	77
4.5- Outils d'optimisation	78
4.5.1- Optimisation mono-objectif	78
4.5.1.1- Méthode de Taguchi	78
4.5.2- Optimisation multi-objectif	78
4.5.2.1- Analyse relationnelle grise	78
4.5.2.2- Approche de la fonction de désirabilité	80
4.5.2.3- Outils d'optimisation de l'informatique douce	82

CHAPITRE 5 : APPLICATION TRIBOLOGIQUE

5.1- Introduction	83
5.2- Matériaux et méthodes	83
5.2.1- Pièce usinée et outil de coupe	83
5.2.2- Opération de coupe et de mesure	84
5.2.3- Calcul de l'énergie de coupe spécifique	86
5.3- Résultats et discussions	86
5.3.1- Résultats expérimentaux	86
5.3.2- Analyse de la variance (ANOVA)	87
5.3.2.1- ANOVA de F_z	87
5.3.2.2- ANOVA de R_a	88
5.3.2.3- ANOVA de T_c	89
5.3.2.4- ANOVA de E_{CS}	89
5.3.3- Méthodologie de surface de réponse	90
5.3.3.1- Régression linéaire multiple	90

5.3.3.2- Comparaison des résultats expérimentaux et prédites	91
5.3.3.3- Surface de réponse 3D	94
5.3.3.4- Validation expérimentale	97
5.3.4- Optimisation multi-objectif	97
5.4- Conclusions	100
Références bibliographiques	102

AVANT-PROPOS

Ce polycopié de cours, intitulé "**Plans d'expériences**" (en anglais : Design of Experiments, DoE), est destiné aux étudiants de Master 1 en génie mécanique, spécialité : Fabrication mécanique et productique (FM&P), durant le semestre 2. L'un des premiers défis pour un expérimentateur est de déterminer le nombre d'expériences nécessaires afin de maximiser les informations tout en tenant compte du budget nécessaire pour mener à bien les essais. A cet égard, les DoE sont essentiels dans divers domaines scientifiques et technologiques, car ils permettent d'organiser les expériences de manière rigoureuse, de collecter des données fiables et d'obtenir des résultats reproductibles. Ce module est crucial dans l'industrie, car il aide à déterminer les conditions optimales pour obtenir une réponse maximale ou minimale, tout en mettant l'accent sur le temps et le coût de production. Le contenu de ce cours a été élaboré conformément aux directives de la dernière mise à jour du programme, en alignement avec les recommandations du comité pédagogique national du domaine des sciences et technologies (CPND-ST).

L'objectif principal de ce module est d'approfondir la compréhension des étudiants en spécialité FM&P sur la planification des expériences. Le cours commence par des définitions clés, en mettant l'accent sur deux méthodes de planification, une liste de contrôle pour effectuer les expériences et une terminologie appropriée. Le chapitre 2 introduit quelques notions de statistiques, telles que les méthodes de régression, le coefficient de détermination (R^2) et les intervalles de confiance. Le chapitre 3 guide les étudiants dans la création de plans d'expériences bien structurés qui maximisent les informations obtenues tout en minimisant le nombre d'essais nécessaires. Le chapitre 4 se concentre sur les méthodes de modélisation mathématique des relations entre les paramètres d'entrée et de sortie, ainsi que sur les outils d'optimisation mono et multi-objectifs, notamment la méthodologie de surface de réponse (RSM), le réseau de neurones artificiels (ANN) et la méthode de la fonction de désirabilité (DF). Enfin, le dernier chapitre présente une étude de cas pour évaluer certains critères d'usinabilité lors du tournage de polypropylène (PP). L'ANOVA, le RSM et la méthode DF sont utilisées pour analyser, modéliser et optimiser les paramètres d'entrée et de sortie.

Ce cours vise à offrir aux étudiants une vision complète et approfondie de la planification des expériences, renforçant ainsi leurs connaissances en modélisation et optimisation dans le domaine de la FM&P.

CHAPITRE 1 :

Historique et objectifs

1.1- Introduction

L'un des premiers défis pour un expérimentateur est de déterminer le nombre d'expériences nécessaires afin de maximiser les informations tout en tenant compte du budget nécessaire pour mener à bien les essais. Dans cette optique, une alternative à cette phrase est la question suivante : "Comment puis-je maximiser les informations avec un budget limité ?". La première étape pour y répondre consiste à poser d'autres questions, telles que : "Quel est l'objectif principal de cette expérience ?" et "Quelles démonstrations espère-je accomplir ?". Typiquement, une expérience peut être réalisée pour une ou plusieurs des raisons suivantes :

- (i) Identifier les causes principales de la variation dans une réponse mesurée,
- (ii) Déterminer les conditions conduisant à une réponse maximale ou minimale,
- (iii) Comparer les réponses obtenues avec différentes configurations des variables contrôlables,
- (iv) Développer un modèle mathématique pour prédire les réponses futures.

L'art de concevoir une expérience et celui d'analyser sont étroitement liés et doivent être abordés de manière conjointe. Lors de la conception d'une expérience, il est crucial de prendre en compte l'analyse qui sera effectuée ultérieurement. À son tour, l'efficacité de l'analyse dépendra du plan expérimental spécifique utilisé pour collecter les données. Sans ces considérations, il est possible d'investir beaucoup de temps, d'efforts et d'argent dans la collecte de données qui semblent pertinentes pour l'objectif visé, mais qui, en réalité, contribuent peu aux questions de recherche posées. Un principe directeur de la conception expérimentale est de "rester simple". En général, l'interprétation et la présentation des résultats des expériences sont plus claires lorsque les expériences sont plus simples.

1.2- Qu'est-ce qu'une expérience

L'expérimentation, également appelée expérience, est une démarche scientifique qui vise à obtenir des résultats expérimentaux. Cette méthode est particulièrement utile pour établir des

relations de cause à effet, il est possible d'observer l'impact d'une cause sur un effet. Les observations peuvent être obtenues à partir d'études observationnelles ainsi que d'expériences, mais seule une expérience permet de tirer des conclusions sur la relation de cause à effet.

1.3- Planification des expériences

Planifier une expérience est passionnant, mais c'est aussi très long. Parfois, on est tenté de collecter des données rapidement sans bien réfléchir à la façon dont on mène l'expérience. Cependant, cette méthode ne garantit pas toujours d'obtenir les bonnes données en quantité suffisante pour les analyser avec précision. La section 1.4 offre un guide détaillé, étape par étape, sur la façon de planifier une expérience.

Lorsqu'il existe plusieurs paramètres d'entrée (notés $X_1, X_2, \dots, X_i, \dots, X_n$) ayant une influence sur un paramètre de sortie Y , la démarche expérimentale peut être utilisée en pratique selon deux méthodes (voir figure 1.1) : la méthode uni-factorielle ou la méthode multifactorielle.



Figure 1.1- Deux méthodes de démarche expérimentale.

1.3.1- Méthode uni-factorielle

La méthode uni-factorielle consiste à ne faire varier qu'un seul facteur X [$X_i \in \{X_{min}, X_{max}\}$] à la fois, tout en maintenant les autres facteurs constants. Pour chaque valeur de $X = X_i$ ($i \in [1, N]$), un essai est réalisé et la valeur correspondante de Y_j est mesurée (voir tableau 1.1).

Tableau 1.1- Méthodologie de réalisation des essais selon la démarche uni-factorielle.

N°	Facteurs	Paramètre de sortie						
	X	Y_1	Y_2	-	Y_j	-	Y_n	Y_{moy}
1	X_1	Y_{11}	Y_{21}	-	Y_{j1}	-	Y_{n1}	Y_{moy1}
2	X_2	Y_{12}	Y_{22}	-	Y_{j2}	-	Y_{n2}	Y_{moy2}
3	X_3	Y_{13}	Y_{23}	-	Y_{j3}	-	Y_{n3}	Y_{moy3}
-	-	-	-	-	-	-	-	-
i	X_i	Y_{1i}	Y_{2i}	-	Y_{ji}	-	Y_{ni}	Y_{moyi}
-	-	-	-	-	-	-	-	-
N	X_N	Y_{1N}	Y_{2N}	-	Y_{jN}	-	Y_{nN}	Y_{moyN}

Cette approche opérationnelle permet de mesurer l'effet de chaque facteur sur le paramètre de sortie Y , en éliminant les interactions possibles avec les autres facteurs. Cependant, elle ne permet pas de prendre en compte les effets combinatoires entre les facteurs.

Il est plus efficace d'examiner simultanément toutes les causes possibles de variation plutôt que de le faire une par une. En général, cela nécessite moins d'observations tout en fournissant plus d'informations sur le système étudié. Cette approche est appelée "expérience multifactorielle".

1.3.2- Méthode multifactorielle (plan d'expériences)

Un plan d'expériences est une méthode statistique qui suit des règles mathématiques pour déterminer le nombre d'essais nécessaires afin de maximiser ou minimiser ces derniers. L'objectif principal de cette approche expérimentale est d'obtenir les résultats les plus fiables et les plus efficaces possible en suivant un programme préétabli.

Le plan d'expériences ou méthode multifactorielle est une recherche expérimentale bien planifiée qui permet d'étudier l'impact de plusieurs paramètres d'entrée (facteurs : variables indépendantes : $X_1, X_2, \dots, X_i, \dots, X_n$) ainsi que de leurs interactions sur un ou plusieurs paramètres de sortie (réponse : variable dépendante : Y) (voir figure 1.2). La réponse dépend de plusieurs facteurs, ces derniers étant modifiés de manière commandée et pouvant avoir une influence sur la variable dépendante.

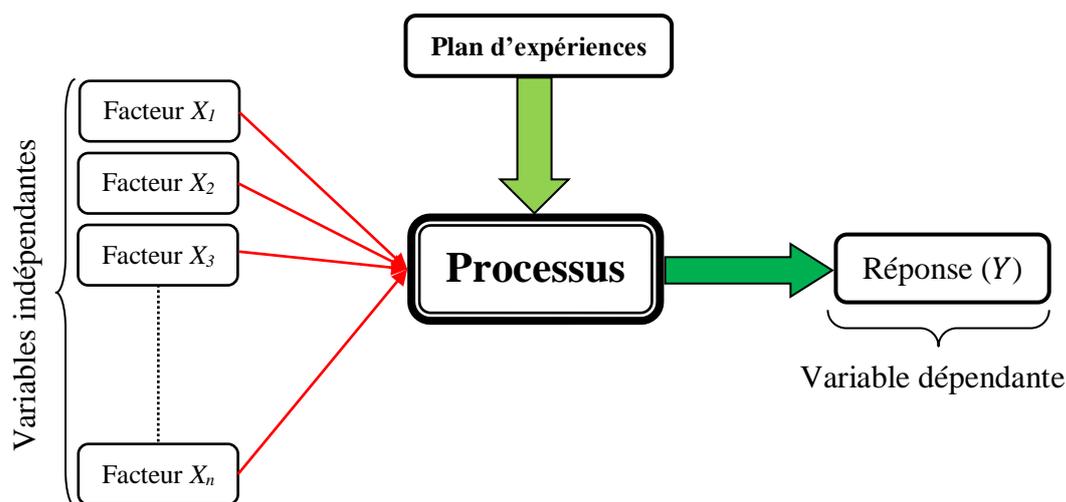


Figure 1.2- Utilité de plan d'expériences : plusieurs paramètres d'entrée (X_i) sur un paramètre de sortie (une grandeur d'intérêt Y).

Il convient également de noter que les termes "facteur" et "réponse" sont les plus couramment utilisés dans le domaine des plans d'expériences, plutôt que le terme "variable".

En utilisant la méthode multifactorielle, il est possible d'étudier l'influence des différents facteurs sur la réponse et de déterminer les interactions possibles entre eux. Cela permet de réaliser une analyse plus complète et d'obtenir des résultats plus précis. De plus, la planification rigoureuse des essais permet de minimiser le nombre d'essais nécessaires tout en maximisant les informations obtenues.

La réponse Y (OUTPUT) d'un processus dépend d'un ou plusieurs facteurs X_i (INPUT), d'où la relation suivante :

$$Y = f(X_i) \dots \dots \dots (1.1)$$

f : la fonction qui relie les variables indépendantes (X_i) et la grandeur d'intérêt (Y).

Pour déterminer le nombre d'essais nécessaires selon la méthode multifactorielle, la formule suivante est utilisée :

$$N = q^k \dots \dots \dots (1.2)$$

Où :

- ✖ N : nombre d'essais,
- ✖ q : nombre de niveaux de variation des facteurs de base,
- ✖ k : nombre de facteurs de base.

1.4- Liste de contrôle

La liste de contrôle suivante résume de nombreuses décisions à prendre à chaque étape du processus de planification expérimentale. Ces étapes ne sont pas indépendantes, et il peut être nécessaire de revenir en arrière et de réviser certaines décisions prises à une étape antérieure à tout moment du processus. Un aperçu simplifié des décisions à prendre à chaque étape de la liste de contrôle est donnée ci-dessous. Les données ne doivent être collectées qu'après avoir pris toutes ces décisions.

1.4.1- Définir les objectifs de l'expérience

Il est crucial d'établir une liste précise des questions à aborder lors de l'expérience, car cette liste guide les décisions à prendre aux étapes ultérieures de la planification. Il est recommandé

de ne lister que les questions essentielles, car des questions secondaires peuvent compliquer inutilement l'expérience, augmentant ainsi les coûts et les risques d'erreurs. Cependant, omettre accidentellement des questions importantes pourrait laisser des lacunes dans les données collectées. Lors de la compilation de la liste des objectifs, il est souvent utile de décrire les conclusions attendues de l'analyse des données. Les objectifs peuvent nécessiter un affinement à mesure que les étapes suivantes de la planification sont complétées.

1.4.2- Identifier toutes les sources de variation

Une source de variation est tout élément pouvant influencer les valeurs numériques des observations. Certaines sources de variation sont mineures et n'entraînent que de légères différences dans les données, tandis que d'autres sont majeures et doivent être prises en compte dans l'expérience. Il est recommandé de dresser une liste exhaustive de chaque source de variation envisageable et de les classer comme majeures ou mineures. Les sources majeures de variation peuvent être divisées en deux catégories : les facteurs de traitement, qui suscitent un intérêt particulier de la part de l'expérimentateur, et les facteurs parasites, qui ne sont pas directement liés à l'objectif de l'expérience.

(i) Facteurs de traitement et leurs niveaux

Le terme "facteur de traitement" est utilisé pour désigner toute source de variation dans une expérience, comme les paramètres de coupe en usinage. À cette étape, nous choisissons les facteurs de traitement et leurs niveaux. Les niveaux sont les différentes valeurs des facteurs que nous allons utiliser dans notre expérience. Par exemple, un facteur de traitement pourrait être la vitesse de coupe ou le type de lubrifiant. Les niveaux pourraient être différentes valeurs de la vitesse de coupe (V_c) telles que 100, 200 et 300 m/min, qui pourraient être codées comme 1, 2, 3 dans le plan d'expérience.

(ii) Unités expérimentales

Les unités expérimentales sont les "matériaux" sur lesquels les niveaux des facteurs de traitement sont testés. Par exemple, en usinage, cela pourrait être de l'acier comme le 100Cr6. Si l'expérience s'étend sur une période de temps avec des observations collectées de manière séquentielle, alors les moments de la journée peuvent également être considérés comme des unités expérimentales. Cependant, elles doivent être représentatives des conditions dans lesquelles les conclusions de l'expérience seront appliquées. Par exemple, les résultats d'une expérience menée sur des étudiants universitaires peuvent ne pas être généralisables à tous les

adultes du pays. De même, les conclusions d'une expérience d'usinage réalisée avec une lubrification spécifique peuvent ne pas s'appliquer dans des conditions de coupe à sec.

(iii) Facteurs de blocage

Lors de la conception d'une expérience, il est crucial de distinguer les effets des facteurs indésirables de ceux des facteurs de traitement. Plusieurs approches peuvent être utilisées pour gérer ces facteurs parasites en fonction de leur nature. Il peut être judicieux de restreindre la portée de l'expérience en fixant le niveau des facteurs parasites, bien que cela puisse nécessiter une révision des objectifs établis à l'étape précédente, car les conclusions pourraient être moins généralisables. Une autre approche consiste à maintenir le niveau des facteurs parasites constant pour un groupe d'unités expérimentales, à le modifier à une valeur fixe différente pour un deuxième groupe, et ainsi de suite. Ce type de facteur parasite est appelé facteur de blocage, et les unités expérimentales mesurées sous le même niveau de ce facteur sont considérées comme étant dans le même bloc. Par exemple, si la température est susceptible d'affecter les observations, mais n'est pas un facteur d'intérêt en soi, toute l'expérience pourrait être réalisée à une seule température pour limiter les conclusions à cette condition spécifique. En revanche, les unités expérimentales pourraient être regroupées en blocs, chaque bloc étant mesuré à une température différente.

1.4.3- Choisir une règle pour attribuer les unités expérimentales aux niveaux des facteurs de traitement

La règle d'attribution, ou la conception expérimentale, détermine quelles unités expérimentales doivent être soumises à quels traitements. Le choix de la conception, qu'il inclue ou non des facteurs de blocage, dépend de toutes les décisions précédentes prises dans la liste de contrôle. Plusieurs conceptions standard sont couramment utilisées en pratique. L'attribution effective des unités expérimentales aux traitements devrait être réalisée de manière aléatoire, tout en respectant les restrictions imposées par la conception choisie.

1.4.4- Spécifier les mesures à effectuer, la procédure expérimentale et les difficultés anticipées

Les données collectées lors d'une expérience sont des mesures d'une variable de réponse, comme la rugosité de surface, les efforts de coupe ou la productivité d'une machine. Les unités de mesure doivent être choisies en fonction des objectifs de l'expérience. Par exemple,

si l'objectif est d'atteindre une haute précision dans la finition d'une pièce mécanique, il est recommandé de prendre des mesures précises au micromètre.

Il est courant de rencontrer des difficultés imprévues lors de la collecte de données, mais celles-ci peuvent souvent être détectées en effectuant quelques mesures de pratique ou en menant une expérience pilote (comme indiqué à l'étape 1.4.5). En dressant une liste des difficultés anticipées, on peut identifier les sources de variation nécessaires à l'étape 1.4.3 de la liste de contrôle. Cela offre également la possibilité de simplifier la procédure expérimentale avant le début de l'expérience.

Il est essentiel d'établir des instructions précises sur la manière de procéder aux mesures. Cela englobe des directives concernant les instruments de mesure à utiliser, le moment des prises de mesures et la méthode d'enregistrement des données. Il est primordial que tous les participants à l'expérience suivent scrupuleusement ces instructions. Il est recommandé de créer une feuille de collecte de données qui spécifie l'ordre des observations à effectuer ainsi que les unités de mesure associées.

1.4.5- Effectuer une expérience pilote

Une expérience pilote est une version réduite de l'expérience principale, comprenant seulement quelques observations. Aucune conclusion définitive n'est nécessairement attendue de cette expérience. Son objectif est d'aider à compléter la liste de contrôle, de permettre la pratique des techniques expérimentales et d'identifier les problèmes imprévus dans la collecte de données. Si l'expérience pilote est assez étendue, elle peut également aider à choisir un modèle approprié pour l'expérience principale. Les erreurs expérimentales observées lors de l'expérience pilote peuvent aider à calculer le nombre d'observations nécessaires pour l'expérience principale (étape 1.4.8). À ce stade, les étapes 1.4.1 à 1.4.4 de la liste de contrôle doivent être réévaluées et des ajustements doivent être apportés si nécessaire.

1.4.6- Spécifier le modèle

Le modèle doit explicitement décrire la relation présumée entre la variable de réponse et les principales sources de variation identifiées à l'étape 1.4.2. Les techniques d'analyse des données expérimentales dépendent de la forme du modèle, il est donc crucial que celui-ci reflète fidèlement la véritable relation. Le modèle le plus courant est le modèle linéaire, où la variable de réponse est exprimée comme une combinaison linéaire des termes représentant les principales sources de variation, ainsi qu'un terme d'erreur qui capture les variations mineures.

Une expérience pilote à l'étape 1.4.5 peut aider à évaluer si les données sont bien décrites par le modèle proposé.

1.4.7- Détailler l'analyse

Le type d'analyse effectuée sur les données expérimentales dépend des objectifs établis à l'étape 1.4.1, de la conception choisie à l'étape 1.4.3 et du modèle associé spécifié à l'étape 1.4.6. L'ensemble de l'analyse devrait être détaillé, incluant les hypothèses à tester et les intervalles de confiance à calculer. Cette analyse détermine non seulement les calculs à effectuer à l'étape 1.4.8, mais vérifie également si la conception est appropriée pour atteindre les objectifs de l'expérience.

Dans l'analyse des données, il est recommandé de fournir à la fois des analyses graphiques et statistiques. Les graphiques illustrant les variations des réponses en fonction des différents réglages des facteurs permettent à l'expérimentateur de saisir les implications pratiques des résultats statistiques et de communiquer efficacement les conclusions de l'expérience à d'autres personnes. De plus, les graphiques de données aident à vérifier la validité du modèle proposé et à repérer d'éventuelles observations atypiques, comme discuté dans le chapitre 5. L'analyse statistique quantifie les variations des réponses en fonction des facteurs, ce qui permet de clarifier des conclusions qui pourraient être ambiguës ou peu évidentes dans les graphiques de données.

Le but d'une expérience peut varier de l'exploratoire, où l'objectif est de découvrir de nouvelles sources de variabilité importantes, au confirmatoire, où l'on cherche à confirmer que les sources de variabilité précédemment identifiées justifient une étude plus approfondie. La philosophie de l'analyse dépend du but de l'expérience. Aux premiers stades de l'expérimentation, l'analyse peut être exploratoire, où les données sont représentées graphiquement et analysées de manière à identifier les sources de variation importantes. Aux stades ultérieurs, l'analyse est généralement confirmatoire, où l'on postule un modèle mathématique de la réponse, où l'on teste des hypothèses et où l'on calcule des intervalles de confiance.

1.4.8- Calculer le nombre d'observations nécessaires

À ce stade de la liste de contrôle, il est crucial de calculer le nombre d'observations nécessaires pour atteindre les objectifs de l'expérience. Un échantillon trop petit peut entraîner

des conclusions peu fiables, tandis qu'un échantillon trop grand gaspille des ressources précieuses.

1.4.9- Examiner les décisions précédentes (réviser si nécessaire)

La révision est indispensable lorsque le nombre d'observations calculé à l'étape 1.4.8 dépasse les limites de temps ou de budget disponibles. Elle doit débiter dès l'étape 1.4.1, en réduisant généralement la portée de l'expérience. Si aucune révision n'est nécessaire, la collecte de données peut débiter. Il est désormais évident que la réflexion préalable à la réalisation d'une expérience est essentielle. La collecte de données est souvent la partie la plus coûteuse et chronophage de la procédure expérimentale. Investir davantage de temps dans la planification contribue à garantir une utilisation optimale des données. Aucune méthode d'analyse ne peut compenser une expérience mal conçue.

Bien qu'un scientifique expérimental compétent en conception et analyse des expériences puisse se passer de l'aide d'un statisticien, il est souvent bénéfique de discuter de la liste de contrôle avec une personne extérieure à l'expérience. L'étape 1.4.1 de la liste de contrôle est souvent la plus délicate à compléter. Lorsqu'un statisticien consulte un client, sa première question est généralement : "Quel est l'objectif précis de votre expérience ? Que cherchez-vous à démontrer ?" Si ces questions ne peuvent pas être clairement répondues, il n'est pas judicieux pour l'expérimentateur de commencer à collecter des données sans savoir exactement pourquoi. De même, il est essentiel qu'un statisticien consultant comprenne non seulement l'objectif de l'expérience, mais aussi la méthode expérimentale. Il n'est pas utile de recommander à un expérimentateur de réaliser une expérience pilote qui absorbe la majeure partie du budget.

1.5- Aperçu historique

- ❧ Dans les années 1925, le biologiste et statisticien britannique Ronald Aylmer FISHER à introduit les bases de la statistique moderne (l'analyse de la variance ANOVA, les plans d'expériences,...etc).
- ❧ Après la deuxième guerre mondiale et exactement entre les années 1950 à 1960, l'ingénieur japonais de statistique Genichi TAGUCHI à développé une méthode statistique pour réaliser des plans d'expériences. Cette technique est appliquée dans l'industrie afin de minimiser le nombre d'essais et améliorer la qualité du processus.

- ❧ Dans les années 1970, les travaux de Taguchi sont diffusés aux Etats Unis et appliqués à grande échelle (agriculture, médecine, biologie, pharmacie, génie civil, chimie, fabrication mécanique,...etc), et notamment dans l'industrie automobile et aéronautique.
- ❧ Dans les années 1980, les grandes et les petites entreprises industrielles en Europe ont commencé l'utilisation des plans d'expériences.
- ❧ Dans les années 1990, c'est l'apparition des plans de mélanges dans les domaines de l'agro-alimentaire afin de les utiliser dans la recherche des compositions optimales de produit.
- ❧ Aujourd'hui, les plans d'expériences sont principalement utilisés en industrie, et en particulier dans le domaine de recherche scientifique, pour l'amélioration de la qualité des produits.

1.6- Objectifs

Les intérêts de la méthode de planification des expériences sont :

- ❖ Optimisation du processus de production,
- ❖ Optimisation le nombre des essais avec le maximum des renseignements,
- ❖ Investigation pour mieux comprendre un phénomène,
- ❖ Efficacité supérieure à toute autres suite d'essais de même taille non planifiée,
- ❖ Etudie l'effet des facteurs et de leur interaction sur les résultats,
- ❖ Facilité d'interprétation des résultats,
- ❖ Diminuer l'effort de réalisation des expériences,
- ❖ Diminuer le temps de réalisation des expériences, par suite le temps machine,
- ❖ Economiser l'échelle (diminution du coût moyen de production).

1.7- Terminologie

1.7.1- Facteurs

Un facteur est une grandeur quantitative ou qualitative que l'on peut modifier facilement d'un essai à un autre et qui influence sur la réponse. Ainsi, il existe deux types de facteurs :

➡ Facteur quantitatif :

Un facteur quantitatif est un paramètre expérimental mesurable exprimé par une valeur numérique, telle que la vitesse, la masse, la force, la dureté, ...etc.

➔ **Facteur qualitatif :**

Un facteur qualitatif est un paramètre non mesurable et qui ne peut pas être exprimé par un chiffre, comme par exemple la couleur, le type de matériau, ...etc.

- ➔ Dans le cas où il est possible de choisir toutes les valeurs possibles dans un intervalle de variation pour un facteur, ce dernier est considéré comme un facteur continu (pression, température, vitesse,...etc). À l'inverse, si le facteur ne peut prendre que des valeurs définies, il est qualifié de facteur discret (couleurs, genre,...etc).

1.7.2- Niveau des facteurs

Un niveau correspond à une valeur donnée à un facteur pour exécuter une expérience. En général, la variation de chaque facteur est limitée entre deux valeurs. La valeur minimale est appelée niveau bas, représenté par "-1", et la valeur maximale est appelée niveau haut, représenté par "+1" (voir figure 1.3). L'ensemble des valeurs comprises entre le niveau bas et le niveau haut est appelé domaine de variation du facteur, ou plus simplement, domaine du facteur.

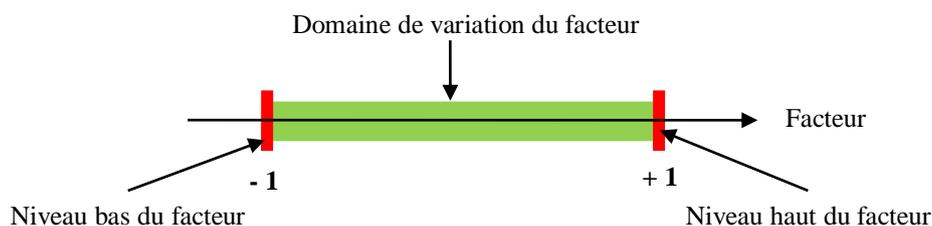


Figure 1.3- Niveau bas du facteur représenté par - 1 et le niveau haut par +1.

1.7.3- Effet

L'effet d'un facteur X sur une réponse Y se mesure en étudiant les variations de Y en fonction des variations de X . Si le facteur X a une interaction avec lui-même, on parle alors d'un effet quadratique. Un effet peut être positif, nul ou négatif, et est exprimé par le coefficient multiplicatif de X dans un modèle mathématique de la réponse. La signification de l'effet de X sur Y est donnée par des tests statistiques.

1.7.4- Interaction

On parle d'interaction entre deux facteurs X_1 et X_2 lorsque l'effet de X_1 dépend du niveau de l'autre facteur X_2 , et vice versa. Cette interaction est notée X_1X_2 ou X_2X_1 .

1.7.5- Espace expérimental

L'espace expérimental est représenté par deux facteurs. Les axes gradués et orientés des deux facteurs sont orthogonaux entre eux. L'espace à deux dimensions ainsi défini est un repère cartésien qui correspond à l'espace expérimental (voir figure 1.4).

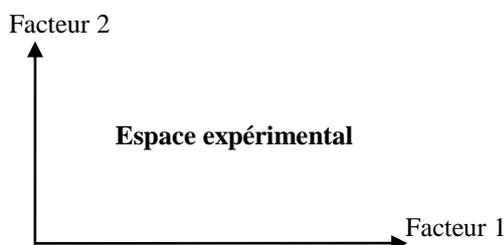


Figure 1.4- Représentation de chaque facteur par un axe gradué et orienté.

Une expérience donnée est représentée par un point dans ce système d'axes. Un plan d'expériences est représenté par un ensemble de points expérimentaux (voir figure 1.5).

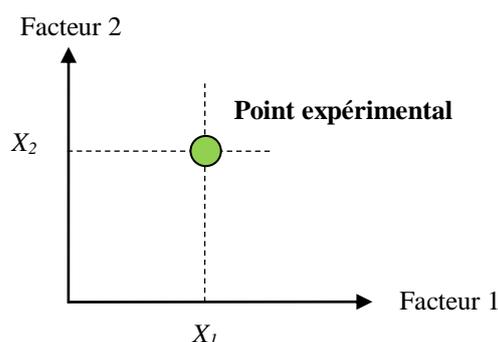


Figure 1.5- Niveaux des facteurs des points expérimentaux.

1.7.6- Domaine d'étude

Le domaine d'étude correspond à la zone de l'espace expérimental choisie par l'expérimentateur pour réaliser les essais. Une étude, qui comprend plusieurs expériences bien définies, est représentée par des points répartis dans le domaine d'étude (voir figure 1.6). Dans le cas d'un domaine d'étude à trois facteurs, un point expérimental (i) défini par l'expérimentateur est positionné dans l'espace.

1.7.7- Réponses

Une réponse est un résultat final mesurable ou observable obtenu à chaque expérience afin de déterminer l'impact des facteurs sur le système étudié. Ainsi, la réponse est une grandeur

d'intérêt mesurée à la fin de chaque essai. Pour garantir une meilleure fiabilité de l'étude, la réponse est caractérisée par la moyenne d'au moins trois mesures effectuées à différents endroits.

