République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ YAHIA FARÈS DE MÉDÉA

Faculté de Technologie
Département de Génie des Procédés et Environnement

MEMOIRE DE FIN D'ETUDES

En vue de l'obtention de diplôme Master en Génie des Procédés

Spécialité : Génie des Procédés Pharmaceutiques

Thème

Étude expérimentale et modélisation du procédé de séchage de quelques poudres pharmaceutiques

Réalisé par

Mr. FEKHAR Khaled

Mr. BOUKHELKHAL Othmane

Proposé et dirigé par

Pr. LAIDI Maamar

Dr. HENTABLI Mohamed

2022/2023

Université Yahia Fares de Médéa, Médéa 26000, Algérie

Remerciements

*Le plus grand merci revient à **ALLAH** qui nous a guidé dans le bon sens durant notre vie qui nous aide à réaliser ce travail.*

*Ce travail a été réalisé sous la direction de Monsieur **Laidi Maamar**, Professeur à l'Université de Yahia Farés de Médéa, qu'il trouve ainsi l'expression de notre profonde reconnaissance pour ses compétences, ses encouragements et ses précieux conseils durant le déroulement de ce travail qui ont éclairé mon chemin vers la réussite de la concrétisation de mon mémoire.*

*Je profite de l'occasion de la présentation de ce travail pour exprimer mes vifs remerciements au Professeur **S. Hanini** à l'université de Médéa, pour avoir accepté de présider le jury de soutenance.*

*Je tiens également à présenter ma profonde gratitude, à Dr **A. S.Rahal**, Maître de Conférences A, à l'université de Médéa, d'avoir accepté de juger ce modeste travail.*

Merci à tous ceux qui m'ont aidé de près ou de loin durant ces années de recherche et d'études. Je ne peux conclure sans un mot de remerciement à tous mes enseignants et à tous mes collègues.

Dédicace

Au-delà des personnes, des lieux et des temps, je dédie ce travail avec ma profonde conviction, à tous ceux qui ont toujours cru et sans doute à la Science.

***À** ceux qui savent concrétiser leurs idées et savent réussir.*

***À** ceux qui ne portent que du bien pour les autres.*

***À** ceux qui ne vivent que dans et par la vérité.*

D'un cœur plein d'affection et de tendresse je dédie ce modeste travail à :

- *Mes parents qui me représentent l'exemple de sacrifice et qui ont fait de moi ce que je suis devenu et qui m'ont toujours encouragé et aidé à surmonter les difficultés en me donnant le plus précieux soutient tout le long de mes études.*
- *Mes sœurs.*
- *A toutes mes enseignants à tous les niveaux.*
- *Aux responsables et les travailleurs du complexe de SAIDAL.*
- *À mes collègues et amis de l'Université Yahia FARES de Médéa.*

Khaled

Tout d'abord, je tiens à remercier DIEU De m'avoir donné la force et le courage de mener à bien ce modeste travail. Je tiens à dédier cet humble travail à :

- *A mes parents qui me représentent l'exemple de sacrifice et qui ont fait de moi ce que je suis devenu et qui m'ont toujours encouragé et aidé à surmonter les difficultés en me donnant le plus précieux soutient tout le long de mes études.*
- *A mes frères et ma chère sœur.*
- *A toutes mes enseignants à tous les niveaux.*
- *Aux responsables et les travailleurs du complexe de SAIDAL*

- *A tous les amis de ma vie. Et toute ma famille.*
- *A tous mes amis de promotion de 2eme Master de génie de procèdes pharmaceutique.*
- *A mon binôme et sa famille celle qui a partagé avec moi les bons et les mauvais moments.*

Othmane

Résume :

ملخص: دراسة تجريبية ونمذجة لعملية تجفيف بعض مساحيق الأدوية

تهدف هذه الدراسة إلى تطوير طريقة حسابية قائمة على شبكات الأعصاب الاصطناعية لتوقع تأثير مختلف المتغيرات المتعلقة بتجفيف مساحيق الأدوية بدقة.

يتمثل المساهمة الأولى لنا في استخدام خطط التجارب لقياس تأثيرات متغيرات الحرارة والضغط ومحتوى الرطوبة والسمك والكتلة على وقت التجفيف بالفرن تحت الفراغ.

تم إجراء هذه العملية على ثلاثة مساحيق دوائية (باراسيتامول ونيستاتين وبولي فينيل بيروليدون) يتم تجفيفها في فرن تحت الفراغ.

تم إجراء الجزء التجريبي على التوالي في مختبر مراقبة الجودة في صايدال في مدينة المدية.

نموذجًا شبه تجريبية من الأدبيات المتعلقة بتجفيف طبقة 52 في المساهمة الثانية ، تم تعديل البيانات المحصلة باستخدام للمسحوق المدروسة ، بالإضافة إلى نموذج تجريبي مقترح في هذه الدراسة ونموذجين (MR) رقيقة من الرطوبة المخفضة للتعلم الآلي يجمعان بين جميع حركات المدروسة. تم اختيار النموذج المقترح كالأكثر ملائمة لوصف حركة التجفيف تتراوح بين 0.0400 و 0.0543 دقيقة RMSE تتراوح بين 0.9728 و 0.9899 وقيم R^2 للمساحيق. عرض النموذج قيم لتسعة حركات مدروس

الكلمات المفتاحية: النمذجة التجريبية، تعلم الالي، التصميم التجريبي، باراسيتامول، نيستاتين، بولي فينيل بيرووليدون، الشبكة العصبية الاصطناعية، التجفيف، حركية التجفيف؛ المساحيق الصيدلانية، تجفيف الفراغ، التجفيف في الفرن

Résume : Étude expérimentale et modélisation du procédé de séchage de quelques poudres pharmaceutiques

L'objectif de ce travail est la mise en œuvre d'une méthode de calcul basée sur les réseaux des neurones artificiels pour prévoir avec précision l'influence de différents paramètres reliés au séchage des poudres pharmaceutiques.

Notre première contribution Nous servirons des plans d'expérience pour quantifier les influences des variables température, pression, la teneur en eau, l'épaisseur et la masse sur le temps de séchage à étuve sous vide.

Cette opération a été effectuée sur trois poudres pharmaceutiques (Paracétamol, nystatine, PVP) sont séchées dans une étuve sous vide.

La partie expérimentale sont réalisés respectivement au sein de laboratoire de contrôle de qualité SAIDAL de Médéa.

Dans la deuxième contribution, les données obtenues ont été ajustées à l'aide de 52 modèles semi-(empiriques) de la littérature de séchage sur couche mince de la teneur en eau réduite (M_R) de la poudre étudiée, un modèle empirique proposé dans cette étude et deux modèles d'apprentissage automatique regroupant toutes les cinétiques étudiées. Le modèle proposé a été choisi comme le modèle le plus approprié à décrire la cinétique de séchage des poudres. Il a présenté respectivement un R^2qui varie entre (0.9728 to 0.9899) et un RMSE varie entre (0.0400 to 0.0543) $_{min}$ pour les neufs cinétiques étudiés.

Mot clés : Modélisation empirique, apprentissage automatique, plan d'expérience, Paracétamol Nystatine, PVP, réseau de neurones artificiel, séchage, cinétique de séchage ; poudres Pharmaceutiques, séchage sous vide, séchage en étuve

Abstract: Experimental Study and Modeling of the Drying Process of Some Pharmaceutical Powders

The objective of this work is the implementation of a calculation method based on artificial neural networks to accurately predict the influence of different parameters related to the drying of pharmaceutical powders.

Our first contribution is the use of experimental designs to quantify the influences of variables such as temperature, pressure, water content, thickness, and mass on the drying time in a vacuum oven. This operation was carried out on three pharmaceutical powders (Paracetamol, Nystatin, PVP) dried in a vacuum oven.

The experimental part was carried out at the SAIDAL quality control laboratory in Médéa.

In the second contribution, the obtained data were adjusted using 52 semi-(empirical) drying models from the literature, considering the reduced water content (MR) of the studied powder. Additionally, an empirical model proposed in this study and two machine learning models were considered, encompassing all the studied kinetics. The proposed model was selected as the most suitable model to describe the drying kinetics of the powders. It exhibited R^2 values ranging from (0.9728 to 0.9899) and RMSE values ranging from (0.0400 to 0.0543) min for the nine studied kinetics.

Key words: Empirical modeling, machine learning , experimental design, Paracetamol, Nystatin, PVP, artificial neural network, drying, drying kinetics; Pharmaceutical powders, vacuum drying, drying in an oven.

Table des Matières

Résume : I
Table des Matières IV
Listes des figures VIII
Listes des Tableaux IX
Abréviations X
Introduction Générale 1
Chapitre 1 Généralités sur Phénomene de Séchage 3
1. Introduction 4
2. Séchage en pharmacie galénique 4
3. Objectif de séchage d'un produit 5
4. Domaine d'utilisation 5
5. Types du séchage 5
6. Humidité 6
7. Taux d'humidité 6
8. Transferts de chaleur 6
8.1. Séchage direct ou par convection 6
8.2. Séchage par contact indirect ou par conduction 8
8.3. Séchage par rayonnement 9
8.4. Séchage par pertes diélectriques 9
8.5. Le séchage par atomisation ou pulvérisation 10
9. Transferts de matière 10
9.1. Transferts d'eau en séchage 11
9.2. La théorie diffusionnelle 11
10. Les formes d'eau dans les produits 12
10.1. Eau d'hydratation 12
10.2. Eau de constitution 12
10.3. Vapeur d'eau 12
11. Une revue critique des modèles de séchage en couche mince. 13
11.1. Un aperçu théorique sur la modélisation de séchage 13
11.2. Modèles semi-théoriques et empiriques 15
11.2.1. Modèle de Lewis 15
11.2.2. Modèle de Page 16
11.2.3. Modèle de Page modifié 16

11.2.4. Modèle de Otsura et al 17
11.2.5. Modèles de Fick simplifiés 17
11.2.6. Modèle de Henderson and Pabis 17
11.2.7. Modèle de Henderson et Pabis modifié - I 17
11.2.8. Modèle de Henderson et Pabis modifié - II 17
11.2.9. Modèle Logarithmique 17
11.2.10. Modèle de Two Term 17
11.2.11. Modèle de Two Term modifié 17
11.2.12. Modèle de Exponentiel Two Term 18
11.2.13. Modèle de Verma et al. 18
11.2.14. Modèle d'approximation de diffusion 18
11.2.15. Modèle de Midilli et al 18
11.2.16. Modèle Midilli et al modifié 18
11.2.17. Modifié Midilli et al Modifié I 18
11.2.18. Modèle de Midilli et al Modifié III [60] 19
11.2.19. Modèle Wang et Singh 19
11.2.20. Modèle Hii et al 19
11.2.21. Modèles de distribution de Weibull I 19
11.2.22. Modèles de distribution de Weibull II [62] 19
11.2.23. Modèles de distribution de Weibull III [63] 19
11.2.24. Modèles de Vega-Galvez et al I 19
11.2.25. Modèles de Vega-Galvez et al II [43] 19
11.2.26. Modèles de Vega-Galvez et al III [64] 19
11.2.27. Modèles de Jena Das 19
11.2.28. Modèles de Wang et al un terme 20
11.2.29. Modèles de Wang et al deux termes 20
11.3. Analyse statistique pour la détermination des modèles appropriés 20
11.3.1. Coefficient de détermination, R2, ... 20
11.3.2. Coefficient de corrélation R : 20
11.3.3. Chi-square 20
11.3.4. Erreur relative moyenne en pourcentage 20
11.4. Root-mean-square error 21
11.4.1. Mean absolute error 21
11.4.2. Erreur quadratique moyenne *MSE* 21
11.4.3. Erreur de biais moyenne, MBE 21
11.4.4. Écart-type 21

11.4.5. *t*-Statistique 22
11.4.6. Coefficient de détermination ajusté 22
12. Conclusion 22
Chapitre 2 Généralités sur quelques méthodes de la modélisation numérique. 23
1. Introduction 24
2. Généralité sur le plan d'expérience 24
2.1. Définition 24
2.2. Objectif de plan d'expérience 25
2.3. Espace expérimental 25
2.4. Domaine d'un facteur 25
2.5. Domaine d'étude 26
2.5.1. Méthodologie des plans sans contraintes 26
2.5.2. Notion de surface de réponse 27
2.6. Les méthodes des plans d'expériences 28
2.7. Les types de plans d'expériences 28
2.7.1. Plans factoriels complets 28
2.7.1.1. Comment fonctionne un plan d'expériences ? 29
2.7.2. Analyse statistique 29
• Statistiques sommaires dans (Logiciel MODDE) 30
3. Chimio-informatique 31
4. RDKit 32
5. Représentations moléculaires 32
5.1. MDL Molfile 33
5.2. Système simplifié d'entrée de ligne moléculaire (smiles) : 33
5.3. Empreintes digitales à connectivité étendue 34
6. Deep learning 35
6.1. Convolution Neural Network (CNN) ou réseaux de neurones convolutifs 36
6.2. Recurrent Neural Network 37
6.3. Auto-encodeur 38
6.3.1. Auto-encodeur variationnel (VAE) 39
6.4. Perceptron multicouche « Multi-Layer Perceptron (MLP) » 39
7. Conclusion 40
Chapitre 3 Etude expérimentale 42
Contribution 1 : Séchage de quelques poudre pharmaceutiques 44
1.1. Introduction 44
2. Application de la méthode de plan d'expérience 44

2.1. Introduction 44
2.2. Définition des objectifs et des réponses 45
2.3. Définition des facteurs 45
2.4. Définition du domaine expérimental 45
Construction du plan d'expériences 46
2.5. Expérimentation 46
3. Matériel utilisé 47
3.1. Séchoir sous vide 47
3.1.1. Les principaux avantages du séchage sous vide 47
3.1.2. La productivité 47
3.1.3. La qualité 47
3.1.4. L'énergie 48
3.2. Les équipements 48
4. Produit utilisé 48
4.1. Définition de la poudre 48
4.1.1. Paracétamol 48
4.1.2. Nystatine 49
4.1.3. PVP 49
4.2. Mode opératoire 49
4.3. Analyse mathématique des résultats d'essais 50
4.4. Analyse statistique du modèle 51
4.4.1. Analyse statistique globale des modèles 51
4.4.2. Analyse graphique du modèle 54
4.4.2.1. Courbe de criblage 54
4.4.2.2. Résumé de l'ajustement (Les statistiques sommaires) 57
4.4.2.3. Représentation graphique de l'effet des facteurs 59
4.4.2.4. Courbes iso-réponse 63
Contribution 2 : Modélisation par l'intelligences artificielle 67
1. Modélisation des données expérimentale par l'intelligences artificielle 67
2. Préparation de la base de données 67
2.1. Calcul de l'empreinte digitale (fingerprint) 67
2.2. Simplified Molecular Input Line Entry System (SMILES) 68
2.3. Analyse statistique descriptive des données 69
2.4. Analyse de multicolinéarité et la corrélation des entrées 69
2.5. Modélisation AE-MLP 72
2.6. Résultats de la modélisation 74

2.7. Analyse de la variance des valeurs prédit et les valeurs de l'expérimentale..... 76
2.8. Validation par Egalité des variances..... 77
Conclusion..... 77
Conclusion Générale..... 79
Référence Bibliographique..... 80

Listes des figures

Figure 1. Etapes de la production des comprimés à différentes vois de granulation. 4
Figure 2. Séchage par contact (conduction) [5]. 8
Figure 3. Espace expérimental de l'étude de deux facteurs pour un plan à deux dimensions. 25
Figure 4 . Le domaine de variation des facteurs. 25
Figure 5. Le domaine d'étude est défini par la réunion des domaines des différents facteurs sans contraintes. 26
Figure 6. Le domaine d'étude sous contraintes est représenté par la partie grisée. 26
Figure 7. Les points expérimentaux sont disposés dans le domaine d'étude défini par l'expérimentateur..... 27
Figure 8. Les points expérimentaux sont disposés dans le domaine d'étude défini par l'expérimentateur..... 27
Figure 9. Exemple d'une table de connexion et d'une ligne d'extrémité dans un molfile MDL V2000 pour. 33
Figure 10.quelques exemples de (smiles)..... 34
Figure 11. Exemple d'empreinte digitale 35
Figure 12.CNN Architecture 37
Figure 13. RNN Architecture. 38
Figure 14. RNN loops. 38
Figure 15. Illustration simplifiée de l'Auto-Encodeur d'architecture..... 39
Figure 16. Multi-Layer Perceptron MLP. 40
Figure 17. Emplacement des bouchons d'aluminium dans l'étuve sous vide pour sécher une couche mince de poudre. 46
Figure 18. Formule 3D semi développée de paracétamol. 48
Figure 19. Formule 3D semi développée de Nystatine. 49
Figure 20. Formule 2D de PVP. 49
Figure 21. La corrélation entre les valeurs prédites et de l'expérience (paracétamol/eau). 54
Figure 22. La corrélation entre les valeurs prédites et de l'expérience (paracétamol/éthanol). 54
Figure 23. La corrélation entre les valeurs prédites et de l'expérience (nystatine/eau). 55
Figure 24. La corrélation entre les valeurs prédites et de l'expérience (nystatine/éthanol). ... 55
Figure 25. La corrélation entre les valeurs prédites et de l'expérience (PVP/eau). 56
Figure 26. La corrélation entre les valeurs prédites et de l'expérience (PVP/éthanol). 56
Figure 27. Résumé de l'ajustement de MODDE (paracétamol/eau). 57
Figure 28. Résumé de l'ajustement de MODDE (paracétamol/éthanol). 57
Figure 29. Résumé de l'ajustement de MODDE (nystatine/eau). 58
Figure 30. Résumé de l'ajustement de MODDE (nystatine/éthanol). 58
Figure 31. Résumé de l'ajustement de MODDE (PVP/eau). 59
Figure 32. Résumé de l'ajustement de MODDE (PVP/éthanol). 59
Figure 33. Effets des facteurs sur les réponses (paracétamol/eau). 60
Figure 34. Effets des facteurs sur les réponses (paracétamol/éthanol). 60

Figure 35. Effets des facteurs sur les réponses (nystatine/eau). 61
Figure 36. Effets des facteurs sur les réponses (nystatine/éthanol). 61
Figure 37. Effets des facteurs sur les réponses (PVP/eau). 62
Figure 38. Effets des facteurs sur les réponses (PVP/éthanol). 62
Figure 39. Courbes iso-réponse (paracétamol/eau)..... 63
Figure 40. Courbes iso-réponse(paracétamol/éthanol). 64
Figure 41. Courbes iso-réponse(nystatine/eau). 64
Figure 42. Courbes iso-réponse (nystatine/éthanol). 65
Figure 43. Courbes iso-réponse (PVP/eau). 65
Figure 44. Courbes iso-réponse (PVP/éthanol). 66
Figure 45: SMILES de Nystatine 68
Figure 46: SMILES Paracétamol..... 68
Figure 47: SMILES de Molécule Crospovidone PVP XL10 69
Figure 48.Multicolinéarité des entrées. 70
Figure 49. Diagramme en violon des entrés..... 71
Figure 50. Modèle AE-MLP..... 73
Figure 51. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase Apprentissage. 74
Figure 52. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase Apprentissage. 75
Figure 53. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase validation. 75
Figure 54. Description statistique des données expérimentale et calculée par le modèle. 77

Listes des Tableaux

Tableau 1. La nature des différents Facteurs. 45
Tableau 2. Le domaine des facteurs étudiés. 45
Tableau 3. Les paramètres fixes. 45
Tableau 4. Les caractéristiques des équipements utilisés dans l'expérience. 48
Tableau 5. Résultat statistique de model (paracétamol/eau). 52
Tableau 6. Résultat statistique de model (paracétamol/éthanol). 52
Tableau 7. Résultat statistique de model (nystatine/eau). 52
Tableau 8 .Résultat statistique de model (nystatine/éthanol). 53
Tableau 9. Résultat statistique de model (PVP/eau). 53
Tableau 10. Résultat statistique de model (PVP/éthanol). 53
Tableau 11. La matrice de valeurs de P Value des entré 71
Tableau 12. Valeurs des critères d'évaluation du modèle optimale à deux phases apprentissage et validation 74
Tableau 13 . Analyse de la variance 76
Tableau 14 . Analyse de teste Tukey HSD / Tukey Kramer 76

Abréviations

A	Surface de séchage ; (m^2).
Ar	Nombre d'Archimède.
a	Constante ; (-).
a_w	Activité de l'eau ; (-).
b	Constante ; (-).
bj	Biais (-).
C	Facteur de compressibilité ; (-).
C_1, C_2 et C_3	Constantes empiriques ; (-).
C_e	Concentration en vapeur d'eau dans le gaz ; (mol/m^3).
C_h	Chaleur massique de l'air ; (J/kg°C).
Cp	Chaleur massique du gaz ; (J/kg°K.
D_e	Coefficient de diffusion moléculaire de la vapeur d'eau dans l'air ; (m^2/s).
D_{eff}	Coefficient de diffusion effectif ; (m^2/s).
d_p	Taille des particules ; (µm).
dp	Diamètre des particules µm.
Gr	Nombre de Gukhman.
H	Enthalpie de l'air humide ; (J/kg).
Le	Nombre de Lewis.
m	Masse Kg ou g
m	Constante ; (-).
M_R	Teneur en eau réduite.
Nu	Nombre de Nusselt.
P	Pression du gaz ; (bar).
Pr	Nombre de Prandtl.
R	Coefficient de régression (-).
r	Rayon ; (cm).
Re	Nombre de Reynolds.
Sh	Nombre de Sherwood.
Sc	Nombre de Schmidt.
T	Température du bulbe humide ; (°C).
t	Temps (s ou min).
Tair	Température de l'air °C.
X	Teneur en eau ; kg eau/kg matière sèche.
X_0	Teneur en eau initiale ; kg eau/kg matière sèche.

Introduction Générale

Le séchage des poudres pharmaceutiques est un processus crucial dans l'industrie pharmaceutique. Il permet de réduire l'humidité présente dans les produits, garantissant ainsi leur stabilité, leur efficacité et leur durée de conservation. Un contrôle précis du processus de séchage est essentiel pour obtenir des produits pharmaceutiques de haute qualité.

Traditionnellement, la modélisation du séchage des poudres pharmaceutiques a été réalisée en utilisant des approches basées sur des plans d'expériences. Cette méthodologie permet d'étudier l'effet de différents paramètres opératoires, tels que la température, la pression, l'humidité, sur le temps de séchage. Les plans d'expériences fournissent une structure systématique pour concevoir et analyser les expériences, ce qui permet de déterminer les conditions optimales de séchage.

Cependant, avec l'avènement de l'intelligence artificielle, de nouvelles approches émergent pour modéliser le séchage des poudres pharmaceutiques. L'une de ces approches prometteuses est le "Deep Learning" ou apprentissage profond. Le Deep Learning est une technique d'intelligence artificielle basée sur des réseaux de neurones artificiels profonds, qui peuvent apprendre à partir de grandes quantités de données et effectuer des prédictions complexes.

Dans ce mémoire, nous nous intéressons à l'étude expérimentale du séchage des poudres pharmaceutiques et à sa modélisation. Nous commencerons par explorer les effets de différents paramètres opératoires sur le temps de séchage en utilisant des plans d'expériences. Nous analyserons les résultats obtenus et déterminerons les conditions optimales pour le processus de séchage.

Ensuite, nous irons plus loin en utilisant la technique du Deep Learning pour modéliser le séchage des poudres pharmaceutiques. Nous construirons un réseau de neurones profond en utilisant une base de données obtenue à partir des expériences réalisées précédemment. Le modèle de Deep Learning nous permettra de prédire avec précision le temps de séchage en fonction des paramètres opératoires.

L'objectif de ce mémoire est de comparer les résultats obtenus à partir des modèles basés sur les plans d'expériences et ceux basés sur le Deep Learning. Nous évaluerons la précision, la fiabilité et l'applicabilité de chaque approche pour la modélisation du séchage des poudres pharmaceutiques.

En conclusion, ce mémoire vise à améliorer notre compréhension du séchage des poudres pharmaceutiques en combinant les approches traditionnelles basées sur les plans d'expériences avec les techniques avancées du Deep Learning. Cette recherche contribuera à l'optimisation des processus de séchage dans l'industrie pharmaceutique, en garantissant des produits de haute qualité et en réduisant les coûts de production.

Chapitre 1 Généralités sur Phénomene de Séchage

1. Introduction

Le Séchage a pour but d'éliminer par vaporisation un corps liquide volatil contenu dans un autre corps non volatil. En pharmacie, le liquide volatil est généralement l'eau et le corps à sécher solide. Cette opération intervient à de nombreux niveaux soit pour transformer un produit liquide ou pâteux en produit solide, soit pour éliminer l'eau, facteur d'instabilité.

2. Séchage en pharmacie galénique

L'industrie pharmaceutique est le secteur économique qui regroupe les activités de recherche, de fabrication et de commercialisation des médicaments pour la médecine humaine ou vétérinaire. Parmi les formes fabriquer, les formes en comprimés. Un comprimé est une forme pharmaceutique solide, destinée à la voie orale, équivalent à une dose (unité de prise) qui peut contenir une ou plusieurs substances actives (principe actif)[1]. Les comprimés sont obtenus en agglomérant par compression d'un volume de particules (poudre ou granule)[2], les différentes étapes de la production des comprimés sont présentées dans la figure 1.

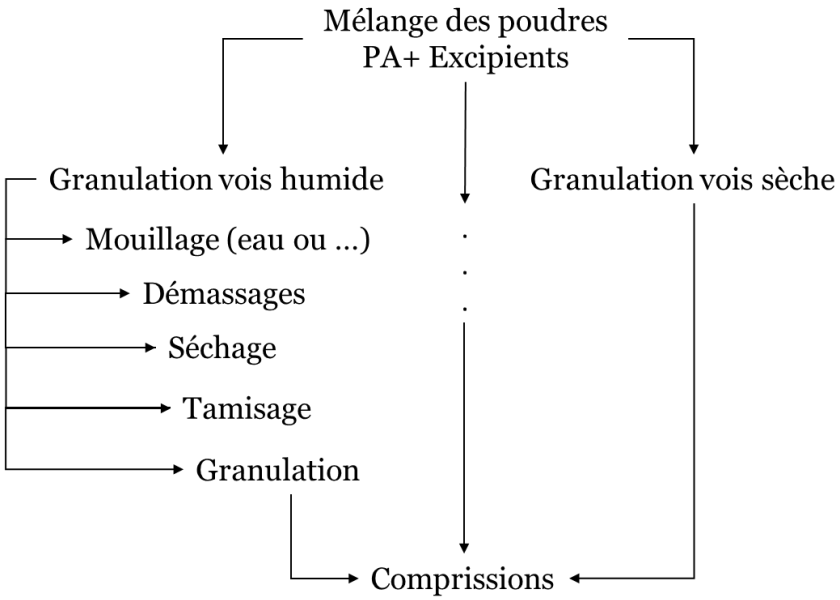


Figure 1. Etapes de la production des comprimés à différentes voient de granulation.

Le séchage c'est l'une des étapes les plus importantes et les plus difficiles de la pharmacie galénique.[3] Il est défini comme séparation thermique consistant à retirer tout ou une partie d'un liquide imprégnant (eau dans la plupart des cas) un corps dit « humide » par vaporisation de ce solvant. Le produit final est un solide, qualifié de « sec » même s'il contient une humidité résiduelle, [4].

3. Objectif de séchage d'un produit

On sèche un produit solide pour diverses raisons :

- Le liquide résiduel est incompatible avec l'utilisation ultérieure du produit, pour des raisons chimiques ou physiques.
- Le produit humide se conserve mal, qu'il subisse une dégradation chimique (hydrolyse de la molécule solide par exemple), ou que son aspect physique soit modifié (mottage d'une poudre par exemple).
- Le coût des manipulations du produit humide est grevé par la présence de liquide résiduel même si ce dernier ne gêne pas l'application (coût de transport par exemple).
- Le séchage, outre son objectif principal qui est d'éliminer le liquide, apporte une modification morphologique du solide intéressante pour son usage ultérieur (création de pores par exemple), [5].

Est très divers selon la filière envisagée, il permet :

- La stabilisation du produit (poudres pharmaceutiques).
- La réduction du volume et la facilité du transport du produit.
- A mise en forme du produit (granulés, comprimé, …), [4,6,7].

4. Domaine d'utilisation

- Industrie agroalimentaire par exemple (les pâtes alimentaires, Certains produits en poudre. Cacao, lait, …).
- Industrie papetière.
- Industrie du bois.
- Bouchons de liège.
- Matériaux de construction par exemple (briques, carrelages…).
- L'industrie céramique par exemple (assiettes, bols, plats…).
- La biotechnologie et l'industrie pharmaceutique par exemple (levure en poudre, antibiotiques, séchage de principes actifs sous forme de poudre avant pastillage…).

5. Types du séchage

- Séchage mécanique : C'est l'élimination du liquide par des forces purement mécaniques.
- Séchage chimique : C'est un procédé basé sur l'utilisation de produits déshydratants (chlorure de calcium) pour extraire l'eau.
- Séchage thermique : Ce type d'opération est essentiellement un transfert de masse nécessitant au préalable une activation de l'eau par une certaine quantité d'énergie apportée par un transfert de chaleur, [8,9].

6. Humidité

Ce terme désigne le liquide contenu dans le corps solide, liquide ou pâteux, et devant être éliminé au cours du séchage, [10].

Elle est caractérisée sur un plan quantitatif global par la teneur en eau (X), Cette quantification ne permet cependant pas d'évaluer les différentes formes que peut prendre l'eau dans un solide. Dans cet ordre d'idée, l'activité de l'eau affine un peu la description, [11].

7. Taux d'humidité

C'est la masse de liquide contenue par unité de masse de matière à sécher. Bien qu'il soit fait très souvent référence à la matière humide, il est préférable d'exprimer le taux d'humidité par rapport à la matière anhydre. La norme expérimentale E 33-060 définit les différentes méthodes permettant d'évaluer le taux d'humidité (elle permet également de caractériser les performances des sécheurs) [10].

8. Transferts de chaleur

Le séchage fait appel aux trois modes de transfert de chaleur : convection - conduction – rayonnement, ceux-ci sont utilisés seuls ou combinés entre eux, [12].

Le bilan thermique se rapportant à un produit soumis au séchage (conditions constantes), pour la première période de séchage sera, [13] :

$$Q_{conv} + Q_{cond} + Q_r = w_{pv} \quad \textbf{(I.1)}$$

Q_{conv} : transfert de chaleur par convection.

Q_{cond} : transfert de chaleur par conduction.

Q_r : transfert de chaleur par rayonnement.

W_{pv} : l'énergie mécanique totale fournie durant le transfert de chaleur.

8.1. Séchage direct ou par convection

C'est le mode de séchage utilisé par la très grande majorité des installations de l'industrie chimique. L'échange thermique est obtenu par contact direct entre le solide humide et des gaz (air) chauds qui agissent comme moyen de transport pour le fluide évaporé [5].

En régime convectif, la chaleur est directement transportée par un fluide caloporteur, qui emmagasine la chaleur (air chaud, vapeur, eau, etc.). La convection est un mode de transfert rapide.

- Si le mouvement de fluide est renforcé par un travail mécanique « artificiel » (ventilation, agitation, pompage d'un fluide), on parle de « convection forcée ».
- Si au contraire, le mouvement du fluide, n'est dû qu'à des courants thermiques ou des mouvements ascensionnels d'ébullition, la convection est dite « naturelle », [14].

Ce type de transfert utilise l'équation suivante :

$$Q_{conv} = \alpha S(T_a - T_p) \quad \textbf{(I.2)}$$

Avec :

α : Le coefficient global d'échange thermique par convection.

S : La surface d'échange thermique entre produit et l'agent séchant.

(Ta – Tp) : L'écart entre la température de l'air (ou gaz de séchage) et la température superficielle du produit à sécher, [13].

La surface d'échange *S* n'est pas toujours très bien définie. Pour les produits granulaires dont la granulométrie n'est pas homogène, on choisit un diamètre moyen des particules. En divisant la répartition granulométrique en tranches, à chaque tranche correspond un diamètre moyen d_{pi} mais aussi un pourcentage x_i de la masse totale de l'échantillon sur laquelle l'analyse a été effectuée. De cette manière, le diamètre moyen des particules sera égal à :

$$d_{pm} = \frac{1}{\sum \frac{x_i}{d_{pi}}} \quad \textbf{(I.3)}$$

En supposant que les particules sont sphériques, la surface spécifique de la couche est alors :

$$a = \frac{6}{d_{pm}} \ (m^2/m^3) \quad \textbf{(I.4)}$$

Dans le cas de particules dont le diamètre est inférieur à 10 mm, le coefficient d'échange par convection α peut être calculé à partir du nombre de Nusselt :

$$Nu = \frac{\alpha \, d_p}{\lambda_a} \quad \textbf{(I.5)}$$

Avec

λ_a : conductivité thermique de l'air, grâce à la relation.

$$Nu = CRe^{0.8} \, Pr^{0.5} \quad \textbf{(I.6)}$$

Où

Re : est le nombre de Reynolds.

Pr : est le nombre de Prandtl, égal à 0,7 dans le cas de l'air.

C : est une constante ajustée expérimentalement (valeur proche de 2).

Pour exprimer le coefficient d'échange par convection dans les sécheurs rotatifs,[15], ont proposé la formule empirique suivante :

$$Ua = 20\left(\eta_f - 1\right)\frac{G^{2/3}}{\mathrm{D}} \quad \textbf{(I.7)}$$

Avec

Ua ($W \cdot m^{-3} \cdot K^{-1}$) : Coefficient moyen volumique d'échange par convection (Exprimé par rapport au volume interne du sécheur rotatif).

η_f : Nombre de pelles.

G ($kg \cdot m^{-2} \cdot s^{-1}$) : Débit spécifique d'air de séchage à travers le sécheur.

D (m) : Diamètre du sécheur.

Dans les sécheurs rotatifs industriels, U_a est voisin de 150 $W \cdot m^{-3} \cdot K^{-1}$, [16].

Le séchage par convection nécessite un apport énergétique important et demande pour être efficace une température relativement élevée quand on élimine de l'eau (60-80°C), ce qui peut poser des problèmes avec les produits thermosensibles dégradables à ces niveaux de température, [17][18]**.**

8.2. Séchage par contact indirect ou par conduction

L'échange thermique est obtenu par contact entre le solide humide et une paroi métallique (figure. 2) réchauffée par un fluide (eau, vapeur, etc.).

La température de la matière est supérieure à la température du bulbe humide correspondant à l'humidité du gaz porteur. La surface d'échange thermique est fonction de la température de la paroi et de la capacité évaporatoire du produit. Plus le gradient de température entre la paroi et le produit sera faible, plus la surface d'échange devra être importante, à même capacité évaporative. En séchage conductif, la diminution de la température de séchage s'obtient sous pression réduite, Parmi les types de séchage, le séchage conductif sous vide est largement répandu dans l'industrie, notamment l'industrie pharmaceutiques. Le séchage sous vide est particulièrement intéressant si l'on traite des produits sensibles à l'oxydation lorsqu'ils sont chauffés, [5].

Le liquide évaporé est éliminé par aspiration sous vide ou bien par émission de gaz de transport. Pour disposer d'un bon transfert de chaleur entre la matière et la paroi chauffante, il faut appliquer cette méthode aux produits liquides, solides pâteux, pulvérulents ou en suspension, ainsi qu'aux matières en ruban : papier, tissu.

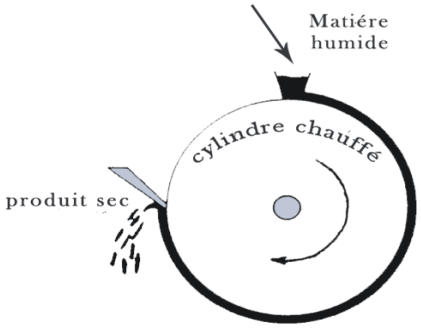


Figure 2. Séchage par contact (conduction) [5].

L'énergie thermique nécessaire au séchage est apportée non pas un gaz en mouvement autour du produit à sécher, mais par contact direct entre le produit et une paroi chauffée. Le flux thermique dans ce mode de transfert est donné par la relation (8) [19] :

$$\frac{Q}{A} = \lambda \frac{dT}{dt} \quad \textbf{(I.8)}$$

Avec λ : Conductivité thermique de la pellicule de produit (W/km).

dT/dt : Gradient de température entre la paroi chaude et le produit.

8.3. Séchage par rayonnement

Ce mode est destiné aux produits en plaque (carton, viande), ou en fibre (tissu, papier), mais aussi aux produits granulaires de faibles épaisseurs (cigarettes). L'énergie est apportée aux produits à sécher par des ondes électromagnétiques (rayonnement), soit par élévation de la température d'un émetteur infrarouge.

La quantité de chaleur Q (en watts) transmise au produit par unité du temps par ce mode de transfert est donnée par la loi de Stefan-Boltzmann, [13] :

$$Q = A_1\, C \left[\left(\frac{T1}{100}\right)^4 - \left(\frac{T2}{100}\right)^4\right]\varphi \quad \textbf{(I.9)}$$

$$C = \frac{1}{\frac{1}{\varepsilon_1} + \frac{A_1}{A_2}\left(\frac{1}{\varepsilon_2} - 1\right)}\,\sigma \quad \textbf{(I.10)}$$

Avec :

T_1 (K) : température de la source émettrice.

T_2 (K) : température du produit.

A_1 (m^2) : surface du corps recevant le rayonnement.

A_2 (m^2) : surface de l'émetteur d'infrarouge.

φ : facteur complexe tenant compte de la position relative de l'émetteur et du produit.

(Lorsque l'émetteur entoure le produit, On a $\varphi = 1$).

ε_1 : émissivité du produit à traiter, (grandeurs sans dimensions).

ε_2 : émissivité de l'émetteur.

σ : constante de Stefan-Boltzmann ($\sigma = 5{,}673 \times 10^{-8} W \cdot m^{-2} \cdot K^{-4}$). [13]

8.4. Séchage par pertes diélectriques

Un diélectrique est un matériau ayant la particularité d'être un isolant électrique. Placé dans un champ électrique, les atomes ou les molécules constituant ce matériau se polarisent. Lorsque la polarité des électrodes est inversée, l'attraction des charges électriques change de

direction, ce qui a comme conséquence une déformation de la matière qui, si elle est entretenue, provoque son échauffement[5]**.** Comme le séchage par micro-ondes et radiofréquences est basé sur la technologie du séchage diélectrique[20]. En chauffant simultanément tout le volume du produit, il est particulièrement recommandé dans les applications où un séchage homogène et doux est requis. Cette technique est la plus adaptée pour sécher les produits thermosensibles dans le domaine alimentaire, industriel et médical, afin de préserver toutes leurs propriétés[21].

8.5. Le séchage par atomisation ou pulvérisation

Depuis plus de quarante ans, le séchage par atomisation (pulvérisation) est la technique de séchage la plus employée dans plusieurs industries, [22]**.**

- **Principes le séchage par pulvérisation**

Souvent appelé séchage par atomisation, est une technique de séchage particulaire. Il consiste à pulvériser le produit à sécher (qui se présente sous forme liquide ou en suspension) dans un courant d'air (ou un autre gaz) suffisamment chaud et sec de manière à obtenir une poudre., il s'établit spontanément entre ce corps et l'air un écart de température et de pression partielle d'eau tel, [22] :

- Qu'un transfert de chaleur s'effectue de l'air vers le produit sous l'effet de l'écart de température.
- Qu'un transfert d'eau s'effectue en sens inverse du fait de l'écart de pression partielle d'eau entre l'air et la surface du produit.
- Transferts couplés de chaleur et d'eau entre une gouttelette et l'air de séchage

9. Transferts de matière

Le transfert de masse joue un rôle très important dans les opérations unitaires de base, telles que le séchage. Dans ces opérations physiques, la résistance au transfert de masse constitue le facteur limitant, quoique le transfert de chaleur et le flux du fluide soient impliqués dans le conditionnement et la conservation où le transfert d'humidité, vapeurs, gaz et composés aromatique, influent sur la qualité de l'aliment, [23].

Les difficultés d'application des théories de transfert de masse, dans les processus de transformation des aliments, résultent des structures physiques complexes et de la composition chimique des denrées alimentaires qui varient pour une même denrée et change pendant la transformation, [23].

Il est à noter que la migration de l'eau (liquide ou vapeur) peut également s'effectuer par filtration à travers le produit poreux sous l'action d'une différence de pression entre l'intérieur et la surface, [24].

L'agitation du solide est un facteur important en permettant le transfert du liquide vers la phase vapeur par évaporation, mais également le transfert thermique entre le moyen de chauffage et le solide humide. Cette agitation peut être réalisée par :

- Agitation externe du sécheur (mouvement de rotation, mouvement alternatif, vibration de basse fréquence).
- Agitation interne (racloir, vis sans fin assurant également le déplacement du solide, agitateur à palettes).
- Fluidisation à l'aide d'un gaz tiers (séchage par entrainement).
- Atomisation (mélange intime du solide et du gaz de séchage dans une turbine d'atomisation), [25].

9.1. Transferts d'eau en séchage

Les corps à sécher sont constitués par une matrice continue, éventuellement interrompue localement, dans le cas d'un solide, par des pores pouvant déboucher à la surface de celui-ci. L'humidité peut se localiser aussi bien à la surface externe du produit que dans les pores ou au sein même de la matière [13].

La vaporisation de l'eau nécessite un apport énergétique de source extérieure au produit à sécher, qui provoque la migration de l'eau vers le milieu ambiant. La vitesse de ce transfert de matière varie au cours du temps. Elle dépend des conditions ambiantes (température, humidité relative, vitesse des gaz en contact avec le produit), mais aussi de la nature même du solide et de l'eau, [13].

Les difficultés sont plus complexes dans les aliments solides que dans les liquides. Le transfert de masse implique un transfert d'une variété de composants avec la phase et en interphase par diffusion moléculaire naturelle ou par convection forcée. La masse est transférée par gradient de concentration ou de pression partielle, [23].

9.2. La théorie diffusionnelle

La diffusivité dépend de nombreux facteurs comme la nature du solide et celle de l'humidité, le taux d'humidité et la température. Quand le solide est homogène et de forme régulière et, si l'on admet que la diffusivité est constante et qu'il y a équilibre immédiat de la surface du solide avec l'air de séchage, l'intégration de la formule (I.11) permet de calculer, pour la phase de séchage à vitesse décroissante, la variation du taux d'humidité en fonction du temps :

$$\vec{\emptyset}_m = -D_{eff}\, grad\, \vec{C} \qquad \textbf{(I.11)}$$

$$M_R = \gamma \sum_{i=0}^{\infty} (2i+1)^{-2}\, exp\left[-(2i+1)^2 \pi^2 D_{eff}\, t/L^2\right] \qquad \textbf{(I.12)}$$

Dans cette équation, L (m) est la longueur caractéristique du produit : le rayon pour une sphère ou l'épaisseur pour une plaque.

γ est un coefficient dépendant de la forme du produit $8/\pi^2$ pour une plaque $6/\pi^2$ pour une sphère,M_R la teneur en eau réduite.

$D_{eff}(m^2s^{-1})$ Est la diffusivité effective de l'humidité dans la matière. Sa valeur est déterminée expérimentalement. Cette diffusivité est considérée comme effective car il n'est fait appel à aucune hypothèse pour expliquer le mouvement de l'humidité au sein du produit.

Lorsque le temps de séchage est suffisamment long, c'est-à-dire quand $M_R < 0.6$ l'équation (III.2.5) peut être simplifiée car la série est pratiquement égale à son premier terme :

$$M_R = \gamma \, exp\left(\frac{\pi^2 D_{eff} t}{r^2}\right) \quad \textbf{(I.13)}$$

La dépendance à la température du coefficient de diffusion effective, Deff, peut être décrite par l'équation Arrhenius [26]

$$D_{eff} = D_0 \, exp\left(-\frac{E_a}{R(T+273.15)}\right) \quad \textbf{(I.14)}$$

Avec $\boldsymbol{E_a}$ Énergie d'activation (J mol^{-1}), $\boldsymbol{T}$ Température(°K), $\boldsymbol{D_0}$ Coefficient de diffusion de référence (m^2s^{-1}) $\boldsymbol{R}$ Constante universelle des gaz parfaits (J mol^{-1} K^{-1}).

10. Les formes d'eau dans les produits

10.1. Eau d'hydratation

L'eau d'hydratation est retenue soit sous forme d'un film adhérant à la surface externe du solide ou dans les espaces interstitiels et les pores par des forces de capillarité liées à la tension superficielle du liquide cette eau est appelée "eau libre'' ; soit d'une façon plus intime, par des forces de cohésion plus importantes avec la matrice solide et/ou avec les autres molécules d'eau ; cette eau est appelée "eau liée".

10.2. Eau de constitution

L'eau de constitution participe à la structure moléculaire intime du solide, comme par exemple l'eau contenue dans les gels. Son élimination, plus ou moins complète, constitue une déshydratation et modifie la structure même du solide, et ce n'est pas notre but recherche lors d'un séchage. La déshydratation est accompagnée d'une modification de la structure moléculaire primitive du produit.

10.3. Vapeur d'eau

La vapeur d'eau mélangée à l'air sec occupe les espaces vides non saturés d'eau liquide

11. Une revue critique des modèles de séchage en couche mince.

L'équation de séchage en couche mince est fondamentale pour la simulation de séchage. L'équation représente l'échange d'humidité entre une fine couche du produit de séchage et son air ambiant. Dans cette étude, des modèles semi-théoriques et empiriques utilisés dans différentes littératures pour différents matériaux ainsi que des méthodes statistiques d'évaluation de ces modèles ont été détaillés.

11.1. Un aperçu théorique sur la modélisation de séchage

Le séchage des produits lors de la chute de la vitesse de séchage est contrôlé par le mécanisme de diffusion liquide et/ou vapeur, à condition que la résistance aux flux d'humidité soit uniformément répartie à l'intérieur du matériau isotrope homogène, Le taux de mouvement de l'humidité est décrit par une valeur de diffusivité effective, Deff, quel que soit le mécanisme de diffusion impliqué [27]. La deuxième loi de diffusion de Fick est souvent utilisée pour décrire un processus de diffusion d'humidité[28].

$$\frac{\partial m}{\partial t} = Deff\nabla^2 m \qquad \textbf{(I.15)}$$

Avec m est la teneur en humidité locale sur une base sèche (d.b.), t est le temps (s) et Deff est la diffusivité de l'humidité ($m^2.s^{-1}$). Cette équation peut être développée en coordonnées cartésiennes, cylindriques ou sphériques par des opérateurs *LAPLACIENS* appropriés [29].

Dans la plupart des situations, le produit alimentaire est considéré comme unidimensionnel et a une teneur en humidité initiale uniforme. Les hypothèses impliquées dans l'analyse de la diffusion sont : le mouvement interne de l'humidité comme résistance principale, aucun retrait dans le produit pendant le séchage et un effet de transfert de chaleur externe et interne négligeable. Les solutions de l'équation fick dans de telles conditions pour différentes géométries ont été présentées par Crank. La solution de l'équation de Fick pour une ailette (slab) infinie est la suivante :

$$MR = \frac{M-M_e}{M_0-M_e} = \frac{8}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\exp\left[-(2n+1)^2\frac{\pi^2 D_{eff}t}{L^2}\right] \qquad \textbf{(I.16)}$$

Avec *M*, est la teneur en humidité d'un produit mesurée ou calculée pendant le séchage, exprimée sur la base sèche, *Mo*, est la teneur en humidité d'un produit avant le début du séchage, exprimée sur la base sèche, *Me*, est la teneur en humidité d'un produit en équilibre exprimées en base sèche et L est l'épaisseur de l'ailette (slab).