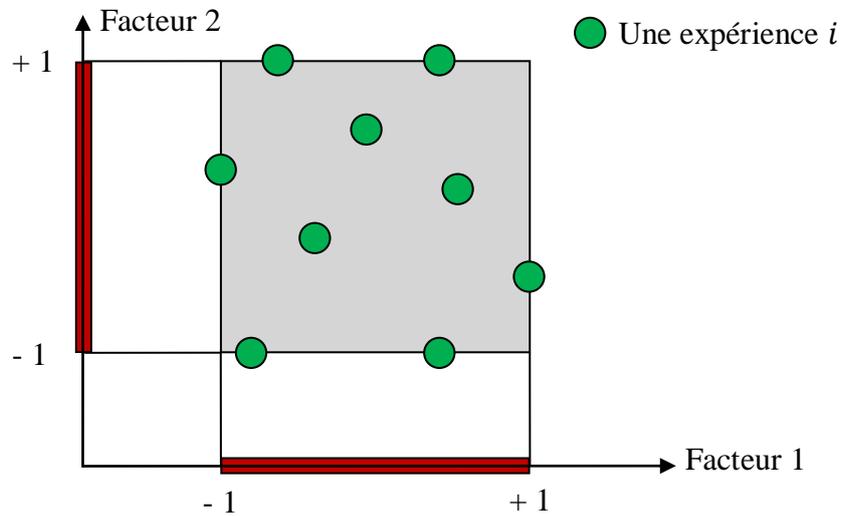


Figure 1.6- Positionnement des points expérimentaux dans le domaine d'étude défini par l'expérimentateur.

CHAPITRE 2 :

Quelques notions de statistiques

2.1- Introduction

La statistique est une discipline scientifique regroupant l'ensemble des méthodes mathématiques permettant de collecter, de traiter et d'analyser des données issues de l'observation de phénomènes aléatoires. Le hasard intervient dans ces phénomènes, dont l'issue est incertaine et méconnue de l'homme. L'analyse des données permet de décrire les phénomènes étudiés, d'extraire les informations contenues dans un ensemble de données, de faire des prévisions et de prendre des décisions éclairées.

La statistique utilise des notions de probabilités et de mathématiques dans le but de fournir une analyse et une interprétation des données permettant de prendre des décisions. Dans ce chapitre, nous nous intéressons aux méthodes statistiques de régression linéaire, les plus utilisées pour modéliser la relation mathématique entre une variable dépendante et des indépendantes (régression simple et régression multiple). Ces deux techniques présentent un grand intérêt pour prédire la valeur de la réponse (Y) en fonction des valeurs connues d'une ou plusieurs variables associées (X_i). Nous présentons les équations de régression, le coefficient de corrélation (R), le coefficient de détermination (R^2), ainsi que les intervalles de confiance et de prévision pour les deux méthodes de modélisation (simple et multiple).

2.2- Terminologie

- ✧ **Statistique descriptive** : La statistique descriptive regroupe un ensemble de méthodes de représentations graphiques et de calculs de caractéristiques numériques permettant de synthétiser les données statistiques. Les données examinées sont généralement issues d'un échantillon.
- ✧ **Univers** : L'univers statistique correspond à l'ensemble des objets étudiés dans une étude statistique.
- ✧ **Variable** : Une variable est une caractéristique qui permet d'étudier l'univers.

- ❖ **Population** : La population correspond à l'ensemble des mesures ou observations de la variable dans l'univers considéré.
- ❖ Une unité expérimentale est un objet de l'univers sur lequel la variable est mesurée.
- ❖ Un échantillon est un sous-ensemble de :
 - ✖ l'univers : s'il est composé d'unités expérimentales,
 - ✖ la population : s'il est composé de mesures de la variable.
- ❖ Un paramètre est une mesure caractéristique de la variable dans la population.
 - ✖ Par exemple : la moyenne de la population. En général, la vraie valeur d'un paramètre est inconnue.
- ❖ Une statistique est une mesure caractéristique de la variable dans un échantillon de la population.
 - ✖ Par exemple : le moyen échantillonnage, une statistique peut être calculée.

2.3- Mesures numériques

- ❖ **Mesures de tendance centrale** : moyenne, médiane, mode.

- **Moyenne** :

Soit $X = \{x_1, x_2, \dots, x_n\}$ un échantillon de n observations d'une population (valeurs numériques).

La moyenne de l'échantillon, ou moyenne échantillonnage est donnée par :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \dots \dots \dots (2.1)$$

Avec :

- ✖ n : nombre de mesure,
- ✖ x : la variable mesurée.

La moyenne d'un ensemble de données n'est pas nécessairement égale à l'une des valeurs de l'ensemble.

- **Médiane** :

La médiane de l'échantillon, notée \tilde{x} , est une valeur telle que 50 % des observations sont supérieures à cette valeur et 50 % sont inférieures à cette valeur.

Si $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ sont les données en ordre croissant, alors :

$$\tilde{x} = \begin{cases} x_{(\frac{n+1}{2})} & \text{si } n \text{ est impair} \\ \frac{1}{2} \left(x_{(\frac{n}{2})} + x_{(\frac{n+1}{2})} \right) & \text{si } n \text{ est pair.} \end{cases} \dots \dots \dots (2.2)$$

Si n est impair, alors la médiane est égale à l'une des données. Si n est pair, elle n'est pas forcément égale à l'une des données.

– **Mode :**

Le mode de l'échantillon est la valeur la plus fréquente parmi les données. Un échantillon peut avoir plusieurs modes. Toutefois, le mode est nécessairement égal à l'une des données. On peut également définir le mode comme étant le point milieu de la classe ayant le plus grand fréquent.

❖ **Mesures de dispersion :** variance, écart-type.

– **Variance :**

Soit x_1, x_2, \dots, x_n un échantillon de n observations d'une population (valeurs numériques).

La variance de l'échantillon, symbolisée par s^2 , est définie par :

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \dots \dots \dots (2.3)$$

– **Ecart-type :**

L'écart-type de l'échantillon, notée s ou σ , est définie par :

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \dots \dots \dots (2.4)$$

2.4- Diagrammes de distribution

Les diagrammes de distribution sont des représentations graphiques des données qui permettent d'avoir une vue d'ensemble et d'interpréter les données. Ainsi, avant toute étude statistique, il est préférable et intéressant de représenter les données par des diagrammes de distribution. Parmi ces diagrammes, on peut citer : l'histogramme, le diagramme à points (ou nuage de points), le diagramme en boîte, le diagramme de Pareto, ...etc. Le plus simple et le

plus utilisé est la représentation graphique en nuage de points. Le diagramme de dispersion, ou le diagramme du nuage de points, permet d'étudier la relation potentielle entre deux variables continues et de savoir si elles sont corrélées.

❖ **Diagramme à points, nuage de points :**

Lors de la création d'un diagramme de dispersion, la variable de réponse est généralement représentée sur l'axe des ordonnées y et la variable de prévision sur l'axe des abscisses x pour chaque observation, dans le cas d'une seule variable. Cependant, dans la pratique, il peut y avoir une ou plusieurs variables à prendre en compte. Il est alors possible d'utiliser des couleurs ou des symboles différents pour distinguer les différentes variables. Cela permet de mieux visualiser les relations entre les variables et de détecter d'éventuelles corrélations. Il est également possible de superposer plusieurs nuages de points pour comparer les relations entre différentes variables. En résumé, la création d'un diagramme de dispersion est une étape importante pour analyser les relations entre les variables et pour effectuer une analyse de corrélation.

- Diagramme à points : une seule variable :

Chaque observation est représentée par un point situé au-dessus de la valeur correspondante sur l'axe horizontal. Si plusieurs observations partagent la même valeur, les points correspondants sont superposés.

- Nuage de points : deux variables :

Les deux axes correspondent aux valeurs des deux variables. Chaque couple d'observations est représenté par un point dans le plan. Si plusieurs observations partagent les mêmes valeurs pour les deux variables, les points correspondants peuvent être superposés ou représentés par un autre symbole. Cette représentation graphique permet de visualiser les relations potentielles entre les variables.

- Nuage de points : trois variables

Ici, chaque point dans l'espace correspond à un triplet de valeurs observées.

Exemple :

Le tableau 2.1 mesure le poids Y et la taille X de 20 individus.

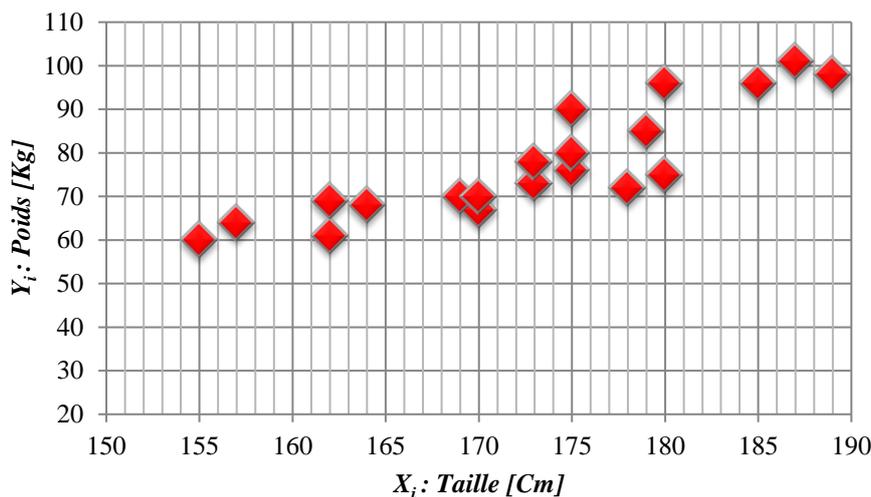
Tableau 2.1- Le poids Y et la taille X de 20 individus.

Individu	1	2	3	4	5	6	7	8	9	10
X_i	155	162	157	170	164	162	169	170	178	173
Y_i	60	61	64	67	68	69	70	70	72	73
Individu	11	12	13	14	15	16	17	18	19	20
X_i	180	175	173	175	179	175	180	185	189	187
Y_i	75	76	78	80	85	90	96	96	98	101

Dans ce cas, chaque couple est composé de deux valeurs numériques. Un couple de nombres (entiers ou réels) peut toujours être représenté comme un point dans un plan.

$$(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_n, y_n).$$

Le diagramme de distribution du poids Y et de la taille X de 20 individus est représenté par un nuage de points dans la figure 2.1.

**Figure 2.1-** Diagramme de dispersion du poids Y et de la taille X de 20 individus.

2.5- Régression linéaire

La régression linéaire est une méthode statistique simple qui permet de décrire et de générer une équation mathématique pour décrire la relation qui existe entre une variable de réponse (Y), dite dépendante et une ou plusieurs variables indépendantes X_i (X_1, X_2, \dots, X_n). L'objectif principal de la régression linéaire est de déterminer une fonction $Y = f(X_i)$ modélisant la relation entre une variable expliquée Y et une ou plusieurs variables explicatives X_i , afin de pouvoir ensuite faire des prévisions sur la variable endogène Y lorsque la variable exogène X est mesurée ou fixée. Cette méthode permet donc d'analyser et de quantifier la

relation entre les variables et de comprendre comment une ou plusieurs variables indépendantes peuvent influencer la variable dépendante.

2.5.1- Régression linéaire simple

La régression linéaire simple est un outil statistique flexible qui relie une variable critère Y à une seule variable prédictive X à l'aide d'une fonction linéaire. Le but est de décrire les données dans le nuage de points par une seule ligne droite, qui exprime la tendance linéaire dans les données recueillies et permet de prédire de nouvelles observations. Mathématiquement, la relation est représentée par l'équation suivante :

$$Y = a_0 + a_1X + \varepsilon \dots \dots \dots (2.5)$$

Où :

- ✧ Y : La valeur prédite de variable à expliquer.
- ✧ X : Est la variable explicative (la variable indépendante).
- ✧ Constant a_0 : L'ordonnée de l'origine : point d'intersection de la droite avec l'axe vertical en $x = 0$ (paramètre à estimer).
- ✧ Coefficient a_1 : La pente de la droite de régression (paramètre à estimer).
- ✧ ε : Est le terme d'erreur aléatoire du modèle, il résume les informations manquantes qui permettraient d'expliquer linéairement les valeurs de Y à l'aide des variables X_i .
- ✧ La désignation "linéaire" correspond au fait que le modèle (2.5) est linéaire en a_0 et a_1 .
- ✧ Le terme "simple" désigne le fait qu'il n'y a qu'une seule variable explicative X , utilisée pour expliquer Y .

2.5.2- Méthode des moindres carrés

La méthode des moindres carrés est une technique utilisée en statistiques pour estimer les paramètres d'un modèle mathématique de manière à minimiser la somme des carrés des écarts entre les valeurs observées et les valeurs prédites par le modèle. Voici une explication simple de cette méthode:

1. Définition du problème :

Supposons des données expérimentales présentées dans un graphique en nuage de points (voir figure 2.2). Chaque point est identifié par un numéro de 1 à n .

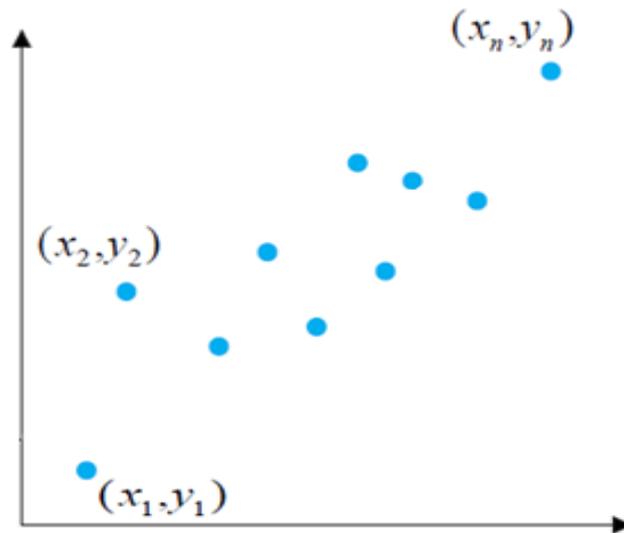


Figure 2.2- Graphique de nuage de points.

2. Choix du modèle :

Nous cherchons maintenant à trouver une équation mathématique (modèle) qui représente au mieux la relation entre les variables et qui passe le plus près possible de tous les points (nous supposons une droite dans ce cas (voir figure 2.3)).

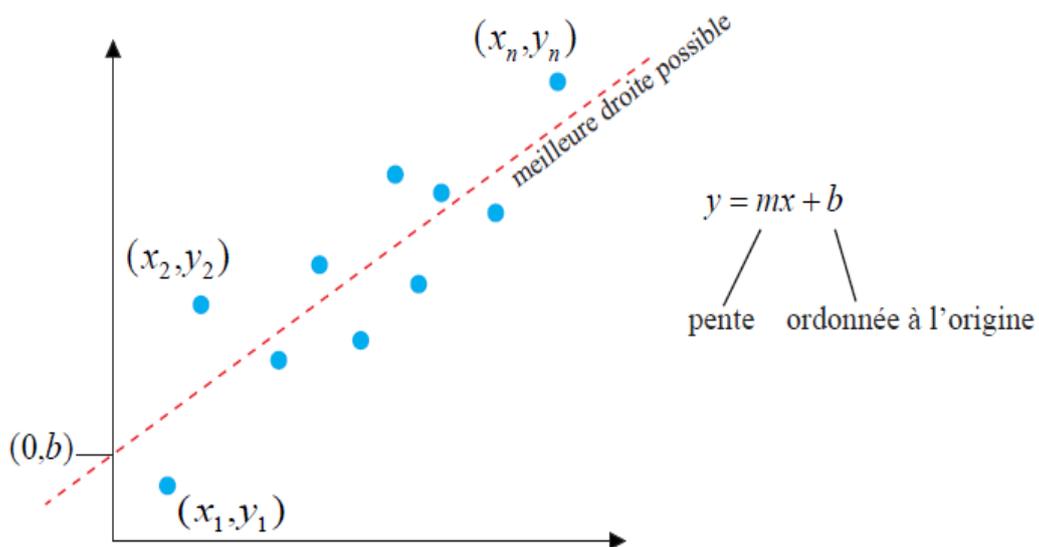


Figure 2.3- Graphique de nuage de points avec meilleure droite possible.

3. Calcul les erreurs :

Nous examinons l'erreur sur l'ordonnée de chaque point par rapport à la meilleure droite possible. La figure 2.4 illustre la méthode de mesure de ces erreurs par rapport à la droite.

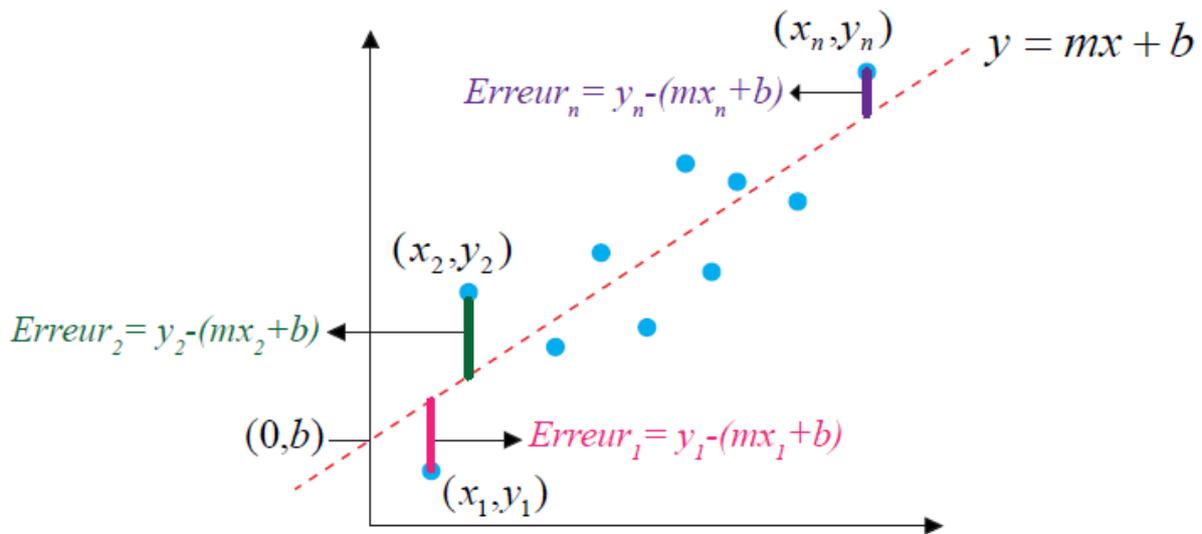


Figure 2.4- Erreurs par rapport à la droite idéale.

4. Carrés des erreurs :

Étant donné que certaines erreurs (ou résidus) sont positives et d'autres sont négatives, il est préférable d'utiliser les carrés des erreurs. Cela permet de rendre toutes les différences positives et de mettre davantage l'accent sur les erreurs importantes.

5. Somme des carrés des écarts :

Pour obtenir une mesure globale de l'erreur entre les valeurs observées et celles prédites par le modèle, nous additionnons tous les carrés des écarts. Cela se traduit par la somme des carrés des erreurs (SCE), que l'on exprime comme suit :

$$SCE = (y_1 - (mx_1 + b))^2 + (y_2 - (mx_2 + b))^2 + \dots + (y_n - (mx_n + b))^2 \dots \dots \dots (2.6)$$

Afin de trouver la meilleure droite, nous devons déterminer les valeurs de m et de b qui minimisent la somme des carrés des erreurs (SCE).

6. Minimisation de la somme des carrés des écarts :

L'objectif est de trouver les valeurs des paramètres du modèle qui minimisent cette somme. En d'autres termes, ajustez les paramètres de manière à rendre la somme des carrés des écarts aussi petite que possible.

D'abord, réécrivons cela de manière plus simple : nous allons développer les carrés des termes dans l'équation 2.6.

$$\begin{aligned}
 SCE &= y_1^2 - 2y_1(mx_1 + b) + (mx_1 + b)^2 \\
 &+ y_2^2 - 2y_2(mx_2 + b) + (mx_2 + b)^2 \\
 &+ \dots \dots \dots (2.7) \\
 &+ y_n^2 - 2y_n(mx_n + b) + (mx_n + b)^2
 \end{aligned}$$

Ensuite, nous développons les carrés des parenthèses restantes.

$$\begin{aligned}
 SCE &= y_1^2 - 2y_1mx_1 - 2y_1b + m^2x_1^2 + 2mx_1b + b^2 \\
 &+ y_2^2 - 2y_2mx_2 - 2y_2b + m^2x_2^2 + 2mx_2b + b^2 \\
 &+ \dots \dots \dots (2.8) \\
 &+ y_n^2 - 2y_nmx_n - 2y_nb + m^2x_n^2 + 2mx_nb + b^2
 \end{aligned}$$

Ensuite, nous regroupons les termes similaires.

$$\begin{aligned}
 SCE &= (y_1^2 + y_2^2 + \dots + y_n^2) - 2m(y_1x_1 + y_2x_2 + \dots + y_nx_n) - 2b(y_1 + y_2 + \dots + y_n) \\
 &+ m^2(x_1^2 + x_2^2 + \dots + x_n^2) + 2mb(x_1 + x_2 + \dots + x_n) + nb^2 \dots \dots \dots (2.9)
 \end{aligned}$$

Le premier terme de l'équation 2.9 peut être reformulé. En fait, la moyenne des carrés des "y" peut être exprimée de la manière suivante :

$$\overline{y^2} = \frac{y_1^2 + y_2^2 + \dots + y_n^2}{n} \Rightarrow n\overline{y^2} = y_1^2 + y_2^2 + \dots + y_n^2$$

Aussi, le deuxième terme de la SCE, la moyenne des produits des "x" par "y", peut être formulé ainsi :

$$\overline{xy} = \frac{y_1x_1 + y_2x_2 + \dots + y_nx_n}{n} \Rightarrow n\overline{xy} = y_1x_1 + y_2x_2 + \dots + y_nx_n$$

En suivant la même méthode pour les autres termes de la somme des carrés des erreurs (SCE), on peut écrire :

$$SCE = n\overline{y^2} - 2mn\overline{xy} - 2bn\overline{y} + m^2n\overline{x^2} + 2mbn\overline{x} + nb^2 \dots \dots \dots (2.10)$$

Maintenant, notre objectif est de minimiser la somme des carrés des erreurs (SCE). Pour ce faire, nous devons trouver les valeurs de "m" et "b" qui rendent la SCE minimale.

Remarque : Dans ce contexte, les valeurs de "x" et "y" sont connues. Les seules variables inconnues sont "m" et "b".

La figure 2.5 présente une droite avec une pente de 5 et une ordonnée à l'origine de 3, accompagnée d'une légère dispersion des points.

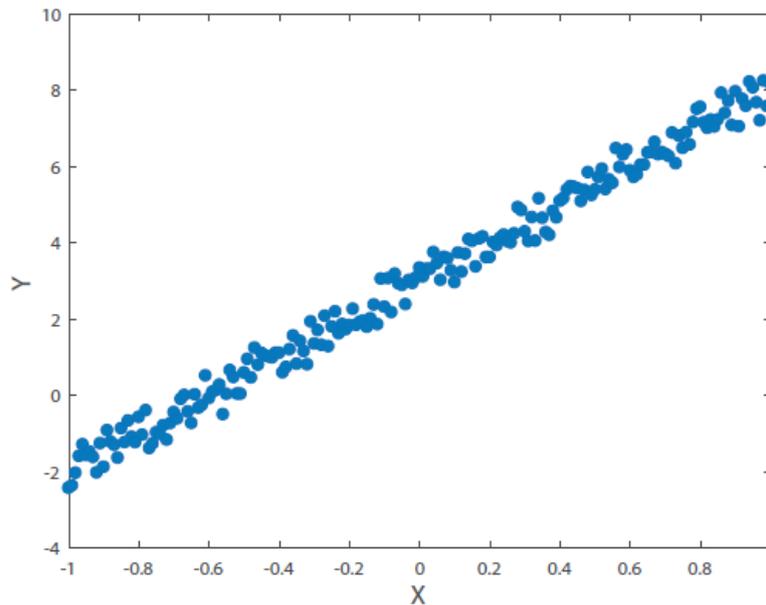


Figure 2.5- Droite de pente égale à 5 et d'ordonnée à l'origine de 3.

Maintenant, si nous traçons la valeur de la somme des carrés des erreurs (SCE) en fonction des paramètres "m" et "b", nous obtenons une surface de la figure 2.6 où le minimum est atteint lorsque "m" vaut 5 et "b" vaut 3.

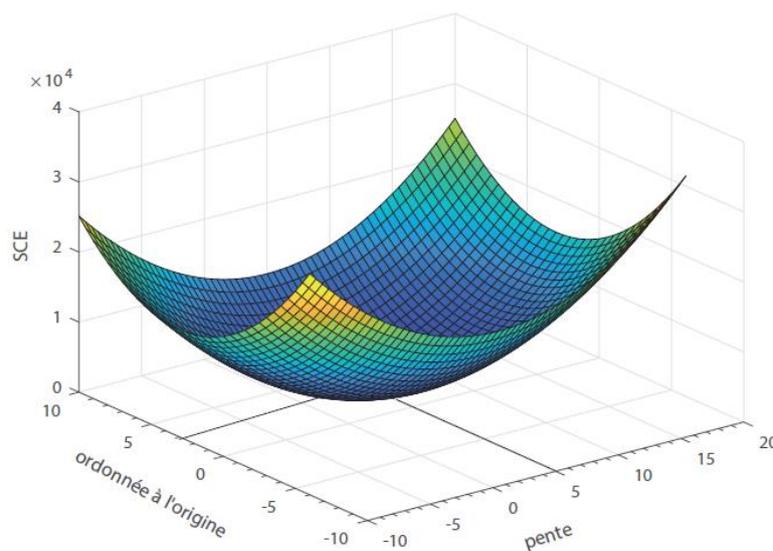


Figure 2.6- Valeur de la SCE en fonction des paramètres m et b .

Maintenant, la question est la suivante : comment trouver mathématiquement les valeurs optimales de "m" et "b" qui minimisent la somme des carrés des erreurs (SCE) ? La réponse réside dans la résolution des deux équations suivantes :

$$\begin{cases} \frac{\partial(SCE)}{\partial m} = 0 \\ \frac{\partial(SCE)}{\partial b} = 0 \end{cases} \dots \dots \dots (2.11)$$

Remarquez que la somme des carrés des erreurs (SCE) atteint un minimum le long de l'axe "m" et un autre le long de l'axe "b". Ainsi, nous exigeons que la dérivée partielle de la SCE par rapport à "m" soit nulle, tout comme celle par rapport à "b". La dérivée partielle signifie que lorsque nous dérivons une fonction de plusieurs variables (ici "m" et "b"), nous dérivons l'équation par rapport à une seule variable tout en maintenant les autres constantes. En résulte:

$$\frac{\partial(SCE)}{\partial m} = -2n\bar{xy} + 2mn\bar{x}^2 + 2bn\bar{x} = 0 \dots \dots \dots (2.12)$$

$$\frac{\partial(SCE)}{\partial b} = -2n\bar{y} + 2mn\bar{x} + 2nb = 0 \dots \dots \dots (2.13)$$

Notez qu'en divisant les deux équations précédentes par 2n, nous obtenons l'équations suivantes :

$$-\bar{xy} + m\bar{x}^2 + b\bar{x} = 0 \dots \dots \dots (2.14)$$

$$-\bar{y} + m\bar{x} + b = 0 \dots \dots \dots (2.15)$$

On peut résoudre directement ce système de deux équations à deux inconnues. Cependant, il est intéressant de les reformuler pour mieux comprendre leur signification physique. Réécrivons-les comme suit :

$$m\bar{x}^2 + b\bar{x} = \bar{xy} \dots \dots \dots (2.16)$$

$$m\bar{x} + b = \bar{y} \dots \dots \dots (2.17)$$

On rappelle maintenant notre objectif : déterminer les paramètres de l'équation $y = mx + b$ qui représente au mieux les points donnés. D'après l'équation 2.17, il est évident que le point $G(\bar{x}, \bar{y})$ situé à $x = \bar{x}$ et $y = \bar{y}$ appartient à cette droite d'ajustement, aussi appelée droite des moindres carrés (voir figure 2.7).

→ $G(\bar{x}, \bar{y}) \in$ droite des moindres carrés

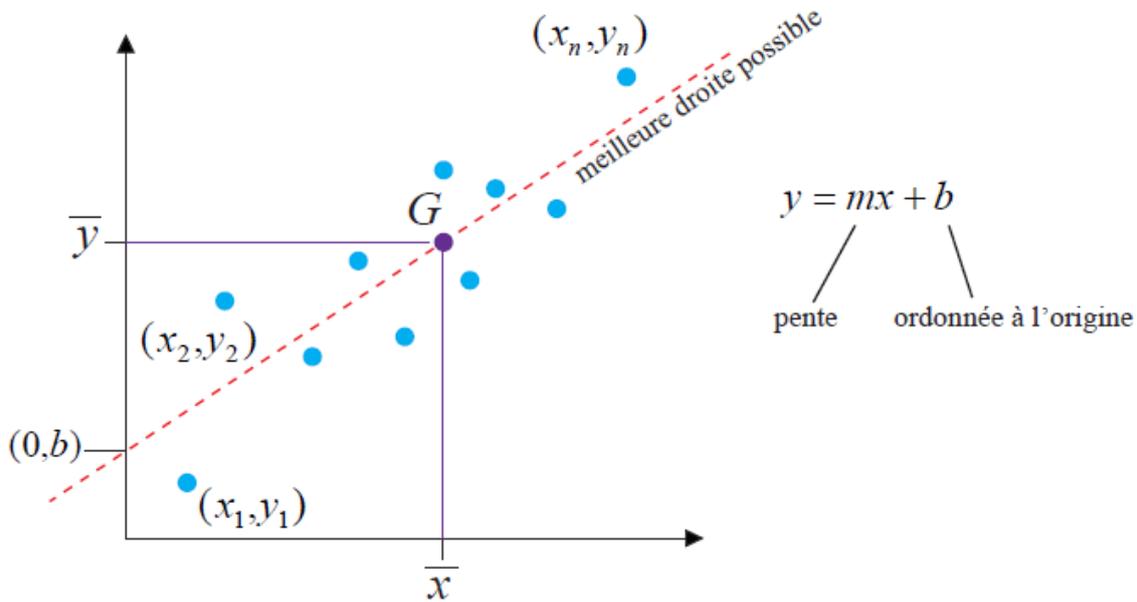


Figure 2.7- Point G fait partie de la droite d'ajustement.

Trouvons maintenant un autre point de la droite en utilisant l'équation 2.16. Pour commencer, réécrivons-la sous la forme :

$$m \frac{\overline{x^2}}{\bar{x}} + b = \frac{\overline{xy}}{\bar{x}} \dots \dots \dots (2.18)$$

Ce qui signifie que le point $H(\frac{\overline{x^2}}{\bar{x}}, \frac{\overline{xy}}{\bar{x}})$ fait également partie de la droite des moindres carrés. Reprenons maintenant les équations 2.18 et 2.17 et apportons quelques modifications. La première transformation consiste à soustraire l'équation 2.17 de l'équation 2.18.

$$\begin{cases} m \frac{\overline{x^2}}{\bar{x}} + b = \frac{\overline{xy}}{\bar{x}} \\ m\bar{x} + b = \bar{y} \end{cases} \Rightarrow m \left(\frac{\overline{x^2}}{\bar{x}} - \bar{x} \right) = \frac{\overline{xy}}{\bar{x}} - \bar{y}$$

Nous avons donc obtenu une nouvelle équation qui nous permettra de déterminer "m" et "b". Prenons cette équation et multiplions-la par " \bar{x}/\bar{x} ". Nous obtenons :

$$m = \frac{\frac{\overline{xy}}{\bar{x}} - \bar{y}}{\frac{\overline{x^2}}{\bar{x}} - \bar{x}} \times \frac{\bar{x}}{\bar{x}} = \frac{\overline{xy} - \bar{y}\bar{x}}{x^2 - (\bar{x})^2} \dots \dots \dots (2.19)$$

et bien sûr, nous pouvons facilement trouver "b" en utilisant la relation :

$$b = \bar{y} - m\bar{x} \dots \dots \dots (2.20)$$

Reprenons maintenant ces résultats en utilisant des indicateurs statistiques tels que la variance et la covariance. La variance des "x" est définie comme suit :

$$\sigma_x^2 = \frac{\sum_{i=1}^n x_i^2}{n} - (\bar{x})^2 = \overline{x^2} - (\bar{x})^2 \dots \dots \dots (2.21)$$

et la covariance entre "x" et "y" est définie par :

$$\text{cov}(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n} \dots \dots \dots (2.22)$$

Et il est possible de démontrer que :

$$\text{cov}(x, y) = \overline{xy} - \bar{x}\bar{y} \dots \dots \dots (2.23)$$

Avec ces définitions, nous pouvons alors écrire :

$$m = \frac{\text{cov}(x, y)}{\sigma_x^2} \quad \text{et} \quad b = \bar{y} - m\bar{x} \dots \dots \dots (2.24)$$

7. Méthode mathématique pour la minimisation:

Des techniques mathématiques, comme les dérivées partielles, peuvent être utilisées pour trouver les valeurs des paramètres qui minimisent la somme des carrés des écarts de manière efficace.

En résumé, la méthode des moindres carrés est une approche qui cherche à ajuster un modèle mathématique aux données en minimisant la somme des carrés des écarts entre les valeurs observées et les valeurs prédites par le modèle. Cela permet d'estimer les paramètres du modèle de manière à rendre les prédictions aussi proches que possible des données réelles.

Exemple d'application :

Donnons maintenant un exemple simple avec un nuage de points ne contenant que trois données. Soient les points (1, 2), (2, 1) et (4, 3) du graphique 2.8.

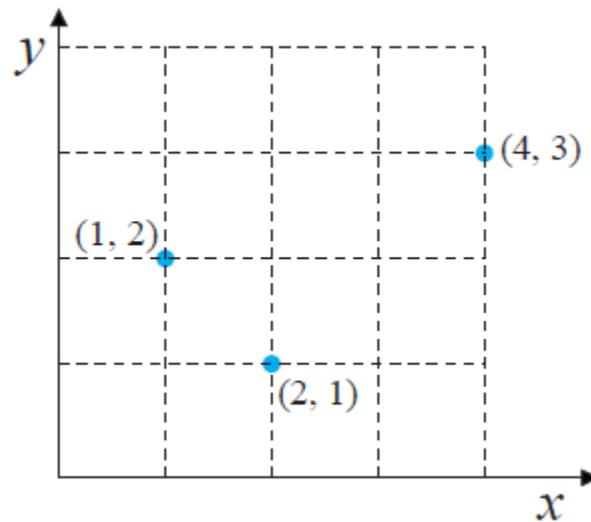


Figure 2.8- Exemple de nuage de points.

Pour déterminer " m " et " b ", nous allons utiliser les relations suivantes :

$$m = \frac{\overline{xy} - \bar{y}\bar{x}}{\overline{x^2} - (\bar{x})^2} \quad \text{et} \quad b = \bar{y} - m\bar{x}$$

Procédons aux calculs des quantités nécessaires

$$\bar{x} = \frac{1 + 2 + 4}{3} = \frac{7}{3}$$

$$\bar{y} = \frac{2 + 1 + 3}{3} = 2$$

$$\overline{xy} = \frac{1 \times 2 + 2 \times 1 + 4 \times 3}{3} = \frac{16}{3}$$

$$\overline{x^2} = \frac{1^2 + 2^2 + 4^2}{3} = \frac{21}{3} = 7$$

Maintenant on peut calculer m et b :

$$m = \frac{\frac{16}{3} - \frac{7}{3} \times 2}{7 - \left(\frac{7}{3}\right)^2} = \frac{\frac{16}{3} - \frac{14}{3}}{7 - \frac{49}{9}} = \frac{\frac{2}{3}}{\frac{63 - 49}{9}} = \frac{2}{3} \times \frac{9}{14} = \frac{3}{7}$$

$$b = 2 - \frac{3}{7} \times \frac{7}{3} = 2 - 1 = 1$$

Le graphique suivant montre le résultat obtenu.

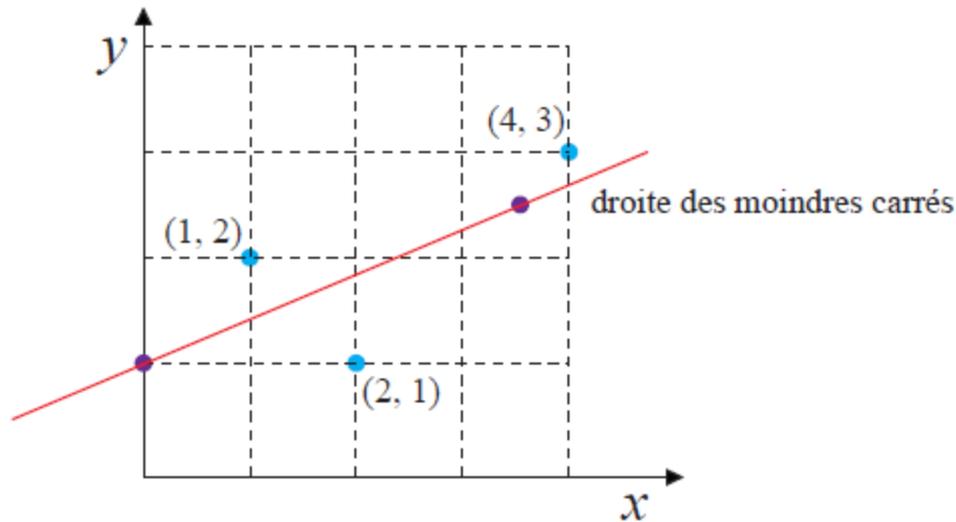


Figure 2.9- Exemple de régression linéaire.

2.5.3- Coefficient de corrélation (r)

Le coefficient de corrélation (CC) de Pearson, noté r , est une mesure statistique qui permet de quantifier l'intensité et le sens de la relation linéaire entre deux variables. En général, sa valeur se situe entre -1 et $+1$, ce qui signifie que $-1 \leq r \leq +1$. Donc, ce coefficient est caractérisé à la fois par sa valeur et par son signe.

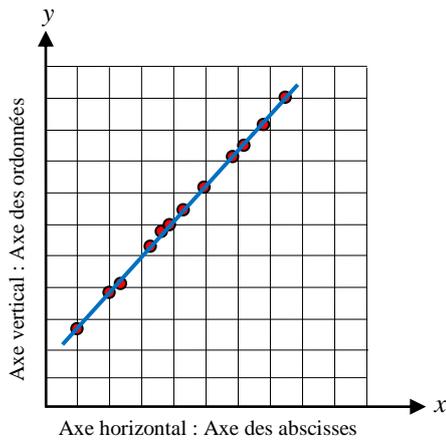
Le signe indique le sens de la corrélation (positive ou négative) (voir figure 2.10).

- ❏ Une valeur dans l'intervalle $[-1, 0[$ → une corrélation négative :
- ❏ Les variables x et y varient en sens opposé ($x \nearrow \Leftrightarrow y \searrow, x \searrow \Leftrightarrow y \nearrow$).
- ❏ Une valeur de 0 → pas de corrélation,
- ❏ Une valeur dans l'intervalle $]0, 1]$ → une corrélation positive :
- ❏ Les variables x et y varient dans le même sens ($x \nearrow \Leftrightarrow y \nearrow, x \searrow \Leftrightarrow y \searrow$).

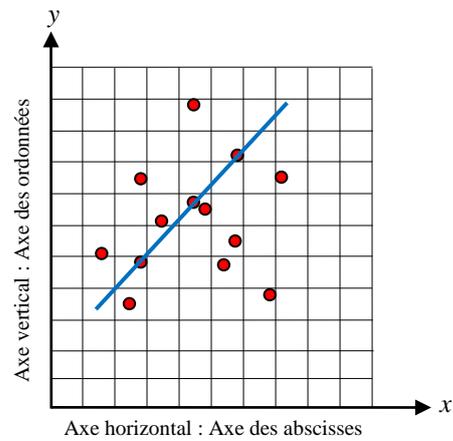
La valeur indique la force (voir figure 2.10):

- ✖ Plus la valeur absolue du coefficient de corrélation ($|r|$) tend vers 1 , plus la corrélation entre les variables est forte. Lorsque $|r|$ est égal à 1 , cela indique une corrélation parfaite entre les x_i et les y_i , ce qui signifie que les points du diagramme de dispersion sont tous alignés sur une même droite.
- ✖ Plus la valeur absolue du coefficient de corrélation ($|r|$) tend vers 0 , plus la corrélation entre les variables est faible. Lorsque le coefficient de corrélation r est égal à 0 , cela

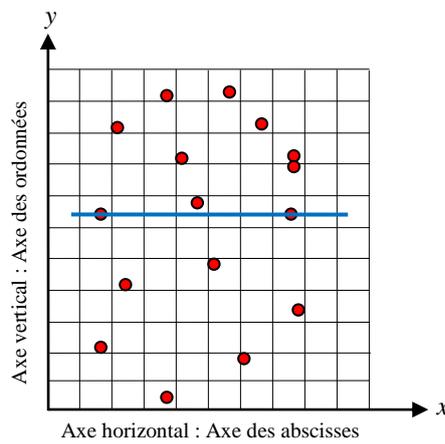
indique l'absence de corrélation linéaire entre les x_i et les y_i . Dans ce cas, les points du diagramme de dispersion sont distribués aléatoirement dans le plan.



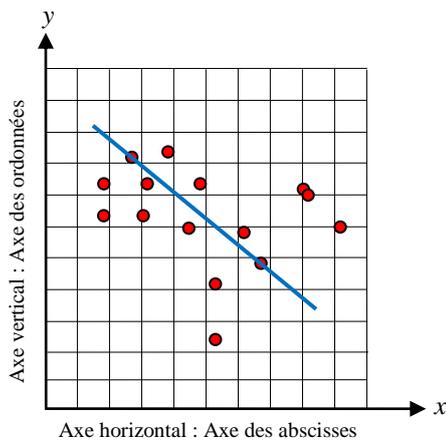
Ici, $r = 1$: corrélation positive parfaite entre les deux variables.



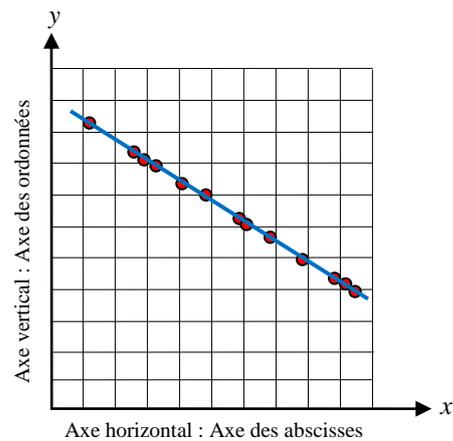
Ici, $r = 0,5$: corrélation positive faible entre les deux variables.



Ici, $r = 0$: absence totale de corrélation, les deux variables sont linéairement indépendantes.



Ici, $r = -0,5$: corrélation négative faible entre les deux variables.



Ici, $r = -1$: corrélation négative parfaite entre les deux variables.

Figure 2.10- Différents cas de corrélation.

Calcul :

Soit n observations de deux variables quantitatives (x_i, y_i) avec $i = 1, 2, \dots, n$. Le coefficient de corrélation de Pearson est une mesure qui permet de quantifier la force et la direction de la relation linéaire entre x et y . Pour calculer ce coefficient, on utilise la formule suivante :

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \dots \dots \dots (2.25)$$

Où :

- ✧ $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$,
- ✧ $S_{yy} = \sum_{i=1}^n (y_i - \bar{y})^2$,
- ✧ $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$.

2.5.4- Coefficient de détermination (R^2)

Le coefficient de détermination (CD) est une mesure statistique qui décrit la proportion de la variance de la réponse (y) qui est expliquée par une ou plusieurs variables d'entrée (x_i) dans un modèle de régression linéaire. Le CD est souvent noté R^2 , car c'est le carré du coefficient de corrélation de Pearson. Le R^2 est toujours compris entre 0 et 1, ou exprimé en pourcentage, entre 0 et 100 % ($CD = CC^2 \times 100$).

Nous allons maintenant expliquer comment est calculé le coefficient de détermination, souvent désigné par R^2 , qui permet de mesurer la qualité de l'ajustement d'une régression linéaire. Cependant, il est important de vérifier visuellement que le nuage de points suit effectivement une tendance linéaire pour interpréter correctement ce coefficient. Imaginons un nuage de points pour lequel nous avons trouvé la droite des moindres carrés $y = mx + b$ (voir figure 2.11).

Les symboles (e_1, e_2, \dots, e_n) représentent les erreurs par rapport à la droite. Les coordonnées \hat{y} (y chapeau) représentent la coordonnée "y" d'un point (x, y) situé sur la droite ayant la même coordonnée "x" que le point expérimental. Nous avons déjà établi qu'il était nécessaire de calculer la somme des carrés des erreurs par rapport à la droite D .

$$SCE_D = e_1^2 + e_2^2 + \dots + e_n^2$$

$$SCE_D = (y_1 - (mx_1 + b))^2 + (y_2 - (mx_2 + b))^2 + \dots + (y_n - (mx_n + b))^2$$

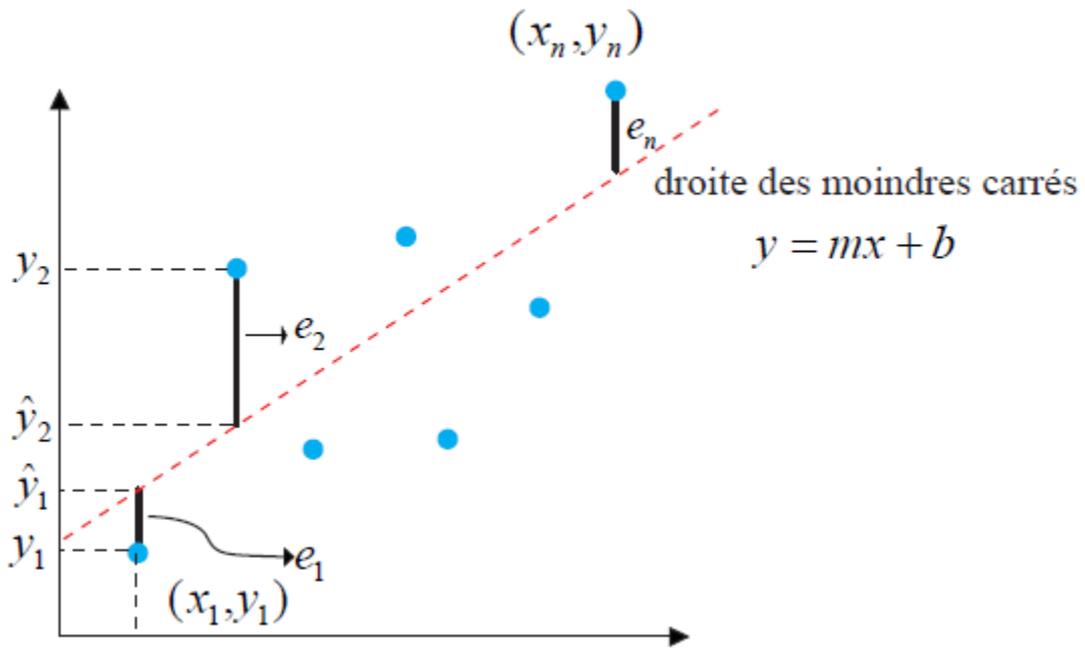


Figure 2.11- Nuage de points où l'on montre les écarts par rapport à la droite D.

En utilisant cette somme, nous pouvons alors déterminer les valeurs de "m" et de "b". À présent, intéressons-nous à une autre mesure : la somme des carrés des erreurs par rapport à la moyenne des "y" (figure 2.12).

$SCE_{\bar{y}} = (y_1 - \bar{y})^2 + (y_2 - \bar{y})^2 + \dots + (y_n - \bar{y})^2$ somme des carrés des erreurs par rapport à la moyenne.

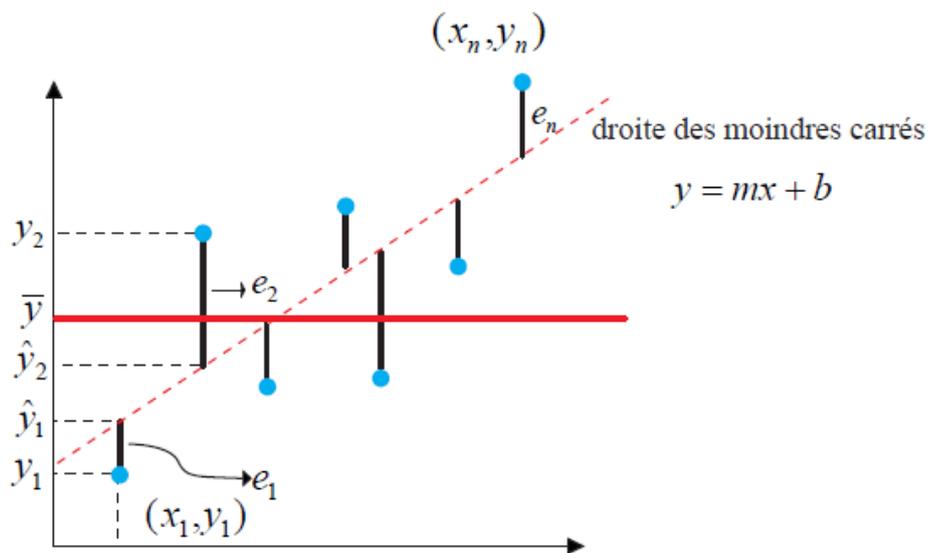


Figure 2.12- Nuage de points avec la moyenne des y.

La figure 2.13 illustre ces écarts par rapport à la moyenne, représentés par des barres vertes.

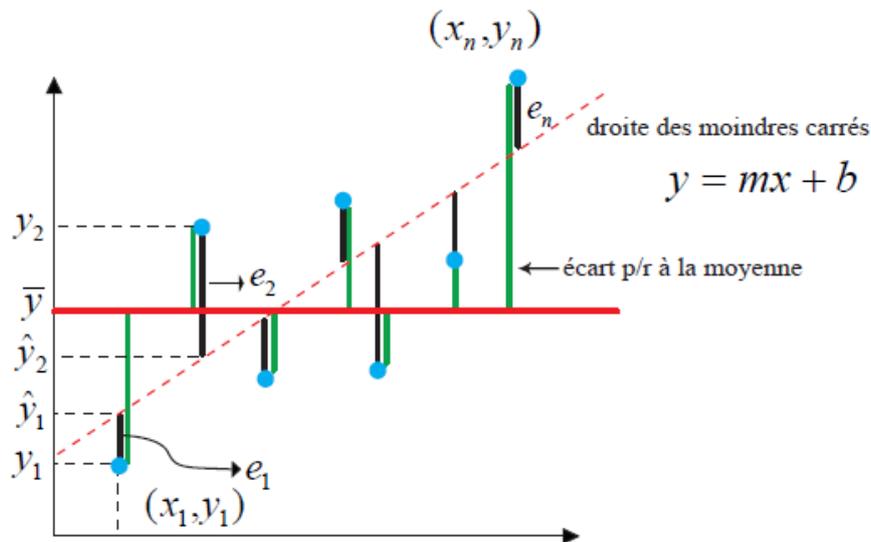


Figure 2.13- Nuage de points où l'on montre les écarts p/r à la moyenne \bar{y} .

Maintenant, la question est la suivante : quelle proportion de la somme des carrés des erreurs par rapport à la moyenne $SCE_{\bar{y}}$ est expliquée par la droite D ? Il est plus facile d'aborder cette question en disant que :

SCE_D : partie de la $SCE_{\bar{y}}$ qui n'est pas expliquée par la droite D

Nous pouvons alors écrire que

$\frac{SCE_D}{SCE_{\bar{y}}}$: proportion de la $SCE_{\bar{y}}$ qui n'est pas expliquée par la droite D

La proportion de la $SCE_{\bar{y}}$ qui est expliquée par D , notée R^2 , et appelée le coefficient de détermination, est donnée par :

$$R^2 = 1 - \frac{SCE_D}{SCE_{\bar{y}}}$$

2.5.5- Régression linéaire multiple

La régression linéaire multiple est une méthode statistique qui permet de modéliser la relation mathématique entre une variable réponse (Y) et deux variables d'entrée ou plus (X_1, X_2, \dots, X_n). Contrairement à la régression linéaire simple, qui ne prend en compte qu'une seule

variable d'entrée, la régression linéaire multiple permet d'explorer l'influence de plusieurs variables d'entrée sur la variable réponse.

Mathématiquement, la relation linéaire entre Y et les variables d'entrée (X_1, X_2, \dots, X_n) peut être représentée par l'équation suivante :

$$Y = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \dots + \beta_p X_{i,p} + \varepsilon \dots \dots \dots (2.4)$$

Où :

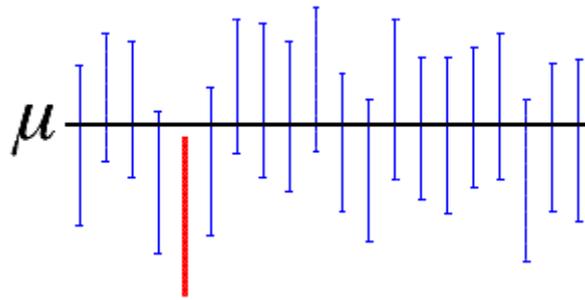
- ❖ $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ sont les coefficients de régression,
- ❖ ε : est le terme d'erreur qui capture la variation non expliquée par le modèle,
- ❖ $i = 1, 2, \dots, n$
- ❖ $n > p + 1$

Les coefficients β peuvent être interprétés comme les effets de chaque variable d'entrée sur la variable réponse, en tenant compte de l'influence des autres variables d'entrée.

La régression linéaire multiple peut être utilisée pour prédire la valeur de la variable réponse à partir des valeurs des variables d'entrée, et pour évaluer l'importance relative de chaque variable d'entrée dans l'explication de la variance de la variable réponse. Les coefficients de régression peuvent être estimés à l'aide de diverses méthodes, notamment la méthode des moindres carrés ordinaires, qui minimise la somme des carrés des écarts entre les valeurs prédites et les valeurs observées de la variable réponse.

2.5.5.1- Intervalle de confiance

Un intervalle de confiance est une plage de valeurs calculée à partir d'échantillons statistiques et ayant de grandes chances de contenir la valeur d'un paramètre inconnu de population. Étant donné que les échantillons sont sélectionnés de manière aléatoire, il est peu probable que deux échantillons différents d'une même population produisent des intervalles de confiance identiques. Cependant, si vous répétez l'échantillonnage à plusieurs reprises, un certain pourcentage des intervalles de confiance obtenus contiendront le paramètre de population inconnu. Ainsi, l'intervalle de confiance fournit une estimation fiable de la valeur du paramètre inconnu, avec un certain niveau de confiance, ce qui est crucial en statistiques pour prendre des décisions éclairées.



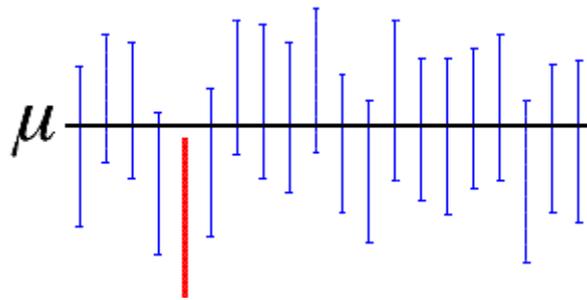
Sur ce graphique, la ligne noire horizontale représente la valeur fixe de la moyenne d'une population inconnue μ . Les intervalles de confiance bleus verticaux qui chevauchent la ligne horizontale incluent la valeur de la moyenne de population. Ce n'est pas le cas de l'intervalle de confiance rouge qui se situe complètement en dessous de la ligne horizontale. Lorsqu'on parle d'un intervalle de confiance à 95 %, cela signifie que si on répétait l'échantillonnage de la même population à plusieurs reprises, environ 19 échantillons sur 20 (soit 95 %) génèreraient des intervalles de confiance qui contiennent le paramètre de population. Cela fournit donc une indication de la précision de l'estimation de la moyenne de la population inconnue μ à partir de l'échantillon donné.

L'intervalle de confiance est un outil important pour évaluer l'estimation d'un paramètre de population. Par exemple, un fabricant peut utiliser l'intervalle de confiance pour déterminer si la longueur moyenne des crayons qu'il produit diffère de la longueur cible. Pour cela, il peut prélever un échantillon aléatoire de crayons et déterminer la longueur moyenne pour cet échantillon, qui est de 52 mm. En calculant l'intervalle de confiance à 95 % pour cette estimation, il obtient (50,54). Par conséquent, il peut être sûr à 95 % que la longueur moyenne de tous les crayons produits est comprise entre 50 et 54 mm. Cela peut aider le fabricant à prendre des décisions éclairées concernant la production et la qualité de ses produits. Il est important de noter que l'intervalle de confiance est calculé à partir de l'échantillon donné et ne fournit qu'une estimation de la valeur du paramètre inconnu pour l'ensemble de la population.

2.5.5.2- Niveau de confiance

Le niveau de confiance représente le pourcentage d'intervalles de confiance qui incluraient le paramètre de population si vous répétiez l'échantillonnage de la même population à plusieurs reprises. Un niveau de confiance de 95 % est généralement considéré comme efficace. Cela signifie que si vous collectez une centaine d'échantillons et que vous générez une centaine d'intervalles de confiance à 95 %, vous pouvez vous attendre à ce que, en moyenne, environ

95 intervalles contiennent le paramètre de population, tel que la moyenne de la population, comme l'illustre la figure suivante. Toutefois, il est important de noter que cela ne garantit pas que chaque intervalle de confiance individuel contiendra le paramètre de population, mais plutôt que sur le long terme, environ 95 % des intervalles de confiance générés incluront le paramètre de population. Le niveau de confiance est donc une mesure de la fiabilité de l'estimation de la valeur du paramètre de population à partir de l'échantillon donné.



Sur ce graphique, la ligne noire horizontale représente la valeur fixe de la moyenne de population inconnue μ . Les intervalles de confiance bleus verticaux qui chevauchent la ligne horizontale comprennent la valeur de la moyenne de population, ce qui indique que ces intervalles sont cohérents avec la vraie valeur de la moyenne de la population. À l'inverse, l'intervalle de confiance rouge situé complètement en dessous de la ligne horizontale ne contient pas la valeur de la moyenne de population et est donc peu probable de correspondre à la vraie valeur du paramètre.

Un niveau de confiance de 95 % signifie qu'environ 19 échantillons sur 20 (soit 95 %) de la même population généreront des intervalles de confiance qui contiennent le paramètre de population, tel que la moyenne de la population. Autrement dit, si vous répétez l'échantillonnage de la même population plusieurs fois et que vous générez des intervalles de confiance à 95 %, vous pouvez vous attendre à ce que la plupart d'entre eux incluent le paramètre de population. Cependant, cela ne garantit pas que chaque intervalle de confiance individuel contiendra le paramètre de population, mais plutôt que sur le long terme, environ 95 % des intervalles de confiance générés incluront le paramètre de population. Le niveau de confiance est donc une mesure de la fiabilité de l'estimation de la valeur du paramètre de population à partir de l'échantillon donné.

CHAPITRE 3 :

Plans optimaux

3.1- Introduction

Dans le contexte des expériences, les "plans optimaux" peuvent se référer à des plans d'expériences (Design of Experiments (DoE)) qui sont soigneusement conçus pour obtenir des informations maximales avec un nombre minimal d'essais. Les plans d'expériences visent à déterminer comment les facteurs influent sur une réponse donnée, tout en optimisant les ressources disponibles.

Les plans d'expériences peuvent être utilisés dans divers domaines tels que la recherche scientifique, l'ingénierie, la fabrication mécanique, la qualité des produits et d'autres disciplines. L'objectif est généralement de comprendre comment les différents paramètres d'entrée (facteurs) affectent le résultat (réponse) d'une expérience, et comment optimiser ces facteurs pour atteindre les meilleurs résultats possibles.

Les plans factoriels sont largement utilisés dans les expériences impliquant plusieurs facteurs, lorsqu'il est nécessaire d'étudier l'effet conjoint des facteurs sur une réponse. L'utilisation de plans d'expériences permet de réduire la variabilité, d'identifier les facteurs clés et d'optimiser les conditions expérimentales.

En résumé, les "plans optimaux" des expériences font souvent référence à des plans d'expériences bien structurés qui maximisent l'information obtenue à partir des essais réalisés, tout en minimisant le nombre d'essais nécessaires. L'objectif est de tirer des conclusions fiables avec un effort expérimental minimal.

3.2- Concepts des plans expérimentaux

Les techniques fondamentales de la conception expérimentale incluent la réplication, le blocage et la randomisation. La réplication et le blocage visent à accroître la précision de l'expérience, tandis que la randomisation est utilisée pour minimiser les biais. Les points suivants résument brièvement ces techniques.

3.2.1- Réplication

La réplication consiste à reproduire les conditions expérimentales pour estimer précisément les effets d'intérêt et évaluer la variabilité associée. Il est crucial de faire la distinction entre la "réplication" et les "mesures répétées". Par exemple, si quatre pièces sont assignées à différentes conditions de coupe avec une mesure prise sur chaque pièce, cela constitue une "réplication". En revanche, si une pièce est assignée à une condition de coupe et mesure quatre fois, les mesures ne sont pas indépendantes, ce sont des "mesures répétées". Il est essentiel de connaître la variation dans les réponses des pièces indépendantes aux conditions de coupe pour généraliser les conclusions de manière pertinente pour toutes les pièces similaires.

3.2.2- Blocage

Une expérience conçue consiste à appliquer des traitements à différentes unités expérimentales afin d'évaluer leur impact sur une réponse spécifique. Les unités expérimentales, qui peuvent être des pièces, des matériaux, des conditions de coupe ou de traitement, des moments dans le temps, ou une combinaison de ces éléments, varieront et influenceront la réponse. Cette variation peut être intentionnelle, car les conditions expérimentales devraient refléter celles auxquelles les conclusions seront appliquées. Afin d'assurer des inférences étendues, les conditions expérimentales doivent être diversifiées de manière appropriée. Le blocage est une technique souvent utilisée pour contrôler et ajuster une partie de la variation au sein des unités expérimentales, permettant ainsi d'améliorer la validité des conclusions.

Bloquer une expérience implique de diviser ou partitionner les unités expérimentales en groupes appelés blocs, de manière à ce que les unités expérimentales au sein de chaque bloc soient délibérément similaires. Ainsi, les traitements assignés aux unités expérimentales dans un même bloc peuvent être comparés dans des conditions expérimentales relativement homogènes. Une mise en œuvre réussie du blocage permet des comparaisons plus précises entre deux traitements ou plus dans l'expérience par rapport à une conception non bloquée. Par exemple, dans une expérience visant à comparer les effets de paramètres de coupe pour un état de surface, les deux traitements peuvent être comparés de manière plus précise sur les pièces du même matériau que sur les pièces de deux matériaux différents. Idéalement, l'une de ces situations peut être reproduite en utilisant des pièces échantillonnées de manière aléatoire à partir de la population d'intérêt, ou du moins représentatifs de celle-ci.