Les équations suivantes sont des solutions pour les cas d'un cylindre et d'une sphère infinis, respectivement :

$$MR = \frac{M-M_e}{M_0-M_e} = \frac{4}{b_n^2}\sum_{n=0}^{\infty}\exp\left[-\frac{b_n^2\pi^2 D_{eff}t}{r_c^2}\right] \qquad \textbf{(I.17)}$$

Avec, bn, n=1,2 … sont des constantes, et rc est le rayon du cylindre (m) et rs est le rayon de la sphère (m) (Hui et al., 2008). Cette équation pourrait également être résolue pour les formes sphéroïdes prolates et oblates, cardioïdes, n hexagones, ondulées et épitrochoïdes (Kahveci et Cihan, 2008). Il est possible d'obtenir des solutions analytiques de l'équation de diffusion pour certaines géométries simples sous certaines hypothèses. Supposer que :

- Le coefficient de diffusion est indépendant du taux d'humidité pour une température donnée.
- Le matériau est isotherme pendant le séchage.
- L'humidité est initialement répartie uniformément dans tout le matériau.
- L'humidité de surface des échantillons atteint instantanément l'équilibre avec les conditions de l'air ambiant.
- La taille et la géométrie du matériau restent constantes pendant le séchage.
- Il n'y a pas de génération de masse ou d'épuisement à l'intérieur du matériau.

Le séchage est un processus de transfert simultané de chaleur et d'humidité. La chaleur est nécessaire pour évaporer l'humidité qui est éliminée de la surface du produit de séchage par le milieu de séchage externe, généralement de l'air. Un certain nombre de produits biologiques, lors du séchage sous forme de particules individuelles dans des conditions externes constantes, présentent une perte d'humidité à taux constant pendant la période de séchage initiale, suivie d'une période de séchage à vitesse décroissante. Ces produits sèchent généralement entièrement pendant la période de taux de chute. Pendant cette période, la surface d'une particule de séchage n'est pas recouverte d'une fine couche d'eau, car la résistance interne au transport de l'humidité est devenue supérieure à la résistance externe. Cela implique que la vitesse de séchage diminue continuellement au cours du séchage[30].

Il existe des relations théoriques, semi-théoriques et empiriques pour prédire le comportement de séchage de différents produits pendant la période de séchage à vitesse décroissante. Seuls les modèles semi-théoriques et empiriques se sont avérés utiles pour les concepteurs de séchoirs [31].

Un modèle de séchage semi-théorique est basé sur la théorie de la diffusion, en supposant que la résistance à la diffusion de l'eau dans un produit se produit dans la couche externe du produit. Ainsi, en utilisant la loi de refroidissement de Newton et en supposant une similitude entre le refroidissement et le séchage d'un corps solide, l'expression suivante peut décrire la vitesse de séchage

$$\frac{\partial M}{\partial t} = -k(M - M_e) \qquad \textbf{(I.18)}$$

Avec *k* est le coefficient de séchage, *Me* est la teneur en humidité à l'équilibre d'un produit et *M* est la teneur en humidité moyenne d'un produit[32].

Les modèles de séchage empiriques donnent souvent les meilleurs résultats pour prédire le comportement de séchage. Les équations peuvent être utilisées en toute confiance dans la plage de température, d'humidité relative, de vitesse d'écoulement de l'air et de teneur en humidité pour laquelle elles ont été dérivées. Ces modèles peuvent être utilisés dans les contrôles automatiques des processus de séchage pour une économie et un temps de calcul court.

Les modèles empiriques constituent une relation directe entre la teneur moyenne en humidité et le temps de séchage. Ils négligent les principes fondamentaux du processus de séchage et leurs paramètres n'ont donc aucune signification physique.

Les modèles semi-théoriques sont généralement dérivés en simplifiant des séries générales de la deuxième loi de Fick ou sont des formes modifiées de modèles simplifiés. Les modèles empiriques et semi-théoriques nécessitent peu de temps par rapport aux modèles théoriques et ne nécessitent pas d'hypothèses sur la géométrie d'un aliment typique ou sa diffusivité et conductivité de masse. Par conséquent, ils sont utiles pour les processus de contrôle automatique.

11.2. Modèles semi-théoriques et empiriques

11.2.1. Modèle de Lewis

Lewis a décrit que le transfert d'humidité des matériaux agricoles peut être considéré comme similaire à la loi de la chaleur d'un corps immergé dans un fluide froid. En comparant ce phénomène avec la loi de refroidissement de Newton, la vitesse de séchage est proportionnelle à la différence entre la teneur en humidité réelle et à l'équilibre.

$$\frac{dM}{dt} = -k(M - M_e) \quad \textbf{(I.19)}$$

La loi de refroidissement de Newton suppose que la résistance interne au mouvement de l'humidité et donc les gradients d'humidité dans le matériau sont négligeables.

La loi de refroidissement de Newton suppose que la résistance interne au mouvement de l'humidité et donc les gradients d'humidité dans le matériau sont négligeables, elle ne considère que la résistance de surface[33].

En supposant une condition aux limites telle que M = M_0 à t = 0, la solution de l'équation ci-dessus est connue sous le nom de modèle de Lewis, Newton, simple ou exponentiel[34]**.**

$$MR = \exp(-kt) \quad \textbf{(I.20)}$$

Avec k est la constante de séchage (s-1).

C'est l'un des modèles les plus simples décrivant le mouvement de l'humidité. L'inconvénient le plus important de ce modèle est qu'il sous-estime généralement les étapes tardives et surestime les premières étapes du processus de séchage[33,35,36]**.**

11.2.2. Modèle de Page

Ce modèle est une modification empirique du modèle de Lewis pour éliminer les défauts de ce modèle en ajoutant une constante empirique sans dimension (n) au terme de temps. Ce paramètre a pour effet de modérer le temps, et le modèle dans ce cas donne de meilleurs résultats pour la prédiction de la perte d'humidité [37].

$$\boldsymbol{MR = \exp(-kt^n)} \quad \textbf{(I.21)}$$

Ce modèle a été utilisé par de nombreux chercheurs pour décrire le taux de perte d'humidité lors du séchage en couche mince de matériaux agricoles et autre dans des conditions de séchage constantes[38].

11.2.3. Modèle de Page modifié

Il y a quelques modifications différentes dans le modèle de Page et elles ont été testées pour décrire le comportement de séchage de différents produits.

Modèle de Page modifié I [39]

$$\mathbf{MR = \exp[(-kt)^n]} \quad \textbf{(I.22)}$$

Modèle de Page modifié II [39]

$$\mathbf{MR = \exp[-(kt)^n]} \quad \textbf{(I.23)}$$

Modèle de Page modifié III [40]

$$\mathbf{MR = \exp[-(-kt)^n]} \quad \textbf{(I.24)}$$

Modèle de Page modifié IV [41]

$$\mathbf{MR = a\ \exp[-(kt^n)]} \quad \textbf{(I.25)}$$

Modèle de Page modifié V [42]

$$\mathbf{MR = \exp[-(kt^n)]} \quad \textbf{(I.26)}$$

Modèle de Page modifié VI [43]

$$\mathbf{MR = \exp(kt^n)} \quad \textbf{(I.27)}$$

Modèle de Page modifié VII [44]

$$\mathbf{MR = \exp\left[-k\left(\frac{t}{L^2}\right)^n\right]} \quad \textbf{(I.28)}$$

Modèle de Page modifié IX [45]

$$\mathbf{MR = k\ \exp\left[\left(\frac{-t}{L^2}\right)^n\right]} \quad \textbf{(I.29)}$$

11.2.4. Modèle de Otsura et al

Otsura et al, modèle était assez similaire à l'équation de Page et utilisé pour le séchage en couche mince du riz brut [46].

$$\mathbf{MR = 1 - exp[-(kt^n)]} \quad \textbf{(I.30)}$$

11.2.5. Modèles de Fick simplifiés

Une solution simplifiée de l'équation de diffusion de Fick valable pour de longs temps de séchage [44].

$$\mathbf{MR = k\ exp\left[-c\left(\frac{t}{L^2}\right)\right]} \quad \textbf{(I.31)}$$

11.2.6. Modèle de Henderson and Pabis

Le modèle de Henderson et Pabis est le premier terme d'une solution générale en série de la deuxième loi de Fick [47].

$$\mathbf{MR = a\ exp(-kt)} \quad \textbf{(I.32)}$$

11.2.7. Modèle de Henderson et Pabis modifié - I

Ce modèle est amélioré en ajoutant le troisième terme de la solution en série générale de la deuxième loi de diffusion de Fick pour corriger les lacunes du modèle de Henderson et Pabis [48]

$$\mathbf{MR = a\ exp(-k_0 t) + b\ exp(-k_1 t) + c\ exp(-k_2 t)} \quad \textbf{(I.33)}$$

11.2.8. Modèle de Henderson et Pabis modifié - II

Modèle de Henderson and Pabis I est amélioré en ajoutant le n comme paramètre [49]

$$\mathbf{MR = a\ exp(-kt^n) + b\ exp(-gt) + c\ exp(-ht)} \quad \textbf{(I.34)}$$

11.2.9. Modèle Logarithmique

Ce modèle est sous forme logarithmique du modèle de Henderson et Pabis avec un ajout d'un terme empirique [50].

$$\mathbf{MR = a\ exp(-kt) + c} \quad \textbf{(I.35)}$$

11.2.10. Modèle de Two Term

Ce modèle est constitué des deux premiers termes des solutions en séries générales de la solution analytique de l'équation de diffusion de la deuxième loi de Fick [51].

$$\mathbf{MR = a\ exp(-k_0 t) + b\ exp(-k_1 t)} \quad \textbf{(I.36)}$$

11.2.11. Modèle de Two Term modifié

Les modifications du modèle à deux termes aux différentes formes ont été données ci-dessous

Modèle de Two Term modifié I [52]

$$\mathbf{MR = a\ exp(k_0 t) + (1 - a)\ exp(-k_1 t)} \quad \textbf{(I.37)}$$

Modèle de Two Term modifié II [43]

$$\mathbf{MR = a\,exp(k_0 t) + (1 - a)\,exp(k_1 t)} \quad \textbf{(I.38)}$$

Modèle de Two Term modifié III [18]

$$\mathbf{MR = a\,exp(-k_0 t) + a\,exp(-k_1 t)} \quad \textbf{(I.39)}$$

Modèle de Two Term modifié IV [43]

$$\mathbf{MR = a\,exp(-k_0 t^n) + b\,exp(-k_1 t)} \quad \textbf{(I.40)}$$

Modèle de Two Term modifié V [55]

$$\mathbf{MR = a\,exp(-k_0 t) + (1 - a)\,exp(-k_1 t)} \quad \textbf{(I.41}$$

11.2.12. Modèle de Exponentiel Two Term

Ce modèle est les deux premiers termes de la solution en série générale de la solution analytique de la deuxième loi de Fick. Le modèle à deux termes a été modifié en diminuant le nombre constant et en organisant le deuxième terme exponentiel. Il a été souligné que le coefficient "b" dans le modèle à deux termes doit être (1 -a) à t = 0 pour obtenir MR = 1 et ce nouveau modèle est proposé comme (Doymaz, 2006) [56].

$$\mathbf{MR = a\,exp(-kt) + (1 - a)\,exp(-kat)} \quad \textbf{(I.42)}$$

11.2.13. Modèle de Verma et al.

Le modèle exponentiel Two Term est modifié en ajoutant une constante empirique [57].

$$\mathbf{MR = a\,exp(-kt) + (1 - a)\,exp(-gt)} \quad \textbf{(I.43)}$$

11.2.14. Modèle d'approximation de diffusion

Après réarrangement du modèle de Verma et al en séparant le terme constant de séchage « k » de « g », le nouveau modèle est obtenu comme (Erbay et Icier, 2009) [58] **:**

$$\mathbf{MR = a\,exp(-kt) + (1 - a)\,exp(-kbt)} \quad \textbf{(I.44)}$$

11.2.15. Modèle de Midilli et al

Le modèle de Midilli et al est composé d'un terme exponentiel et d'un terme linéaire décrivant le taux d'humidité en fonction du temps de séchage[59].

$$\mathbf{MR = a\,exp(-kt^n) + bt} \quad \textbf{(I.45)}$$

11.2.16. Modèle Midilli et al modifié

Lorsque la teneur en humidité instantanée est égale à la teneur en humidité d'équilibre au début du processus de séchage, le coefficient du modèle de Midilli et al doit être égal à 1. Ainsi, le modèle modifié de Midilli et al est testé pour la fibre de lin et est donné comme suit[36] :

$$\mathbf{MR = exp(-kt^n) + bt} \quad \textbf{(I.46)}$$

11.2.17. Modifié Midilli et al Modifié I

The shape term “a” of the Midilli et al model had to be 1 at t=0 and proposed a modification [36] as:

$$\mathbf{MR = exp(-kt^n) + bt} \quad \textbf{(I.47)}$$

Modèle de Midilli et al Modifié II [43]

$$\mathbf{MR = exp(-kt) + bt} \quad \textbf{(I.48)}$$

11.2.18. Modèle de Midilli et al Modifié III [60]

$$\mathbf{MR = a\ exp(-kt) + bt} \quad \textbf{(I.49)}$$

11.2.19. Modèle Wang et Singh

Wang et Singh ont proposé une nouvelle équation quadratique pour ajuster les données de séchage[55] :

$$\mathbf{MR = 1 + at + bt^2} \quad \textbf{(I.50)}$$

11.2.20. Modèle Hii et al

Hii et al ont proposé un nouveau modèle composé de modèles de Page et de séchage à deux termes.

$$\mathbf{MR = a\ exp(-kt^n) + c\ exp(-gt^n)} \quad \textbf{(I.51)}$$

11.2.21. Modèles de distribution de Weibull I

Ce modèle n'a aucune signification physique ; c'est une approche purement statistique. Un modèle basé sur la distribution de Weibull est présenté ci-dessous [61]

$$\mathbf{MR = a - b\ exp[-(kt^n)]} \quad \textbf{(I.52)}$$

11.2.22. Modèles de distribution de Weibull II [62]

$$\boldsymbol{MR = a - b\ exp[-kt^n]} \quad \textbf{(I.53)}$$

11.2.23. Modèles de distribution de Weibull III [63]

$$\boldsymbol{MR = exp[-(t/a)^n]} \quad \textbf{(I.54)}$$

11.2.24. Modèles de Vega-Galvez et al I

Vega-Galvez et al. (2008) ont proposé trois modèles empiriques différents pour décrire le séchage

$$\boldsymbol{MR = n + k\sqrt{t^n}} \quad \textbf{(I.55)}$$

11.2.25. Modèles de Vega-Galvez et al II [43]

$$\mathbf{MR = exp(n + kt)} \quad \textbf{(I.56)}$$

11.2.26. Modèles de Vega-Galvez et al III [64]

$$\mathbf{MR = (a + bt)^2} \quad \textbf{(I.57)}$$

11.2.27. Modèles de Jena Das

Dans ce modèle, deux termes exponentiels de temps de séchage ont été incorporés. Les valeurs de l'exposant dans le premier et deuxième terme du modèle ont été maintenues constantes à 1,0 et 0,5 pour plus de simplicité, respectivement. Ce modèle est proposé pour surmonter l'effet négatif de la variation de la diffusivité sur la teneur en humidité [65].

$$\mathbf{MR = a\ exp\left(-kt + b\sqrt{t}\right) + c} \quad \textbf{(I.58)}$$

11.2.28. Modèles de Wang et al un terme

Ils ont proposé trois modèles différents

$$\mathbf{MR = a\,exp(bkt) + (1 - a)} \quad \textbf{(I.59)}$$

11.2.29. Modèles de Wang et al deux termes

$$\mathbf{MR = (1 - a)\,exp(bkt) + aexp(ckt)} \quad \textbf{(I.60)}$$

11.3. Analyse statistique pour la détermination des modèles appropriés

Afin de trouver le modèle le mieux adapté pour expliquer le comportement de séchage de tout produit avec différentes méthodes de séchage ou différentes conditions, des méthodes statistiques ont généralement été utilisées. Les principales méthodes utilisées pour les études de séchage dans la littérature sont présentées ci-dessous.

11.3.1. Coefficient de détermination, R2,

Est utilisé dans le contexte de modèles statistiques dont le but principal est la prédiction de résultats futurs sur la base d'autres informations connexes. C'est la proportion de variabilité dans un ensemble de données qui est prise en compte par le modèle statistique. Il fournit une mesure de la façon dont les résultats futurs sont susceptibles d'être prédits par le modèle. Le coefficient de détermination n'est probablement pas 0 ou 1, mais plutôt quelque part entre ces limites. Plus il est proche de 1, plus la relation entre les valeurs expérimentales et prédites est grande[66].

$$\boldsymbol{R^2 = 1 - \frac{SS\ résiduel}{SS\ Total\ corrigé}} \quad \textbf{(I.61)}$$

11.3.2. Coefficient de corrélation R :

Est la racine carrée de R2 (Neter et al., 1990). Il s'agit d'une mesure de la corrélation (dépendance linéaire) entre deux variables, donnant une valeur comprise entre +1 et −1 inclus. Il est largement utilisé dans les sciences comme mesure de la force de la dépendance linéaire entre deux variables.

$$\boldsymbol{r = \frac{\sum_{i=1}^{N}(MR_{exp,i} - MR_{exp,ave})(MR_{pre,i} - MR_{pre,ave})}{\sqrt{\sum_{i=1}^{N}(MR_{exp} - MR_{exp,ave})^2 \sum_{i=1}^{n}(MR_{pre} - MR_{pre,ave})^2}}} \quad \textbf{(I.62)}$$

11.3.3. Chi-square

Est le carré moyen des écarts entre les valeurs expérimentales et prédites pour les modèles et utilisé pour évaluer l'accord d'ajustement de chaque modèle [67].

$$\boldsymbol{X^2 = \frac{\sum_{i=1}^{n}(M_{R\,exp,i} - M_{R\,pre,i})^2}{N - n}} \quad \textbf{(I.63)}$$

11.3.4. Erreur relative moyenne en pourcentage

Est la moyenne de la différence relative en pourcentage entre les valeurs expérimentales et prédites. Il compare la différence absolue entre la teneur en humidité prévue et la teneur en

humidité expérimentale tout au long du processus de séchage. Il est utilisé pour évaluer la précision prédictive des modèles. Les valeurs de P inférieures à 5 % indiquent un excellent ajustement, moins de 10 % un bon ajustement, tandis que les valeurs supérieures à 10 % indiquent un mauvais ajustement[68].

$$\boldsymbol{P = \frac{100}{N} \sum_{i=1}^{N} \frac{|Mr_{exp,i} - Mr_{Per,i}|}{Mr_{exp,i}}} \quad \textbf{(I.64)}$$

11.4. Root-mean-square error

Est une mesure fréquemment utilisée des différences entre les valeurs prédites par un modèle ou un estimateur et les valeurs réellement observées à partir de la chose modélisée ou estimée

$$\mathbf{RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^{N} \left(MR_{exp,i} - MR_{pre,i} \right)^2}} \quad \textbf{(I.65)}$$

11.4.1. Mean absolute error

Est une erreur de mesure dont l'amplitude reste constante pour toutes les observations ; une sorte d'erreur systématique

$$\mathbf{MAE = \frac{1}{N} \sum_{i=1}^{N} \left| MR_{exp,i} - MR_{pre,i} \right|} \quad \textbf{(I.67)}$$

11.4.2. Erreur quadratique moyenne *MSE*

Est l'une des nombreuses façons de quantifier la différence entre les valeurs implicites d'un estimateur et les vraies valeurs de la quantité estimée. MSE est une fonction de risque, correspondant à la valeur attendue de la perte d'erreur quadratique ou de la perte quadratique. MSE mesure la moyenne des carrés des "erreurs". L'erreur est le montant par lequel la valeur impliquée par l'estimateur diffère de la quantité à estimer. La différence se produit en raison du caractère aléatoire ou parce que l'estimateur ne tient pas compte des informations qui pourraient produire une estimation plus précise.

$$\mathbf{MSE = \frac{1}{N} \sum_{i=1}^{N} \left(MR_{exp,i} - MR_{pre,i} \right)^2} \quad \textbf{(I.68)}$$

11.4.3. Erreur de biais moyenne, MBE

Est un moyen global de l'erreur, ce dernier calculé par la formule suivante [69].

$$\mathbf{MBE = \frac{1}{N} \sum_{i=1}^{N} \left(MR_{pre,i} - MR_{exp,i} \right)} \quad \textbf{(I.69)}$$

11.4.4. Écart-type

Écart-typemontre combien de variation ou de "dispersion" existe par rapport à la moyenne (moyenne ou valeur attendue). Un écart-type faible indique que les points de données ont tendance à être très proches de la moyenne, tandis qu'un écart-type élevé indique que les points

de données sont répartis sur une large plage de valeurs.

$$\mathbf{SD} = \sqrt{\frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i=1}}^{\mathbf{N}}\left(\mathbf{MR_{exp,i}} - \mathbf{MR_{pre,i}}\right)} \quad \textbf{(I.70)}$$

11.4.5. *t*-Statistique

Est utilisé pour déterminer si les estimations de l'équation sont ou non statistiquement significatives, c'est-à-dire non significativement différentes de leurs homologues réels, à un niveau de confiance particulier[70].

$$\boldsymbol{t - stat} = \left[\frac{(\boldsymbol{n-1})\boldsymbol{MBE^2}}{\boldsymbol{RMSE^2 - MBE^2}}\right]^{\frac{1}{2}} \quad \textbf{(I.71)}$$

11.4.6. Coefficient de détermination ajusté

Est utilisé pour différence entre les modèles [71]

$$\mathbf{R^2_{ajd}} = \mathbf{1} - (\mathbf{1} - \mathbf{R^2})\left(\frac{\mathbf{n-1}}{\mathbf{n-(k+1)}}\right) \quad \textbf{(I.72)}$$

12. Conclusion

En règle générale, en pharmacie, le séchage entre principalement en jeu dans le cadre de la fabrication de formes comprimés ou granulés après l'étape de granulation humide. C'est dans ce cas une opération destinée à éliminer les solvants, le plus fréquemment l'eau ou l'éthanol jusqu'à atteindre une humidité résiduelle adaptée à la suite du procédé (compression-conditionnement). De nombreuses autres applications du séchage sont exploitées en pharmacie, parmi lesquelles l'atomisation d'excipients, la préparation d'extraits de poudres, le séchage des extraits de plantes et la lyophilisation.

En physique, le séchage est l'opération unitaire consistant à extraire d'un matériau tout ou partie d'un solvant. Le cas de l'élimination de l'eau est le cas le plus couramment rencontré dans notre domaine.

Chapitre 2 Généralités sur quelques méthodes de la modélisation numérique.

1. Introduction

Le présent chapitre offre une vue d'ensemble des méthodes de modélisation numérique utilisées dans la recherche scientifique et l'industrie. Plus précisément, nous nous concentrerons sur deux approches largement utilisées : le plan d'expérience et le machine learning.

La modélisation numérique joue un rôle crucial dans la compréhension et l'optimisation des processus complexes. Elle permet d'explorer et de prédire le comportement des systèmes en utilisant des techniques basées sur les mathématiques et l'informatique. Dans ce chapitre, nous présenterons brièvement ces deux méthodes puissantes et leurs applications dans différents domaines.

Le plan d'expérience est une approche systématique pour concevoir et exécuter des expériences de manière à recueillir des données pertinentes et informatives. Cette méthodologie permet d'analyser les effets des variables indépendantes sur une variable dépendante donnée, tout en contrôlant les facteurs de confusion. Nous explorerons les principes de base du plan d'expérience, tels que la sélection des niveaux des variables, la randomisation des essais et l'analyse des résultats. Nous discuterons également des avantages de cette méthode pour optimiser les processus et prendre des décisions éclairées.