3.2.3- Randomisation

La randomisation vise à éviter les biais dans une expérience. En assignant de manière aléatoire le matériel expérimental ou les conditions de traitement avant le début de l'expérience, on s'assure que les observations ne sont pas influencées de manière systématique par des facteurs inconnus. Ainsi, les observations sont choisies de manière aléatoire plutôt que selon un schéma prédéfini, ce qui renforce la fiabilité des résultats. La randomisation des essais dans un plan d'expériences vise à réduire l'impact de facteurs non contrôlés en répartissant aléatoirement les essais. Cela contribue à rendre les conclusions plus robustes et généralisables.

Les autres concepts associés aux plans d'expériences sont :

1. **Plan Factoriel** : Dans un plan factoriel, les niveaux de plusieurs facteurs sont combinés de manière systématique pour examiner les effets principaux de chaque facteur ainsi que les interactions entre les facteurs.
2. **Surface de Réponse** : La surface de réponse est une représentation graphique tridimensionnelle qui montre comment la réponse change en fonction de deux facteurs tout en maintenant les autres facteurs à des niveaux spécifiques. Cela aide à identifier les conditions optimales.
3. **Optimisation** : L'optimisation dans le contexte des plans d'expériences implique la recherche des conditions expérimentales qui maximisent ou minimisent la réponse, en utilisant des méthodes statistiques ou mathématiques.
4. **Analyse de Variance (ANOVA)** : L'ANOVA est une technique statistique couramment utilisée pour analyser les résultats des plans d'expériences. Elle permet de déterminer quels facteurs ont des effets significatifs sur la réponse.

Les plans optimaux sont une approche puissante pour concevoir des essais de manière efficace, extraire des informations utiles et prendre des décisions éclairées en minimisant les ressources nécessaires. Ils sont largement utilisés dans la recherche scientifique, l'industrie et d'autres domaines pour optimiser les processus et améliorer les performances.

3.3- Plan optimal de premier ordre

Les plans factoriels sont couramment utilisés dans les expériences impliquant plusieurs facteurs pour étudier l'effet combiné de ceux-ci sur une réponse. Certains types spécifiques de ces plans sont particulièrement importants, non seulement parce qu'ils sont largement

appliqués en recherche, mais aussi parce qu'ils servent de base à d'autres plans ayant une grande valeur pratique.

Le plus important de ces cas particuliers concerne k facteurs, chacun évalué à seulement deux niveaux. Ces niveaux peuvent être quantitatifs, comme deux valeurs distinctes de température de coupe, de rugosité de surface ou de l'usure d'outil ; ou qualitatifs, tels que deux machines différentes, deux opérateurs, des niveaux élevés et bas d'un facteur, ou la présence et l'absence d'un facteur. Une réplication complète de ce plan nécessite $2 \times 2 \times \dots \times 2 = 2^k$ observations et est connue sous le nom de plan factoriel 2^k .

Le plan 2^k est particulièrement utile dans les premières étapes du travail expérimental, lorsque de nombreux facteurs sont susceptibles d'être étudiés. Il offre le plus petit nombre d'essais permettant d'examiner k facteurs dans un plan factoriel complet. Par conséquent, ces plans sont largement utilisés dans les expériences de criblage de facteurs.

3.3.1- Plan optimal à deux niveaux de type 2^k

Les plans factoriels complets à deux niveaux sont simples et essentiels, constituant ainsi le point de départ fondamental de toute étude. Les résultats initiaux qu'ils produisent peuvent toujours être enrichis par de nouvelles expériences, permettant ainsi d'atteindre le niveau de précision et d'information recherché.

Le premier plan de la série 2^k comprend seulement deux facteurs, disons A et B, chacun testé à deux niveaux. Ce plan est appelé plan factoriel 2^2 . Les niveaux des facteurs peuvent être arbitrairement nommés «bas -1» et «haut +1». Dans ce plan, k représente le nombre de facteurs examinés, tandis que 2 indique le nombre de niveaux par facteur. Chaque facteur peut ainsi être réglé à deux niveaux, généralement désignés comme -1 et +1, représentant respectivement des réglages bas et haut pour chaque facteur. Pour un plan factoriel à deux facteurs, l'espace d'étude est représenté sous forme de carré. Par exemple, dans la figure 3.1, on visualise un exemple de plan factoriel complet à deux facteurs. Les spécifications et les détails particuliers d'un plan optimal à deux niveaux de type 2^k incluent :

1. Nombre de facteurs (k)

- ✦ k : représente le nombre de facteurs,
- ✦ Chaque facteur peut être réglé à deux niveaux, généralement -1 et +1.

2. Nombre total de combinaisons (2^k)

- ✗ Le nombre total de combinaisons expérimentales est 2^k , ce qui signifie que chaque facteur est testé à tous les niveaux possibles (-1 et +1).
- ✗ Le nombre total d'essais augmente exponentiellement avec le nombre de facteurs, ce qui peut rendre ce type de plan excessif pour un grand nombre de facteurs.

3. Analyse des effets principaux et des interactions

- ✗ Ce plan permet d'analyser les effets principaux de chaque facteur, c'est-à-dire comment le changement d'un facteur affecte la réponse.
- ✗ Il permet également d'étudier les interactions entre les facteurs, c'est-à-dire comment les variations conjointes de deux ou plusieurs facteurs influent sur la réponse.

4. Réplication

- ✗ La réplication des essais est souvent incluse dans le plan pour estimer la variabilité expérimentale et améliorer la fiabilité des résultats.

5. Avantages et limitations

- ✗ Avantages : Offre une vue exhaustive des effets principaux et des interactions, ce qui le rend adapté aux études exploratoires où tous les facteurs sont jugés influents.
- ✗ Limitations : Le nombre d'essais augmente rapidement avec le nombre de facteurs, ce qui peut rendre la réalisation du plan coûteuse en termes de temps et de ressources. Par exemple, pour étudier 8 facteurs, cela nécessiterait initialement 256 essais. Cependant, afin de réduire le nombre d'essais tout en conservant la capacité d'examiner tous les facteurs, il est possible de considérer des plans factoriels fractionnaires à deux niveaux avec 2^{k-p} .

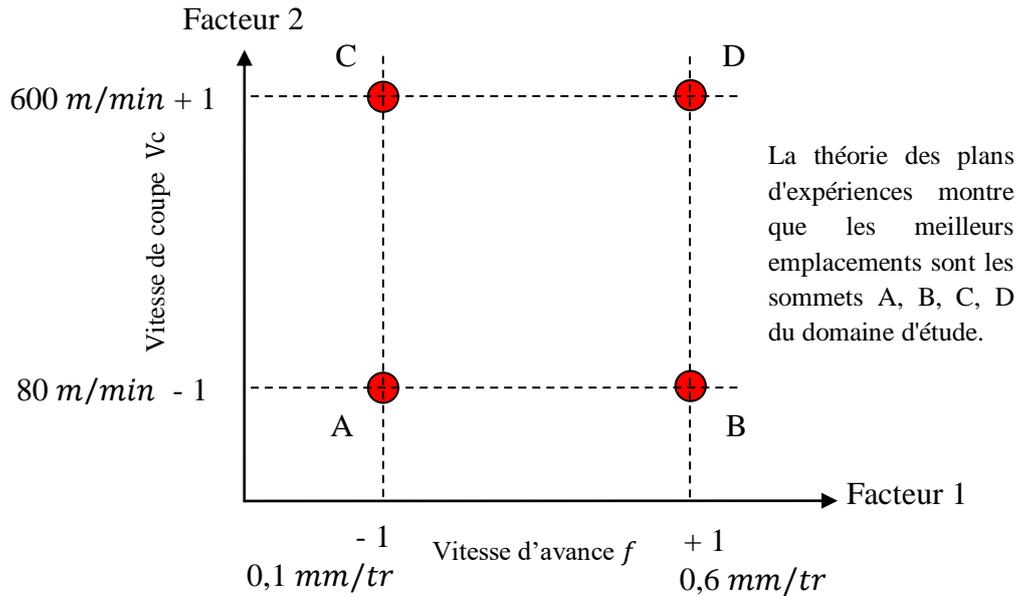


Figure 3.1- Meilleurs emplacements des points expérimentaux dans le cadre du modèle du premier degré.

Par exemple, considérons une étude portant sur l'effet de la concentration du réactif et de la quantité de catalyseur sur la conversion (rendement) dans un processus chimique. L'objectif de l'expérience est de déterminer si des ajustements de l'un ou l'autre de ces deux facteurs pourraient augmenter le rendement. Supposons que la concentration du réactif soit le facteur A, avec deux niveaux d'intérêt, soit 15 % et 25 %. Le catalyseur est le facteur B, où le niveau élevé correspond à l'utilisation de 2 catalyseurs et le niveau bas à l'utilisation d'un seul catalyseur. L'expérience est répétée trois fois, ce qui donne un total de 12 essais. L'ordre des essais est aléatoire, ce qui en fait une expérience entièrement randomisée. Les données obtenues sont présentées dans le tableau 3.1.

Tableau 3.1- Données obtenues.

Facteurs		Combinaison de traitement	Réplication			Total
A	B		1	2	3	
-	-	A bas, B bas	28	25	27	80
+	-	A haut, B bas	36	32	32	100
-	+	A bas, B haut	18	19	23	60
+	+	A haut, B haut	31	30	29	90

Les quatre combinaisons de traitement dans ce plan sont illustrées graphiquement à la figure 3.2. Ainsi, "A" se réfère à l'effet du facteur A, "B" à l'effet du facteur B, et "AB" à l'interaction entre A et B. Dans le plan 2^2 , les niveaux bas et élevés de A et B sont notés

respectivement par “-” et “+” sur les axes A et B. Ainsi, “-” sur l’axe A représenté le niveau bas de concentration (15 %), tandis que “+” représente le niveau élevé (25 %). De même, “-” sur l’axe B représente le niveau bas du catalyseur, et “+” indique le niveau élevé.

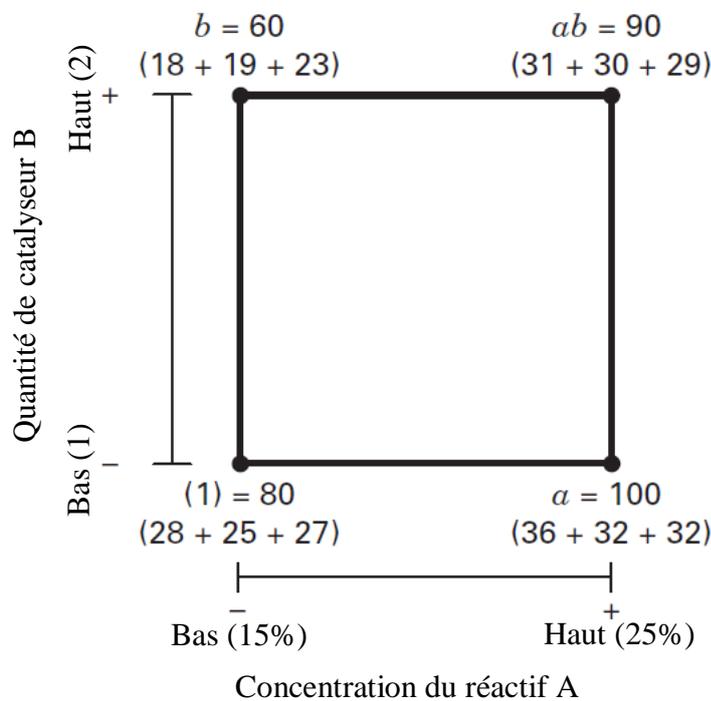


Figure 3.2- Combinaisons de traitement dans la conception 2^2 .

Les quatre combinaisons de traitement dans le plan sont également représentées par des lettres minuscules, comme indiqué à la figure 3.2. D’après la figure, le niveau élevé de tout facteur dans la combinaison de traitement est désigné par la lettre minuscule correspondante, tandis que le niveau bas est représenté par l’absence de la lettre correspondante.

Ainsi, **a** représenté la combinaison de traitement avec A au niveau élevé et B au niveau bas, **b** représente A au niveau bas et B au niveau élevé, et **ab** représente les deux facteurs au niveau élevé. Par convention, (1) est utilisé pour désigner les deux facteurs au niveau bas. Cette notation est utilisée tout au long de la série 2^2 .

Dans un plan factoriel à deux niveaux, nous pouvons définir l’effet moyen d’un facteur comme le changement de réponse produit par une variation du niveau de ce facteur, en moyennant les niveaux de l’autre facteur. De plus, les symboles (1), a, b et ab représentent maintenant la somme des observations de la réponse pour les n répétitions effectuées pour chaque combinaison de traitement, comme illustré à la figure 3.2. Ainsi, l’effet de A au niveau bas de

B est $[a - (1)]/n$, et l'effet de A au niveau élevé de B est $[ab - b]/n$. La moyenne de ces deux valeurs donne l'effet principal de A :

$$\begin{aligned} A &= \frac{1}{2n} \{[ab - b] + [a - (1)]\} \\ &= \frac{1}{2n} [ab + a - b - (1)] \end{aligned} \quad (3.1)$$

L'effet moyen principal de B est déterminé à partir de l'effet de B au niveau bas de A (c'est-à-dire $[b - (1)]/n$) et au niveau élevé de A (c'est-à-dire $[ab - a]/n$) comme suit :

$$\begin{aligned} B &= \frac{1}{2n} \{[ab - a] + [b - (1)]\} \\ &= \frac{1}{2n} [ab + b - a - (1)] \end{aligned} \quad (3.2)$$

Nous définissons l'effet d'interaction AB comme la différence moyenne entre l'effet de A au niveau élevé de B et l'effet de A au niveau bas de B. Ainsi,

$$\begin{aligned} AB &= \frac{1}{2n} \{[ab - b] - [a - (1)]\} \\ &= \frac{1}{2n} [ab + (1) - a - b] \end{aligned} \quad (3.3)$$

Alternativement, nous pouvons définir AB comme la différence moyenne entre l'effet de B au niveau élevé de A et l'effet de B au niveau bas de A. Cela conduit également à l'équation 3.3.

Les formules pour les effets de A, B et AB peuvent être dérivées par une autre méthode. L'effet de A peut être déterminé comme la différence entre la réponse moyenne des deux combinaisons de traitement situées sur le côté droit du carré de la figure 3.2 (appelons cette moyenne \bar{y}_{A+} car il s'agit de la réponse moyenne pour les combinaisons de traitement où A est au niveau élevé) et celle des deux combinaisons de traitement situées sur le côté gauche (ou \bar{y}_{A-}). C'est-à-dire :

$$\begin{aligned} A &= \bar{y}_{A+} - \bar{y}_{A-} \\ &= \frac{ab + a}{2n} - \frac{b + (1)}{2n} \end{aligned}$$

$$= \frac{1}{2n} [ab + a - b - (1)]$$

Ceci donne exactement le même résultat que dans l'équation 3.1. L'effet de B, équation 3.2, est obtenu comme la différence entre la moyenne des deux combinaisons de traitement situées en haut du carré (\bar{y}_{B^+}) et la moyenne des deux combinaisons de traitement situées en bas (\bar{y}_{B^-}), soit :

$$\begin{aligned} B &= \bar{y}_{B^+} - \bar{y}_{B^-} \\ &= \frac{ab + b}{2n} - \frac{a + (1)}{2n} \\ &= \frac{1}{2n} [ab + b - a - (1)] \end{aligned}$$

Enfin, l'effet d'interaction AB est la différence entre la moyenne des combinaisons de traitement situées sur la diagonale de droite à gauche dans le carré [ab et (1)] et la moyenne des combinaisons de traitement situées sur la diagonale de gauche à droite (a et b), soit :

$$\begin{aligned} AB &= \frac{ab + (1)}{2n} - \frac{a + b}{2n} \\ &= \frac{1}{2n} [ab + (1) - a - b] \end{aligned}$$

Ce qui est identique à l'équation 3.3.

En utilisant l'expérience illustrée dans la figure 3.2, nous pouvons estimer les effets moyens comme suit :

$$A = \frac{1}{2(3)} (90 + 100 - 60 - 80) = 8,33$$

$$B = \frac{1}{2(3)} (90 + 60 - 100 - 80) = -5,00$$

$$AB = \frac{1}{2(3)} (90 + 80 - 100 - 60) = 1,67$$

L'effet de A (concentration du réactif) est positif ; cela suggère qu'augmenter A du niveau bas (15 %) au niveau élevé (25 %) augmentera le rendement. L'effet de B (catalyseur) est négatif ;

cela indique qu'augmenter la quantité de catalyseur ajoutée au processus réduira le rendement. L'effet d'interaction semble être faible par rapport aux deux effets principaux.

Dans les expériences impliquant des conceptions 2^k , il est toujours important d'examiner l'ampleur et la direction des effets des facteurs pour déterminer quelles variables sont susceptibles d'être importantes. L'analyse de variance (ANOVA) peut généralement être utilisée pour confirmer cette interprétation (des tests t peuvent également être utilisés). L'ampleur et la direction des effets doivent toujours être prises en compte en conjonction avec l'ANOVA, car l'ANOVA seule ne transmet pas cette information. Il existe plusieurs excellents logiciels statistiques qui sont utiles pour configurer et analyser les conceptions 2^k . Il existe également des méthodes spéciales permettant d'effectuer les calculs manuellement de manière à gagner du temps.

Considérons la détermination des sommes des carrés pour A , B et AB . Notez, à partir de l'équation 3.1, qu'un contraste est utilisé pour estimer A , à savoir :

$$\text{Contraste}_A = ab + a - b - (1) \quad (3.4)$$

Nous appelons généralement ce contraste l'effet total de A . D'après les équations 3.2 et 3.3, nous voyons que des contrastes sont également utilisés pour estimer B et AB . De plus, ces trois contrastes sont orthogonaux. La somme des carrés pour un contraste est égale au carré du contraste divisé par le nombre d'observations dans chaque total du contraste multiplié par la somme des carrés des coefficients du contraste. Par conséquent, nous avons :

$$SS_A = \frac{[ab + a - b - (1)]^2}{4n} \quad (3.5)$$

$$SS_B = \frac{[ab + b - a - (1)]^2}{4n} \quad (3.6)$$

Et

$$SS_{AB} = \frac{[ab + (1) - a - b]^2}{4n} \quad (3.7)$$

Comme les sommes des carrés pour A , B et AB . Remarquez à quel point ces équations sont simples. Nous pouvons calculer les sommes des carrés en ne mettant qu'un nombre au carré.

En utilisant l'expérience illustrée dans la figure 3.2, nous pouvons trouver les sommes des carrés à partir des équations 3.5, 3.6 et 3.7 comme suit :

$$\begin{aligned} SS_A &= \frac{(50)^2}{4(3)} = 208,33 \\ SS_B &= \frac{(-30)^2}{4(3)} = 75,00 \end{aligned} \quad (3.8)$$

Et

$$SS_{AB} = \frac{(10)^2}{4(3)} = 8,33$$

La somme totale des carrés est trouvée de la manière habituelle, c'est-à-dire :

$$SS_T = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^n y_{ijk}^2 - \frac{y_{\dots}^2}{4n} \quad (3.9)$$

En général, SS_T a $4n - 1$ degrés de liberté. La somme des carrés de l'erreur, avec $4(n - 1)$ degrés de liberté, est généralement calculée par soustraction comme suit :

$$SS_E = SS_T - SS_A - SS_B - SS_{AB} \quad (3.10)$$

Pour l'expérience illustrée dans la figure 3.2, nous obtenons :

$$\begin{aligned} SS_T &= \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^n y_{ijk}^2 - \frac{y_{\dots}^2}{4n} \\ &= 9398,00 - 9075,00 \\ &= 323,00 \end{aligned}$$

Et

$$\begin{aligned} SS_E &= SS_T - SS_A - SS_B - SS_{AB} \\ &= 323,00 - 208,33 - 75,00 - 8,33 \\ &= 31,34 \end{aligned}$$

En utilisant SS_A , SS_B et SS_{AB} provenant des équations 6.8. L'ANOVA complète est résumée dans le tableau 3.2. Sur la base des valeurs P , nous concluons que les effets principaux sont statistiquement significatifs et qu'il n'y a pas d'interaction entre ces facteurs. Cela confirme notre interprétation initiale des données basée sur les amplitudes des effets des facteurs.

Tableau 3.2- Analyse de la variance pour l'expérience de la figure 3.2.

Source de variation	Somme des carrés	Degrés de liberté	Moyenne des carrés	Valeur F	Valeur P
A	208.33	1	208.33	53.15	0.0001
B	75.00	1	75.00	19.13	0.0024
AB	8.33	1	8.33	2.13	0.1826
Erreur	31.34	8	3.92		
Total	323.00	11			

Il est souvent pratique de noter les combinaisons de traitement dans l'ordre (1), a, b, ab. Cet ordre est appelé ordre standard (ou ordre de Yates, du nom de Frank Yates, qui était l'un des collaborateurs de Fisher et qui a apporté de nombreuses contributions importantes à la conception et à l'analyse des expériences). En utilisant cet ordre standard, nous voyons que les coefficients de contraste utilisés pour estimer les effets sont :

Effets	(1)	<i>a</i>	<i>b</i>	<i>ab</i>
<i>A</i>	-1	+1	-1	+1
<i>B</i>	-1	-1	+1	+1
<i>AB</i>	+1	-1	-1	+1

Notez que les coefficients de contraste pour estimer l'effet d'interaction sont simplement le produit des coefficients correspondants pour les deux effets principaux. Le coefficient de contraste est toujours soit +1 soit -1, et un tableau de signes plus et moins, comme dans le tableau 3.3, peut être utilisé pour déterminer le signe approprié pour chaque combinaison de traitement. Les en-têtes de colonne dans le tableau 3.3 représentent les effets principaux (*A* et *B*), l'interaction *AB*, et *I*, qui représente le total ou la moyenne de l'ensemble de l'expérience. Remarquez que la colonne correspondant à *I* n'a que des signes positifs. Les désignateurs de ligne sont les combinaisons de traitement. Pour trouver le contraste permettant d'estimer un effet quelconque, il suffit de multiplier les signes de la colonne appropriée du tableau par la combinaison de traitement correspondante et d'additionner. Par exemple, pour estimer *A*, le contraste est $-(1) + a - b + ab$, ce qui concorde avec l'équation 3.1. Notez que les contrastes pour les effets *A*, *B* et *AB* sont orthogonaux. Ainsi, le 2^2 (et tous les designs 2^k) est un design

orthogonal. Le codage ± 1 pour les niveaux bas et élevés des facteurs est souvent appelé codage orthogonal ou codage des effets.

Tableau 3.3- Signes algébriques pour le calcul des effets dans le plan 2^2 .

Combinaisons de traitement	Effet factoriel			
	<i>I</i>	<i>A</i>	<i>B</i>	<i>AB</i>
(1)	+	–	–	+
<i>a</i>	+	+	–	–
<i>b</i>	+	–	+	–
<i>ab</i>	+	+	+	+

Le Modèle de Régression. Dans un design factoriel 2^2 , il est facile d'exprimer les résultats de l'expérience en termes d'un modèle de régression. Étant donné que le 2^2 est simplement un design factoriel, nous pourrions également utiliser un modèle par effets ou par moyennes, mais l'approche du modèle de régression est beaucoup plus naturelle et intuitive. Pour l'expérience de processus chimique illustrée dans la figure 3.2, le modèle de régression est :

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \quad (3.11)$$

Où x_1 est une variable codée qui représente la concentration du réactif, x_2 est une variable codée qui représente la quantité de catalyseur, et les β sont les coefficients de régression. La relation entre les variables naturelles, la concentration du réactif et la quantité de catalyseur, et les variables codées est :

$$x_1 = \frac{Conc - (Conc_{bas} + Conc_{haut})/2}{(Conc_{haut} - Conc_{bas})/2}$$

Et

$$x_2 = \frac{Cataly - (Cataly_{bas} + Cataly_{haut})/2}{(Cataly_{haut} - Cataly_{bas})/2}$$

Lorsqu'il n'y a que deux niveaux pour les variables naturelles, ce codage produira la notation familière de ± 1 pour les niveaux des variables codées. Pour illustrer cela dans notre exemple, notez que :

$$\begin{aligned} x_1 &= \frac{Conc - (15 + 25)/2}{(25 - 15)/2} \\ &= \frac{Conc - 20}{5} \end{aligned}$$

Ainsi, si la concentration est au niveau élevé (Conc 25 %), alors $x_1 = +1$; si la concentration est au niveau bas (Conc 15 %), alors $x_1 = -1$. De plus,

$$\begin{aligned} x_2 &= \frac{\text{Cataly} - (1 + 2)/2}{(2 - 1)/2} \\ &= \frac{\text{Cataly} - 1,5}{0,5} \end{aligned}$$

Ainsi, si le catalyseur est au niveau élevé (Catalyseur 2), alors $x_2 = +1$; si le catalyseur est au niveau bas (Catalyseur 1), alors $x_2 = -1$.

Le modèle de régression ajusté est

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)x_1 + \left(\frac{-5,00}{2}\right)x_2$$

Où l'ordonnée à l'origine est la moyenne générale des 12 observations, et les coefficients de régression $\widehat{\beta}_1$ et $\widehat{\beta}_2$ sont égaux à la moitié des estimations des effets des facteurs correspondants. Le coefficient de régression est égal à la moitié de l'estimation de l'effet, car il mesure l'effet d'un changement d'une unité de x sur la moyenne de y , alors que l'estimation de l'effet est basée sur un changement de deux unités (de -1 à $+1$). Cette méthode simple d'estimation des coefficients de régression aboutit à des estimations des paramètres par moindres carrés.

Résidus et adéquation du modèle. Le modèle de régression peut être utilisé pour obtenir la valeur prédite ou ajustée de y aux quatre points du plan d'expérience. Les résidus sont les différences entre les valeurs observées et les valeurs ajustées de y . Par exemple, lorsque la concentration du réactif est au niveau bas ($x_1 = -1$) et que le catalyseur est au niveau bas ($x_2 = -1$), le rendement prédit est

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)(-1) + \left(\frac{-5,00}{2}\right)(-1) = 25,835$$

Il y a trois observations pour cette combinaison de traitement, et les résidus sont :

$$e_1 = 28 - 25,835 = 2,165$$

$$e_2 = 25 - 25,835 = -0,835$$

$$e_3 = 27 - 25,835 = 1,165$$

Les valeurs prédites restantes et les résidus sont calculés de manière similaire. Pour le niveau élevé de la concentration du réactif et le niveau bas du catalyseur,

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)(+1) + \left(\frac{-5,00}{2}\right)(-1) = 34,165$$

Et

$$e_4 = 36 - 34,165 = 1,835$$

$$e_5 = 32 - 34,165 = -2,165$$

$$e_6 = 32 - 34,165 = -2,165$$

Pour le niveau bas de la concentration du réactif et le niveau élevé du catalyseur,

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)(-1) + \left(\frac{-5,00}{2}\right)(+1) = 20,835$$

Et

$$e_7 = 18 - 20,835 = -2,835$$

$$e_8 = 19 - 20,835 = -1,835$$

$$e_9 = 23 - 20,835 = 2,165$$

Enfin, pour le niveau élevé des deux facteurs,

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)(+1) + \left(\frac{-5,00}{2}\right)(+1) = 29,165$$

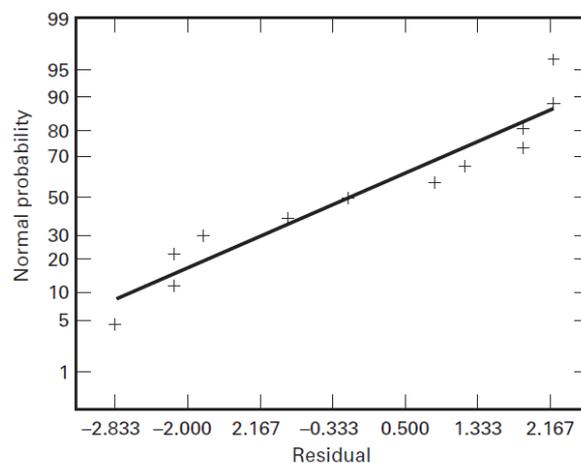
Et

$$e_{10} = 31 - 29,165 = 1,835$$

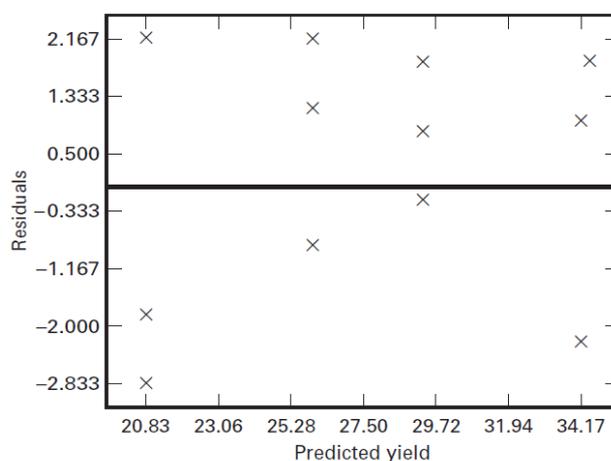
$$e_{11} = 30 - 29,165 = 0,835$$

$$e_{12} = 29 - 29,165 = -0,165$$

La figure 3.3 présente un graphique de probabilité normale de ces résidus ainsi qu'un graphique des résidus en fonction du rendement prédit. Ces graphiques semblent satisfaisants, et nous n'avons donc aucune raison de soupçonner des problèmes quant à la validité de nos conclusions.



(a) Normal probability plot



(b) Residuals versus predicted yield

Figure 3.3- Graphiques résiduels pour l'expérience du processus chimique.