Ensuite, nous aborderons le machine learning, qui est un domaine de l'intelligence artificielle axé sur la construction de modèles et d'algorithmes qui permettent aux ordinateurs d'apprendre à partir de données et de prendre des décisions ou d'effectuer des prédictions sans être explicitement programmés. Nous présenterons les concepts clés du machine learning, tels que les réseaux de neurones, les algorithmes d'apprentissage supervisé et non supervisé, et les techniques d'évaluation des performances des modèles. Nous explorerons également les applications courantes du machine learning, notamment la classification, la régression et le clustering.

2. Généralité sur le plan d'expérience

2.1. Définition

L'approche Plan d'expériences (DOE) cherche à déterminer la relation entre deux grandeurs :

- La réponse correspondant à la grandeur physique étudiée.
- Les facteurs, qui correspondent à des grandeurs physiques modifiables par l'expérimentateur et supposent des changements affectant la réponse**.**

La quantité d'intérêt, généralement notée y, est nommée d'après la réponse, et la variable qui modifie la réponse est appelée un facteur x, on parle alors de facteur affectant la réponse. Les termes facteur et réponse sont couramment utilisés dans le domaine de la conception expérimentale.

2.2. Objectif de plan d'expérience

C'est un problème courant dans les affaires. En R&D, souvent la moitié des ressources est consacrée à la résolution de problèmes d'optimisation. Avec l'augmentation rapide du coût des expériences, l'optimisation doit être effectuée avec le moins d'expériences possible. C'est l'une des raisons importantes pour lesquelles un plan d'expériences est utilisé.

- Développement de nouveaux produits et procédés
- Optimiser la qualité et les performances des produits existants
- Optimisation des procédés industriels dans les domaines suivants (chimie, polymères, matériaux, pharmacie et pharmacie, agro-alimentaire, cosmétique, peintures).

2.3. Espace expérimental

Pour démontrer l'espace expérimental, nous utiliserons un espace à deux dimensions, ce qui facilitera une représentation graphique. Il est alors aisé d'étendre les concepts introduits à des espaces multidimensionnels. Les facteurs continus peuvent être représentés par des axes orientés avec des échelles. Le deuxième axe est disposé orthogonalement au premier axe. Ainsi, nous obtenons un repère cartésien définissant un espace euclidien à deux dimensions. Cet espace est appelé espace expérimental.

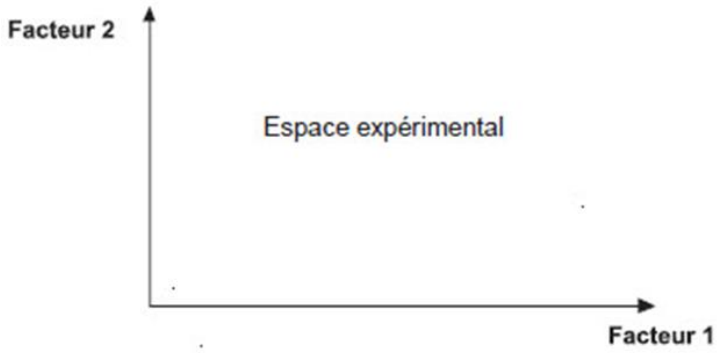


Figure 3. Espace expérimental de l'étude de deux facteurs pour un plan à deux dimensions.

2.4. Domaine d'un facteur

La valeur donnée à un facteur pour réaliser une expérience est appelée niveau. Lorsqu'on étudie l'influence d'un facteur, en général, on limite ses variations entre deux bornes. La borne inférieure est le niveau bas. La borne supérieure est le niveau haut (figure 4).

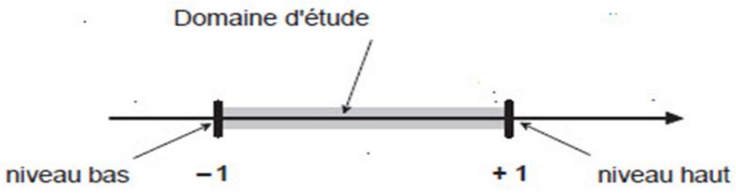


Figure 4 . Le domaine de variation des facteurs.

Le niveau bas du facteur est noté par - 1 et le niveau haut par +1. Le domaine de variation du facteur est constitué de toutes les valeurs comprises entre le niveau bas et le niveau haut. L'ensemble de toutes les valeurs que peut prendre le facteur entre le niveau bas et le niveau haut, s'appelle le domaine de variation du facteur ou plus simplement le domaine du facteur. Les variables centrées réduites idéique zéro [-1 0 1].

2.5. Domaine d'étude

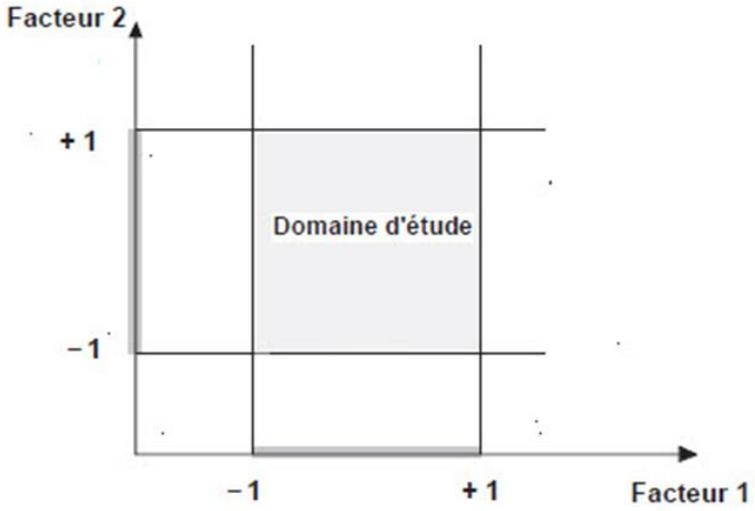


Figure 5. Le domaine d'étude est défini par la réunion des domaines des différents facteurs sans contraintes.

Il peut y avoir des contraintes sur le domaine d'étude La figure (6) illustre une réduction possible du domaine d'étude initial. Une zone du domaine d'étude initial échappe aux expériences.

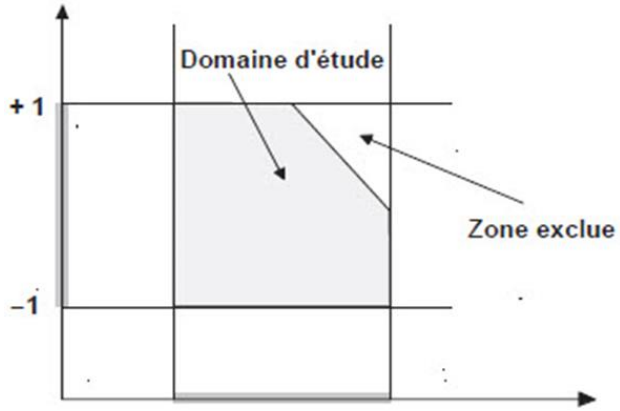


Figure 6. Le domaine d'étude sous contraintes est représenté par la partie grisée.

2.5.1. Méthodologie des plans sans contraintes

Le regroupement des domaines des facteurs définit le « domaine d'étude ». Ce domaine d'étude est la zone de l'espace expérimental choisie par l'expérimentateur pour faire ses essais. Une étude, c'est-à-dire plusieurs expériences bien définies, est représentée par des points

répartis dans le domaine d'étude. Cette façon de représenter une expérimentation par des points dans un espace cartésien est une représentation géométrique de l'étude.

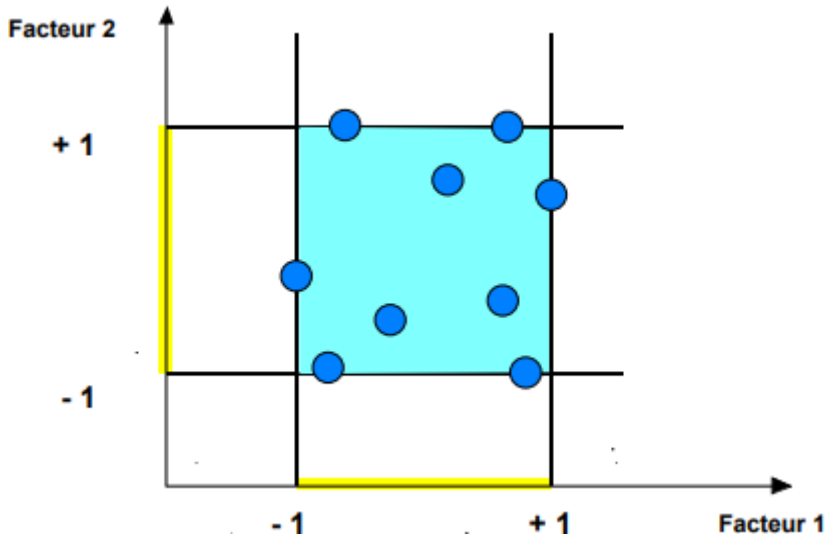


Figure 7. Les points expérimentaux sont disposés dans le domaine d'étude défini par l'expérimentateur

2.5.2. Notion de surface de réponse

À chaque point du domaine d'étude correspond une réponse. À l'ensemble de tous les points du domaine d'étude correspond un ensemble de réponses qui se localise sur une surface appelée surface de réponse, [3]

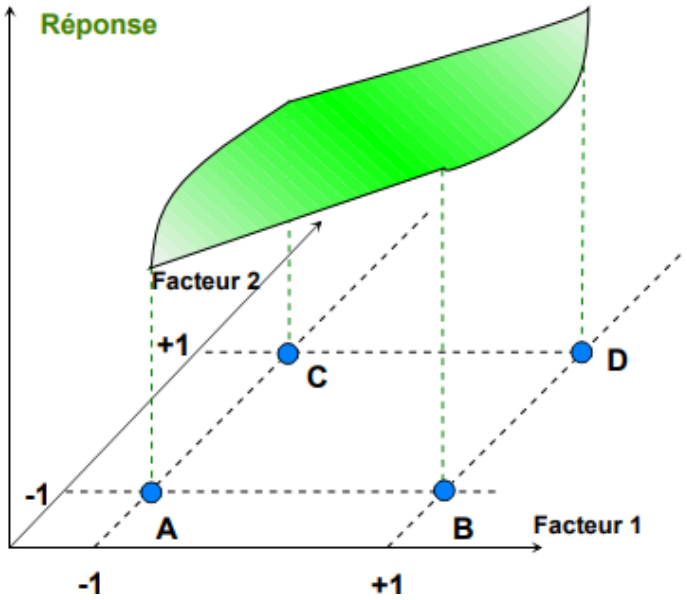


Figure 8. Les points expérimentaux sont disposés dans le domaine d'étude défini par l'expérimentateur.

2.6. Les méthodes des plans d'expériences

Les méthodes de plans d'expérience sont utilisées pour trois types d'investigations :

- **Les plans de criblage ou screening** sont utilisés au début d'une investigation. Ici, l'objectif premier est de réduire l'ensemble initial des facteurs qui peuvent affecter le résultat à un petit ensemble de facteurs qui ensemble influencent le plus le résultat.
- **Les plans en surface de réponse (MSR)** sont utilisés ensuite dans l'investigation pour l'optimisation finale, et lorsque l'on veut développer un modèle reliant les facteurs dominants aux réponses permettant de quantifier les résultats (performance, rendement, qualité).
- **Les plans spéciaux** sont nécessaires pour les mélanges. Ce type de problème est courant en chimie, dans les industries alimentaires, en cosmétique et en pharmacie. Ils sont nécessaires car les facteurs sont exprimés en pourcentages dont la somme fait 100%. Cela introduit une contrainte sur le plan et cela doit être résolu par des outils et des modèles spéciaux.

2.7. Les types de plans d'expériences

Nous allons maintenant appliquer les notions et les propriétés que nous venons de décrire aux plans d'expériences les plus classiques. Il y a les plans suivants :

- Plans factoriels complets.
- Plans factoriels fractionnaires.
- Plans à plusieurs niveaux.
- Plans pour surfaces de réponse.
- Plans de mélanges.

2.7.1. Plans factoriels complets

Une première catégorie de plans d'expériences est destinée à fournir une information la plus complète possible sur des systèmes présentant relativement peu de facteurs. Ces plans consistent à tester toutes les combinaisons possibles, en faisant varier tous les facteurs à tous leurs niveaux de manière exhaustive, d'où leur nom de plans factoriels complets ou plus simplement plans complets. Pour ce faire, *Yates et Hunter* ont proposé une technique simple pour le cas où chaque facteur n'a que deux niveaux. Elle consiste à numéroter les facteurs et à faire varier successivement leurs niveaux de la façon suivante :

Au premier essai, tous les facteurs sont au niveau bas, on change le niveau du premier facteur à chaque essai..., celui du deuxième facteur tous les 2 essais ..., et plus généralement, celui du Kème facteur tous les 2k-1 essais. La méthode peut s'adapter au cas où les facteurs ont plus de deux niveaux, [4].

2.7.1.1. Comment fonctionne un plan d'expériences ?

Les plans d'expériences (DOE) constituent la meilleure approche pour organiser un travail expérimental. Ils permettent de sélectionner un ensemble varié et représentatif d'expériences dans lequel tous les facteurs sont indépendants les uns des autres même si l'expérimentateur les fait varier façon simultanée. Le résultat est un modèle prédictif indiquant les importances des facteurs et de leurs interactions. Ce modèle peut être résumé sous la forme de graphiques de contours mettant en évidence la combinaison optimale des niveaux des facteurs. Les plans d'expériences sont utilisés pour trois objectifs, [5]:

- **Criblage :** quels sont les facteurs les plus influents et dans quelles plages
- **Optimisation :** comment trouver le réglage optimal prenant en compte de multiples réponses
- **Test de robustesse :** une fois le point optimal trouvé, peut-on garantir la robustesse au voisinage de ce point ou faut-il modifier les spécifications pour obtenir cette robustesse

2.7.2. Analyse statistique

Cette étape de l'analyse statistique aboutit à la construction du tableau d'analyse de régression et à la détermination de la qualité descriptive du modèle [6].

- Le Q^2 estime la capacité de prédiction du modèle, à savoir, sa capacité à généraliser à de nouvelles données invisibles. Le Q2 est défini comme :

$$Q^2 = 1 + \frac{PRESS}{SS_{tot}} \quad \textbf{(II.1)}$$

Où $$SS_{tot} = \sum_{i=1}^{n}(y_i - \bar{y})^2 \quad \textbf{(II.2)}$$

SS_{tot} : c'est la somme totale des carrés (total sum of squares) des échantillons de formation autour de la moyenne (proportionnelle à la variance de la réponse)

$\bar{y}$: c'est la réponse moyenne

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \textbf{(II.3)}$$

- Lorsque le modèle de régression est un modèle MLR, la « *PRESS* » est calculée comme dans généralisée croisée validation

$$PRESS_{MLR} = \sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{1 - x_i\,(X'X)^{-1}\,x'_i}\right)^2 \quad \textbf{(II.4)}$$

Lorsque le modèle est un modèle de régression PLS, la presse est calculée dans une croisée procédure de validation.

$$PRESS_{PLS} = \sum_{i=1}^{n}(y_i - \hat{y}_I^{CV})^2 \quad \textbf{(II.5)}$$

$\hat{y}_I^{CV}$: Est le prédites -out gauche réponse des valeurs de la validation croisée procédure.

- Les modèles de régression sont également validés par la R2, Le coefficient de détermination (aussi parfois appelé le coefficient de corrélation multiple). Le R2 désigne la fraction de la réponse qui est expliquée par le modèle et est défini comme

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y})^{2\prime}} \quad \textbf{(II.6)}$$

- L'ajusté R2 est la fraction de la variance de la réponse expliquée par le modèle, mais ajusté pour tenir compte des degrés de liberté du modèle, c'est-à-dire

$$R^2_{adj} = 1 - \frac{SS_{res}/DF_{res}}{SS_{tot}/DF_{tot}} = 1 - (1 - R^2)\frac{n-1}{DF_{res\prime}} \quad \textbf{(II.7)}$$

Où $DF_{tot} = n - 1$: est les degrés de liberté de la variance de la population de la variable dépendante.

DF_{res} : est les degrés de liberté de l'estimation de la variance de l'erreur de la population. Donc R^2_{adj} peut être négatif.

La valeur de $R^2_{adj} \leq R2 \leq 1$ n'augmente lorsque de nouvelles variables explicatives expliquent les données plus que ce qui peut être attribué au hasard, contrairement à la R2 ce qui augmente lorsque de nouvelles variables explicatives sont ajoutées au modèle.

Avec

y_i La réponse mesurée au point i,

$\hat{y}$ La réponse calculée par le modèle au point i de tous les échantillons,

- **Statistiques sommaires dans (Logiciel MODDE)**

Les statistiques sommaires sont présentées dans quatre paramètres (R2, Q2, validité du modèle, et reproductibilité) où 1, ou 100%, est parfait. R2 et Q2, les deux premières colonnes, devrait être proche de la taille. La différence ne devrait pas être plus de 20% dans la plupart des situations, [5].

R^2 : Montre l'ajustement du modèle, un modèle avec R^2 de 0,5 est un modèle avec assez faible importance.

Q2 : Affiche une estimation de la précision de prévision de l'avenir. Q2 doit être supérieure à 0,1 pour un modèle important et supérieur à 0,5 pour un bon modèle. Q2 est le meilleur indicateur.

Validité du modèle

Un test de problèmes modèles divers, une valeur inférieure à 0,25 pour modèle validité indique les problèmes modèles statistiquement significatifs, comme la présence de valeurs aberrantes, un incorrect modèle, ou un problème de transformation. Une faible valeur ici peut également indiquer qu'un terme, comme une interaction ou carrée est manquante. Lorsque l'erreur pure est très faible (réplique presque identique), la validité du modèle peut être faible, même si le modèle est bon et complet. Lorsque l'erreur pure est si faible que les répétitions sont considérés comme identiques par MODDE, la validité du modèle est étiquetée manquant car il ne peut être calculé. Validité du modèle pourrait être faible dans de très bons modèles (Q2 > 0,9) en raison de la haute sensibilité du test ou extrêmement bonnes répliques.

Reproductibilité

La variation des répliques par rapport à la variabilité globale, la reproductibilité doit être supérieure à 0,5.

3. Chimio-informatique

Le présent chapitre se concentre sur un domaine interdisciplinaire en pleine expansion appelé chimio-informatique. Cette discipline combine la chimie et l'informatique pour résoudre des problèmes liés à la découverte, la conception et l'optimisation de molécules à des fins pharmaceutiques, médicales et chimiques. La chimio-informatique joue un rôle essentiel dans le processus de développement de médicaments et dans la compréhension des interactions entre les molécules et les cibles biologiques. Elle exploite les avancées de l'informatique, de la modélisation moléculaire, de l'apprentissage automatique et de la bioinformatique pour accélérer et rationaliser les différentes étapes de la découverte de médicaments. Dans cette partie, nous explorerons les principes fondamentaux de la chimio-informatique. Nous aborderons les bases de données chimiques, qui recensent un large éventail de molécules et de structures chimiques. Nous discuterons des techniques de recherche et de filtrage de ces bases de données pour identifier des candidats potentiels pour la conception de médicaments.

Nous examinerons également les méthodes de modélisation moléculaire, qui permettent de prédire les propriétés physico-chimiques des molécules, leur activité biologique et leur interaction avec des cibles spécifiques. Nous discuterons des techniques de modélisation moléculaire telles que le docking moléculaire, la simulation de dynamique moléculaire et la QSAR (Quantitative Structure-Activity Relationship). Une autre facette importante de la chimio-informatique est l'utilisation de l'apprentissage automatique et de l'intelligence artificielle pour analyser les grandes quantités de données chimiques et biologiques. Nous explorerons les algorithmes d'apprentissage automatique utilisés pour la prédiction de l'activité biologique, la classification des molécules et la découverte de nouveaux composés. Enfin, nous

discuterons des applications concrètes de la chimio-informatique dans la recherche pharmaceutique, la conception de médicaments et la chimie médicinale. Nous aborderons les défis et les perspectives futures de ce domaine en pleine évolution. [72].

4. RDKit

RDKit est une boîte à outils open source largement utilisée en chimie computationnelle et en chimio-informatique. Il s'agit d'une collection de bibliothèques logicielles écrites en langage de programmation Python, conçue pour la manipulation, l'analyse et la visualisation de données chimiques. La chimie computationnelle joue un rôle essentiel dans la découverte et la conception de nouveaux médicaments, la prédiction des propriétés chimiques et biologiques des molécules, ainsi que dans la compréhension des interactions moléculaires. RDKit offre une gamme complète d'outils et de fonctionnalités pour faciliter ces tâches. L'une des principales forces de RDKit réside dans sa capacité à représenter et à manipuler des structures moléculaires. Il permet de générer, de convertir et de comparer différents formats de fichiers chimiques, tels que les fichiers SMILES et les fichiers SDF, facilitant ainsi l'importation et l'exportation de données chimiques. De plus, RDKit offre des fonctionnalités avancées pour la recherche de similarité moléculaire, la génération de conformères et la prédiction de propriétés moléculaires. RDKit propose également des outils pour l'analyse et la visualisation de données chimiques. Il permet d'effectuer des calculs de descripteurs moléculaires, tels que la masse moléculaire, la polarité et la flexibilité, qui fournissent des informations importantes sur les propriétés physico-chimiques des molécules. De plus, RDKit offre des fonctionnalités de visualisation graphique permettant de représenter les structures moléculaires et d'explorer les interactions atomiques. Grâce à sa flexibilité et à sa facilité d'utilisation, RDKit est devenu un outil essentiel pour de nombreux chercheurs et praticiens de la chimie computationnelle. Il est utilisé dans divers domaines, tels que la découverte de médicaments, la chimie médicinale, la toxicologie et la modélisation moléculaire.

5. Représentations moléculaires

Une représentation robuste des molécules [72] est nécessaire pour le fonctionnement précis de les ML modèles. Une représentation moléculaire idéale devrait être unique, invariante avec respect différent opérations de symétrie, inversible, efficace pour obtenir, et capturer la physique, la chimie stéréo, et le motif structurel. Certains d'entre eux peuvent être obtenus en utilisant le physique, chimique, et propriétés structurales [73].

Les molécules peuvent être représentées sur une feuille de papier à l'aide d'un schéma bidimensionnel (2D) comme celui du chlorhydrate de dicycloverine dans le schéma 1, mais il y a beaucoup de différents options pour représenter [74] composés par calcul. Lorsque les ordinateurs étaient commercialisés, des chaînes de caractères alphanumériques ont été

préférées, car elles nécessitent moins mémoire à stocker et moins de puissance de calcul à traiter, mais que les ordinateurs développés et la mémoire/puissance de traitement est devenue moins coûteuse, les représentations moins compactes (plus flexibles et moins ambiguës) sont devenues plus répandues. Un aperçu des différentes classes de représentation moléculaire.

5.1. MDL Molfile

Les fichiers fondus MDL se composent de trois sections principales : un bloc d'en-tête (contenant le titre, l'horodatage et un commentaire facultatif), une CCT et une ligne finale qui doit se lire « M END » [75][76] Le CTab se compose d'un certain nombre de sections, mieux comprises en considérant la **figure10**.