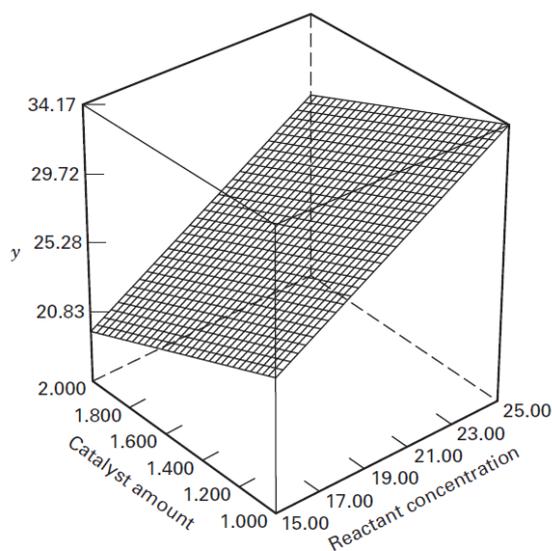
La surface de réponse. Le modèle de régression

$$\hat{y} = 27,5 + \left(\frac{8,33}{2}\right)x_1 + \left(\frac{-5,00}{2}\right)x_2$$

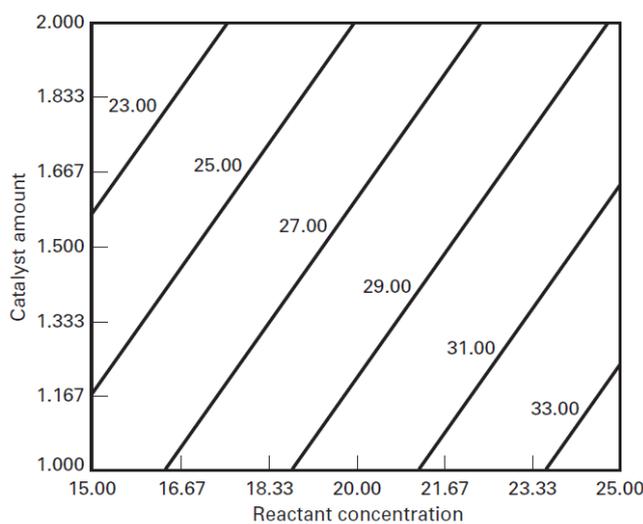
Peut être utilisé pour générer des graphiques de surface de réponse. S'il est souhaitable de construire ces graphiques en fonction des niveaux de facteurs naturels, nous substituons simplement les relations entre les variables naturelles et codées que nous avons données précédemment dans le modèle de régression, ce qui donne

$$\begin{aligned}\hat{y} &= 27,5 + \left(\frac{8,33}{2}\right)\left(\frac{Conc - 20}{5}\right) + \left(\frac{-5,00}{2}\right)\left(\frac{Cataly - 1,5}{0,5}\right) \\ &= 18,33 + 0,8333Conc - 5,00Cataly\end{aligned}$$

La figure 3.4a présente le graphique de surface de réponse tridimensionnelle du rendement provenant de ce modèle, et la figure 3.4b est le graphique des contours. Étant donné que le modèle est de premier ordre (c'est-à-dire qu'il ne contient que les effets principaux), la surface de réponse ajustée est un plan. En examinant le graphique des contours, nous observons que le rendement augmente à mesure que la concentration du réactif augmente et que la quantité de catalyseur diminue. Souvent, nous utilisons une surface ajustée comme celle-ci pour trouver une direction d'amélioration potentielle pour un processus.



(a) Response surface



(b) Contour plot

Figure 3.4- Graphique de surface de réponse et graphique de contour du rendement de l'expérience de processus chimique

Dans le contexte des plans factoriels complets à deux facteurs (2^2), l'équation linéaire utilisée est la suivante :

$$Y = a_0 + a_1X_1 + a_2X_2 + a_{12}X_1X_2 + \varepsilon \dots \dots \dots (3.1)$$

Avec :

- ✧ Y : réponse,
- ✧ X_i : niveau attribué au facteur i ,
- ✧ a_0 : valeur de la réponse au centre du domaine d'étude,
- ✧ a_1 : effet (ou effet principal) du facteur 1,
- ✧ a_2 : effet (ou effet principal) du facteur 2,
- ✧ a_{12} : interaction entre les facteurs 1 et 2,
- ✧ ε : erreur (résidu).

Dans le cas des plans impliquent l'exploration de k facteurs, chacun avec deux niveaux distincts. Le modèle mathématique sous-jacent est un polynôme englobant la moyenne, les effets individuels de chaque facteur, ainsi que toutes les interactions possibles entre les facteurs, qu'ils soient considérés deux par deux, trois par trois, jusqu'à k par k .

$$Y = a_0 + \sum a_iX_i + \sum a_{ij}X_iX_j + \sum a_{ijl}X_iX_jX_l + \dots + \sum a_{ij\dots k}X_iX_j\dots X_k \dots \dots \dots (3.1)$$

Exemple 2 :

Considérons un exemple simple avec trois facteurs ($k = 3$) pour illustrer un plan optimal à deux niveaux de type 2^k . Imaginons que nous étudions l'effet de trois facteurs sur le rendement d'un processus de fabrication, et chaque facteur peut être réglé à deux niveaux (-1 et +1).

Les trois facteurs sont A, B et C, et nous avons les niveaux -1 et +1 pour chacun d'eux. Un plan 2^3 impliquerait toutes les combinaisons possibles de ces niveaux, soit $2^3 = 8$ essais expérimentaux (voir tableau 3.1).

Tableau 3.1- Configuration du plan 2^3 .

Essai	Niveau A	Niveau B	Niveau C
1	- 1	- 1	- 1
2	- 1	- 1	+ 1
3	- 1	+ 1	- 1

4	- 1	+ 1	+ 1
5	+ 1	- 1	- 1
6	+ 1	- 1	+ 1
7	+ 1	+ 1	- 1
8	+ 1	+ 1	+ 1

Chaque essai représente une combinaison spécifique des niveaux des facteurs. Par exemple, l'essai 1 signifie que les facteurs A, B et C sont tous réglés à leur niveau -1. Ce plan permet d'observer comment les variations des niveaux de ces facteurs influent sur le rendement du processus.

L'analyse des résultats de ces essais, souvent effectuée à l'aide de méthodes statistiques telles que l'analyse de variance (ANOVA), permettra de déterminer les effets principaux de chaque facteur, les interactions éventuelles, et d'optimiser les niveaux des facteurs pour atteindre un rendement optimal.

3.3.2- Plan optimal simplifié de type 2^{k-p}

Considérons une situation dans laquelle trois facteurs, chacun à deux niveaux, sont d'intérêt, mais les expérimentateurs ne peuvent pas se permettre de réaliser les 8 combinaisons de traitement du plan 2^3 . Ils peuvent cependant se permettre de réaliser quatre essais. Cela suggère une fraction d'un demi du plan 2^3 . Comme le plan contient $2^{3-1} = 4$ combinaisons de traitement, une fraction d'un demi du plan 2^3 est souvent appelée un plan 2^{3-1} .

Le tableau des signes plus et moins pour le plan 2^3 est présenté dans le tableau 3.2. Supposons que nous sélectionnions les quatre combinaisons de traitement a, b, c et abc comme notre fraction d'un demi. Ces essais sont montrés dans la première moitié du tableau 3.2 et dans la figure 3.5a.

Tableau 3.2- Signes plus et moins pour le plan factoriel 2^3 .

Combinaisons de traitement	Effet factoriel							
	<i>I</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
c	+	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+		+		-	-
ac	+	+	-	+	-	+	-	-
bc	+	-	+	+	-	-	+	-
(1)	+	-	-	-	+	+	+	-

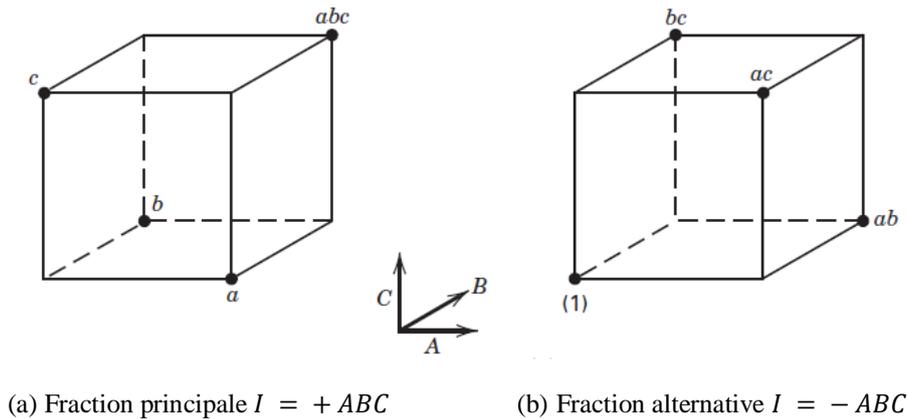


Figure 3.5- Deux demi-fractions du dessin 2^3 .

Remarquez que le plan 2^{3-1} est formé en sélectionnant uniquement les combinaisons de traitement ayant un signe plus dans la colonne ABC. Ainsi, ABC est appelé le générateur de cette fraction particulière. Habituellement, nous désignons un générateur comme ABC par le terme *mot*. De plus, la colonne d'identité I est toujours avec un signe plus, donc nous l'appelons :

$$I = ABC$$

La relation définissant de notre plan. En général, la relation définissant pour un plan factoriel fractionnaire sera toujours l'ensemble de toutes les colonnes qui sont égales à la colonne d'identité I .

Les combinaisons de traitement dans le plan 2^{3-1} fournissent trois degrés de liberté que nous pouvons utiliser pour estimer les effets principaux. En se référant au tableau 3.2, nous remarquons que les combinaisons linéaires des observations utilisées pour estimer les effets principaux de A , B et C sont :

$$[A] = \frac{1}{2}(a - b - c + abc)$$

$$[B] = \frac{1}{2}(-a + b - c + abc)$$

$$[C] = \frac{1}{2}(-a - b + c + abc)$$

La notation $[A]$, $[B]$ et $[C]$ est utilisée pour indiquer les combinaisons linéaires associées aux effets principaux. Il est également facile de vérifier que les combinaisons linéaires des observations utilisées pour estimer les interactions entre deux facteurs sont :

$$[BC] = \frac{1}{2}(a - b - c + abc)$$

$$[AC] = \frac{1}{2}(-a + b - c + abc)$$

$$[AB] = \frac{1}{2}(-a - b + c + abc)$$

Ainsi, $[A] = [BC]$, $[B] = [AC]$, et $[C] = [AB]$; par conséquent, il est impossible de différencier entre A et BC, B et AC, et C et AB. En fait, lorsque nous estimons A, B et C, nous estimons en réalité $A + BC$, $B + AC$, et $C + AB$. Deux effets ou plus ayant cette propriété sont appelés *alias*. Dans notre exemple, A et BC sont des alias, B et AC sont des alias, et C et AB sont des alias. Nous indiquons cela par la notation $[A] \rightarrow A + BC$, $[B] \rightarrow B + AC$, et $[C] \rightarrow C + AB$.

La structure des alias pour ce modèle peut être facilement déterminée en utilisant la relation définissant $I = ABC$. En multipliant n'importe quelle colonne (ou effet) par la relation définissante, on obtient les alias pour cette colonne (ou effet). Dans notre exemple, cela donne l'alias de A.

$$A \cdot I = A \cdot ABC = A^2BC$$

Ou, parce que le carré de n'importe quelle colonne est simplement l'identité I ,

$$A = BC$$

De même, nous trouvons les alias de B et C comme

$$B \cdot I = B \cdot ABC$$

$$B = AB^2C = AC$$

Et

$$C \cdot I = C \cdot ABC$$

$$C = ABC^2 = AB$$

Cette fraction d'un demi, avec $I = ABC$, est généralement appelée la fraction principale.

Supposons maintenant que nous ayons choisi l'autre fraction d'un demi, c'est-à-dire les combinaisons de traitements dans le tableau 3.2 associées au signe moins dans la colonne ABC . Cette fraction alternative, ou complémentaire, d'un demi (composée des essais (1), ab , ac , et bc) est illustrée dans la figure 3.2b. La relation définissant pour ce modèle est :

$$I = -ABC$$

La combinaison linéaire des observations, disons $[A]'$, $[B]'$ et $[C]'$, provenant de la fraction alternative nous donne :

$$[A]' \rightarrow A - BC$$

$$[B]' \rightarrow B - AC$$

$$[C]' \rightarrow C - AB$$

Ainsi, lorsque nous estimons A , B et C avec cette fraction particulière, nous estimons en réalité $A - BC$, $B - AC$ et $C - AB$.

En pratique, il n'importe pas laquelle des fractions est utilisée. Les deux fractions appartiennent à la même famille ; c'est-à-dire que les deux fractions d'un demi forment un plan complet 2^3 . Cela est facilement visible en se référant aux parties a et b de la figure 3.5.

Supposons qu'après avoir exécuté l'une des fractions d'un demi du plan 2^3 , l'autre fraction ait également été exécutée. Ainsi, les huit essais associés au plan complet 2^3 sont maintenant disponibles. Nous pouvons maintenant obtenir des estimations désaïlées de tous les effets en analysant les huit essais comme un plan 2^3 complet en deux blocs de quatre essais chacun. Cela pourrait également être réalisé en ajoutant et en soustrayant la combinaison linéaire des effets des deux fractions individuelles. Par exemple, considérons $[A] \rightarrow A + BC$ et $[A]' \rightarrow A - BC$. Cela implique que :

$$\frac{1}{2}([A] + [A]') = \frac{1}{2}(A + BC + A - BC) \rightarrow A$$

Et cela

$$\frac{1}{2}([A] - [A]') = \frac{1}{2}(A + BC - A + BC) \rightarrow BC$$

Ainsi, pour les trois paires de combinaisons linéaires, nous obtiendrions ce qui suit :

i	Depuis $\frac{1}{2}([i] + [i]')$	Depuis $\frac{1}{2}([i] - [i]')$
A	A	BC
B	B	AC
C	C	AB

De plus, en assemblant le plan complet 2^3 de cette manière avec $I = +ABC$ dans le premier groupe d'essais et $I = -ABC$ dans le second, le plan 2^3 confond ABC avec les blocs.

Le plan optimal de premier ordre 2^{k-p} est une variante du plan à deux niveaux de type 2^k , où k représente le nombre de facteurs et p représente le nombre d'interactions d'ordre supérieur incluses dans le plan. L'idée est de simplifier le plan pour réduire le nombre d'essais tout en conservant l'information cruciale sur les effets principaux et les interactions. Les détails supplémentaires sur le plan optimal de premier ordre 2^{k-p} sont :

1. Nombre de facteurs (k)

✧ k : représente le nombre de facteurs que vous souhaitez étudier.

2. Nombre d'interactions d'ordre supérieur incluses (p)

✧ p : représente le nombre d'interactions d'ordre supérieur incluses dans le plan.

✧ Les interactions d'ordre supérieur sont les interactions entre trois facteurs ou plus. En incluant certaines de ces interactions, le plan devient plus complet que les plans plus simples à deux niveaux.

3. Nombre total de combinaisons (2^{k-p})

✧ Le nombre total de combinaisons expérimentales est 2^{k-p} , ce qui signifie que le plan est simplifié par rapport au plan 2^k , en éliminant certaines interactions d'ordre supérieur.

4. Avantages et limitations

✧ Avantages : Réduit le nombre d'essais nécessaires par rapport au plan 2^k complet tout en fournissant des informations importantes sur les effets principaux et les interactions.

✧ Limitations : Peut ne pas prendre en compte toutes les interactions d'ordre supérieur, ce qui peut être un compromis acceptable dans certaines situations pour réduire le coût expérimental.

5. Analyse des effets principaux et des interactions

- ✧ Comme dans le plan 2^k , ce plan permet d'analyser les effets principaux de chaque facteur et les interactions entre les facteurs, mais il offre une réduction du nombre total d'essais.

Exemple simplifié :

- ✧ Supposons que $k = 4$ (quatre facteurs) et $p = 2$ (deux interactions d'ordre supérieur incluses). Le plan optimal de premier ordre 2^{4-2} comprendrait $2^{4-2} = 14$ essais expérimentaux, éliminant deux interactions d'ordre supérieur par rapport au plan 2^4 complet.

L'utilisation d'un tel plan est souvent justifiée lorsque le nombre de facteurs est élevé et qu'il est important de réduire le coût expérimental tout en maintenant une certaine robustesse dans l'exploration des effets principaux et des interactions.

3.4- Plan optimal de deuxième ordre

Le plan optimal de deuxième ordre (ou plan optimal quadratique) est un concept utilisé principalement en optimisation et en théorie du contrôle. Il concerne des systèmes où la fonction de coût, ou critère d'optimisation, est de nature quadratique.

CHAPITRE 4 :

Optimisation des phénomènes

4.1- Introduction

En règle générale, les ingénieurs et les chercheurs réalisent des expériences pour les raisons suivantes :

- 1) Pour évaluer les performances, c'est-à-dire la production,
- 2) Pour déterminer l'impact des différentes combinaisons de variables de conception ou d'influence présentes dans le système,
- 3) Pour créer ou développer des modèles mathématiques qui lient les entrées et les sorties des processus,
- 4) Pour continuellement améliorer les performances du processus ou de la conception.

En conséquence, de nombreux efforts de recherche sont déployés dans le monde entier pour développer des techniques ou des approches permettant d'étudier les variables qui améliorent les performances d'usinabilité. Il est donc essentiel de bien connaître ces techniques.

4.2- Approche expérimentale classique de l'ingénierie

Les ingénieurs et les scientifiques utilisent l'approche expérimentale classique, également appelée "un facteur à la fois" (OFAT), pour étudier l'influence des variables sur les performances. Dans cette approche, on fait varier un facteur à la fois tout en maintenant les autres facteurs à une valeur fixe. Cette procédure est répétée pour chaque facteur que l'on souhaite étudier. L'expérimentateur obtient ainsi des informations sur l'influence d'un facteur à la fois sur les performances.

Cette technique uni-factorielle permet aux ingénieurs de déterminer l'effet individuel de chaque facteur sur le processus ou le système étudié. Cependant, elle peut ne pas prendre en compte les interactions entre les facteurs, ce qui peut conduire à des conclusions biaisées ou à des résultats peu fiables.

Les avantages significatifs de l'approche "un facteur à la fois" (OFAT) sont les suivants :

- 1) Elle permet de se concentrer sur un espace de conception limité.
- 2) Elle fournit des informations à l'expérimentateur qui peut rapidement prendre des décisions concernant la plage de fonctionnement optimale pour chaque facteur (c'est-à-dire la valeur maximale et minimale).
- 3) Elle permet de déterminer la sensibilité de chaque facteur indépendant, ce qui aide à filtrer les facteurs insignifiants et à les classer en fonction de leur importance.

Cependant, l'approche OFAT présente également des limites importantes qui ont été discutées dans la littérature, telles que :

- 1) Elle nécessite un grand nombre d'essais expérimentaux pour obtenir le même degré de précision que d'autres méthodes d'optimisation et d'estimation des effets des facteurs.
- 2) Elle ne permet pas d'évaluer l'interaction entre les facteurs, ce qui peut conduire à des conclusions biaisées ou incomplètes.
- 3) Les conclusions tirées de l'analyse peuvent ne pas être généralisables,
- 4) Le réglage optimal recommandé peut ne pas toujours fournir une solution globale, ce qui peut limiter son utilité dans certaines situations.
- 5) L'approche OFAT ne permet pas d'établir des équations prédictives, ce qui est souvent utilisé dans la modélisation et l'optimisation.
- 6) Elle n'est pas adaptée pour étudier les interactions non linéaires entre les facteurs, limitant ainsi la possibilité d'obtenir des informations précises sur le processus.

4.3- Approche de modélisation numérique

La difficulté pratique à obtenir des données sur différents paramètres de l'intégrité de surface (tels que la texture de surface, les contraintes résiduelles, la couche blanche, etc.), ainsi que des forces de coupe, de température et d'usure des outils pendant les expériences, limite les chercheurs dans le développement d'outils plus précis et fiables basés sur l'expérimentation. Les méthodes numériques basées sur l'analyse par éléments finis ((Finite Element Analysis FEA) offrent une voie efficace pour étudier les facteurs influençant les procédés d'usinage.

Jusqu'à aujourd'hui, de nombreux efforts ont été faits pour utiliser les techniques de simulation numérique (telles que FEA, ABAQUS ou ANSYS) pour étudier la mécanique et la dynamique impliquées dans les procédés d'usinage.

Les outils de modélisation numérique présentent plusieurs limites importantes, notamment :

- 1) L'usinage est supposé se produire sous des conditions de déformation plane, alors qu'en réalité la largeur de coupe peut être supérieure à l'épaisseur du copeau non déformé.
- 2) Dans la simulation, l'outil de coupe est considéré comme parfaitement élastique, tandis qu'en réalité, l'élasticité de l'outil est comparativement plus grande que celle de la pièce.
- 3) La simulation suppose que le tranchant de l'outil est parfaitement affûté, alors qu'en réalité, il y a une perte progressive de la netteté du tranchant lors des expériences.
- 4) L'interface entre l'outil et la pièce est modélisée avec certaines hypothèses sur le frottement.
- 5) L'approche de modélisation numérique ne permet pas de dériver des équations prédictives généralisées et nécessite une expertise pour interpréter les résultats.
- 6) Les logiciels de simulation sont limités à tester les effets des facteurs individuels et ne peuvent pas prendre en compte les fortes interactions qui existent entre les facteurs dans les procédés d'usinage par enlèvement de matière.

4.4- Modélisation des relations entre les paramètres d'entrée-sortie

La modélisation consiste à établir une relation entre les entrées et les sorties d'un processus, ce qui est essentiel pour optimiser les processus. On utilise généralement des modèles mécaniques ou empiriques pour développer un modèle mathématique explicite, qui est régi par les paramètres du processus. Les modèles numériques basés sur l'analyse par éléments finis fonctionnent selon les principes des concepts analytiques ou numériques, qui utilisent de nombreuses hypothèses pour établir la relation entrée-sortie. Cependant, ces modèles ne fournissent pas une compréhension détaillée de la physique du processus, comme l'influence des propriétés des matériaux, de l'environnement de coupe et de la géométrie sur la mécanique de coupe. C'est pourquoi les modèles empiriques, qui se basent sur les résultats d'expériences pratiques, sont souvent préférés à la modélisation numérique.

Beaucoup de recherches ont été menées pour développer des outils de modélisation qui donnent des résultats satisfaisants. Cependant, ces outils ont différentes contraintes et limitations qui peuvent limiter leur utilité pratique. Il est donc important de bien comprendre les forces et les faiblesses d'un modèle avant de choisir l'outil approprié. Dans cette optique,

nous allons maintenant examiner les modèles empiriques bien connus qui sont utilisés pour établir des relations entrées-sorties.

4.4.1- Méthode Taguchi

Genichi Taguchi a développé des concepts de conception robuste visant à améliorer la productivité, la fiabilité et l'usinabilité d'un processus. Sa technique, appelée méthode Taguchi (Taguchi Method), est une méthodologie systématique utilisée dans la conception robuste. Elle consiste à utiliser des données expérimentales planifiées pour effectuer une analyse significative et obtenir des conclusions objectives. En 1980, l'application de la méthode Taguchi a permis à de nombreux fabricants américains, tels que American Telephone and Telegraph et Ford Motor, d'améliorer la qualité de leurs produits et de leurs processus.

La méthode Taguchi utilise une conception de réseau orthogonal flexible bien définie, ce qui permet à l'expérimentateur de connaître les contributions des différents facteurs sur les résultats cibles, tout en minimisant les coûts et les ressources nécessaires pour réaliser les essais expérimentaux. Cette méthode a suscité beaucoup d'intérêt chez les chercheurs et les ingénieurs, qui l'ont utilisée pour résoudre des problèmes liés à de nombreux processus de coupe de métaux. Elle a permis de réduire la complexité et le temps de calcul nécessaires pour la simulation, comme FEA et ABAQUS. Les paramètres de coupe (vitesse de coupe, vitesse d'avance, profondeur de coupe et angle de coupe) ont été intentionnellement modifiés pour observer leur effet sur la force de coupe, la température et l'usure des outils avec différents matériaux de coupe. De nombreux chercheurs ont utilisé la méthode Taguchi pour déterminer les paramètres et leurs niveaux de fonctionnement optimaux pour obtenir une finition de surface minimale, une usure minimale des outils et des forces de coupe minimales.

Malgré ses nombreux avantages, la méthode Taguchi est critiquée par certains chercheurs pour les raisons suivantes :

1. La méthode Taguchi utilise des tableaux orthogonaux limités, ce qui limite la possibilité de tester toutes les interactions potentielles d'un processus,
2. La méthode Taguchi peut aboutir à une solution sous-optimale car elle se concentre principalement sur la réduction du nombre d'essais expérimentaux et des coûts associés,

3. La méthode Taguchi n'utilise pas de modèles empiriques ou mécaniques pour exécuter les expériences et fournir des connaissances détaillées ou une vue d'ensemble des paramètres nécessaires pour surveiller ou contrôler le processus,
4. La méthode Taguchi est adaptée pour optimiser une seule réponse, tandis que d'autres méthodes sont essentielles pour optimiser simultanément plusieurs réponses,
5. La méthode Taguchi ne permet pas d'établir des équations de réponse, c'est-à-dire des relations entrée-sortie, qui sont essentielles pour effectuer une analyse significative et optimiser le processus.

4.4.2- Méthodologie de surface de réponse

La méthodologie de surface de réponse (RSM) est un ensemble d'outils mathématiques et statistiques visant à résoudre des problèmes d'ingénierie complexes impliquant plusieurs facteurs, tout en limitant les efforts, les coûts et le temps nécessaires pour mener des essais. La RSM permet de cartographier la relation entre les entrées et les sorties d'un processus à partir des données expérimentales collectées. Contrairement à la méthode Taguchi, qui peut échouer à tester tous les effets des facteurs d'interaction et à identifier les relations non linéaires, la RSM peut surmonter efficacement ces limites et permettre de dériver des équations de régression. Pour résoudre les problèmes multifactoriels liés aux procédés d'usinage, de nombreux chercheurs ont utilisé des outils de modélisation et d'optimisation tels que la conception composite centrale (CCD) ou la conception Box-Behnken (BBD).

La figure 4.1 illustre les étapes de la modélisation mathématique par cette technique, qui comprend la planification des expériences, la collecte des données, la construction, l'analyse, et la validation du modèle. La RSM, est couramment utilisée dans divers domaines tels que l'ingénierie, la chimie, la biotechnologie, et d'autres domaines où l'optimisation des processus est cruciale. Elle permet de réduire le nombre d'expérimentations nécessaires tout en fournissant des informations précieuses sur la manière d'ajuster les paramètres du processus pour obtenir des résultats optimaux.

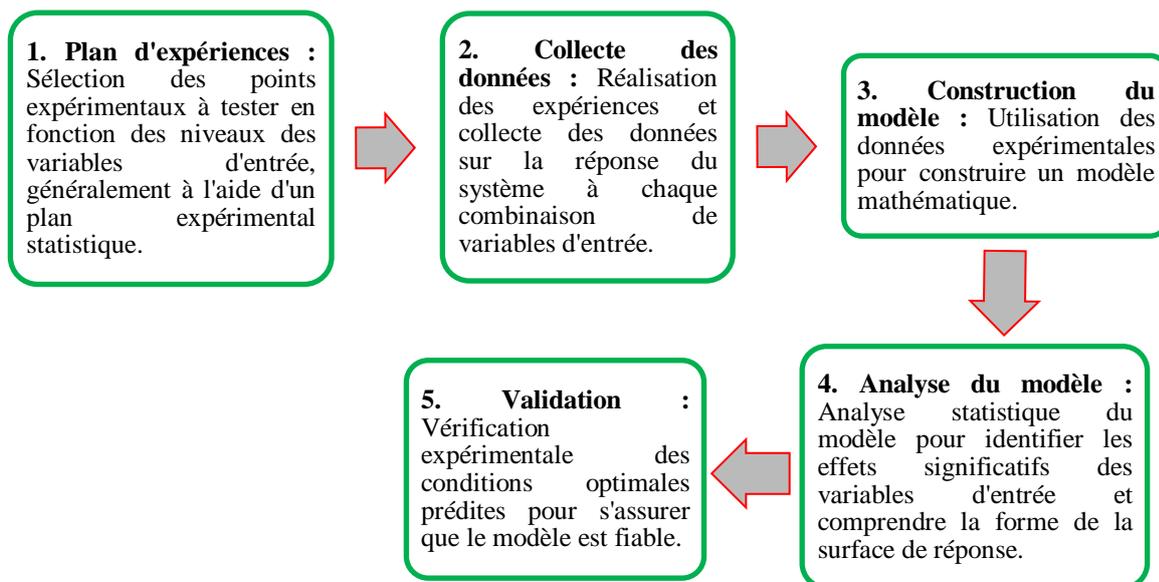


Figure 4.1- Etapes de modélisation mathématique par RSM.

Le plan d'expérience (Design of Experiment, DoE) est un ensemble de méthodes et de procédures principalement utilisées pour analyser des données relatives à des variables spécifiques dans le cadre d'un problème de recherche particulier. Il existe plusieurs types de DoE, tels que le plan factoriel complet, le plan factoriel fractionné, le plan de Plackett-Burman, le plan composite central rotatif et le plan de Box-Behnken, ...etc. Chacun de ces plans présente des avantages et des inconvénients, et le choix du plan à utiliser dépend principalement de l'objectif de la recherche et des ressources disponibles. La première étape dans l'utilisation du DoE consiste à définir les variables indépendantes et dépendantes. Les variables indépendantes, ou facteurs, peuvent être modifiées à différents niveaux ou valeurs, tandis que la variable dépendante, ou réponse, est influencée par les facteurs. Pour valider le modèle mathématique obtenu avec le DoE et pour déterminer un ensemble de conditions, il est nécessaire de calculer le résidu, qui est la différence entre le résultat expérimental et le résultat estimé. Un bon modèle mathématique présente une valeur de résidu faible.

Lors de l'évaluation de la signification du modèle mathématique, le test t et l'analyse de la variance (ANOVA) sont les méthodes statistiques les plus couramment utilisées. Le test t est employé pour comparer deux échantillons ou traitements, tandis que l'ANOVA est utilisée lorsqu'il y a plus de deux traitements. Pour les deux analyses, le niveau de signification doit être fixé, par exemple à 5 % ($p < 0,05$), pour indiquer que les traitements présentent une différence statistiquement significative. Le niveau de signification est également appelé « α », qui représente la probabilité de rejeter une hypothèse nulle. Ainsi, lorsque la valeur de « p »

calculée par l'ANOVA est inférieure au niveau de signification, les résultats peuvent être considérés comme significatifs, tandis qu'une valeur de « p » supérieure au niveau de signification signifie que l'hypothèse nulle est valide et que les résultats ne sont pas significativement différents, ce qui implique qu'il n'y a pas de preuves pour rejeter l'hypothèse nulle.

4.4.2.1- Plan factoriel complet

Le plan factoriel complet (FFD : Full Factorial Design) est généralement utilisé dans la méthodologie de surface de réponse (RSM). Les facteurs (k) et leurs niveaux sont combinés de manière à ce que le plan inclue toutes les combinaisons possibles. Ce type de DoE fournit les effets principaux ainsi que les effets d'interaction. Le facteur est le paramètre que le chercheur souhaite évaluer expérimentalement, et le niveau est une valeur qu'un facteur peut prendre. Le FFD peut être symétrique, lorsque le nombre de niveaux est identique pour tous les facteurs étudiés, ou asymétrique, lorsque chaque facteur a un nombre distinct de niveaux.

Les plans à deux niveaux, 2^k , sont les formes simples d'un plan orthogonal. Ils sont généralement utilisés pour sélectionner des variables et analyser leurs facteurs, et ils comprennent l'analyse de k facteurs à deux niveaux : haut (+) et bas (-). Le FFD à deux niveaux est bénéfique dans la phase initiale de l'expérience, en particulier lorsque le nombre de facteurs est ≤ 4 . Une hypothèse pour les facteurs à 2 niveaux est que la réponse est presque linéaire avec une faible et une valeur élevée pour chaque facteur. Une représentation graphique du FFD est présentée dans la figure 4.2. Les plans à trois niveaux ou plus sont plus fréquemment utilisés pour construire la surface de réponse, car ils montrent les effets non seulement des facteurs linéaires, mais aussi des facteurs quadratiques.

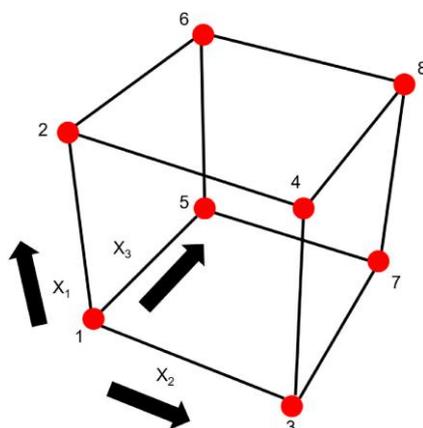


Figure 4.2- Représentation graphique du plan FFD (2^3).