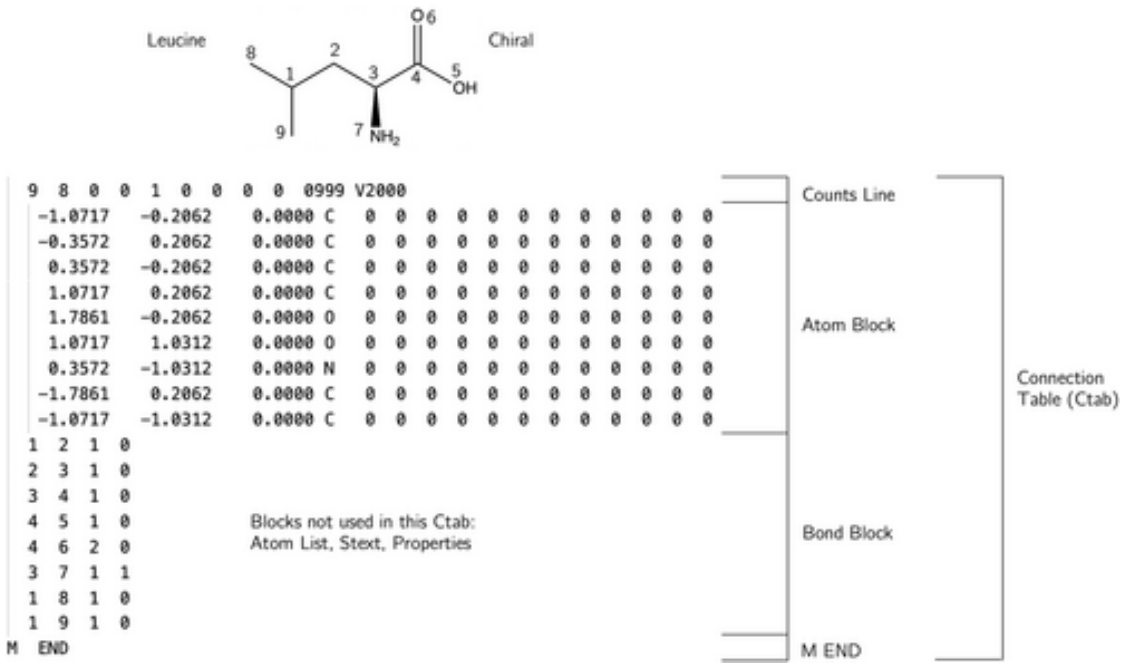


Figure 9. Exemple d'une table de connexion et d'une ligne d'extrémité dans un molfile MDL V2000 pour.

5.2. Système simplifié d'entrée de ligne moléculaire (smiles) :

SMILES [77] représente des molécules (organiques) avec une chaîne de caractères ASCII. Les atomes sont simplement représenté par le même symbole d'une ou deux lettres utilisé dans le tableau périodique. Les liaisons simples peuvent être implicites ou représentées avec, et double, triple et quadruple liaisons avec =, #, et $, respectivement. Les anneaux sont représentés avec un numéro après le (arbitrairement choisi) l'atome initial dans l'anneau et l'atome de fermeture (p. ex., C1CCNCC1 et N1CCCCC1 sont équivalent). Le branchement est représenté par des parenthèses autour de la branche, par exemple, 4-éthylheptane (une colonne vertébrale à 7-C avec une chaîne latérale à deux carbone sur le quatrième carbone) serait CCCC(CC)CCC.

Les succursales peuvent également être imbriquées dans d'autres succursales en ajoutant plus paires de parenthèses. L'aromatique peut être représenté soit par des liaisons simples/doubles, ou en écrivant des atomes liés aromatiquement en minuscules. Pour illustrer ceci, considérez

tous ces représentations équivalentes du benzène :

c1ccccc1, C1═C_C═C_C═C1, C1═CC═CC═C1.

Bien qu'il soit facile d'écrire une chaîne SMILES [78] qui est syntaxiquement correcte puisqu'il y a plusieurs façons d'écrire la même chose, cette nature non canonique de SMILES [79] peut rendre difficiles les recherches (sous)structures dans une base de données. différents algorithmes ont été développé pour canoniser SMILES [80] cordes dont Universal SMILES, Rdkit SMILES et CANGEN.

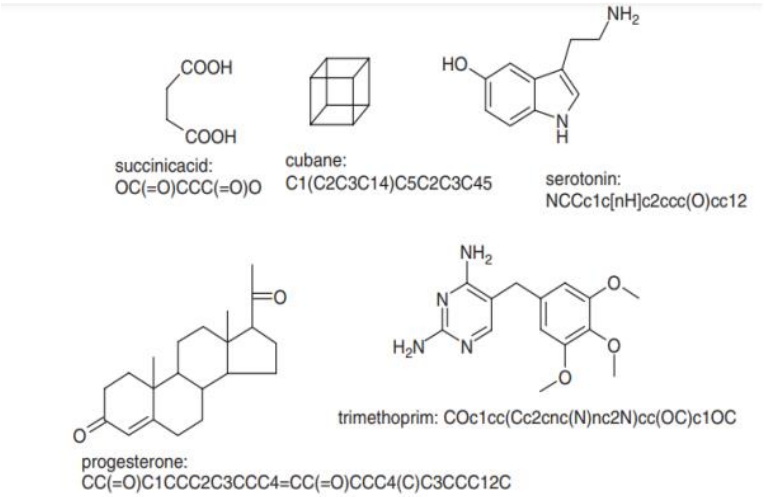


Figure 10.quelques exemples de (smiles)

5.3. Empreintes digitales à connectivité étendue

Empreintes moléculaires [81] sont destinées à représenter la présence (ou l'absence) de sous-structures au sein de molécules souvent dans un vecteur clairsemé, et ils tombent généralement dans l'un des deux catégories : faire correspondre les sous-structures d'une molécule à celles d'un ensemble défini par un expert, ou l'énumération algorithmique et le hachage des sous-structures d'une molécule.

L'ECFP est l'un des les techniques d'empreinte chimique les plus largement utilisées en raison en grande partie du paquet python populaire RDKit [82], appelé « Morgan » Toutefois, les techniques d'empreinte chimique existaient avant le PFCE, voir exemple HOSE et FREL. Comparer et contraster les méthodes d'empreintes digitales la documentation, donc ce texte donnera plutôt un bref aperçu de la façon dont les PFCA sont générés et peut être utilisé dans les modèles ML.

En générant un ECFP utilise des informations sur chaque atome dans une molécule pour produire un descripteur de l'atome et de son environnement immédiat qui est invariable avec la façon dont les atomes dans la molécule est numérotée [83]. Tout ensemble de propriétés qui remplissent ces critères pourrait être utilisé.

Pour les ECFP, la propriété définie provient de la règle des invariants atomiques de jour nombre de connexions, nombre de liaisons non hydrogène, numéro atomique, signe de charge, charge absolue, et nombre d'hydrogènes attachés.

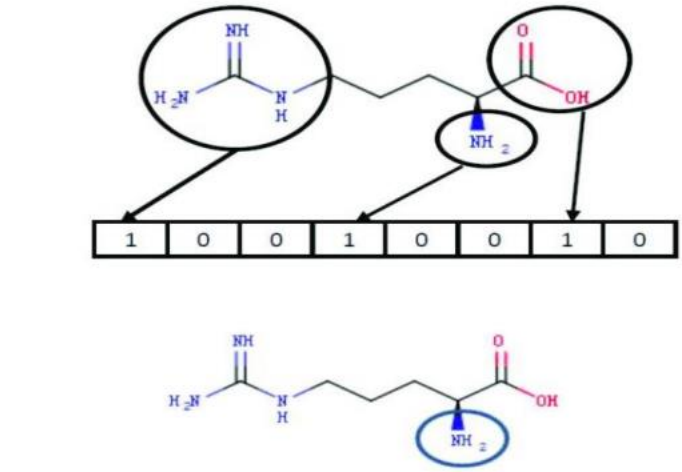


Figure 11. Exemple d'empreinte digitale

6. Deep learning

Deep learning est un type d'apprentissage automatique qui permet aux ordinateurs d'apprendre de leurs erreurs et de comprendre le monde comme une hiérarchie de concepts [84]. Il n'est pas nécessaire qu'un opérateur d'ordinateur humain décrive expressément toute l'information dont l'ordinateur a besoin puisque l'ordinateur apprend par expérience. Un graphique de ces hiérarchies serait de nombreuses couches profondes, permettant à la machine de comprendre des concepts sophistiqués en les construisant à partir de concepts plus simples. Ce livre couvre une grande variété de sujets liés à l'apprentissage profond. Le travail fournit une base mathématique et conceptuelle en algèbre linéaire, théorie des probabilités et théorie de l'information, calcul numérique et apprentissage automatique, entre autres sujets. Il couvre des applications telles que le traitement du langage naturel, la reconnaissance de la parole, la vision par ordinateur, les systèmes de recommandation en ligne, la bio-informatique et les jeux vidéo, ainsi que les techniques d'apprentissage profond utilisées par les praticiens de l'industrie, telles que les réseaux de flux avant profond, régularisation, algorithmes d'optimisation, réseaux convolutionnels, modélisation de séquences et méthodologie pratique [85]. Enfin, le livre fournit des vues de recherche sur les modèles de facteurs linéaires, les encodeurs automatiques, l'apprentissage de la représentation, les modèles probabilistes structurés, les techniques de

Monte Carlo, la fonction de partition, l'inférence d'approximation et les modèles génératifs profonds, entre autres sujets théoriques. Les étudiants de premier cycle et des cycles supérieurs intéressés par un emploi dans les affaires ou la recherche, ainsi que les développeurs de logiciels intéressés à intégrer l'apprentissage profond dans leurs produits ou plateformes, peuvent bénéficier de l'apprentissage profond. Un de ces domaines est la chimie, qui traite d'un large éventail de défis extrêmement difficiles tels que la rétro synthèse, l'optimisation des réactions et la création de médicaments. Historiquement, ils ont fait face à d'intenses critiques de la part des chimistes, qui ont vu les techniques informatiques basées sur l'heuristique et les règles codées à la main avec méfiance. Ces questions sont très semblables à ceux que l'apprentissage profond a grandis pour dominer, comme la vision par ordinateur traitement du langage naturel. 12 En conséquence, l'adoption de ces technologies en chimie a régulièrement augmenté, beaucoup affichant des augmentations considérables de la précision prévue et capacité d'émuler la prise de décision humaine. Compte tenu de l'utilisation généralisée de l'apprentissage profond dans chimie, il est essentiel pour les chercheurs en chimie qui ne sont pas conscients des complexités des techniques pour se familiariser avec leur fonctionnement.

Apprentissage profond en chimie, expliquant les idées de base dans le sens le plus large possible, sans accent sur les détails mathématiques, et en référence aux applications chimiques. Cette la compréhension sera ensuite utilisée pour fournir un aperçu général des influences de l'apprentissage profond a jusqu'ici eu sur la chimie appliquée et théorique.

6.1. Convolution Neural Network (CNN) ou réseaux de neurones convolutifs

Les CNN désignent une sous-catégorie de réseaux de neurones et sont à ce jour un des modèles de classification d'images réputés être les plus performants.

Leur mode de fonctionnement est à première vue simple : l'utilisateur fournit en entrée une image sous la forme d'une matrice de pixels. Celle-ci dispose de 3 dimensions :

Deux dimensions pour une image en niveaux de gris.

Une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales (Rouge, Vert, Bleu).

Contrairement à un modèle MLP (Multi Layers Perceptron) classique qui ne contient qu'une partie classification, l'architecture du Convolutional Neural Network dispose en amont d'une partie convolutive et comporte par conséquent deux parties bien distinctes :

Une partie convolutive : Son objectif final est d'extraire des caractéristiques propres à chaque image en les compressant de façon à réduire leur taille initiale. En résumé, l'image fournie en entrée passe à travers une succession de filtres, créant par la même occasion de nouvelles images appelées cartes de convolutions. Enfin, les cartes de convolutions obtenues sont concaténées dans un vecteur de caractéristiques appelé code CNN.

Une partie classification : Le code CNN obtenu en sortie de la partie convolutive est fourni en entrée dans une deuxième partie, constituée de couches entièrement connectées appelées perceptron multicouche (MLP pour Multi Layers Perceptron). Le rôle de cette partie est de combiner les caractéristiques du code CNN afin de classer l'image

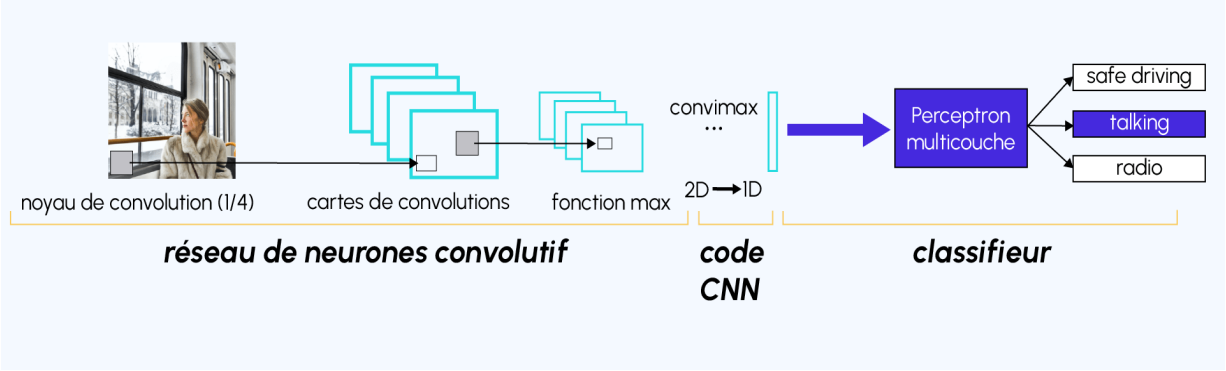


Figure 12.CNN Architecture

6.2. Recurrent Neural Network

Un Réseau de Neurones Récurent (RNN) est un type de réseau de neurones artificiels qui peut traiter des données séquentielles en maintenant un état interne (mémoire) des informations qu'il a traitées jusqu'à présent. Cela rend les RNN adaptés aux tâches telles que le traitement du langage naturel, la reconnaissance vocale et la prédiction de séries temporelles.

Dans un RNN, chaque neurone (unité) a une connexion récurrente avec lui-même, ce qui permet au réseau de conserver les informations des entrées précédentes tout en traitant les nouvelles. La sortie de chaque neurone est une fonction de son entrée actuelle et de son état interne (sortie de l'étape de temps précédente).

Le type le plus courant de RNN est le réseau LSTM (Long Short-Term Memory), qui dispose de mécanismes supplémentaires pour stocker et récupérer sélectivement des informations de la mémoire. Les LSTM sont particulièrement efficaces pour traiter le problème de gradient évanouissant qui peut survenir dans les RNN traditionnels, où les gradients deviennent très petits ou disparaissent complètement à mesure qu'ils se propagent dans le temps.

Dans l'ensemble, les RNN et les LSTM se sont révélés être des outils puissants pour le traitement de données séquentielles et ont atteint des résultats de pointe dans diverses applications.

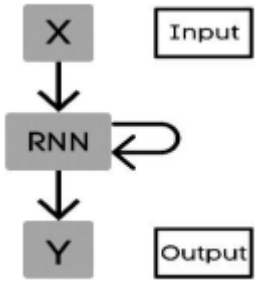


Figure 13. RNN Architecture.

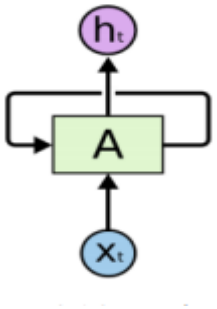


Figure 14. RNN loops.

6.3. Auto-encodeur

Un auto-encodeur est un type de réseau de neurones artificiels qui est utilisé pour la réduction de la dimensionnalité et la reconstruction d'informations. Il s'agit d'un réseau de neurones à architecture symétrique qui prend en entrée des données brutes et les comprime en une représentation latente de plus petite dimension. Ensuite, il décompresse cette représentation pour reconstruire les données d'origine.

L'auto-encodeur est composé de deux parties principales : un encodeur et un décodeur. L'encodeur prend en entrée les données et les transforme en une représentation de plus petite dimension. Le décodeur prend ensuite cette représentation et la transforme en une reconstruction des données d'origine.

L'objectif principal de l'auto-encodeur est de minimiser l'erreur de reconstruction entre les données d'origine et les données reconstruites par le décodeur. En minimisant cette erreur, l'auto-encodeur peut apprendre à extraire les caractéristiques importantes des données et à les représenter de manière plus compacte.

Les auto-encodeurs peuvent être utilisés pour diverses tâches, telles que la réduction de la dimensionnalité, la détection d'anomalies, la génération de données, la compression de données et la préparation de données pour l'apprentissage supervisé.

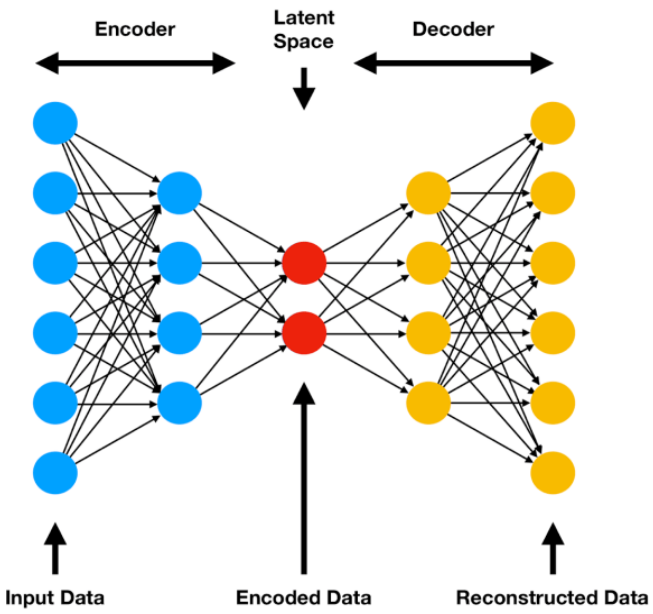


Figure 15. Illustration simplifiée de l'Auto-Encodeur d'architecture.

- Encodeur : Il s'agit de la partie du réseau qui comprime l'entrée en un nombre réduit de bits appelé espace latent, parfois appelé goulot d'étranglement. Cette représentation d'espace latent est appelée "encodage" de l'entrée.
- Décodeur : C'est la partie du réseau qui reconstruit l'image d'entrée à partir de la représentation compressée.

6.3.1. Auto-encodeur variationnelle (VAE)

Auto-encodeur variationnelle (VAE) peut être considéré comme une version probabiliste d'un auto-encodeur (AE), où la sortie du décodeur n'est pas directement une valeur de x mais les paramètres d'une distribution de probabilité de x. Comme nous le verrons ci-dessous, la même formulation probabiliste s'applique à l'encodage de z. Le modèle probabiliste résultant peut être utilisé pour:

- Générer de nouvelles données (à partir de valeurs de z non observées).
- Transformer des données existantes dans un schéma d'encodage-modification-décodage.

6.4. Perceptron multicouche « Multi-Layer Perceptron (MLP) »

Le MLP, abréviation de Multi-Layer Perceptron, est une architecture couramment utilisée dans l'apprentissage automatique pour les réseaux de neurones artificiels (RNA). Il se compose de plusieurs couches de neurones qui sont entièrement connectées les unes aux autres. La première couche, appelée couche d'entrée, reçoit les données d'entrée, tandis que la dernière couche, appelée couche de sortie, produit les prédictions du modèle. Les couches intermédiaires, connues sous le nom de couches cachées, sont chargées d'extraire les caractéristiques des données d'entrée. Chaque neurone dans une couche du MLP est connecté à

tous les neurones de la couche précédente et de la couche suivante. Chaque connexion possède un poids associé, qui est ajusté pendant l'entraînement du modèle afin de minimiser l'erreur de prédiction. Le MLP utilise des fonctions d'activation non linéaires pour introduire de la non-linéarité dans le modèle, améliorant ainsi sa capacité à modéliser des relations complexes entre les entrées et les sorties. Le MLP est un modèle largement utilisé pour la classification et la régression. Il a été appliqué avec succès dans divers domaines, tels que la reconnaissance vocale, la vision par ordinateur, la détection de fraude et la prédiction de séries temporelles.

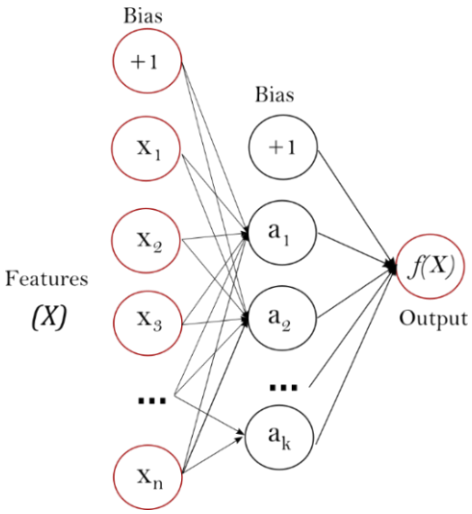


Figure 16. Multi-Layer Perceptron MLP.

7. Conclusion

En conclusion, ce chapitre a présenté un aperçu des différentes méthodes de modélisation numérique utilisées dans divers domaines tels que la chimie, les sciences pharmaceutiques et l'apprentissage automatique.

Nous avons examiné les plans d'expérience, une approche statistique permettant de concevoir des expériences de manière systématique afin de recueillir des données pertinentes et d'optimiser les processus. Les plans d'expérience offrent une approche efficace pour étudier et modéliser les relations entre les facteurs et les réponses, permettant ainsi d'optimiser les paramètres et de prendre des décisions éclairées. Ensuite, nous nous sommes intéressés à la chimio-informatique, une discipline interdisciplinaire qui associe la chimie et l'informatique pour la découverte et la conception de molécules à des fins pharmaceutiques. Nous avons exploré les bases de données chimiques, les méthodes de modélisation moléculaire et l'utilisation de l'apprentissage automatique pour analyser les données chimiques massives. La chimio-informatique offre des outils puissants pour accélérer le processus de découverte de médicaments et optimiser la conception de molécules. Enfin, nous avons abordé le domaine du deep learning, une branche de l'apprentissage automatique basée sur des réseaux de neurones artificiels profonds. Nous avons exploré les architectures de réseaux neuronaux profonds, telles

que le Multi-Layer Perceptron (MLP) et les réseaux de neurones convolutifs (CNN), qui ont montré des performances remarquables dans des domaines tels que la reconnaissance d'images et le traitement du langage naturel. En intégrant ces différentes méthodes de modélisation numérique, il devient possible d'optimiser les processus, de découvrir de nouvelles molécules prometteuses et d'exploiter des relations complexes entre les données d'entrée et les sorties souhaitées. L'utilisation de méthodes de modélisation numérique telles que les plans d'expérience, la chimio-informatique et le deep learning offre des opportunités passionnantes pour la recherche et le développement dans divers domaines scientifiques et pharmaceutiques. En combinant ces approches, nous pouvons accélérer la découverte de connaissances, améliorer les performances des processus et faire progresser les avancées scientifiques. Il reste encore de nombreux défis à relever, mais ces méthodes offrent un potentiel considérable pour l'innovation et la découverte de nouvelles solutions.

Chapitre 3 Etude expérimentale

Introduction

Le chapitre expérimental de cette étude se concentre sur le processus de séchage de trois poudres pharmaceutiques spécifiques. L'objectif principal est d'explorer et d'optimiser ce processus en combinant la planification de plans d'expérience et les techniques avancées de deep learning.

Le séchage des poudres pharmaceutiques est une étape cruciale dans la production de médicaments, car il influence directement la qualité et la stabilité du produit final. Une compréhension approfondie de ce processus est essentielle pour améliorer l'efficacité et la rentabilité de la production pharmaceutique.

Pour atteindre cet objectif, nous utilisons une approche en deux volets. Tout d'abord, nous mettons en place une planification de plans d'expérience. Cette méthodologie nous permet de concevoir et de conduire des expériences de séchage en manipulant de manière systématique les variables clés telles que la température, la pression, l'humidité relative et le temps. Les données expérimentales recueillies sont ensuite utilisées pour établir des relations entre ces variables et les performances de séchage des poudres pharmaceutiques.

Ensuite, nous exploitons les techniques avancées de deep learning pour modéliser et prédire le processus de séchage. Le deep learning, en particulier les réseaux de neurones, est un puissant outil d'apprentissage automatique capable de découvrir des motifs complexes dans les données. En utilisant des architectures de réseaux de neurones adaptées, nous développons des modèles prédictifs qui peuvent estimer les résultats du séchage en fonction des différentes conditions expérimentales.

L'utilisation de la planification de plans d'expérience combinée au deep learning offre un cadre méthodologique complet pour l'étude du séchage des poudres pharmaceutiques. Cette approche permet de maximiser l'efficacité des expériences et d'obtenir des modèles de prédiction précis, ce qui facilite l'optimisation des paramètres de séchage et la prise de décisions éclairées dans l'industrie pharmaceutique.

Dans les prochaines sections, nous détaillerons les plans d'expérience mis en place, les méthodes de collecte de données, les architectures de réseaux de neurones utilisées et les résultats obtenus. Cette approche combinée de planification de plans d'expérience et de deep learning ouvre de nouvelles perspectives pour améliorer le procédé de séchage des poudres pharmaceutiques et optimiser la production de médicaments de haute qualité.

Contribution 1 : Séchage de quelques poudre pharmaceutiques

1.1. Introduction

Le but de cette contribution est d'étudier expérimentalement l'effet de quelques paramètres opératoires sur le séchage sous vide de trois poudres pharmaceutiques. Les deux principes actifs sont : " paracétamol, nystatine " et l'excipient :"PVP (polyplasdone) ".

La partie expérimentale et modélisation seront réalisé respectivement au sein de laboratoire de contrôle qualité SAIDAL de Médéa.

On va étudier expérimentalement l'effet de quelques paramètres opératoires {Température T, Épissure Ep, Pression P, masse m, taux d'humidité X0} sur le temps de séchage. La modélisation de temps de séchage repose sur l'utilisation de la méthode de plan d'expériences (par Logiciel MODDE®) qui permis d'obtenir un modèle empirique quadratique en n'effectuant qu'un nombre minimal d'expérimentations.

2. Application de la méthode de plan d'expérience

2.1. Introduction

Nous nous servirons des plans d'expérience pour quantifier les influences des variables température, pression, la teneur en eau et la masse sur le temps de séchage. Toutes les planifications des expériences et l'analyse statistique des résultats sont faites par le logiciel MODDE (version 6).

MODDE est un logiciel fonctionnant sous Windows permettant de définir des plans d'expériences, de les analyser et de les optimiser, le tout dans un unique produit convivial. Le logiciel est conçu pour les techniciens et les ingénieurs de tous les secteurs de l'industrie.

Une dépendance fondamentale existe entre l'objectif recherché et la définition du plan d'expériences. Cependant, dans les deux cas, les étapes de la démarche se déroulent dans un ordre similaire à savoir :

- Définition des objectifs et des réponses.
- Choix d'une stratégie expérimentale.
- Définition des facteurs.
- Définition du domaine expérimental.
- Expression du modèle.
- Construction du plan d'expériences.
- Expérimentation.
- Analyse globale des résultats d'essais.
- Analyse mathématique des résultats d'essais.
- Analyse statistique du modèle.

- Analyse graphique du modèle.
- Validation du modèle et des informations obtenues.

2.2. Définition des objectifs et des réponses

Le but de cette étude est d'estimer l'effet des facteurs identifiés comme potentiellement influent par les physiciens[23], et d'identifier ceux ayant une influence statistiquement significative sur les performances observées. Les interactions seront également observées afin de compléter l'estimation des effets moyens des facteurs.

La réponse de cette étude identifiée comme le temps nécessaire pour sécher les poudres pharmaceutiques choisi dans le domaine expérimental.

2.3. Définition des facteurs

Cette étape définir 7 facteurs potentiellement influents sur la réponse observée, pour deux natures différentes ont été identifiés dans le Tableau 1.

Tableau 1. La nature des différents Facteurs.

Natures	**Facteurs**
Qualitative	Le mode de séchage
	La matière de support
	La poudre
	Solvant de mouillage
Quantitative	La teneur en eau
	L'épaisseur
	La surface de contacte
	La température
	La pression

2.4. Définition du domaine expérimental

Le domaine expérimental est constitué de l'ensemble des combinaisons des facteurs qu'il est possible de réaliser. Dans notre cas, nous avons étudié les facteurs quantitatifs et fixe les facteurs qualitatifs plus la surface de contacte, le tableau 1 identifiés le domaine de ces facteurs et le tableau 2 résume les facteurs fixés.

Tableau 2. Le domaine des facteurs étudiés.

Facteurs	Unité	Domaine
Température	°C	[40-80]
Pression	bar	[0.3-1]
La teneur en eau initial	%	[15-25]
L'épaisseur	Cm	[0.45-1]

Tableau 3. Les paramètres fixes.

La surface de contacte	3.14 cm^2
Le mode de séchage	Séchage sous vide
La matière de support	Aluminium.



Figure 17. Emplacement des bouchons d'aluminium dans l'étuve sous vide pour sécher une couche mince de poudre.

Construction du plan d'expériences

Nous avons choisi de réaliser un plan factoriel mixé pour déterminer les nombres d'expériences. Nous avons choisi, la température(T), la pression(P), épaisseur (Ep), l'humidité initial(H) comme des facteurs et le temps final de séchage comme réponse. Le plan choisi donne par la méthode D-Optimale, 52 expériences plus 3 expériences pour la reproductibilité, (les niveaux de la température {40, 50,60,70 et 80 °C}, de pression {0.3, 0.6 et 1}, de teneur en eau initial {15 et 25%} l'épaisseur {0.45 – 1cm}) ce plan a été appliquée dans les diffèrent poudres et dans différente solvant de mouillage utilisées.

2.5. Expérimentation

Cette étude rassemble des résultats de l'étude expérimentale effectuée au sein du laboratoire de contrôle qualité, site de production de Médéa, groupe SAIDAL. Les essais de séchage des poudres pharmaceutiques (paracétamol, nystatine, PVP) ont été réalisés dans une étuve sous vide.

Nous présenterons tout d'abord les caractéristiques de la poudre, puis nous détaillons les procédures expérimentales concernant les mesures de la teneur en eau de poudre à sécher, la détermination de la quantité d'eau évaporée, et étudier la cinétique du séchage.

3. Matériel utilisé

3.1. Séchoir sous vide

De nombreux échantillons sont très sensibles et ont tendance à être endommagés par le processus de séchage. L'opération de séchage dans l'étuve sous vide diminue ce risque. Le but est de sécher en douceur des produits de haute qualité contenant de l'eau ou des solvants sans modifier leurs propriétés. Le séchage sous vide entraîne une diminution de la pression dans l'étuve de séchage, permettant l'évaporation de l'eau ou du solvant même à basses températures et favorisant par ailleurs le transport par capillarité.

3.1.1. Les principaux avantages du séchage sous vide

De par ses caractéristiques de fonctionnement, le séchage sous vide comporte des avantages non négligeables de points de vue de la productivité, de la qualité et de la consommation énergétique voyons de façon plus détaillée chacun de ces facteurs, (Allemands, 2006).

3.1.2. La productivité

Les temps de séchage des procédés sous vide sont plus courts, peuvent s'expliquer par trois principales raisons :

- La température d'ébullition de l'eau est abaissée sous vide, d'où l'intensification de l'évaporation en surface (la température d'ébullition de l'eau à une pression de 101.3KPa est de 100°C tandis qu'elle est de 50°C à une pression de 12.3 KPa.
- La conductivité hydrique effective du matériau (facilité de l'eau à circuler à l'intérieur de la poudre) s'accroit avec la réduction de la pression, d'où un transfert de masse interne (déplacement de l'eau à l'intérieur de la poudre) s'accroit sous vide.
- L'action de vide se traduit par une diminution de la résistance au transfert de masse au niveau de la couche limite (évaporation en surface). Habituellement le séchage sous vide en présence de vapeur surchauffée peut être jusqu'à 5 fois plus rapide que le séchage conventionnel. Le facture de rapidité est variable en fonction de l'essence et de la dimension des particules. Il est important de rappeler que les temps de séchage ne peuvent pas être toujours en fonction de la teneur en humidité initial de produit, da la teneur en humidité finale visée, des caractéristiques de la ressource séchée et surtout de la quantité finale désirée.

3.1.3. La qualité

Le séchage sous vide possède des caractéristiques favorables à l'obtention d'une bonne qualité des produits séchés, le vide provoquant un abaissement du point d'ébullition de l'eau, permet un séchage à des températures relativement basses, le vide amélioré également la

migration de l'eau dans la poudre favorisant ainsi un séchage à plus faible gradient de teneur en humidité, le développement de contraintes de séchage en est alors diminué.

3.1.4. L'énergie

Les temps de séchage plus courts et les températures d'opération plus basses minimisent les pertes de chaleurs des séchoirs sous vide. De plus aucun échange d'air avec l'extérieur n'est nécessaire car l'opération de séchage est réalisée en circuit fermé.

3.2. Les équipements

Durant notre stage à SAIDAL, on a employé les équipements résumés dans le tableau 4 qui donne la marque, la précision et la plage de marche de chaque équipement utilisé dans l'opération de séchage à l'étuve sous vide.

Tableau 4. Les caractéristiques des équipements utilisés dans l'expérience.

Equipment	Précision
Balance	m= ±0.001g
Etuve sous vide	T= ± 0.1°C P= ± 5 m bar

4. Produit utilisé

4.1. Définition de la poudre

4.1.1. Paracétamol

Le **paracétamol**, dont la dénomination commune internationale (DCI) est l'acétaminophène, **est l'antalgique le plus utilisé**. Il permet de soulager de nombreuses manifestations douloureuses (maux de tête, douleurs dentaires ou articulaires, états grippaux, règles douloureuses, etc.). Il est également **efficace contre la fièvre**.

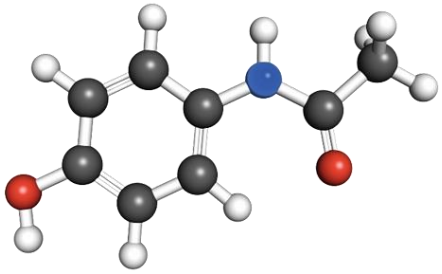


Figure 18. Formule 3D semi développée de paracétamol.

- **Propriétés physiques**

  Masse molaire : 151,163 g/mol

  Température de fusion : 170°C

  Densité : d21 = 1,293

  Solubilité dans Eau est 14 g/L à 20°C (peu soluble) et dans Éthanol est très soluble.

4.1.2. Nystatine

La nystatine est un antibiotique antifongique de contact de la famille des polyènes, extrait de culture de Streptomyces noursei. Active in vitro sur une large variété de champignons levuriformes et filamenteux, son action in vivo s'exerce essentiellement sur le genre candida et geotrichum. Son activité est essentiellement locale (de contact).

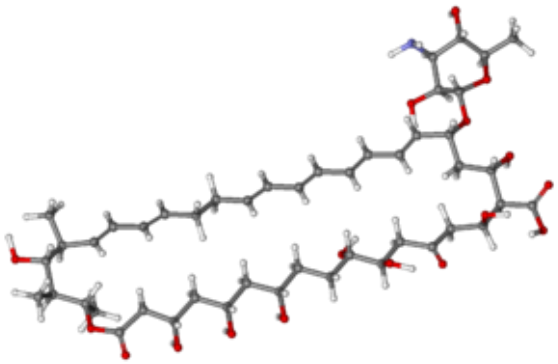


Figure 19. Formule 3D semi développée de Nystatine.

4.1.3. PVP

La PVP ou polyvinylpyrrolidone est un polymère hydrosoluble. Elle est très polyvalente en cosmétique et peut être utilisée en tant que liant, filmogène, stabilisateur d'émulsion, agent de suspension ou fixateur capillaire. Elle est principalement employée dans des produits tels que les mascaras, l'eye-liner, les produits capillaires ainsi que les shampooings. Elle est interdite en bio.

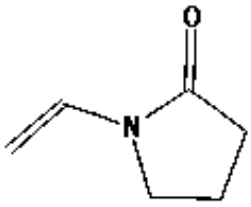


Figure 20. Formule 2D de PVP.

4.2. Mode opératoire

Les étapes à suivre pour augmenter la teneur en humidité de chaque poudre sont indiquées ci-dessous :

- Déterminer la teneur en humidité initial de chaque poudre

  Pour la teneur en humidité initial est égal 15% :

- Afin d'augmenter le pourcentage d'eau dans la poudre, nous utilisons les relations de perte de dessiccation suivante, supposons que la masse m_s = 9g et le pourcentage d'eau requis 15%.

$$X = \frac{m_i - m_s}{m_i} \times 100 \quad \text{(III. 1)}$$

Avec m_s =8.5g

$$0.15m_i - m_i = -8.5 \;\Rightarrow\; -0.85m_i = -8.5 \quad \text{(III. 2)}$$

$$m_i = \frac{-8.5}{-0.85} = 10\ g \quad \text{(III. 3)}$$

En ajoute de l'eau jusqu'à que la masse m_i = 10g.

Avec :

X (%) : la teneur en humidité recherchée (%),

X_0 (%) : la teneur en humidité initial (%).

m_e : la masse de l'eau en (g) $m_e = m_i - m_s$.

m_i : la masse totale de poudre (g), elle est égale à 10g.

- On mélange la masse d'eau calculé avec 9g de la poudre sèche dans un Bechar pour bien homogénéisée la poudre.
- On refaire les mêmes étapes pour la teneur en humidité égale 25 %.
- Le mélange est prêt pour faire les expériences.

4.3. Analyse statistique des résultats d'essais

Chacun des essais a été réalisé avec précision. Les valeurs relevées lors des mesures étant des valeurs particulières de variables aléatoires, En présence d'une valeur suspecte, l'essai est réalisé de nouveau et la valeur suspecte est écartée. L'analyse et les calcules mathématique consiste essentiellement à déterminer les coefficients des modèles à partir des résultats des *52* expériences réalisées. L'utilisation de la méthode des moindres carrés sur les résultats des tests a permis d'obtenir des modèles de calcul du temps final de séchage pour les trois poudres, à savoir le paracétamol, la nystatine et le PVP xL10, qui sont représentés par les modèles suivants.

- **Coefficients de modèles (paracétamol/eau)**

$$\text{Temps} = -266{,}718 - 1{,}59515T + 99{,}7122P + 27{,}0523H + 141{,}55Ep + 0{,}0317822 \times T \times T + 13{,}8463P \times P - 0{,}623555 \times H \times H - 46{,}7889Ep \times Ep - 2{,}24224 \times T \times P - 0{,}0415429 \times T \times H - 1{,}09246 \times T \times Ep + 1{,}14659 \times P \times H + 19{,}3525 \times P \times Ep + 0{,}405944 \times H \times Ep. \quad \text{(III.4)}$$

- **Coefficients de modèles (paracétamol/éthanol)**

$$Temps = -358{,}74 - 3{,}20121\text{T} + 97{,}8861\text{P} + 30{,}0072\text{H} + 479{,}189\text{Ep} + 0{,}0381406 \times \text{T} \times \text{T} - 31{,}6708 \times \text{P} \times \text{P} - 0{,}752348 \times \text{H} \times \text{H} - 300{,}908 \times \text{Ep} \times \text{Ep} - 1{,}36708 \times \text{T} \times \text{P} - 0{,}0227228 \times \text{T} \times \text{H} - 0{,}929495 \times \text{T} \times \text{Ep} + 1{,}61039 \times \text{P} \times \text{H} + 27{,}1393 \times \text{P} \times \text{Ep} + 1{,}6118 \times \text{H} \times \text{Ep}. \quad \text{(III.5)}$$

- **Coefficients de modèles (nystatine/eau)**

$$\text{Temps} = -111{,}807 - 7{,}1172\text{T} + 115{,}773\text{P} + 18{,}3465\text{H} + 448{,}931\text{Ep} + 0{,}0769974 \times \text{T} \times \text{T} + 89{,}1327\text{P} \times \text{P} - 0{,}388481 \times \text{H} \times \text{H} - 233{,}161\text{Ep} \times \text{Ep} - 3{,}82723 \times \text{T} \times \text{P} - 0{,}018673 \times \text{T} \times \text{H} - 1{,}1502 \times \text{T} \times \text{Ep} - 0{,}00555841 \times \text{P} \times \text{H} + 61.9945 \times \text{P} \times \text{Ep} - 1{,}70516 \times \text{H} \times \text{Ep}. \quad \text{(III.6)}$$

- **Coefficients de modèles (nystatine/éthanol)**

$$Temps = -61{,}2056 - 9{,}26914T + 173{,}696P + 19{,}882H + 411{,}61Ep + 0{,}0946056 \times T \times T + 66{,}0548P \times P - 0{,}474286 \times H \times H - 217{,}691Ep \times Ep - 4{,}43542 \times T \times P - 0{,}002816 \times T \times H - 1{,}25392 \times T \times Ep - 0{,}876905 \times P \times H + 53{,}0048 \times P \times Ep - 1{,}13139 \times H \times Ep. \quad \text{(III.7)}$$

- **Coefficients de modèles (PVP/eau)**

$$\boldsymbol{Temps} = -101{,}561 + 1{,}32038\text{T} + 51{,}3086\text{P} + 14{,}3505\text{H} - 115{,}964\text{Ep} - 0{,}0151699 \times \text{T} \times \text{T} - 8{,}00374 \times \text{P} \times \text{P} - 0{,}376276 \times \text{H} \times \text{H} + 45{,}3886 \times \text{Ep} \times \text{Ep} - 0{,}554817 \times \text{T} \times \text{P} - 0{,}00212367 \times \text{T} \times \text{H} + 0{,}353757 \times \text{T} \times \text{Ep} - 0{,}935575 \times \text{P} \times \text{H} + 39{,}5833 \times \text{P} \times \text{Ep} + 1{,}50216 \times \text{H} \times \text{Ep}. \quad \text{(III.8)}$$

- **Coefficients de modèles (PVP/éthanol)**

$$Temps = 604{,}585 - 0{,}956775T + 55{,}7056P - 4{,}42527H - 1622{,}05Ep + 0{,}00428493 \times T \times T - 41{,}5967P \times P + 0{,}137554 \times H \times H + 1120{,}17Ep \times Ep + 0{,}0265264 \times T \times P - 0{,}0109577 \times T \times H - 0{,}0758486 \times T \times Ep - 0{,}73999 \times P \times H + 39{,}0266 \times P \times Ep - 0{,}0865417 \times H \times Ep. \quad \text{(III.9)}$$

4.4. Analyse statistique des modèles

L'analyse statistique a été réalisée en trois parties. Tout d'abord, les modèles ont été analysés dans leur ensemble pour vérifier si l'hypothèse de sélection de modèles de forme quadratique était bien validée. Ensuite, une analyse statistique des composantes des modèles a été effectuée afin d'identifier le meilleur facteur pour chacun des modèles. Enfin, la normalité de la distribution des résidus a été vérifiée pour tous les modèles retenus.

4.4.1. Analyse statistique globale des modèles

Un test ANOVA (Analyse de la variance) peut être réalisé pour évaluer la différence entre les valeurs expérimentales et les valeurs calculées par le modèle. Ce test permet de déterminer si les différences observées sont statistiquement significatives. Le test ANOVA implique de comparer la variance entre les valeurs expérimentales et les valeurs calculées par le modèle avec la variance à l'intérieur de chaque groupe. Si la variance entre les groupes est significativement plus grande que la variance à l'intérieur des groupes, cela suggère qu'il y a une différence statistiquement significative entre les valeurs expérimentales et les valeurs calculées par le modèle. Il est important de noter que le test ANOVA suppose certaines

conditions et hypothèses, telles que la normalité des résidus et l'homogénéité des variances. Si ces hypothèses ne sont pas satisfaites, des tests alternatifs peuvent être utilisés. Le test ANOVA permet d'évaluer statistiquement la différence entre les valeurs expérimentales et les valeurs calculées par le modèle. Il fournit une indication de la qualité de l'ajustement du modèle aux données expérimentales. Cette analyse donne les résultats présentés dans les tableaux (5-10).

Nous obtenons pour la réponse une valeur nul de *P* ce qui conforte l'hypothèse de l'utilisation des modèles quadratiques pour modéliser la réponse.

Tableau 5. Résultat statistique de modèle (paracétamol/eau).

Temps	**DF**	**SS**	**MS**	**F**	**p**	**SD**
			(variance)			
Total	55	65222	1185,85			
Constant	1	37388,3	37388,3			
Total corrigé	54	27833,7	515,439			22,7033
Régression	14	22528,2	1609,16	12,132	**0,000**	40,1143
Résiduel	40	5305,48	132,637			11,5168
Manque de concordance	38	5299,48	139,46	46,4867	**0,021**	11,8093
Erreur pure	2	6	3			1,73205

N = 55 Q2 =0,513 Cond. no. = Infinite
DF = 40 R2 = 0,793 Y-miss = 0
Comp. = 2 R2 Adj. = 0,743 RSD = 11, 5168

Tableau 6. Résultat statistique de modèle (paracétamol/éthanol).

Temps	**DF**	**SS**	**MS**	**F**	**p**	**SD**
			(variance)			
Total	55	62729	1140,53			
Constant	1	37128	37128			
Total corrigé	54	25601	474,092			21,7737
Régression	14	20293	1449,5	10,9231	**0,000**	38,0723
Résiduel	40	5308,01	132,7			11,5196
manque de concordance	38	5270,01	138,684	7,29918	**0,128**	11,7764
Erreur pure	2	38	19			4,3589

N = 55 Q2 = 0,487 Cond. no. = Infinite
DF = 40 R2 = 0,793 Y-miss = 0
Comp. = 2 R2 Adj. = 0,720 RSD = 11, 5196

Tableau 7. Résultat statistique de modèle (nystatine/eau).

Temps	**DF**	**SS**	**MS**	**F**	**p**	**SD**
			(variance)			
Total	55	208574	3792,25			
Constant	1	113092	113092			
Total corrigé	54	95482,4	1768,19			42,0499
Régression	14	74504,6	5321,76	10,1474	**0,000**	72,9504
Résiduel	40	20977,8	524,446			22,9008
manque de concordance	38	20945,8	551,206	34,4504	**0,029**	23,4778
Erreur pure	2	32	16			4

N = 55 Q2 = 0,458 Cond. no. = Infinite
DF = 40 R2 = 0,780 Y-miss = 0
Comp. = 2 R2 Adj. = 0,703 RSD = 22.9008

Tableau 8 .Résultat statistique de modèle (nystatine/éthanol).

Temps	**DF**	**SS**	**MS**	**F**	**p**	**SD**
			(variance)			
Total	55	209762	3813,85			
Constant	1	98537,9	98537,9			
Total corrigé	54	111224	2059,71			45,384
Régression	14	86981,2	6212,94	10,2512	**0,000**	78,8222
Résiduel	40	24242,9	606,073			24,6185
manque de concordance	37	23990,4	648,389	7,70364	**0,058**	25,4635
Erreur pure	3	252,5	84,1667			9,17424

N = 55 Q2 = 0,494 Cond. no. = Infinite
DF = 40 R2 =0,782 Y-miss = 0
Comp. = 2 R2 Adj. =0.706 RSD = 24.6185

Tableau 9. Résultat statistique de modèle (PVP/eau).

Temps	**DF**	**SS**	**MS**	**F**	**P**	**SD**
			(variance)			
Total	55	95451	1735,47			
Constant	1	80258,2	80258,2			
Total corrigé	54	15192,8	281,348			16,7734
Régression	14	12536,2	895,444	13,4827	**0,000**	29,924
Résiduel	40	2656,57	66,4144			8,1495
manque de concordance	37	2638,07	71,2993	11,5621	**0,033**	8,44389
Erreur pure	3	18,5	6,16667			2,48328

N = 55 Q2 = 0,609 Cond. no. = Infinite
DF = 40 R2 = 0.825 Y-miss = 0
Comp. = 2 R2 Adj. = 0,764 RSD = 8.1495

Tableau 10. Résultat statistique de modèle (PVP/éthanol).

Temps	**DF**	**SS**	**MS**	**F**	**p**	**SD**
			(variance)			
Total	55	97488	1772,51			
Constant	1	79268,1	79268,1			
Total corrigé	54	18219,9	337,406			18,3686
Régression	14	10424,4	744,598	3,82063	**0,000**	27,2873
Résiduel	40	7795,55	194,889			13,9603
manque de concordance	37	7683,05	207,65	5,53733	**0,091**	14,4101
Erreur pure	3	112,5	37,5			6,12372

N = 55 Q2 = 0,317 Cond. no. = Infinite
DF = 40 R2 = 0,572 Y-miss = 0
Comp. = 2 R2 Adj. = 0,422 RSD = 13.9603

Avec **DF** dégrée de liberté**. SS** (Sum Square) La somme des carrés des écarts SCE.