Le temps, l'effort et le coût des expériences augmentent avec le nombre de facteurs à étudier. De plus, l'augmentation du nombre de paramètres et de niveaux rend les expériences plus complexes, ce qui peut rendre un FFD impraticable. Dans ces cas, un plan factoriel fractionné peut être une alternative viable. Des notes concises sur ce plan (FFD) sont présentées dans la figure 4.3.

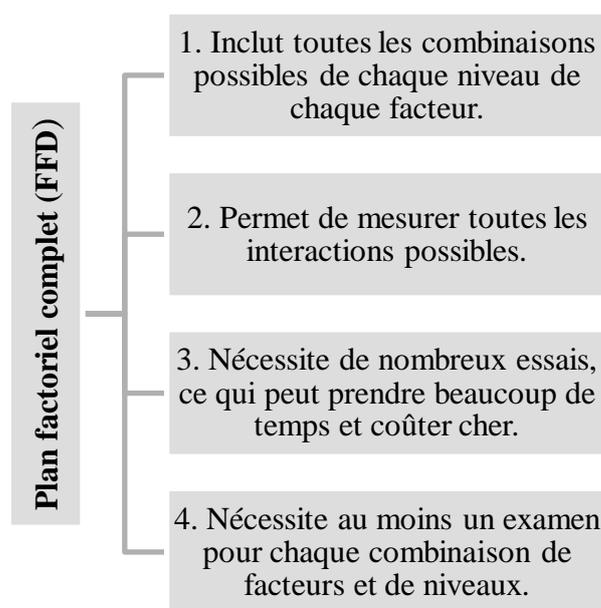


Figure 4.3- Un bref résumé du plan factoriel complet (FFD).

Les conceptions factorielles fractionnaires, telles que CCD, BBD et Taguchi, sont dérivées à partir de matrices expérimentales factorielles complètes en utilisant une structure d'alias sélectionnée. Par exemple, pour étudier trois facteurs (X_1 , X_2 et X_3) à trois niveaux chacun (faible, moyen et élevé), une matrice de taille expérimentale de 27 ($3^3 = \text{niveaux}^{\text{Facteurs}}$) peut être utilisée. Les points expérimentaux peuvent être représentés géométriquement sur les sommets d'un cube, les coins ou les bords, ainsi que sur l'origine et le milieu des faces.

4.4.2.2- Plan factoriel fractionnaire

Un plan factoriel fractionnaire (Fractional Factorial Design) utilise un sous-ensemble sélectionné des conditions expérimentales d'un plan factoriel complet, ce qui le rend plus économique en réduisant le nombre d'expériences. Cependant, cette réduction limite la capacité à distinguer certains effets des facteurs étudiés. Ainsi, le degré de fractionnement dépend de la résolution souhaitée : les plans de faible résolution se concentrent sur les effets

principaux en négligeant les interactions, tandis que les plans de haute résolution permettent d'analyser à la fois les effets principaux et les interactions.

La notation fréquemment utilisée pour ce type de plan est q^{k-p} , où q représente le nombre de niveaux analysés pour chaque facteur ; k est le nombre de facteurs utilisés dans l'analyse, et p est la taille de la fraction du plan factoriel complet appliqué.

4.4.2.3- Conception composite centrale

Un plan factoriel complet courant est le plan factoriel à trois niveaux, où chaque facteur peut prendre trois valeurs : bas, central et élevé. Pour k facteurs à trois niveaux, le nombre total d'expériences nécessaires est de 3^k . Cependant, un inconvénient majeur de ce plan est le grand nombre d'essais expérimentaux nécessaire, qui peuvent souvent générer des interactions d'ordre supérieur inutiles. Il est généralement plus avantageux d'utiliser le plan factoriel à trois niveaux lorsque le nombre de facteurs ne dépasse pas cinq. De plus, les modèles du second ordre peuvent considérablement améliorer le processus d'optimisation, surtout dans le cadre de ce plan. Pour éviter les difficultés liées à l'ajustement des modèles de polynômes de second ou d'ordre supérieur dans les plans factoriels classiques, Box et Wilson ont introduit le plan composite central (CCD : Central Composite Design) en 1951.

Un plan composite central est le design de surface de réponse le plus fréquemment utilisé. Le CCD est particulièrement utile dans le cadre d'expérimentations chronologiques, car il peut être utilisé pour s'appuyer sur des expériences factorielles antérieures avec l'ajout de points axiaux et de points centraux. Un CCD peut être utilisé pour (i) approximer efficacement les termes du premier et du deuxième ordre ; et (ii) modéliser une variable de réponse en utilisant la courbure grâce à l'ajout de points centraux et axiaux à un plan factoriel, comme le montre la figure 4.4.

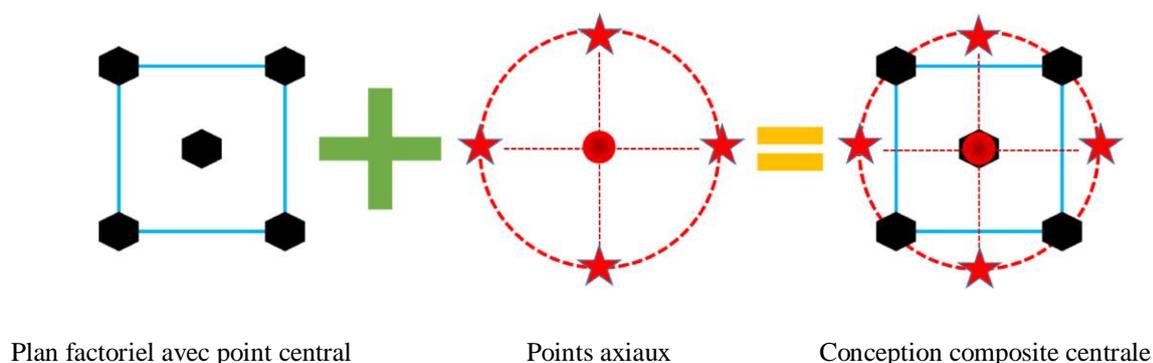


Figure 4.4- Représentation graphique de CCD.

Le nombre d'expériences pour ce design peut être obtenu à partir de l'équation 4.1.

$$N = K^2 + 2K + n \quad (4.1)$$

Où k représente le nombre de facteurs, et n représente le nombre de répétitions du point central.

La courbure de la surface de réponse peut être estimée à partir des points situés au centre du domaine expérimental et des points “étoiles” en dehors de ce domaine. Dans un plan factoriel, les niveaux des points sont ± 1 , tandis que dans un design “étoiles”, ils sont $\pm\alpha$, où $|\alpha| \geq 1$. Le CCD se divise en trois catégories, comme illustré dans la figure 4.5. La valeur du paramètre α est déterminée en fonction des possibilités de calcul et de la précision requise pour estimer la surface de réponse. La performance des prédictions dépend de la position des points. Il est crucial de noter que le réglage de la valeur α et le nombre d'essais au centre du domaine influencent la précision de l'estimation.

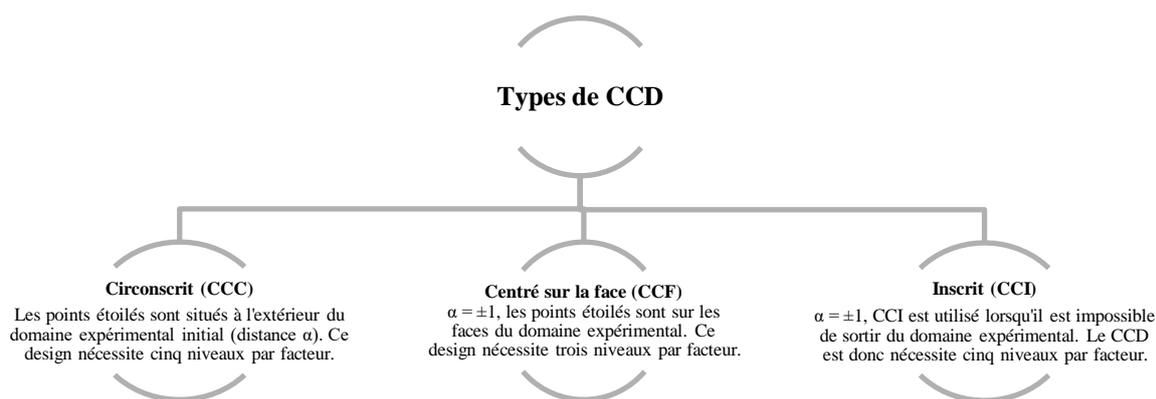


Figure 4.5- Trois types de CCD.

4.4.2.4- Plan de Plackett-Burman

Une expérience avec un plan de Plackett-Burman (PBD) utilise un nombre N d'essais expérimentaux et permet de tester jusqu'à $N - 1$ facteurs (N étant défini comme un multiple de 4). Le PBD ne prend pas en compte les interactions possibles entre les facteurs. Par conséquent, cet outil est uniquement utile pour estimer l'effet principal des facteurs impliqués dans le processus, et il ne peut pas être utilisé pour obtenir la surface de réponse lors de l'optimisation dudit processus.

Comme le plan 2^k , le plan de Plackett-Burman (PBD) permet deux niveaux pour chacun des facteurs de contrôle k ($-1, +1$), mais il nécessite un nombre d'expériences beaucoup plus

réduit. Pour construire un PBD avec k facteurs, la première ligne où les éléments sont égaux à -1 ou $+1$ doit être sélectionnée de manière à ce que le nombre de 1 soit égal à $(k + 1)/2$ et le nombre de -1 soit égal à $(k - 1)/2$. Ensuite, $k - 1$ lignes sont générées à partir de la première ligne en effectuant des décalages cycliques vers la droite $k - 1$ fois. Enfin, la dernière ligne du plan est ajoutée avec uniquement des -1 . Pour que le PBD soit utilisé correctement, il suggéré de choisir une matrice avec quatre essais ou plus que le nombre de facteurs à étudier.

4.4.2.5- Plan Box-Behnken

Un autre type de plan de surface de réponse est connu sous le nom de conception de box-Behnken (BBD). Un BBD est une catégorie de plan de surface de réponse qui ne comprend pas un plan factoriel. Par conséquent, le BBD n'est pas basé sur des plans factoriels complets ou fractionnaires. Cependant, le plan de Box-Behnken permet une prédiction efficace des coefficients d'ordre un et deux, car il comporte souvent un nombre plus réduit de points de plan. En conséquence, le BBD peut être moins chronophage et plus économique. Toutefois, comme le BBD ne possède pas de plan factoriel intégré, contrairement au CCD, il n'est pas adapté aux travaux expérimentaux chronologiques.

Bien que le BBD couvre mal les coins de l'espace de conception non linéaire, il est considéré comme plus performant et efficace que d'autres plans de surface de réponse, tels que le FFD à trois niveaux et le CCD. Le BBD est utilisé pour produire une réponse d'ordre supérieur avec moins d'essais qu'une méthode factorielle classique. Le BBD, ainsi que le CCD, supprime certains essais sélectionnés afin de maintenir la surface d'ordre supérieur.

Le BBD est rotatif et nécessite trois niveaux par facteur. À l'instar du CCD, le BBD peut s'adapter au modèle quadratique complet du plan de surface de réponse [58]. Dans le BBD, les combinaisons de traitements sont situées aux points médians des arêtes du cube et au centre, comme illustré à la figure 4.6. Ce plan de surface de réponse doit être pris en compte pour les travaux expérimentaux comportant plus de deux facteurs et lorsque l'optimum est attendu au milieu des intervalles des facteurs.

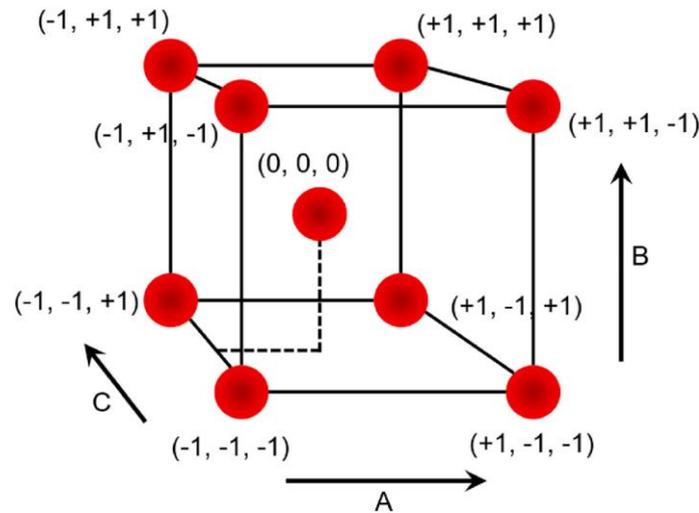


Figure 4.6- Représentation graphique de BBD.

L'équation 4.2 donne le nombre d'expériences (N) requis pour le plan de Box-Behnken, où k est le nombre de facteurs et C est le nombre de réplicats du point central :

$$N = 2 \times K(K - 1) + C \quad (4.2)$$

Ce plan est également utilisé dans la méthodologie de surface de réponse (RSM) pour obtenir le modèle mathématique du second ordre, ce qui permet d'optimiser le processus étudié. L'avantage du plan de Box-Behnken est qu'il identifie les problèmes liés aux limites expérimentales et évite des combinaisons de traitements inutiles.

Bien que l'utilisation de la méthodologie de surface de réponse (RSM) présente des avantages pour l'analyse des données expérimentales, cette technique ne précise pas quel type de plan doit être utilisé pour collecter et analyser les informations. Le BBD ne nécessite que trois niveaux pour chaque facteur : 1, 0, et +1, tandis que les CCD nécessitent cinq niveaux ($-\alpha$, 1, 0, +1, et α). Tous les points de plan (à l'exception du centre) ont une longueur de deux ; c'est-à-dire qu'ils se trouvent tous sur la même sphère. Ces plans sont particulièrement adaptés aux régions sphériques. En raison de cette propriété sphérique, il est recommandé d'effectuer au moins trois à cinq essais au point central. Par exemple, dans un BBD avec $k = 3$, tous les points se trouvent sur les arêtes, à une distance de $\sqrt{2}$ du centre. Il n'y a ni points factoriels ni points sur les faces. Bien que le BBD couvre les arêtes, il ne prend pas en compte l'ensemble du cube, car aucun point ne se trouve aux coins, même à une distance de $\sqrt{3}$ du centre. En raison de sa structure sphérique et de la disposition des plans rotatifs (ou quasi-rotatifs), il est

recommandé d'effectuer plusieurs essais au centre. Pour $k = 3$, trois à cinq essais sont nécessaires au point central.

4.4.3- Réseau de neurones artificiel

La modélisation mathématique basée sur des données expérimentales, également appelée modélisation empirique, est une approche précieuse pour comprendre un phénomène étudié, développer des modèles prédictifs, optimiser les paramètres du système et prendre des décisions éclairées dans divers domaines de l'ingénierie, y compris l'industrie, la finance, la santé, la recherche scientifique, ...etc. Elle permet d'économiser du temps et des ressources tout en fournissant des informations pour améliorer la performance des processus étudiés. A cet égard, le réseau de neurones artificiel (ANN) est un outil puissant pour la modélisation empirique. Le ANN est un modèle de machine learning qui peuvent apprendre à partir de données d'entrée et découvrir des modèles complexes dans les données. Cette approche non conventionnelle de modélisation basée sur l'intelligence artificielle comprend les éléments suivants : les valeurs d'entrée (x_i), les poids (W_i), la fonction de sommation (Σ), le biais (b), la fonction d'activation (f) et la sortie (Y).

La figure 4.2 illustre la structure du réseau de neurones artificiels de type 5-3-1.. Cela signifie qu'il comporte une couche d'entrée avec 5 nœuds (représentant le nombre de facteurs), une seule couche cachée avec 3 nœuds, et une couche de sortie avec 1 nœud pour la réponse.

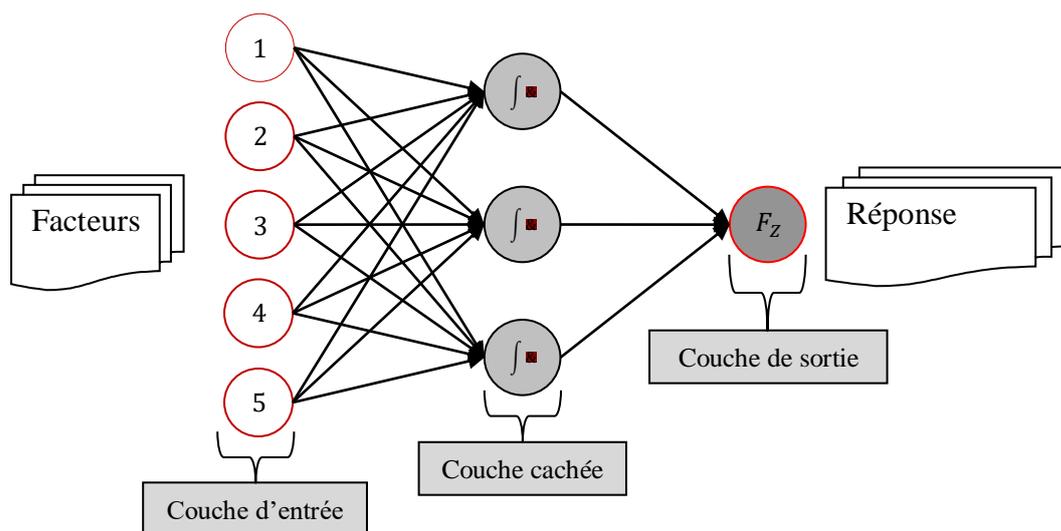


Figure 4.2- Représentation graphique de la structure des modèles ANN.

L'algorithme de rétro-propagation basé sur le gradient descendant avec un taux d'apprentissage de $n = 0,1$ est généralement utilisé lors de cette étape, où le comportement

du réseau est modifié progressivement jusqu'à ce qu'il atteigne le comportement souhaité. Pour évaluer la performance et la généralisation du réseau de neurones, le coefficient de détermination R^2 est analysé au niveau de confiance de 95% ($\alpha = 0,05$). Une fois que le réseau de neurones est entraîné, il est essentiel de le valider sur des données non vues auparavant afin d'évaluer sa performance et son niveau de généralisation.

4.5- Outils d'optimisation

4.5.1- Optimisation mono-objectif

4.5.1.1- Méthode de Taguchi

Afin d'analyser les résultats expérimentaux, le rapport signal/bruit (S/N) est utilisé parce que c'est le critère le plus essentiel de la méthode de Taguchi. Le rapport signal/bruit S/N de cette technique permet d'optimiser les facteurs de contrôle. Ces derniers sont les variables qui peuvent être contrôlées de manière pratique et économique. Le but est de minimiser ou de maximiser un paramètre de sortie afin d'améliorer la qualité de la réponse. Selon cette méthode pour obtenir les conditions optimales, le rapport S/N doit avoir une valeur maximale ou minimale selon le cas. Les rapports S/N sont généralement divisés en trois catégories données par les trois équations suivantes :

Nominal est le meilleur :

$$S/N = 10 * \log\left(\frac{\bar{y}}{s_y^2}\right) \quad (4.3)$$

Le plus petit est le meilleur (minimiser) :

$$S/N = -10 * \log\left(\frac{1}{n} \sum_{i=1}^n y_i^2\right) \quad (4.4)$$

Le plus grand est le meilleur (maximiser) :

$$S/N = -10 * \log\left(\frac{1}{n} \sum_{i=1}^n \frac{1}{y_i^2}\right) \quad (4.5)$$

4.5.2- Optimisation multi-objectif

4.5.2.1- Analyse relationnelle grise

L'analyse relationnelle grise (GRA), est une approche robuste d'analyse multicritère utilisée pour évaluer les relations entre des ensembles de données. La méthode GRA s'adapte bien à des situations où les données sont incomplètes ou incertaines car elle repose sur la théorie des

ensembles flous et de la relation grise. L'objectif principal de GRA est de mesurer le degré de relation entre différentes séries de données. Cela peut inclure la comparaison de performances entre différents facteurs ou la recherche de relations entre des variables dans des ensembles de données. Le GRA figure parmi les techniques d'optimisation multi-objectif les plus couramment adoptées dans ces dernières années. GRA est couramment utilisée dans des domaines de l'ingénierie mécanique tels que le soudage, le frittage, le tournage d'alliage de magnésium, la rectification, ...etc. Elle peut aider à la prise de décision et à l'optimisation. GRA implique plusieurs étapes, dont la normalisation des données, le calcul des relations grises entre les séries de données, la détermination des poids attribués à chaque série, et enfin, la synthèse des résultats pour obtenir une mesure globale de la relation. Les étapes mentionnées ci-dessous offrent une orientation générale pour comprendre le processus de base de l'analyse relationnelle grise.

a) Normalisation des données

Avant d'appliquer la méthode GRA, les données expérimentales doivent souvent être normalisées pour les rendre comparables. Cette normalisation vise à ajuster les différentes échelles de mesure des paramètres dépendants dans l'intervalle. L'objectif est d'assurer une influence égale de chaque réponse dans l'analyse, permettant ainsi une évaluation équitable. Cette étape est réalisée en fonction de l'objectif d'optimisation (maximisation ou minimisation).

- ✧ Pour la minimisation d'un paramètre de sortie, tels que R_a , T_c , et E_{cs} dans notre étude, la séquence originale est normalisée à l'aide de l'équation 4.6 :

$$x_i(k) = \frac{\max(x_i^0(k)) - x_i^0(k)}{\max(x_i^0(k)) - \min(x_i^0(k))} \quad (4.6)$$

- ✧ Pour la maximisation d'un paramètre de sortie, bien qu'aucun cas ne soit étudié ici, la séquence originale de normalisation est réalisée par l'équation 4.7 :

$$x_i(k) = \frac{x_i^0(k) - \min(x_i^0(k))}{\max(x_i^0(k)) - \min(x_i^0(k))} \quad (4.7)$$

Où :

- $x_i(k)$: la valeur normalisée,
- $\max(x_i^0(k))$: la plus grande valeur de $x_i^0(k)$ pour la $k^{\text{ième}}$ réponse,
- $\min(x_i^0(k))$: la plus petite valeur de $x_i^0(k)$ pour la $k^{\text{ième}}$ réponse.

b) Calcul du GRC

Afin de calculer les degrés de similarité entre chaque paire de séries de données, les données normalisées de l'étape précédente sont utilisées. Plusieurs méthodes peuvent être utilisées pour calculer la similarité, telles que le coefficient relationnel gris (GRC). La conversion des degrés de similarité en GRC se fait en utilisant la fonction de transfert spécifiée dans l'équation 4.8. Sa valeur est située dans l'intervalle de 0 à 1.

$$\eta_i(k) = \frac{\Delta_{\min} + \omega \Delta_{\max}}{\Delta_{0i}(k) + \omega \Delta_{\max}} \quad (4.8)$$

Où :

- * $\Delta_{0i}(k)$: la différence en valeur absolue entre $x_0^k(k)$ et $x_i^k(k)$ (équation 4.9),
- * Δ_{\min} : la plus petite valeur de $\Delta_{0i}(k)$ (équation 4.10),
- * ω : le coefficient de la fonction de transfert généralement égal à 0.5,
- * Δ_{\max} : la plus grande valeur de $\Delta_{0i}(k)$ (équation 4.11),

$$\Delta_{0i} = \|x_0(k) - x_i(k)\| \quad (4.9)$$

$$\Delta_{\min} = \min_{\forall j \in i} \min_{\forall k} \|x_0(k) - x_i(k)\| \quad (4.10)$$

$$\Delta_{\max} = \max_{\forall j \in i} \max_{\forall k} \|x_0(k) - x_i(k)\| \quad (4.11)$$

c) Calcul de GRG

Pour identifier la série de données de référence, soit la meilleure solution, on calcule les degrés de relation entre chaque série de données à l'aide du grade relationnel gris (GRG (α_i)) de l'équation 4.12. La série ayant le GRG le plus élevé est considérée comme la meilleure solution. Ceci permet d'identifier les séries représentant la meilleure solution par rapport à la moins performante.

$$\alpha_i = \frac{1}{n} \sum_{i=1}^n \eta_i(k) \quad (4.12)$$

Où :

- n : le nombre des caractéristiques de performance (dans notre cas $n = 3$).

4.5.2.2- Approche de la fonction de désirabilité

En 1980, Derringer et Suich ont créé une méthode d'optimisation pour plusieurs sorties, appelée approche de fonction de désirabilité. Les méthodes BBD, CCD et Taguchi sont

couramment utilisées pour établir une relation entre les entrées et les sorties, mais elles ne sont pas idéales pour trouver les valeurs extrêmes de la réponse (c'est-à-dire l'optimisation). La méthode DF (Desirability Function) surmonte ces lacunes en matière d'optimisation. Elle permet d'estimer séparément les surfaces de réponse pour chaque variable de sortie. Les valeurs prédites pour chaque surface de réponse sont converties en une valeur sans dimension appelée " d_i ", qui varie entre zéro (réponse indésirable) et un (réponse totalement souhaitable). Pour l'optimisation de plusieurs réponses, la valeur composite de désirabilité (D) est calculée à l'aide de l'équation 4.13.

$$D = (d_1^{w_1} \times d_2^{w_2} \times \dots \times d_m^{w_m})^{1/m} \quad (m : \text{nombre de réponses})$$

Désirabilité composite

$$= (\text{désirabilité}_1^{\text{pondération}_1} \times \text{désirabilité}_2^{\text{pondération}_2} \times \dots \times \text{désirabilité}_m^{\text{pondération}_m})^{1/m} \quad (4.13)$$

Il est fréquent d'utiliser l'optimisation multi-objectif pour résoudre des problèmes où les réponses sont en conflit, c'est-à-dire qu'elles nécessitent à la fois une maximisation et une minimisation. Pour trouver une solution optimale qui réponde à ces deux objectifs, il est nécessaire de formuler une équation mathématique. Si la fonction de sortie est de type maximisation, alors l'équation 4.14 sera utilisée.

$$Y_x = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (4.14)$$

De manière similaire, si la fonction de réponse est de type minimisation, alors l'équation 4.15 sera utilisée pour l'optimisation multi-objectif.

$$Y_x = \frac{x_{\max} - x}{x_{\max} - x_{\min}} \quad (4.15)$$

Les termes x_{\max} et x_{\min} sont les valeurs maximale et minimale de la réponse x . Le terme Y_x est la valeur de sortie correspondant à la valeur souhaitable de la réponse.

La méthode DF est un outil simple et flexible qui utilise des concepts intuitifs pour résoudre l'optimisation à objectif unique et à objectifs multiples. Cependant, cette méthode présente certaines limitations, qui sont énumérées ci-dessous :

- ✧ Dans de nombreux cas, la DF attribue une importance égale (c'est-à-dire une fraction pondérale) à toutes les réponses. Cependant, dans le monde réel, il est rare que toutes les réponses soient également importantes. Par conséquent, attribuer une importance égale à toutes les réponses peut conduire à des résultats trompeurs.
- ✧ La DF utilise un chemin de recherche déterministe avec des règles de transition spécifiques pour localiser les solutions. Toutefois, le déplacement d'une solution vers une autre peut conduire à de nombreuses solutions locales, ce qui peut entraîner une perte de diversité dans les résultats finaux.

4.5.2.3- Outils d'optimisation de l'informatique douce

Les outils informatiques doux (Soft Computing Tools) sont de nature stochastique, fonctionnant avec des ensembles de règles probabilistes souhaités pour estimer des solutions optimales dans un espace de recherche multidimensionnelle. Cette méthode permet de cibler les solutions globales tout en réduisant les efforts et les coûts de calcul. Contrairement aux techniques de recherche mathématique itérative (telles que la programmation linéaire ou d'objectifs) et aux outils d'optimisation statistique (DFA, GRA, etc.), les algorithmes de l'informatique douce sont exempts de dérivées et ne dépendent pas des relations d'entrée-sortie. Il est intéressant de noter que l'optimisation basée sur l'enseignement et l'apprentissage (TLBO) et l'algorithme Jaya ne nécessitent pas de réglage de paramètres spécifiques pour atteindre les solutions globales. Les outils informatiques doux tels que GA, PSO, TLBO, et Jaya sont considérés comme des outils appropriés pour l'optimisation des processus de fabrication. Par conséquent, il existe de nombreuses possibilités d'application des outils informatiques doux dans la modélisation et l'optimisation du processus de tournage dur. Cependant, certaines limites de l'optimisation basée sur l'informatique douce sont à noter, notamment :

1. Les logiciels d'optimisation ne garantissent pas toujours de trouver la solution globale, même s'ils convergent.
2. Il n'existe pas de normes spécifiques pour sélectionner les paramètres d'un algorithme d'optimisation tels que le poids d'inertie, la taille de l'essaim et la génération de PSO.
3. Plus il y a d'étapes et de paramètres dans un algorithme, plus cela peut augmenter la complexité ainsi que le coût, l'effort et le temps requis.

CHAPITRE 5 :

Application tribologique

5.1- Introduction

L'objectif principal de ce chapitre est d'évaluer les critères d'usinabilité lors du tournage de polypropylène non renforcé (PP), notamment la force tangentielle (F_z), la rugosité de surface (Ra), la température de coupe (T_c), et l'énergie de coupe spécifique (E_{cs}). Les essais expérimentaux ont été réalisés à sec et avec lubrification à quantité minimale (MQL). Les facteurs modifiés comprenaient la vitesse de coupe (V_c), la vitesse d'avance (f), la profondeur de coupe (ap) et le rayon du bec (r_ϵ). L'analyse de variance (ANOVA), la méthodologie de surface de réponse (RSM) et la fonction de désirabilité (DF) sont utilisées pour analyser, modéliser et optimiser les paramètres dépendantes et indépendantes. A cet égard, ce chapitre offre une référence précieuse pour le développement durable de l'usinage MQL des matériaux polymères.

5.2- Matériaux et méthodes

5.2.1- Pièce usinée et outil de coupe

Les échantillons étudiés sont de tiges pleines en polypropylène non renforcé (PP) produites par GEHR, une entreprise allemande. Ces tiges ont un diamètre de 80 mm et une longueur de 150 mm. Le choix du PP pour cette étude s'explique par le fait qu'il y a peu d'études sur l'usinage de ce matériau, malgré ses excellentes propriétés mécaniques et thermiques, comme le démontre le tableau 5.1.

Pour les essais de coupe, nous avons utilisé un tour CNC (GA-230) et un porte-outil (PSBNL 2525M 12) avec trois plaquettes en carbure revêtu : SNMG 120404, SNMG 120408 et SNMG 120412. Les opérations de coupe ont été effectuées en utilisant la technique du tournage avec lubrification à quantité minimale (MQL). Plus précisément, nous avons utilisé un appareil WERTE avec les paramètres suivants : débit de 26 ml/h, pression de 5 bars, distance de buse de 70 mm, angle de buse de 75 degrés et diamètre d'extrémité de buse de 1 mm. Le liquide de coupe pour légumes Erroil KT/2000 commercial a été comme lubrifiant, présente les

caractéristiques suivantes à une température ambiante de 20°C : une densité de 0,85 g/cm³ et une viscosité de 11 cst. Le schéma du dispositif expérimental est illustré dans la figure 5.1.