MS (Mean Square) carré des moyennes ou CM. **F** teste de fisher. **P-Value**, **SD** écart type.

Les qualités descriptives (R^2ajusté) et prédictives (Q^2) ont été calculées pour ces modèles.

4.4.2. Analyse graphique du modèle

4.4.2.1. Courbe de criblage

Les figures suivantes illustrent les courbes de corrélation entre les valeurs expérimentales et les valeurs calculées par les modèles. La valeur élevée des coefficients de détermination R^2 associée aux modèles des Temps ses traduit par des graphiques d'adéquation, fournis en figures (21-26), sur lesquels l'alignement des nuages de points est très proche de la première bissectrice**.** Les parcelles avec les points proches d'une droite indiquent de bons modèles.

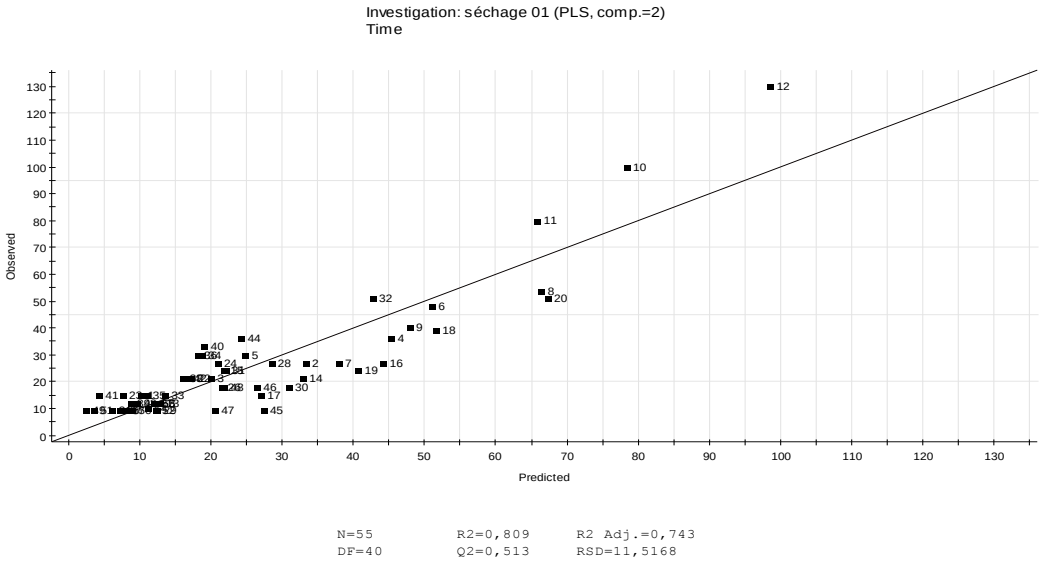


Figure 21. La corrélation entre les valeurs prédites et de l'expérience (paracétamol/eau).

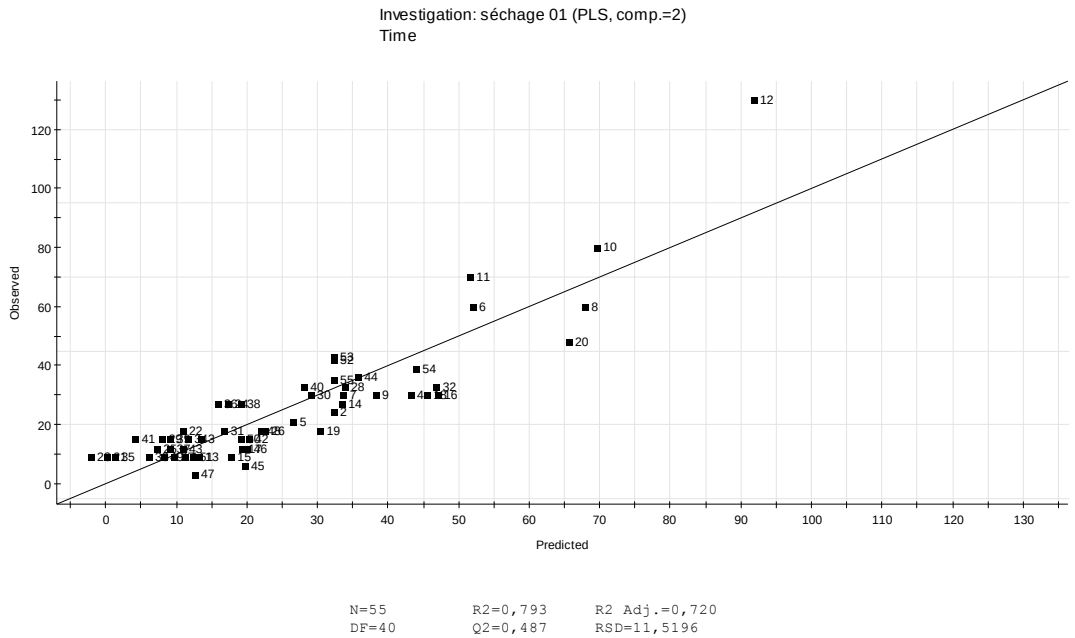


Figure 22. La corrélation entre les valeurs prédites et de l'expérience (paracétamol/éthanol).

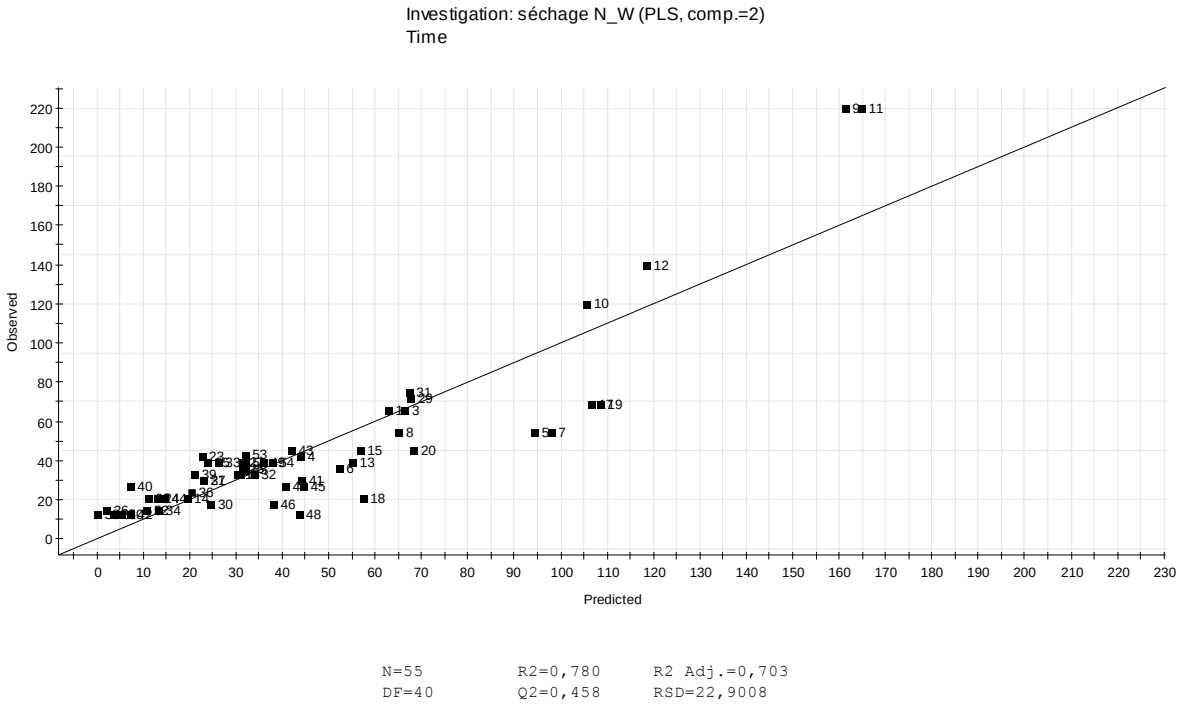


Figure 23. La corrélation entre les valeurs prédites et de l'expérience (nystatine/eau).

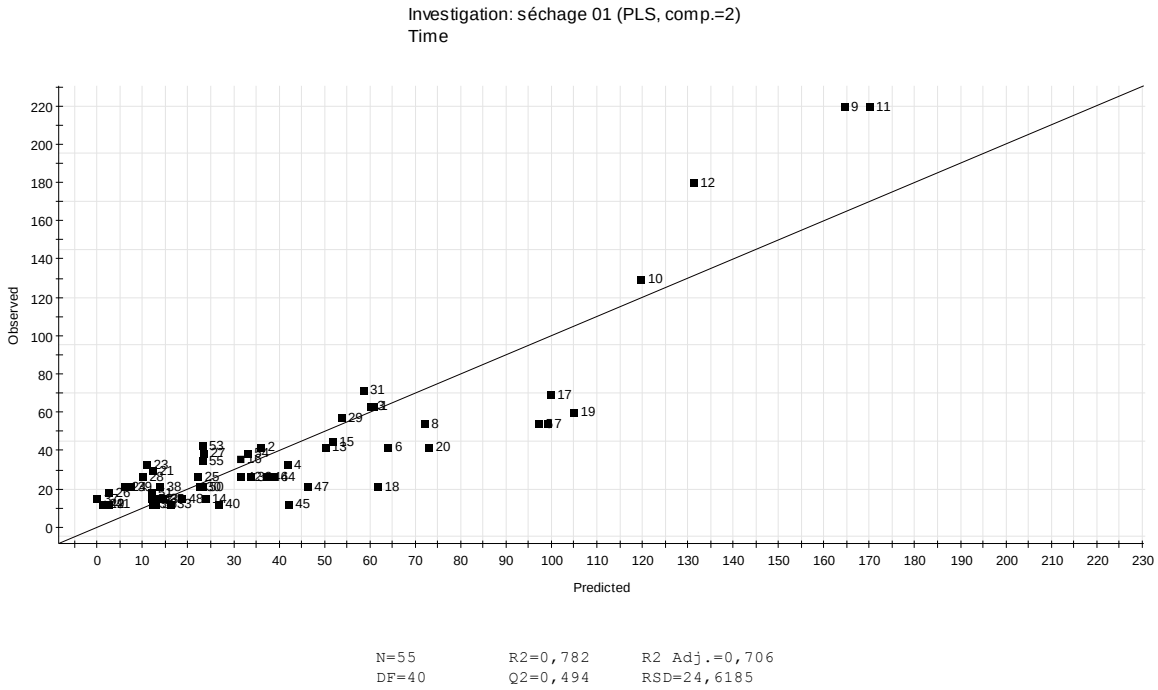


Figure 24. La corrélation entre les valeurs prédites et de l'expérience (nystatine/éthanol).

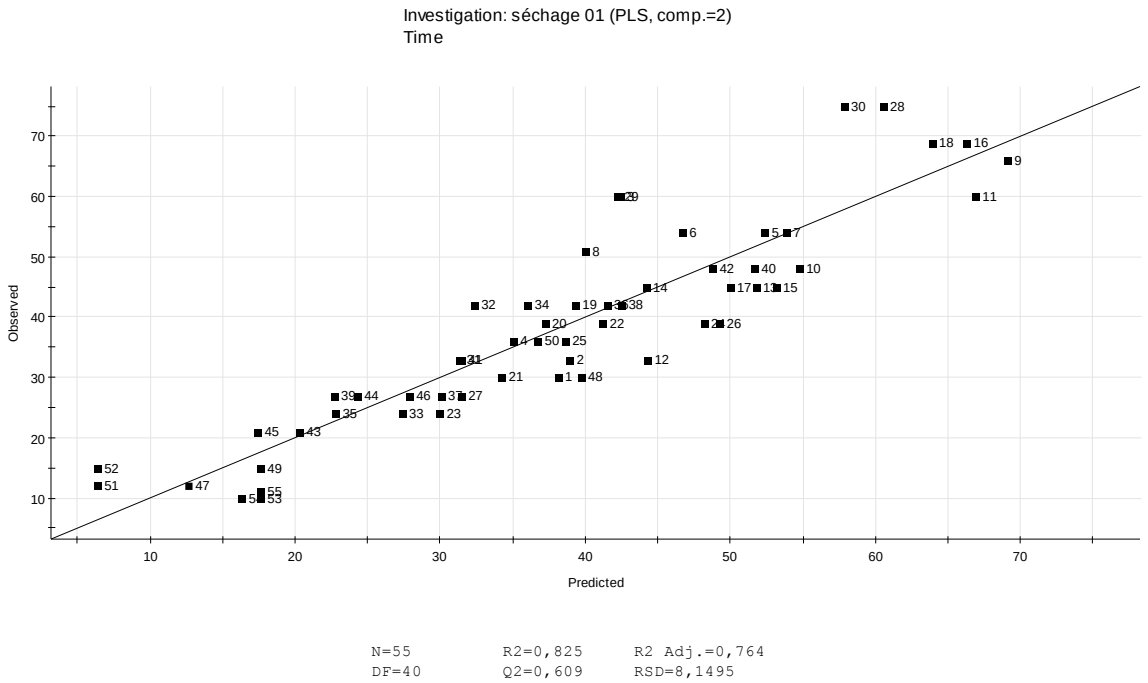


Figure 25. La corrélation entre les valeurs prédites et de l'expérience (PVP/eau).

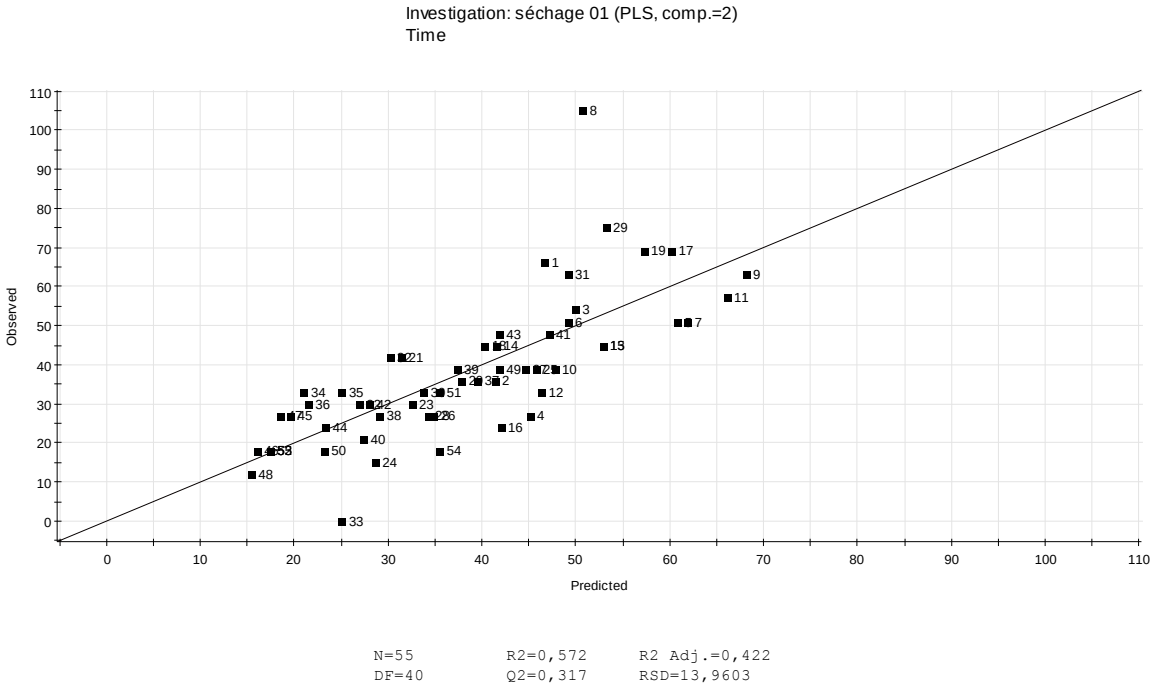


Figure 26. La corrélation entre les valeurs prédites et de l'expérience (PVP/éthanol).

4.4.2.2. Résumé de l'ajustement (Les statistiques sommaires)

Les valeurs significatives des coefficients d'ajustement de R2 et Q2, montrent que les deux modèles obtenus sont représentatifs de point de vue ajustement et prédiction. L'histogramme suivant illustre les valeurs de ces coefficients donnés par les figures (27-32).

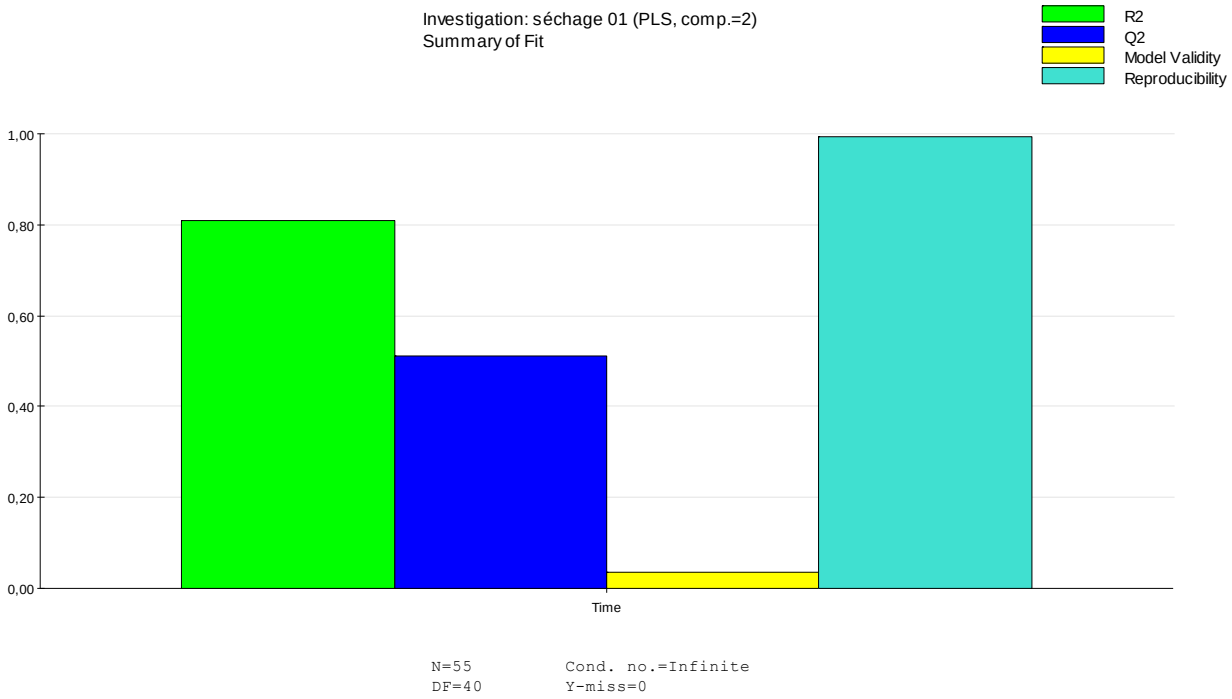


Figure 27. Résumé de l'ajustement de MODDE (paracétamol/eau).

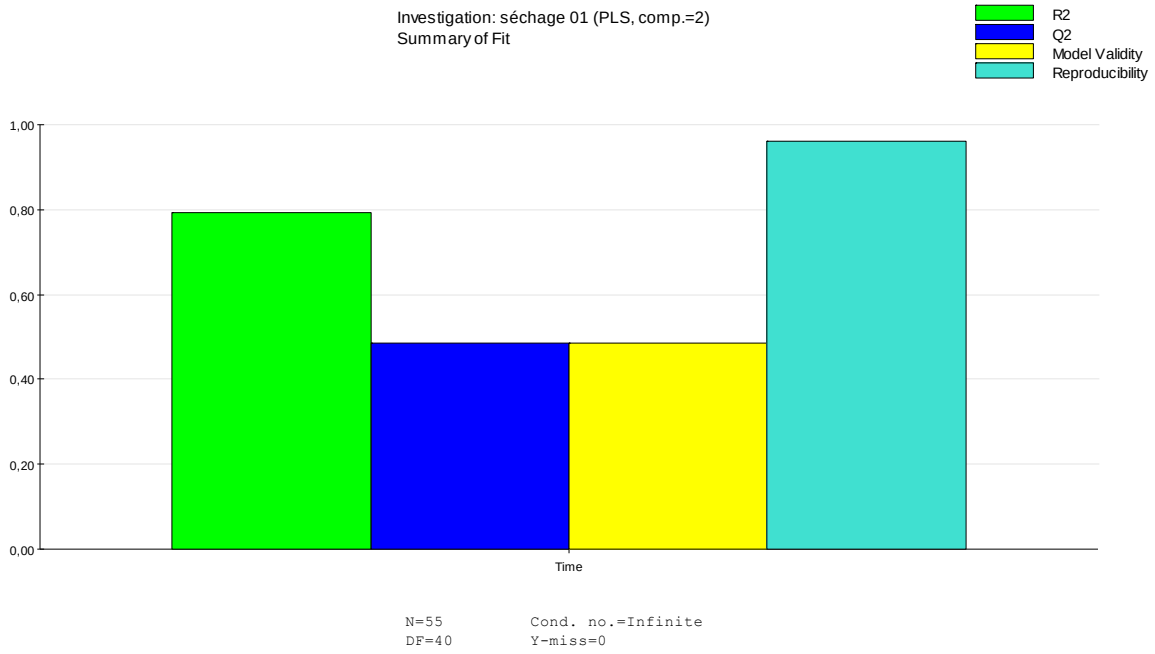


Figure 28. Résumé de l'ajustement de MODDE (paracétamol/éthanol).

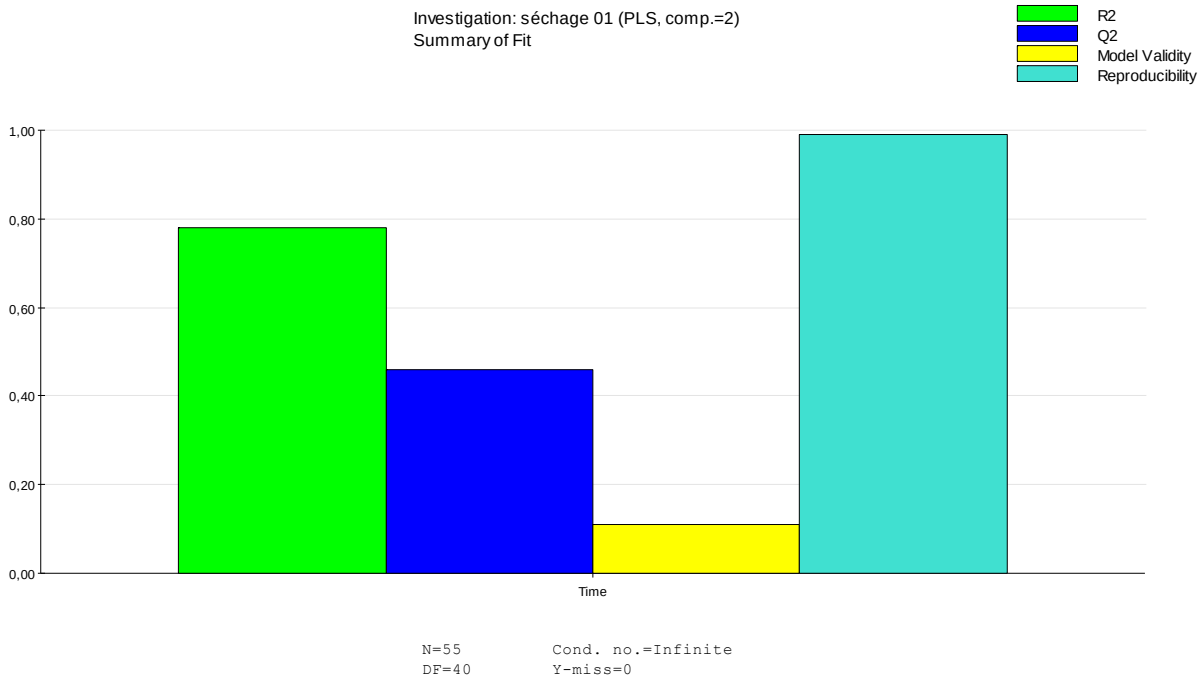


Figure 29. Résumé de l'ajustement de MODDE (nystatine/eau).