Tableau 5.1- Caractéristiques mécaniques et thermiques du PP.

1. Propriétés mécaniques	
Résistance de la limite d'élasticité à la traction (MPa)	36
Allongement à la limite d'élasticité (%)	9
Allongement à la rupture (%)	≥ 70
Dureté Shore-D	76
Module d'élasticité (E _i) (MPa)	1725
2. Propriétés thermiques	
Point de ramollissement Vicat VST/A/50 (°C)	150
Température de transition vitreuse (T _g) (°C)	- 20
Température de fusion (T _m) (°C)	165
Température de service maximale admissible (°C)	100

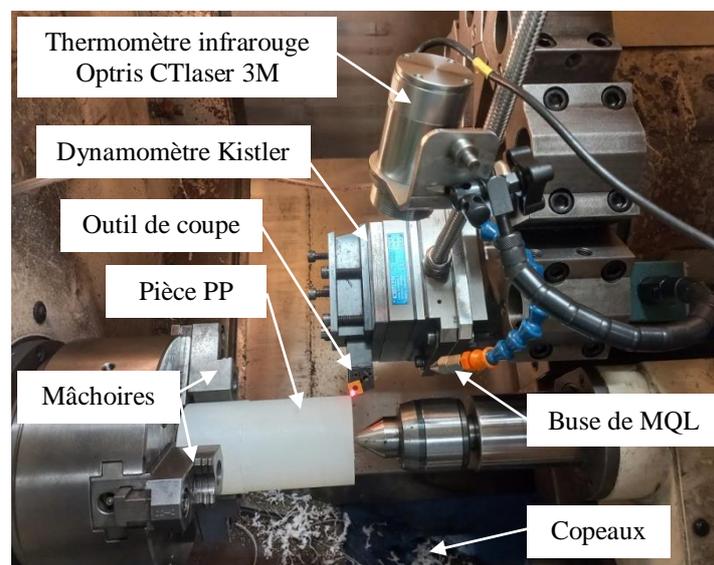


Figure 5.1- Etape expérimentale.

5.2.2- Opération de coupe et de mesure

Les variables indépendantes prises en compte dans cette étude sont les suivantes : la condition de refroidissement (cooling condition : CC), la vitesse de coupe (V_c), la vitesse d'avance (f), la profondeur de coupe (a_p) et le rayon du bec (r_ϵ), comme indiqué dans le tableau 5.2. Les plages de ces paramètres de coupe sont recommandées par GEHR, le fabricant du polymère PP, ainsi que par Sandvik Coromant, le fabricant des outils de coupe. Pour réduire le nombre d'expériences, nous avons choisi d'utiliser la méthode de Taguchi avec un plan orthogonal L₁₈ ($2^1 \times 3^4$), ce qui nous permet d'évaluer deux niveaux pour un facteur et trois niveaux pour

quatre facteurs. Cela nous permet de réduire le nombre d'expériences par rapport à un plan factoriel complet ($FFD = 2^1 \times 3^4 = 162$).

Tableau 5.2- Paramètres de coupe et leurs niveaux.

Niveau	CC	V_c (m/min)	f (mm/tr)	ap (mm)	r_e (mm)
1	Dry	250	0,05	1	0,4
2	MQL	375	0,1	1,8	0,8
3	–	500	0,15	2.6	1,2

Dans cette étude, la force de coupe tangentielle (F_z) a été mesurée à l'aide d'un dynamomètre Kistler de type 9257B équipé de quatre capteurs piézoélectriques (voir figure 5.2). Les signaux analogiques provenant des capteurs ont été convertis en signaux numériques à l'aide d'une carte d'acquisition, et les mesures ont été effectuées et analysées à l'aide du logiciel DynoWare.

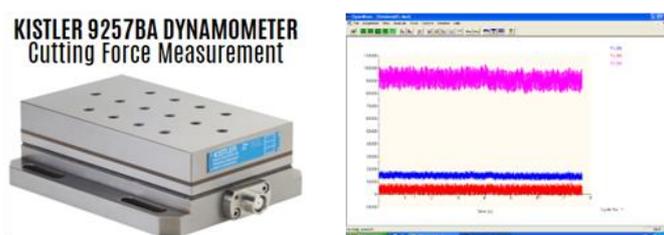


Figure 5.2- Dynamomètre Kistler 9257BA.

La température a été mesurée sans contact à l'aide du thermomètre infrarouge Optris CTlaser 3M H1 modèle double laser (voir figure 5.3). Cet appareil est capable de mesurer la température jusqu'à 1800 °C, utilisant une plage spectrale de 2,3 μm . De plus, il offre un temps de réponse rapide de 1 ms. Le double laser était focalisé sur la face de coupe pendant le processus d'usinage. Les données ont été collectées et analysées par le logiciel CompactConnect.



Figure 5.3- Thermomètre infrarouge Optris CTlaser 3M.

La rugosité moyenne arithmétique (R_a) a été mesurée à l'aide d'un rugosimètre SJ-210 équipé d'un palpeur de pointe en diamant ayant un rayon de courbure de $5 \mu\text{m}$ (voir figure 5.4). Le palpeur se déplaçait à une vitesse de $0,5 \text{ mm/s}$ lors des mesures. La longueur d'échantillonnage (L_n) était fixée à 5 mm , tandis que la longueur de coupe (cut-off, λ_c) était établie à $0,8 \text{ mm}$.



Figure 5.4- Rugosimètre digital SurfTest SJ 210.

5.2.3- Calcul de l'énergie de coupe spécifique (SCE)

Le calcul le plus simple de l'énergie de coupe spécifique (E_{cs}) est directement lié à la force tangentielle (F_z) sans nécessiter le calcul de la puissance de coupe (P_c) ni du taux d'enlèvement de matière (MRR). Cette relation est exprimée par l'équation 5.1.

$$E_{cs} = \frac{F_z}{f \cdot ap} \times 10^{-3} \quad (5.1)$$

Dans cette équation, E_{cs} représente l'énergie réelle (SCE) en J/mm^3 , F_z est la force de coupe tangentielle (principale) en N , f est la vitesse d'avance en mm/rev , et ap est la profondeur de coupe en mm .

5.3- Résultats et discussions

5.3.1- Résultats expérimentaux

Les résultats expérimentaux de la force de coupe tangentielle, de la rugosité de surface, de la température de coupe et de l'énergie de coupe spécifique sont présentés dans le tableau 5.3. La valeur minimale et maximale de l'effort tangentiel (F_z) sont respectivement $24,41 \text{ N}$ et $103,93 \text{ N}$. Le facteur de la qualité de surface (R_a) varie de $0,52 \mu\text{m}$ à $5,45 \mu\text{m}$. Le troisième facteur étudié (T_c) varie de $33,25 \text{ }^\circ\text{C}$ à $53,28 \text{ }^\circ\text{C}$. Enfin, la dernière réponse (E_{cs}) varie de $0,234 J/mm^3$ à $0,662 J/mm^3$.

Tableau 5.3- Résultats des paramètres de sortie (F_Z , R_a , T_c et E_{cs}).

N°	Facteurs					Réponses d'usinabilité			
	CC	V_c (m/min)	f (mm/tr)	a_p (mm)	r_ϵ (mm)	F_Z (N)	R_a (μm)	T_c (°C)	E_{cs} (J/mm ³)
1	1	250	0,05	1,0	0,4	33,10	4,52	45,98	0,662
2	1	250	0,10	1,8	0,8	55,10	4,74	48,35	0,306
3	1	250	0,15	2,6	1,2	103,9	5,16	49,01	0,266
4	1	375	0,05	1,0	0,8	31,10	4,11	53,28	0,622
5	1	375	0,10	1,8	1,2	68,65	4,36	49,22	0,381
6	1	375	0,15	2,6	0,4	95,24	5,45	46,27	0,244
7	1	500	0,05	1,8	0,4	51,74	3,59	47,93	0,575
8	1	500	0,10	2,6	0,8	88,27	3,32	45,56	0,340
9	1	500	0,15	1,0	1,2	64,90	3,45	45,95	0,433
10	2	250	0,05	2,6	1,2	56,32	1,4	39,46	0,433
11	2	250	0,10	1,0	0,4	24,41	2,53	34,54	0,244
12	2	250	0,15	1,8	0,8	63,25	2,51	35,13	0,234
13	2	375	0,05	1,8	1,2	36,82	1,01	38,91	0,409
14	2	375	0,10	2,6	0,4	66,66	1,79	35,26	0,256
15	2	375	0,15	1,0	0,8	43,08	2,19	35,57	0,287
16	2	500	0,05	2,6	0,8	57,10	0,52	37,59	0,439
17	2	500	0,10	1,0	1,2	33,31	1,09	37,9	0,333
18	2	500	0,15	1,8	0,4	63,15	1,98	33,25	0,234

5.3.2- Analyse de la variance (ANOVA)

5.3.2.1- ANOVA de F_Z

Les efforts de coupe jouent un rôle essentiel dans le processus d'usinage par enlèvement de matière. Leur impact peut être considérable sur la qualité, la précision, la durée de vie de l'outil et la productivité globale du processus. Le tableau 5.4 présente les résultats de l'analyse ANOVA de la force tangentielle (F_Z) lors de l'usinage du PP. Les valeurs de probabilité (P) et les pourcentages de contribution montrent que la profondeur de coupe (a_p) a le plus grand impact sur la variation de F_Z . Ensuite, la vitesse f est le deuxième facteur le plus important, avec une contribution de 27,45%. La lubrification MQL a également un effet significatif, avec une contribution de 14,30%. Par ailleurs, le rayon du bec (r_ϵ) a le moins d'influence parmi les paramètres étudiés. En résumé, a_p et f sont les deux facteurs les plus déterminants pour F_Z , suivis de près par MQL, tandis que r_ϵ et V_c ont un impact négligeable.

Tableau 5.4- ANOVA de F_z .

Source	SC	DF	MC	F-value	P-value	Cont. %	Remarques
Modèle	8370,23	5	1674,05	151,01	<0,0001	98,44	Significatif
CC	1215,74	1	1215,74	109,66	<0,0001	14,30	Significatif
V_c	41,66	1	41,66	3,76	0,076	0,49	Insignificatif
f	2334,39	1	2334,39	210,57	<0,0001	27,45	Significatif
ap	4705,27	1	4705,27	424,43	<0,0001	55,33	Significatif
r_ε	73,16	1	73,16	6,60	0,025	0,86	Significatif
Residual	133,03	12	11,09				
Cor total	8503,26	17					

5.3.2.2- ANOVA de R_a

La qualité de surface est un facteur crucial dans les critères d'usinabilité. Pour améliorer le fonctionnement d'une pièce mécanique de précision, il est essentiel de maximiser l'amélioration de sa texture de surface. Dans cette optique, la rugosité moyenne arithmétique (R_a) est le paramètre le plus déterminant pour évaluer l'état de surface obtenu par un procédé d'usinage avec une arête de coupe bien définie. L'ANOVA de cette variable dépendante (R_a) est présentée par le tableau 5.5. Dans cette analyse, tous les paramètres d'entrée sont significatifs car les valeurs de probabilité (P-value) sont inférieures à 0,05, à l'exception de la profondeur de coupe (ap), qui sa valeur de probabilité dépasse 0,05. Cela indique que ce facteur n'a pas d'impact significatif sur la variation de Ra . De plus, la condition de coupe (CC) est le facteur la plus importante, ce qui signifie qu'elle joue un rôle primordial dans la variation de la rugosité de surface, avec une contribution de 78.52%. Ensuite, la vitesse de coupe (V_c) est la deuxième variable indépendante qui influence la finition de la pièce PP, suivie de la vitesse d'avance (f) et enfin le rayon du bec (r_ε).

Tableau 5.5- ANOVA de R_a .

Source	SC	DF	MC	F-value	P-value	Cont. %	Remarques
Modèle	38,6983	5	7,7397	94,93	< 0,0001	97,53	Significatif
CC	31,1524	1	31,1524	382,09	< 0,0001	78,52	Significatif
V_c	3,9790	1	3,9790	48,80	< 0,0001	10,03	Significatif
f	2,6040	1	2,6040	31,94	< 0,0001	6,56	Significatif
ap	0,0052	1	0,0052	0,0639	0,805	0,01	Insignificatif
r_ε	0,9577	1	0,9577	11,75	0,005	2,41	Significatif
Residual	0,9784	12	0,0815				
Cor total	39,6766	17					

5.3.2.3- ANOVA de T_c

La température de coupe joue un rôle crucial dans l'usinage des polymères. Elle peut avoir un impact sur les propriétés mécaniques, la stabilité dimensionnelle et la qualité des pièces usinées. De plus, Les caractéristiques physiques des polymères, telles que leur point de fusion et leur conductivité thermique inférieures par rapport aux métaux, nécessitent une minimisation de la production de chaleur pendant la coupe. La tableau 5.6 présente ANOVA de cet aspect d'usinabilité (T_c). Il est clair que la vitesse V_c et la profondeur ap sont des facteurs insignifiants pour la variation de la température de coupe, car leurs valeurs de probabilité (P-value) sont supérieures à 0,05. En revanche, les trois autres paramètres (CC , f , et r_ε) sont significatifs, bien que leur importance varie. Notamment, le facteur CC est le plus déterminant dans la variation de T_c , avec le plus grand pourcentage de contribution égal à 87,72%.

Tableau 5.6- ANOVA de T_c .

Source	SC	DF	MC	F-value	P-value	Cont. %	Remarques
Modèle	653,351	5	130,670	50,78	< 0.0001	95,49	Significatif
CC	600,196	1	600,196	233,24	< 0.0001	87,72	Significatif
V_c	1,534	1	1,534	0,60	0,455	0,22	Insignificatif
f	26,910	1	26,910	10,46	0,007	3,93	Significatif
ap	0,000	1	0,000	0,00	0,990	0,00	Insignificatif
r_ε	24,711	1	24,711	9,60	0,009	3,61	Significatif
Residual	30,880	12	2,573				
Cor total	684,230	17					

5.3.2.4- ANOVA de E_{cs}

L'énergie de coupe spécifique (E_{cs}) est un indicateur clé utilisé dans l'analyse de l'usinabilité lors du tournage des matériaux. Le tableau 5.7 présente les résultats de test statistique ANOVA pour évaluer la signification des différentes variables indépendantes sur SCE. Dans ce tableau, la vitesse f est de probabilité inférieure à 0,05 et de contribution la plus grande égale à 55,78%. Ce qui signifie l'impact statistiquement significatif de ce facteur et leur importance à la variation de cette réponse. Ainsi, la condition de refroidissement de coupe (cooling condition CC) est identifiée comme le deuxième facteur le plus influent, tandis que la profondeur de coupe (ap) est le troisième paramètre d'entrée affectant l'énergie de coupe spécifique. Les deux autres facteurs (V_c et r_ε) ne sont pas significatifs, car leur probabilité est supérieure à 0,05 et leur contribution inférieure cumulée est inférieure à 1,5%.

Tableau 5.7- ANOVA de E_{cs} .

Source	SC	DF	MC	F-value	P-value	Cont. %	Remarques
Modèle	0,258201	5	0,051640	11,84	<0.0001	83,15	Significatif
CC	0,051066	1	0,051066	11,71	0,005	16,45	Significatif
V_c	0,003574	1	0,003574	0,82	0,383	1,15	Insignificatif
f	0,173221	1	0,173221	39,73	<0.0001	55,78	Significatif
ap	0,030203	1	0,030203	6,93	0,022	9,73	Significatif
r_ε	0,000137	1	0,000137	0,03	0,862	0,04	Insignificatif
Residual	0,052316	12	0,004360				
Cor total	0,310517	17					

5.3.3- Méthodologie de surface de réponse

La méthodologie de surface de réponse (RSM) est une méthode statistique qui permet de modéliser et d'explorer la relation mathématique entre les variables indépendantes et la réponse souhaitée, en utilisant des expérimentations systématiques. Elle est utilisée comme un outil empirique puissant de conception d'expériences (DOE) pour comprendre l'influence des paramètres indépendants sur les critères d'usinabilité. La méthode RSM améliore les résultats et évite les expériences répétées, ce qui est à la fois coûteux et chronophage. A cet égard, l'évaluation du coefficient de détermination (R^2) permet de mesurer l'importance et la précision du modèle obtenu grâce à cette technique.

5.3.3.1- Régression linéaire multiple

Les résultats des investigations du tableau 5.3 sont utilisés pour appliquer la régression linéaire multiple afin de déterminer les modèles de sortie. L'équation 5.2 représente le modèle empirique de la force principale, F_Z . Le coefficient de détermination R^2 de ce modèle est de 0,9844, ce qui signifie que 98,44% des variations de l'effort sont expliquées par ce modèle, laissant seulement 1,66% inexpliquées. Ce coefficient démontre une bonne corrélation entre le modèle et les données expérimentales.

$$F_Z = -0,76 - 16,44 CC + 0,01491V_c + 279,0f + 24,75ap + 6,17r_\varepsilon \quad (5.2)$$

$$R_{F_Z}^2 = 98,44\%$$

L'équation 5.3 représente le modèle empirique de la rugosité de surface, Ra . Le coefficient de détermination R^2 de ce modèle est de 0,9753, ce qui signifie que 97,53% des variations de rugosité sont expliquées par ce modèle, laissant seulement 2,47% inexpliquées. Ce coefficient démontre une bonne corrélation entre le modèle et les données expérimentales.

$$R_a = 8,339 - 2,631CC - 0,004607V_c + 9,32f - 0,026ap - 0,706r_\varepsilon \quad (5.3)$$

$$R_{R_a}^2 = 97,53\%$$

L'équation 5.4 représente le modèle empirique de la température de coupe, T_c . Son coefficient de détermination R^2 est de 0,9549, ce qui signifie que 95,49% des variations de température sont capturées par ce modèle, laissant seulement 4,51% non capturées. Ce coefficient témoigne d'une corrélation significative entre le modèle et les données expérimentales.

$$T_c = 60,71 - 11,549CC - 0,00286V_c - 29,95f - 0,007ap + 3,59r_\varepsilon \quad (5.4)$$

$$R_{T_c}^2 = 95,49\%$$

L'équation 5.5 représente le modèle empirique de l'énergie de coupe spécifique (SCE), E_{cs} . Son coefficient de détermination R^2 est de 0,8315, indiquant ainsi que 83,15% des variations de SCE sont expliquées par ce modèle, laissant 16,85% non capturées. Ce coefficient démontre une corrélation entre le modèle de régression et les données expérimentales.

$$E_{cs} = 0,827 - 0,1065CC + 0,000138V_c - 2,403f - 0,0627ap + 0,0084r_\varepsilon \quad (5.5)$$

$$R_{E_{cs}}^2 = 83,15\%$$

5.3.3.2- Comparaison des résultats expérimentaux et prédites

Les figures 5.5-5.8 comparent les résultats expérimentaux et les résultats des modèles empiriques de régression des réponses F_Z , R_a , T_c et E_{cs} respectivement. La barre noire représente les valeurs expérimentales qui se rapprochent les barres rouges des valeurs estimées. Les grandeurs d'intérêt F_Z , R_a , T_c et E_{cs} présentent respectivement des coefficients de détermination (R^2) de 0,984, 0,975, 0,954 et 0,831. De plus, leurs erreurs absolues moyennes en pourcentage (MAPE) sont de 3,92%, 7,85%, 1,91% et 12,96% respectivement. Ces résultats démontrent la précision et la capacité des quatre modèles empiriques à prédire de nouvelles valeurs dans l'intervalle de variation des facteurs décrits dans le tableau 5.2.

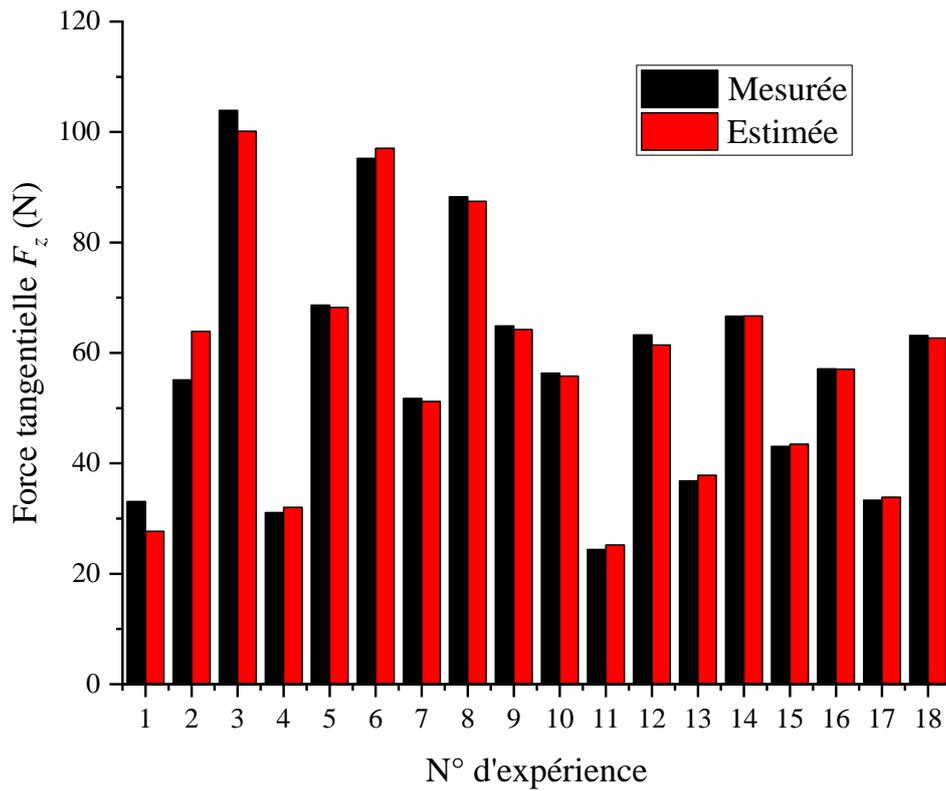


Figure 5.5- Comparaison entre les résultats expérimentales et prédites de F_z .

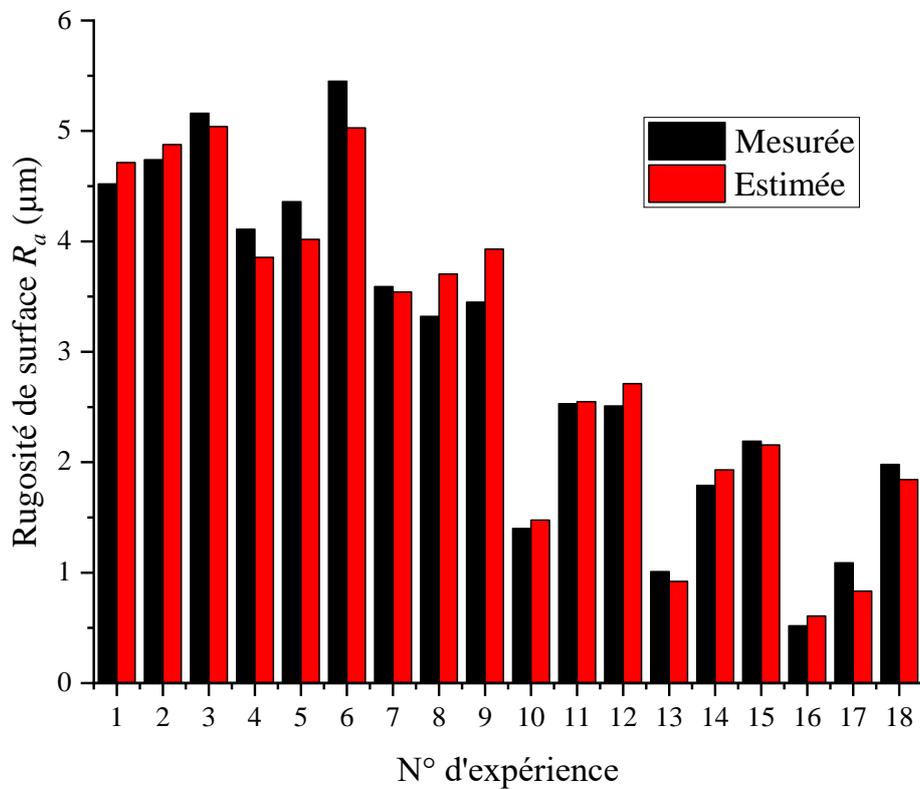


Figure 5.6- Comparaison entre les résultats expérimentales et prédites de R_a .

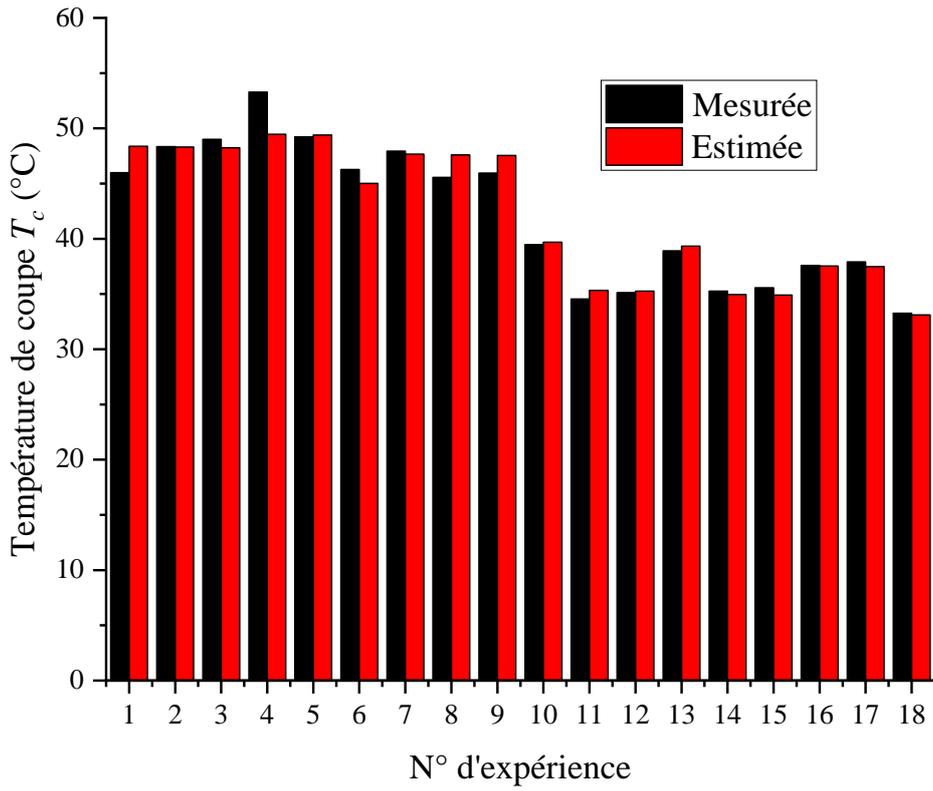


Figure 5.7- Comparaison entre les résultats expérimentales et prédites de T_c .

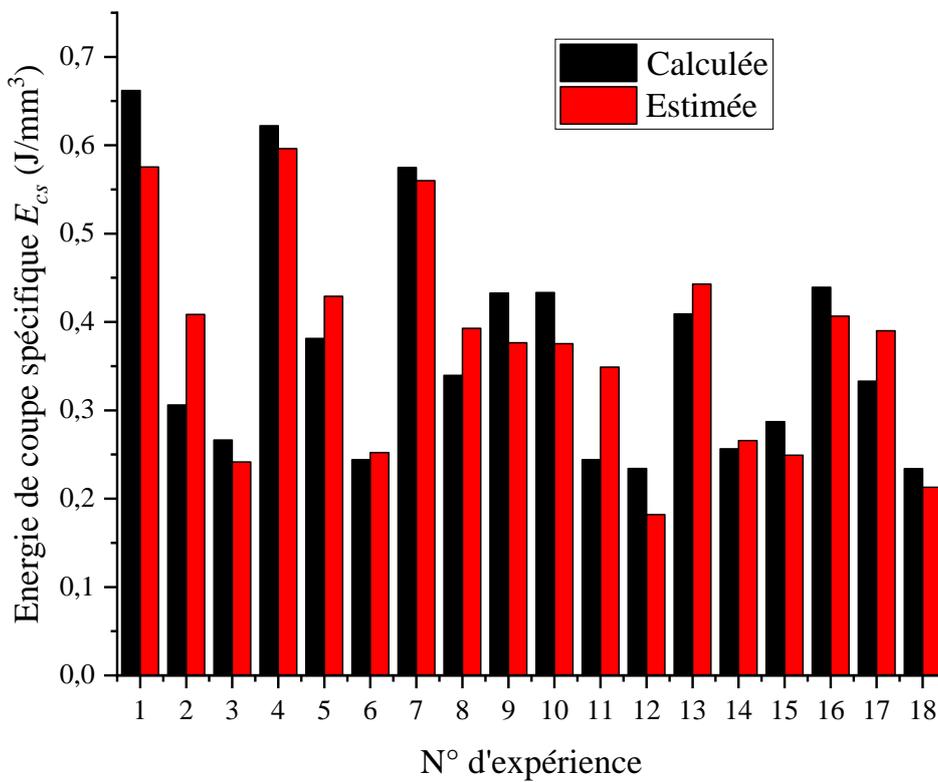


Figure 5.8- Comparaison entre les résultats expérimentales et prédites de E_{cs} .

5.3.3.3- Surface de réponse 3D

Les figures 5.9-5.12 présentent les surfaces de réponse en 2D et 3D des variables endogènes (F_Z , R_a , T_c et E_{cs}) en fonction des variables exogènes (CC , V_c , f , ap et r_ε). Seules les deux interactions les plus significatives sont représentées, tandis que les facteurs non représentés sont maintenus constants à leur niveau intermédiaire ($CC = 2$, $V_c = 375 \text{ m/min}$, $f = 0.1 \text{ mm/tr}$, $ap = 1.8 \text{ mm}$ et $r_\varepsilon = 0.8 \text{ mm}$).

Les réponses 2D et 3D peuvent être interprétées comme suit :

- ✖ Selon la figure 5.9, la profondeur de passe (ap) est le premier facteur qui influence la variable dépendante F_Z , car sa pente est la plus importante, suivie par la vitesse d'avance (f). Cependant, pour réduire la force tangentielle, il est recommandé d'adopter le MQL, une vitesse d'avance réduite et une faible profondeur de coupe.

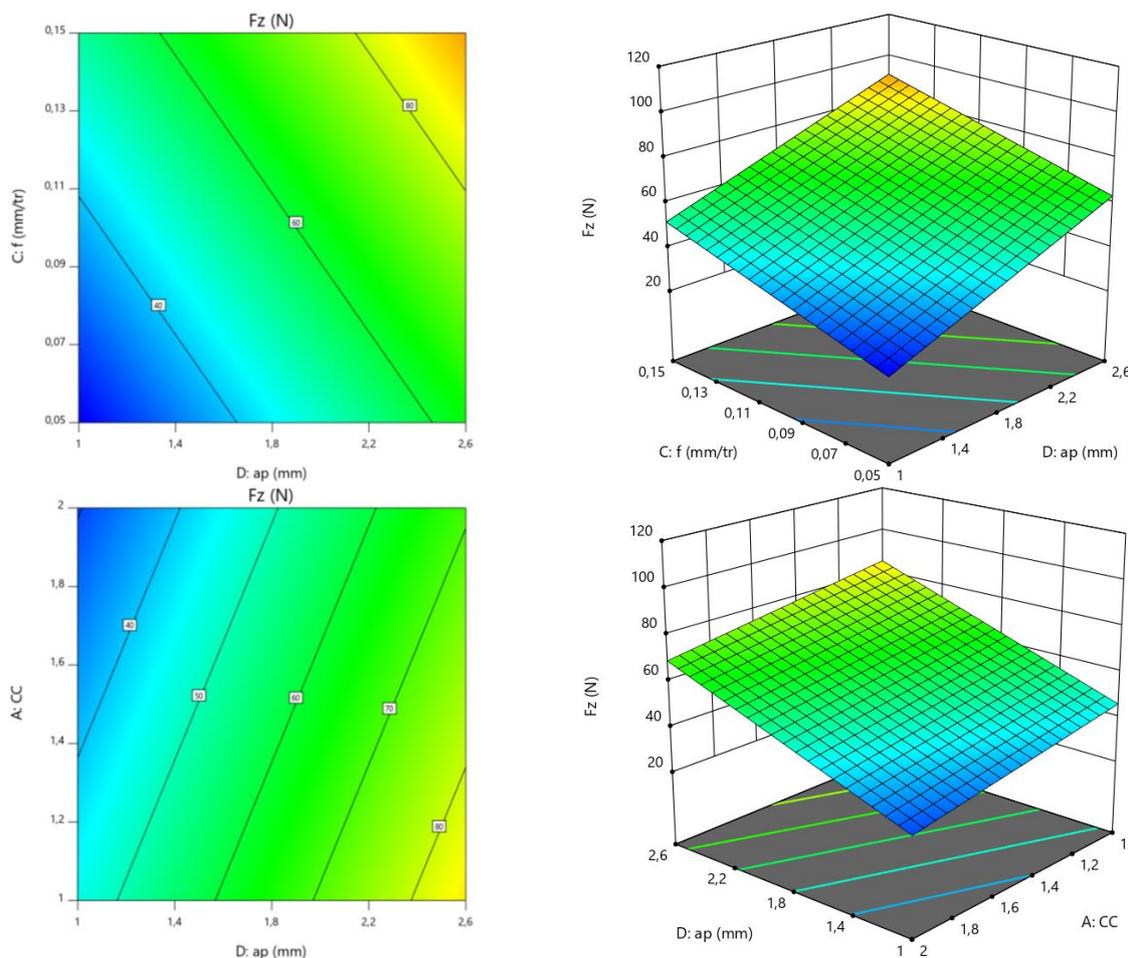


Figure 5.9- Surface de réponse 2D et 3D de F_Z en tenant compte deux interactions les plus significatives.

- ✘ En se référant à la figure 5.10, on constate que la pente associée à la condition de refroidissement (CC) est la plus importante, suivie par celle de la vitesse de coupe (V_c), et enfin celle de la vitesse d'avance (f). Par conséquent, pour obtenir une meilleure qualité de pièce usinée, il est recommandé d'opter pour le tournage avec lubrification à quantité minimale (MQL), une vitesse de coupe élevée, et une vitesse d'avance réduite.

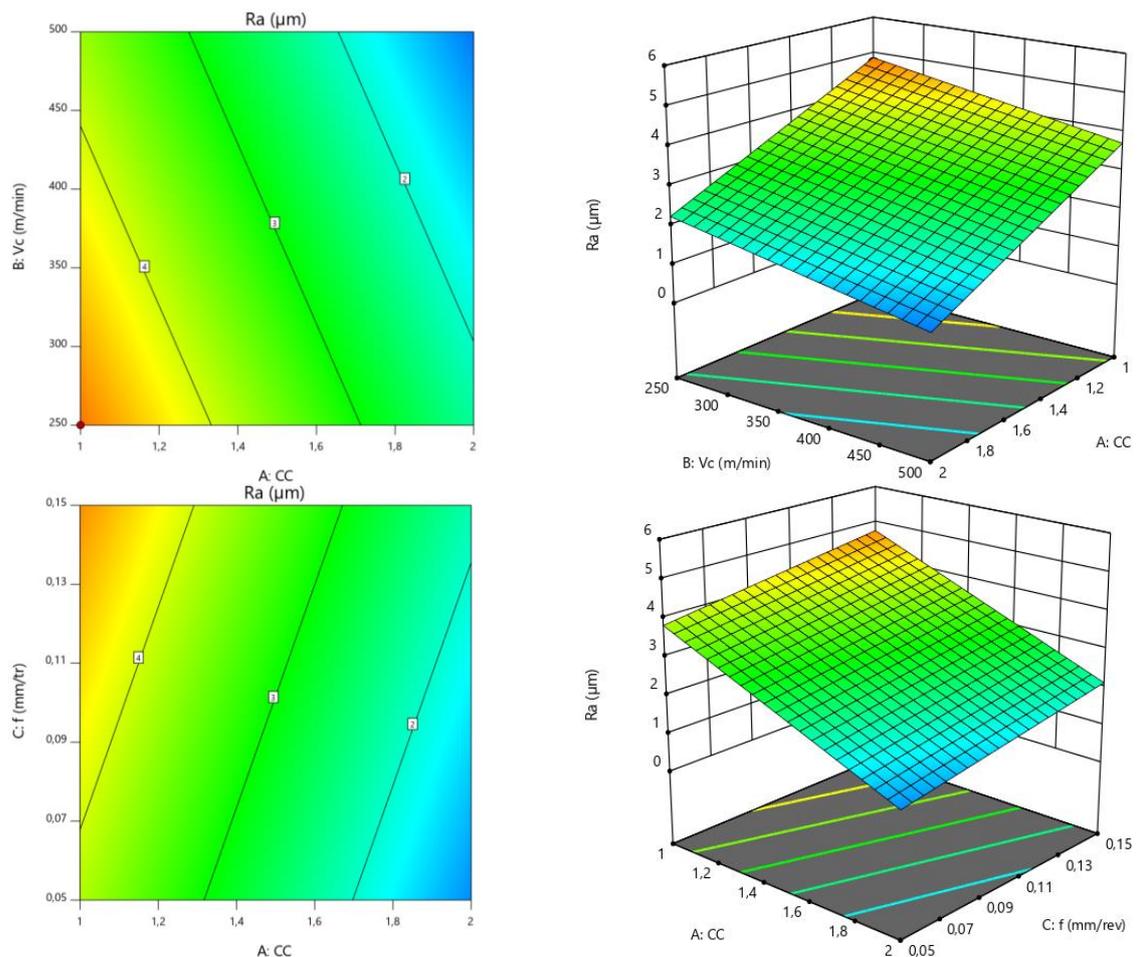


Figure 5.10- Surface de réponse 2D et 3D de R_a en tenant compte deux interactions les plus significatives.

- ✘ En observant la figure 5.11, on constate que la pente associée à la condition de refroidissement (CC) reste la plus importante, suivie de la pente de la vitesse d'avance (f) et enfin du rayon du bec (r_ϵ). Par conséquent, pour maintenir une température basse et éviter le ramollissement du matériau PP, il est recommandé d'utiliser le tournage MQL, une vitesse f réduite et un rayon de bec réduit.

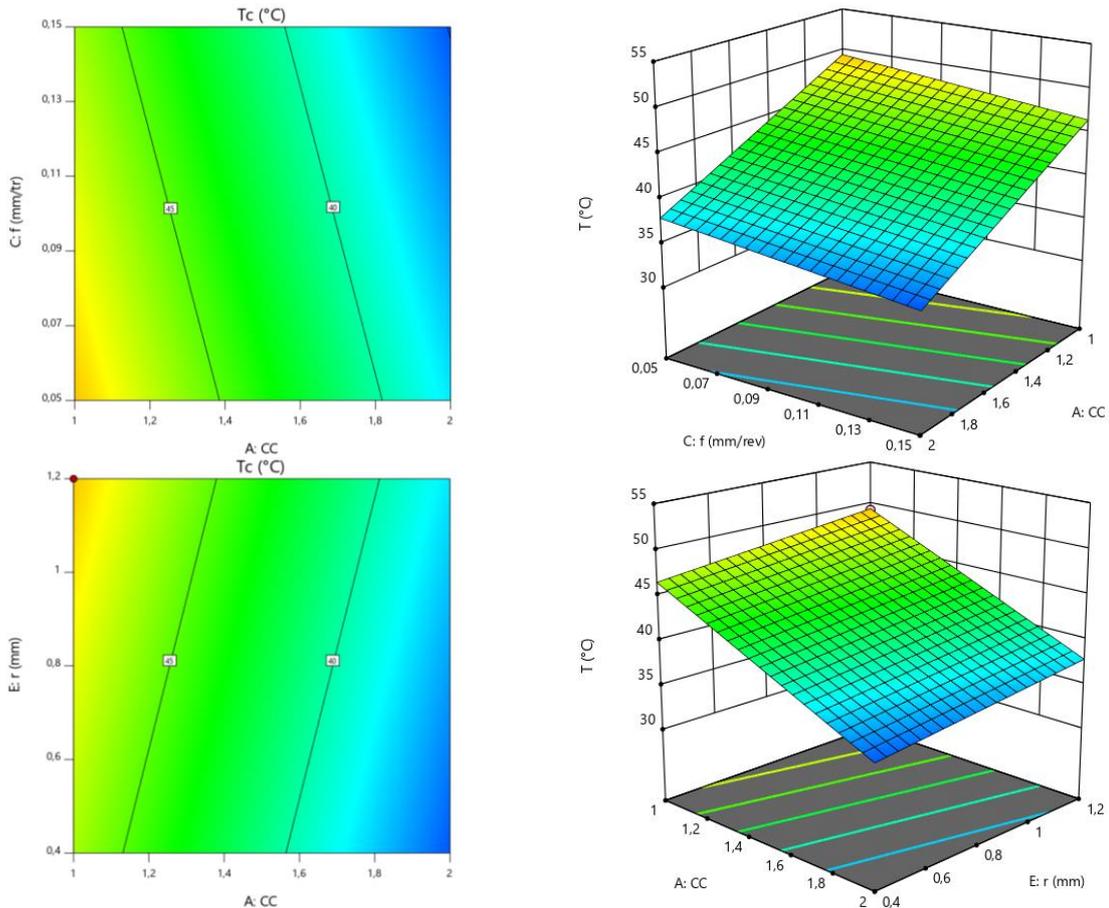


Figure 5.11- Surface 2D et 3D de T_c en fonction des deux interactions les plus significatives.

- ✘ Selon la figure 5.12, la vitesse d'avance (f) est le premier facteur qui influence la variable dépendante E_{CS} , car sa pente est la plus importante. L'utilisation de MQL reste bénéfique dans l'usinage du matériau PP, suivie par la profondeur de coupe (ap). Cependant, pour réduire la consommation d'énergie de coupe spécifique, il est recommandé d'adopter le MQL, une vitesse f réduite et une ap importante.

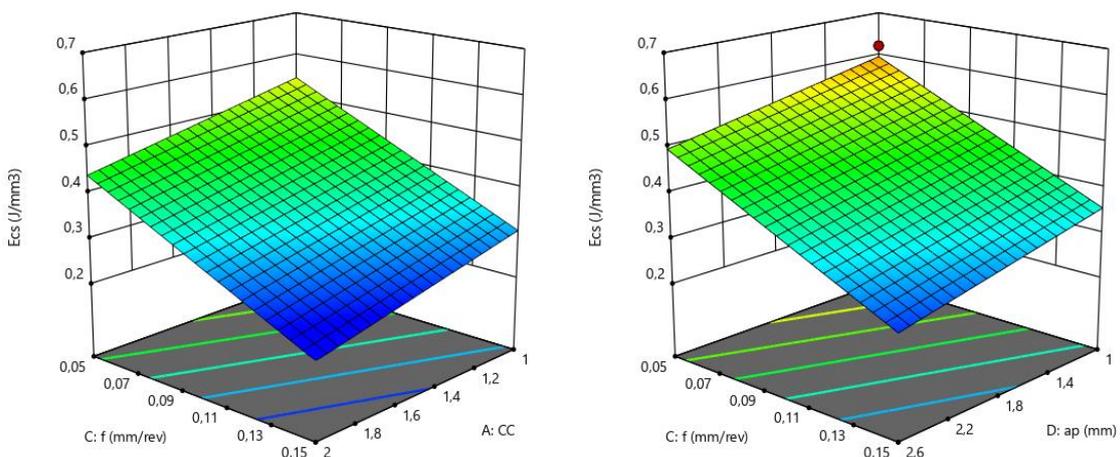


Figure 5.12- Surface 2D et 3D de E_{CS} en fonction des deux interactions les plus significatives.

5.3.3.4- Validation expérimentale

La validation expérimentale est essentielle pour garantir la fiabilité, la validité et l'applicabilité des modèles empiriques. Elle implique de confronter le modèle à des données réelles afin de vérifier sa précision. Cela permet de calculer l'erreur absolue en pourcentage entre les prédictions du modèle et les mesures observées. Le tableau 5.8 présente les valeurs mesurées, les prédictions des modèles 5.2, 5.3, 5.4 et 5.5, ainsi que l'erreur absolue en pourcentage (APE) pour quatre cas de conditions de coupe sélectionnés au hasard et différents des conditions du tableau 5.2. Le APE de la rugosité de surface varie de 2,78% à 5,85%, tandis que celle de la température varie de 2,13% à 4,15%. Pour ce qui est du SCE, l'erreur absolue en pourcentage varie de 3,82% à 9,33%. Ces résultats montrent que les modèles empiriques des paramètres R_a , T_c et E_{CS} sont capables de reproduire fidèlement les résultats observés et de prédire de manière précise de nouveaux résultats.

Tableau 5.8- Résultats de validation expérimentale.

N°	CC	V_c (m/min)	f (mm/tr)	ap (mm)	r_ϵ (mm)	Résultats expérimentales	Résultats prédites	APE (%)
(a) Rugosité R_a (μm)								
1	Dry	350	0.07	1.4	0.4	4.17	4.43	5.85
2	MQL	400	0.09	2.2	0.8	1.4	1.45	3.51
3	Dry	300	0.12	1.6	1.2	4.75	4.56	4.27
4	MQL	450	0.14	1.2	1.2	1.47	1.43	2.78
(b) Température de coupe T_c ($^\circ\text{C}$)								
1	Dry	350	0.07	1.4	0.4	45.52	47.49	4.15
2	MQL	400	0.09	2.2	0.8	35.85	36.63	2.13
3	Dry	300	0.12	1.6	1.2	50.66	49.01	3.38
4	MQL	450	0.14	1.2	1.2	35.13	36.43	3.57
(c) Energie de coupe spécifique E_{CS} (J/mm^3)								
1	Dry	350	0.07	1.4	0.4	0.468	0.516	9.33
2	MQL	400	0.09	2.2	0.8	0.334	0.322	3.82
3	Dry	300	0.12	1.6	1.2	0.358	0.383	6.60
4	MQL	450	0.14	1.2	1.2	0.263	0.275	4.20

5.3.4- Optimisation multi-objectif

La fonction de désirabilité (DF) est un outil d'optimisation mathématique largement utilisé dans divers domaines de recherche, ce qui témoigne de son efficacité. La DF permet de combiner plusieurs critères ou objectifs en une seule fonction, permettant ainsi de trouver une solution optimale qui maximise ou minimise ces critères simultanément. Notre objectif principal est d'optimiser le processus de tournage du polypropylène non renforcé (PP) en

utilisant la DF afin de trouver les valeurs optimales pour les paramètres indépendants. L'optimisation vise à minimiser toutes les réponses. A cet égard, une surface de haute qualité (faible rugosité) améliore la fonctionnalité de la pièce, tandis qu'une température basse réduit le risque de ramollissement du polymère PP. De plus, une faible consommation d'électricité est obtenue grâce à une énergie de coupe spécifique réduite. Les conditions d'optimisation (les contraintes) sont résumées dans le tableau 5.9.

Tableau 5.9- Conditions de l'optimisation multi-objectifs (DF).

Conditions	Goal	Lower	Upper
Cooling condition CC	In range	1	2
Vitesse de coupe V_c (m/min)	In range	250	500
Vitesse d'avance f (mm/tr)	In range	0,05	0,15
Profondeur de coupe ap (mm)	In range	1	2,6
Rayon du bec r_ϵ (mm)	In range	0,4	1,2
Force tangentielle F_z (N)	In range	24,41	103,93
Rugosité R_a (μm)	Minimize	0,52	5,45
Température T_c ($^\circ\text{C}$)	Minimize	33,25	53,28
SCE E_{cs} (J/mm^3)	Minimize	0,234	0,662

Le tableau 5.10 présente les 10 premiers résultats de l'optimisation des paramètres indépendantes et dépendantes (CC , V_c , f , ap et $r_\epsilon + F_z$, R_a , T_c , et E_{cs}) en utilisant la fonction de désirabilité (DF). De plus, la figure 5.13 montre la désirabilité 3D globale de toutes les réponses en fonction de la vitesse de coupe (V_c) et de la condition de refroidissement (CC). Par ailleurs, la figure 5.14 représente les barres des désirabilités des paramètres indépendantes et dépendantes, ainsi que la désirabilité combinée $DF = 0,85$. Enfin, pour compléter, la figure 5.15 présente l'optimisation globale.

Tableau 5.10- Résultats d'optimisation des trois réponses F_z , R_a , T_c , et E_{cs} .

Sol n°.	CC	V_c (m/min)	f (mm/tr)	ap (mm)	r_ϵ (mm)	F_z (N)	R_a (μm)	T_c ($^\circ\text{C}$)	E_{cs} (J/mm^3)	DF
1	2,00	499,18	0,149	1,00	0,42	42,82	1,83	33,25	0,265	0,850
2	1,99	499,94	0,150	1,00	0,42	43,02	1,85	33,25	0,264	0,849
3	2,00	499,99	0,149	1,00	0,47	43,18	1,80	33,41	0,26	0,849
4	2,00	499,99	0,138	1,00	0,40	39,48	1,74	33,48	0,29	0,849
5	1,99	499,99	0,146	1,01	0,40	42,11	1,82	33,27	0,272	0,849
6	1,99	499,39	0,148	1,00	0,40	42,43	1,86	33,25	0,268	0,848
7	2,00	499,97	0,150	1,00	0,52	43,58	1,77	33,58	0,265	0,848
8	1,99	499,99	0,150	1,00	0,46	43,53	1,82	33,38	0,263	0,848
9	2,00	499,75	0,146	1,05	0,40	43,34	1,82	33,25	0,268	0,848
10	2,00	496,98	0,146	1,05	0,40	43,01	1,83	33,25	0,269	0,847

Design-Expert® Software
Factor Coding: Actual

Desirability
0 1

X1 = A: CC
X2 = B: Vc

Actual Factors
C: f = 0,149165
D: ap = 1
E: r = 0,428345

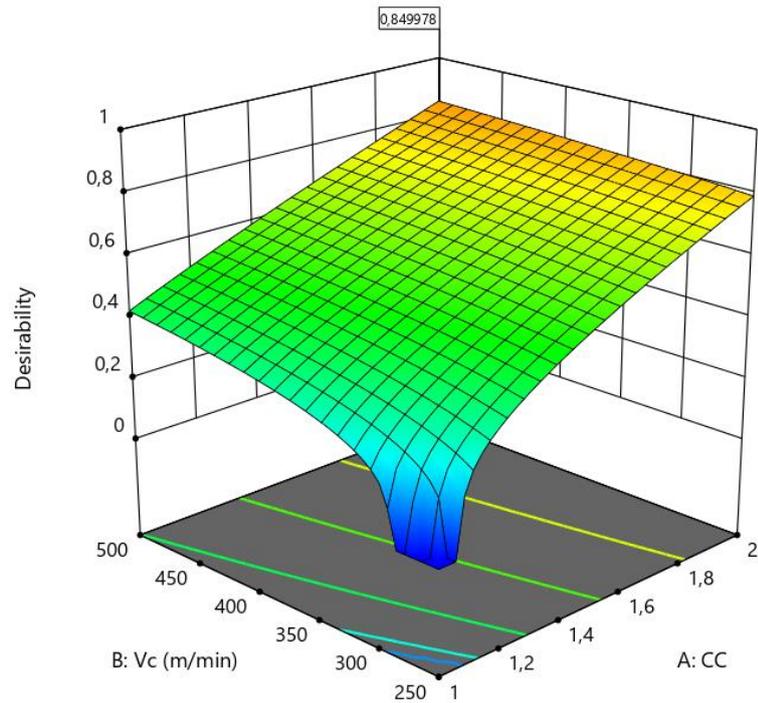


Figure 5.13- Désirabilité 3D combinée pour les réponses F_z , R_a , T_c , et E_{cs} .

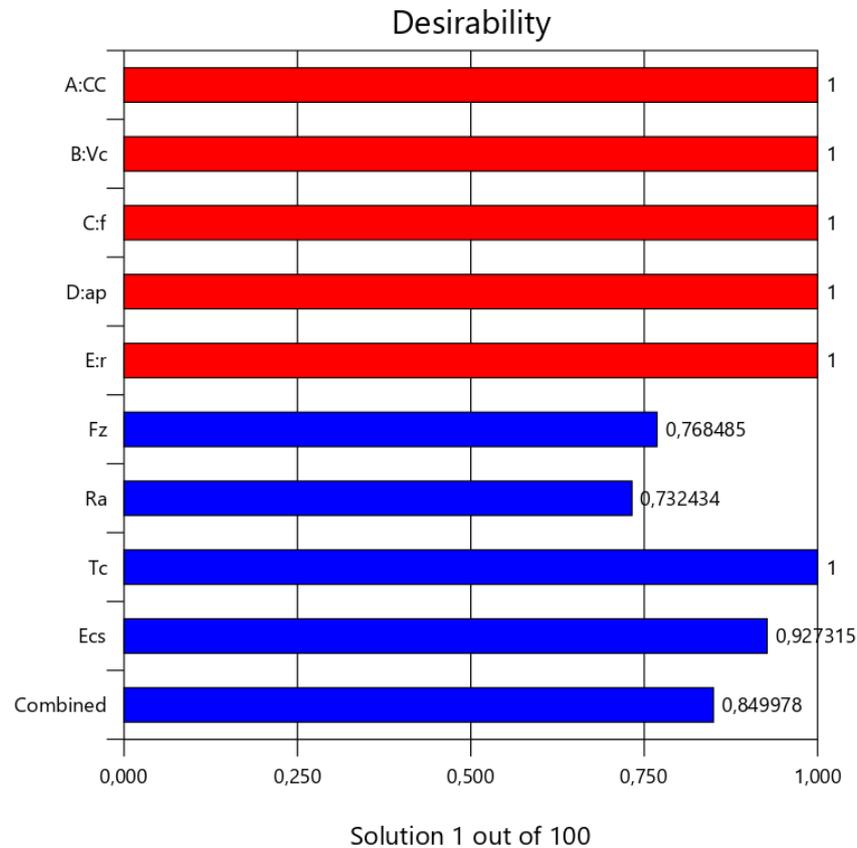


Figure 5.14- Niveaux de désirabilité pour les paramètres indépendantes et dépendantes

$$(DF_{combinée} = 0,85).$$

Sur la base du tableau 5.10 et les figures 5.13-5.15, les valeurs des paramètres de coupe optimaux pour obtenir une meilleure qualité de surface, une température basse et une consommation d'énergie de coupe spécifique réduite sont les suivantes : le tournage MQL, $V_c = 500 \text{ m/min}$, $f = 0,15 \text{ mm/tr}$, $a_p = 1,0 \text{ mm}$, $r_e = 0,4 \text{ mm}$. Toutefois, pour ce régime de coupe, les paramètres de sortie sont les suivants : $F_z = 42,82 \text{ N}$, $R_a = 1,83 \mu\text{m}$, $T_c = 33,25 \text{ }^\circ\text{C}$, et $E_{cs} = 0,265 \text{ J/mm}^3$.

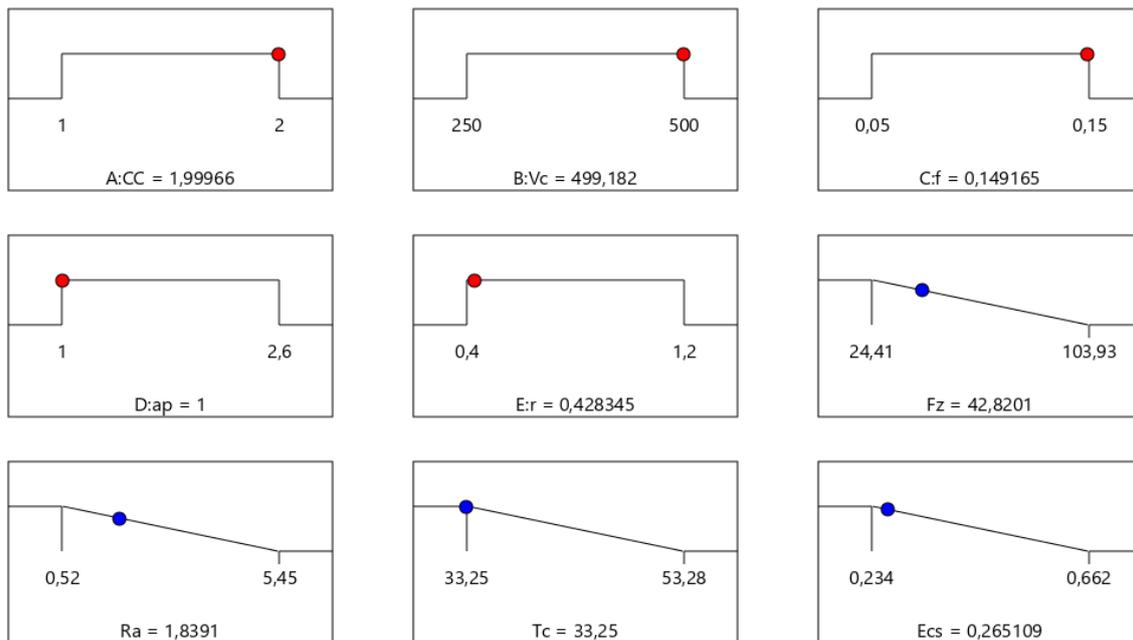


Figure 5.15- Graphique de l'optimisation combinée des variables indépendantes et dépendantes.

5.4- Conclusions

Dans ce chapitre, une analyse approfondie à été menée pour évaluer l'usinabilité du polypropylène non renforcé (PP) lors du tournage à sec et de lubrification à quantité minimale (MQL) avec divers paramètres indépendants. Les critères d'évaluation incluaient la force tangentielle, la rugosité de surface, la température de coupe et l'énergie de coupe spécifique (SCE). Les résultats obtenus ont permis de tirer les conclusions principales suivantes :

- La lubrification à quantité minimale (MQL) offre de nombreux avantages lors du tournage du matériau PP. Elle réduit tous les critères d'usinabilité considérés dans cette étude, tels que l'effort de coupe principal, la rugosité de surface, la température de coupe et la consommation de l'énergie de coupe spécifique. Donc, le MQL contribue à améliorer l'efficacité énergétique des opérations d'usinage.

- L'analyse de variance révèle que le MQL et la vitesse f sont les deux facteurs les plus significatifs pour expliquer les variations des variables d'intérêt F_z , R_a , T_c , et E_{cs} . La contribution du MQL est de 14,30%, 78,52%, 87,72% et 16,45% respectivement, tandis que la contribution de la vitesse d'avance est de 27,45%, 6,56 %, 3,93 % et 55,78 % respectivement. À l'exception, la vitesse de coupe est le deuxième facteur le plus important avec une contribution 10.03% dans la variation de rugosité du PP et la profondeur de coupe est premier facteur affecte la force tangentielle avec une contribution 55,33%.
- Les modèles empiriques obtenus grâce à la méthodologie de surface de réponse (RSM) des paramètres dépendants tels que la force tangentielle (F_z), la rugosité de surface (R_a), la température de coupe (T_c) et l'énergie de coupe spécifique (E_{cs}) sont en mesure de reproduire fidèlement les résultats observés et de prédire de manière précise de nouveaux résultats.
- La méthode de DF a permis d'obtenir la combinaison optimale des paramètres indépendants suivante : le tournage MQL, une vitesse de coupe élevée de $V_c = 500 \text{ m/min}$, une faible vitesse d'avance de $f = 0,15 \text{ mm/tr}$, une profondeur de coupe faible de $ap = 1,0 \text{ mm}$, et un faible rayon de bec de $r_\epsilon = 0,4 \text{ mm}$. En ce qui concerne les paramètres dépendants, les valeurs obtenues sont les suivantes : $F_z = 42,82 \text{ N}$, $R_a = 1,83 \text{ }\mu\text{m}$, $T_c = 33,25 \text{ }^\circ\text{C}$, et $E_{cs} = 0,265 \text{ J/mm}^3$.

Référence bibliographie :

- [1] J. Goupy : Les plans d'expériences. Numéro 34, Revue MODULAD, 2006.
- [2] D.C. Montgomery: Design and Analysis of Experiments. Eighth Edition, ISBN 978-1-118-14692-7, Arizona State University, John Wiley & Sons, Inc, 2013.
- [3] M.S. Daas : Plans d'expériences. Cours, Université Frères Mentouri - Constantine 1, Algérie, 2019/2020.
- [4] P. Alexis : Cours de plan d'expérience (sous forme littérale). 15/12/2015.
- [5] S. Vivier : Stratégies d'optimisation par la méthode des plans d'expériences, et Application aux dispositifs électrotechniques modélisés par éléments finis. Thèse de doctorat, Ecole Centrale de Lille Université des Sciences et Technologies de Lille, France, 11/07/2002.
- [6] G.W. Oehlert: A First Course in Design and Analysis of Experiments. ISBN 0-7167-3510-5, University of Minnesota, 2010.
- [7] R.W. Mee: A Comprehensive Guide to Factorial Two-Level Experimentation. ISBN 978-0-387-89102-6, e-ISBN 978-0-387-89103-3, The University of Tennessee, 2009.
- [8] J. Goupy: Plans d'expériences. Technique de l'ingénieur, Réf. : P230 V2, 2020.
- [9] R. Sabre : Plans d'expériences - Méthode de Taguchi. Technique de l'ingénieur, Réf. : F1006 V1, 2007.
- [10] A. Dean, D. Voss, D. Draguljić: Design and analysis of experiments. Second Edition, ISBN 978-3-319-52248-7, Switzerland, 2017.
- [11] A. Kamoun, M.M. Chaabouni, H.F. Ayedi : Plans d'expériences et traitements de surface - Méthodologie des surfaces de réponses (MSR). Technique de l'ingénieur, Réf. : M1429 V1, 2020.
- [12] J. Goupy : Modélisation par les plans d'expériences. Technique de l'ingénieur, Réf. : R275 V1, 2000.
- [13] D. Mathieu, R. Phan-tan-luu : Planification d'expériences en formulation : optimisation. Technique de l'ingénieur, Réf. : J2241 V1, 2001.
- [14] J. Paulo Davim: Design of experiments in production engineering. ISBN 978-3-319-23837-1, University of Aveiro, 2016.
- [15] M. Pillet : Les plans d'expériences par la méthode TAGUCHI. ISBN - 978-2-9539428-0-4, Université de Savoie, Edition 2011.

- [16] M. Tyagi, A. Sachdeva, V. Sharma: Optimization methods in engineering. ISBN 978-981-15-4549-8, Select Proceedings of CPIE 2019, 2021.
- [17] S. Banerjee: Mathematical modeling Models, Analysis and Applications. ISBN-13 : 978-1-4822-2916-5, CRC Press Taylor & Francis Group, 2014.