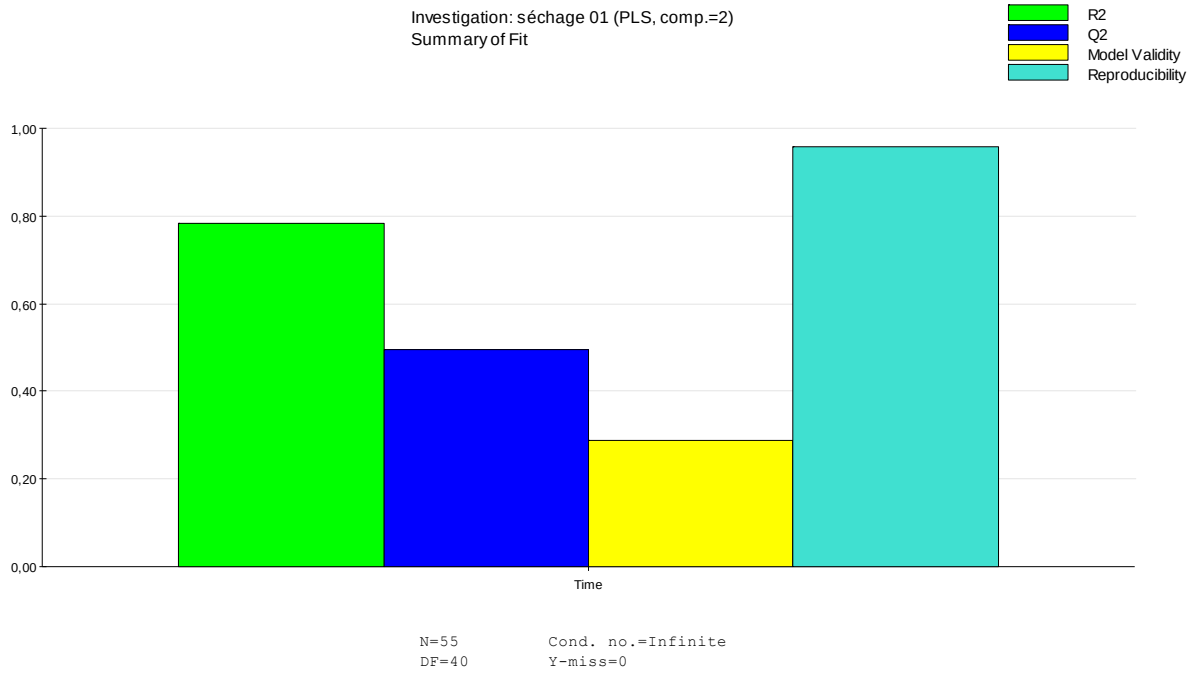


Figure 30. Résumé de l'ajustement de MODDE (nystatine/éthanol).

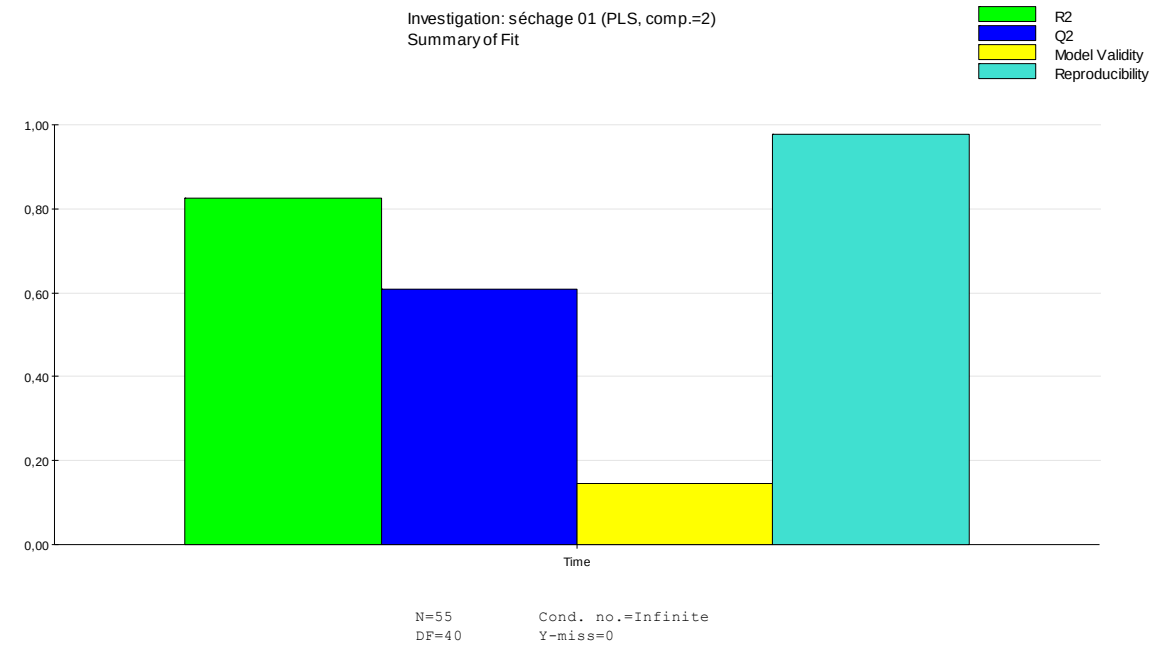


Figure 31. Résumé de l'ajustement de MODDE (PVP/eau).

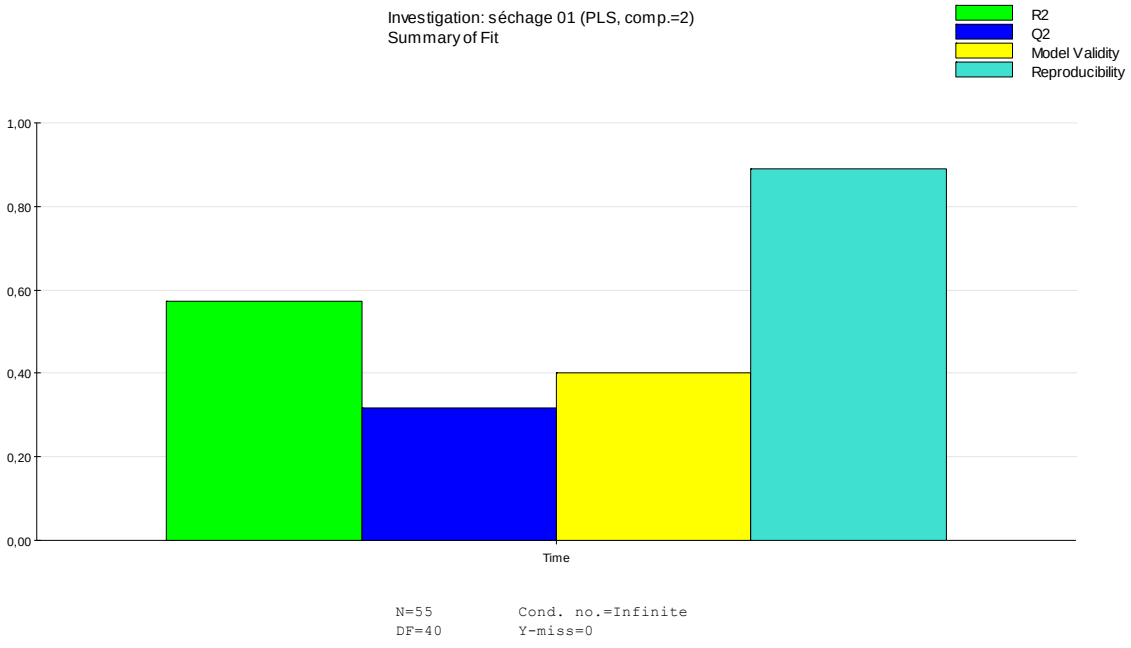


Figure 32. Résumé de l'ajustement de MODDE (PVP/éthanol).

4.4.2.3. Représentation graphique de l'effet des facteurs

L'étude a permis d'analyser l'effet des facteurs tels que la température, l'épaisseur, le taux d'humidité et la pression sur les temps finaux de séchage. Les résultats ont démontré que ces facteurs ont une influence significative sur les temps de séchage des échantillons. La température et la pression ont été identifiées comme les principaux facteurs ayant un impact majeur sur les temps de séchage. Une augmentation de la température a conduit à des temps de séchage plus courts, tandis qu'une diminution de l'épaisseur et de la pression et du taux

d'humidité a également réduit les temps de séchage. Ces conclusions sont importantes pour la compréhension et l'optimisation du processus de séchage des poudres pharmaceutiques, ce qui peut conduire à une amélioration de l'efficacité et de la qualité du séchage dans l'industrie pharmaceutique. Les figures (33-38) expriment les effets de interactions des facteurs et facteurs sur le temps de séchage.

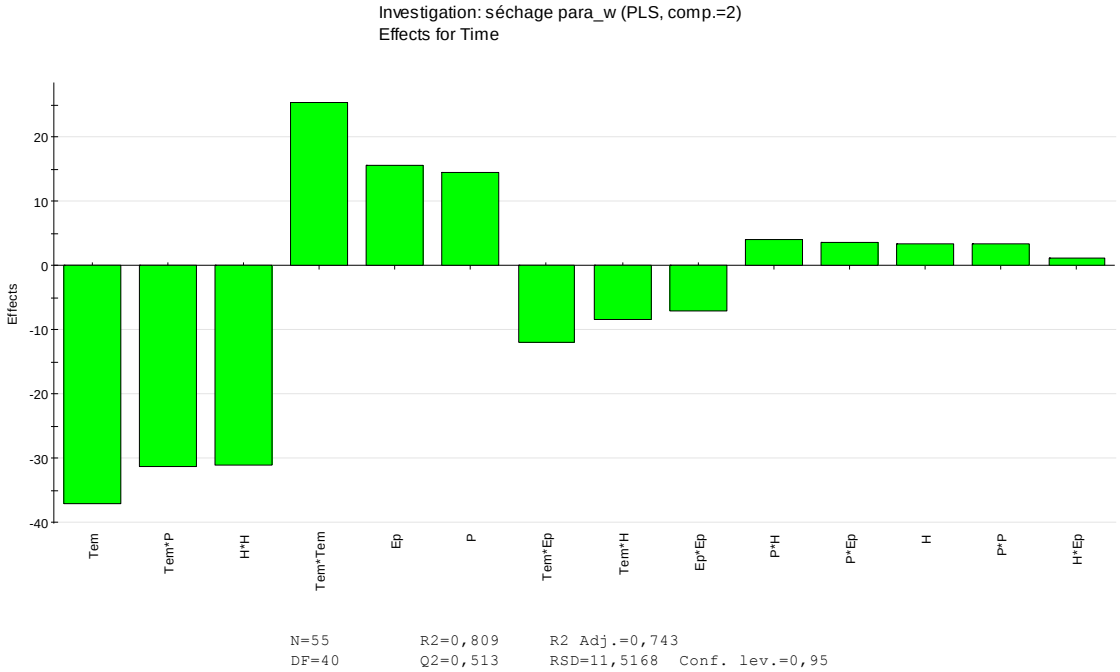


Figure 33. Effets des facteurs sur les réponses (paracétamol/eau).

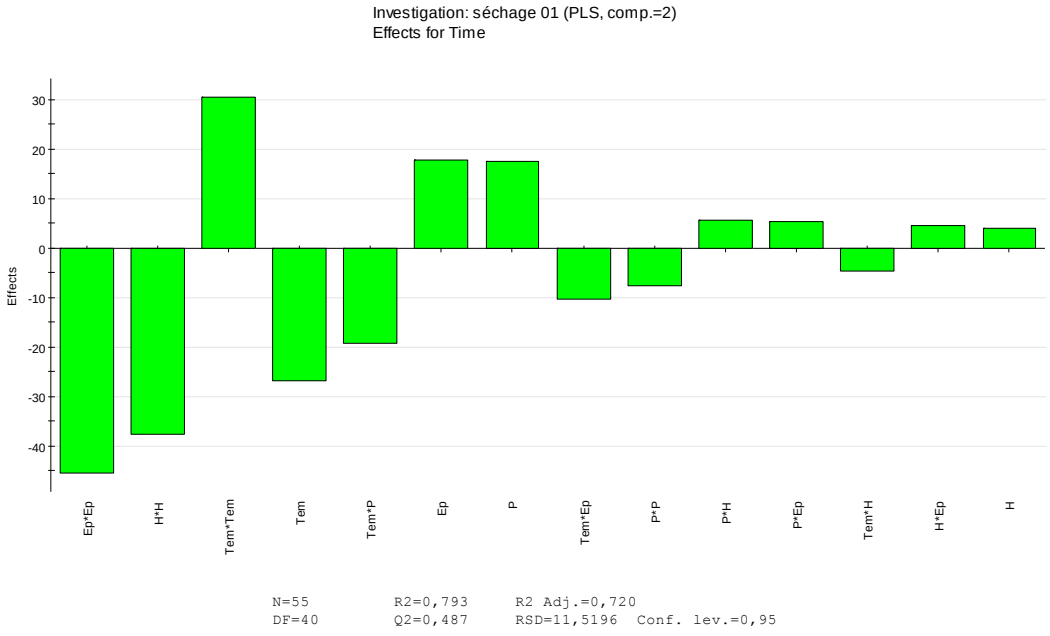


Figure 34. Effets des facteurs sur les réponses (paracétamol/éthanol).

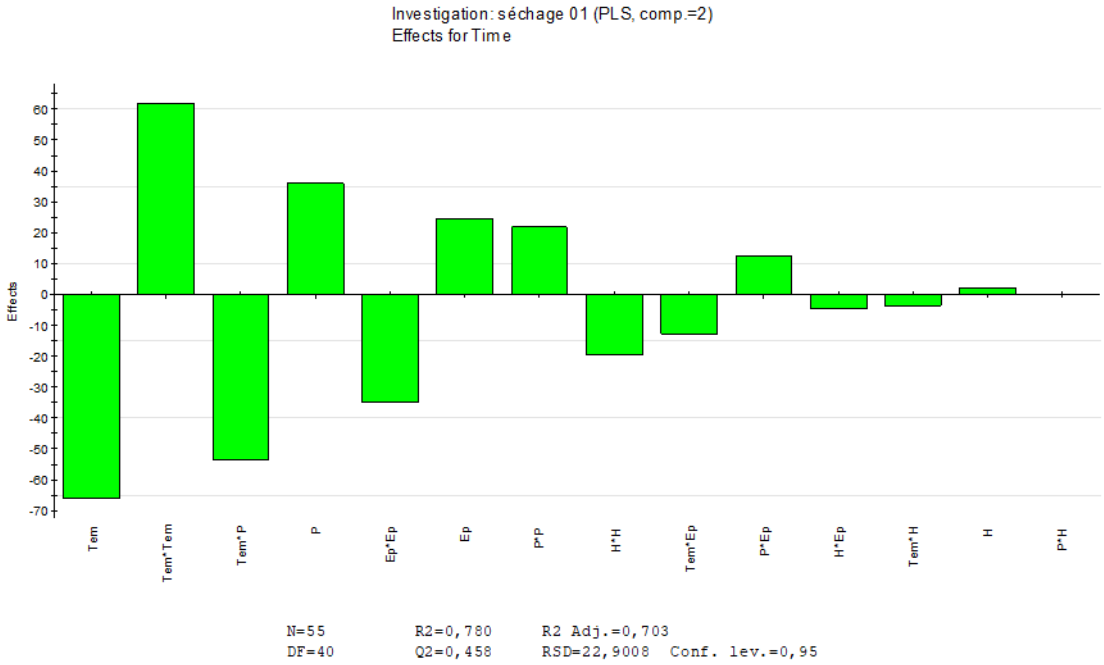


Figure 35. Effets des facteurs sur les réponses (nystatine/eau).

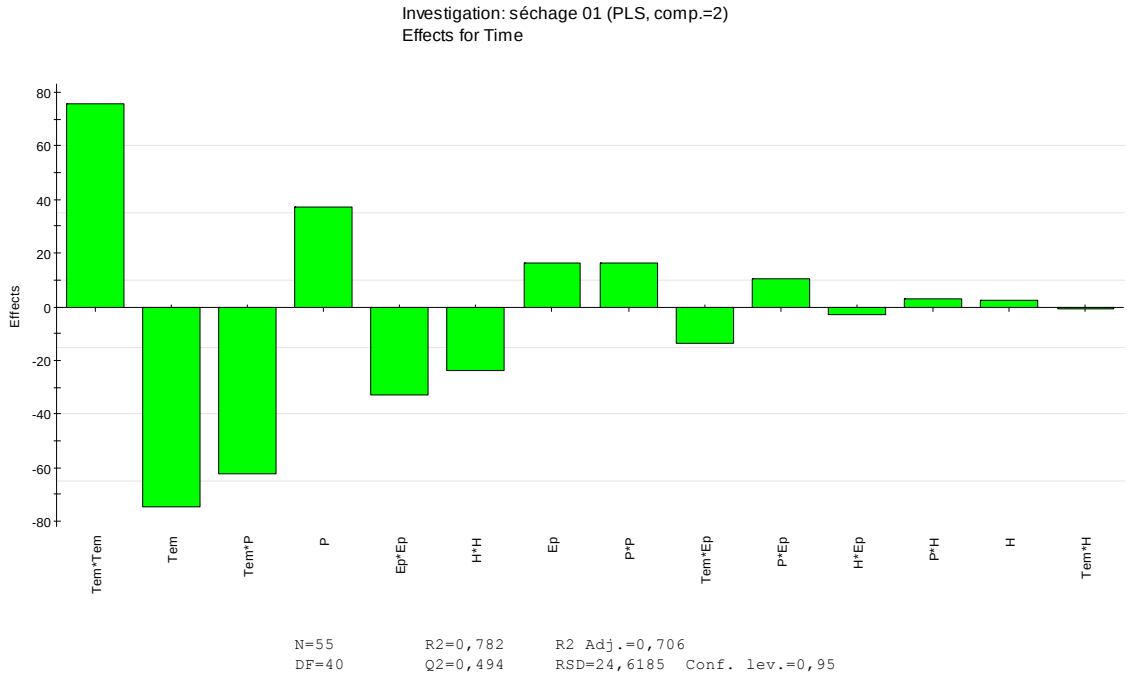


Figure 36. Effets des facteurs sur les réponses (nystatine/éthanol).

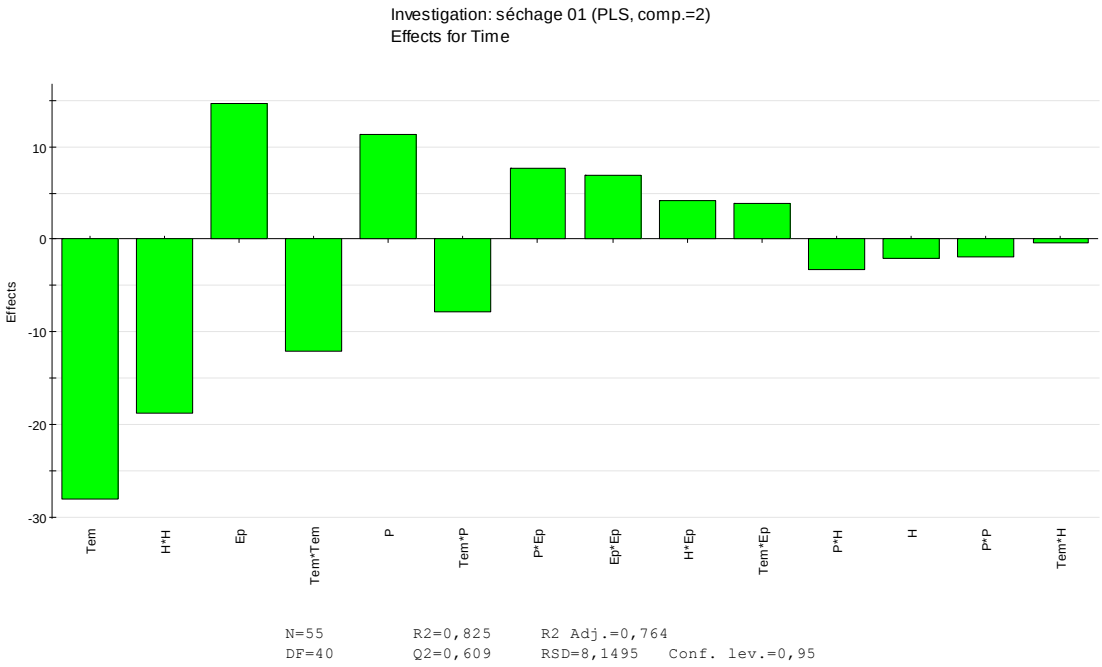


Figure 37. Effets des facteurs sur les réponses (PVP/eau).

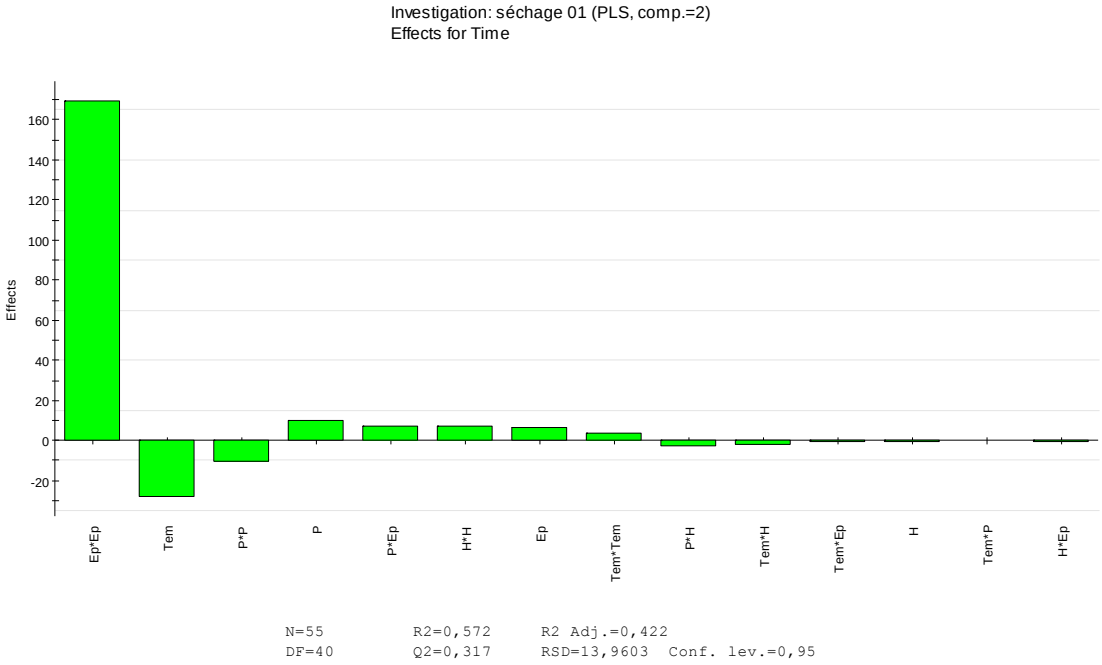


Figure 38. Effets des facteurs sur les réponses (PVP/éthanol).

D'après les figures (33-38), on tirer ces deux conclusions :

- Le plus longtemps est le temps de séchage, le moins sera l'effet de la température T et les interactions entre P×T, m_0×T et P².
- Le plus longtemps est le temps de séchage, le plus sera l'effet de la pression P, la masse initiale m_0, la teneur initiale X_0 et les interactions entre T×T, m_0×P, m_0× m_0, m_0×X_0 et P×T².

4.4.2.4. Courbes iso-réponse

Les figures (39-44) montrent les variations des facteurs sur la réponse et monte les minimas locaux de chaque répartition.

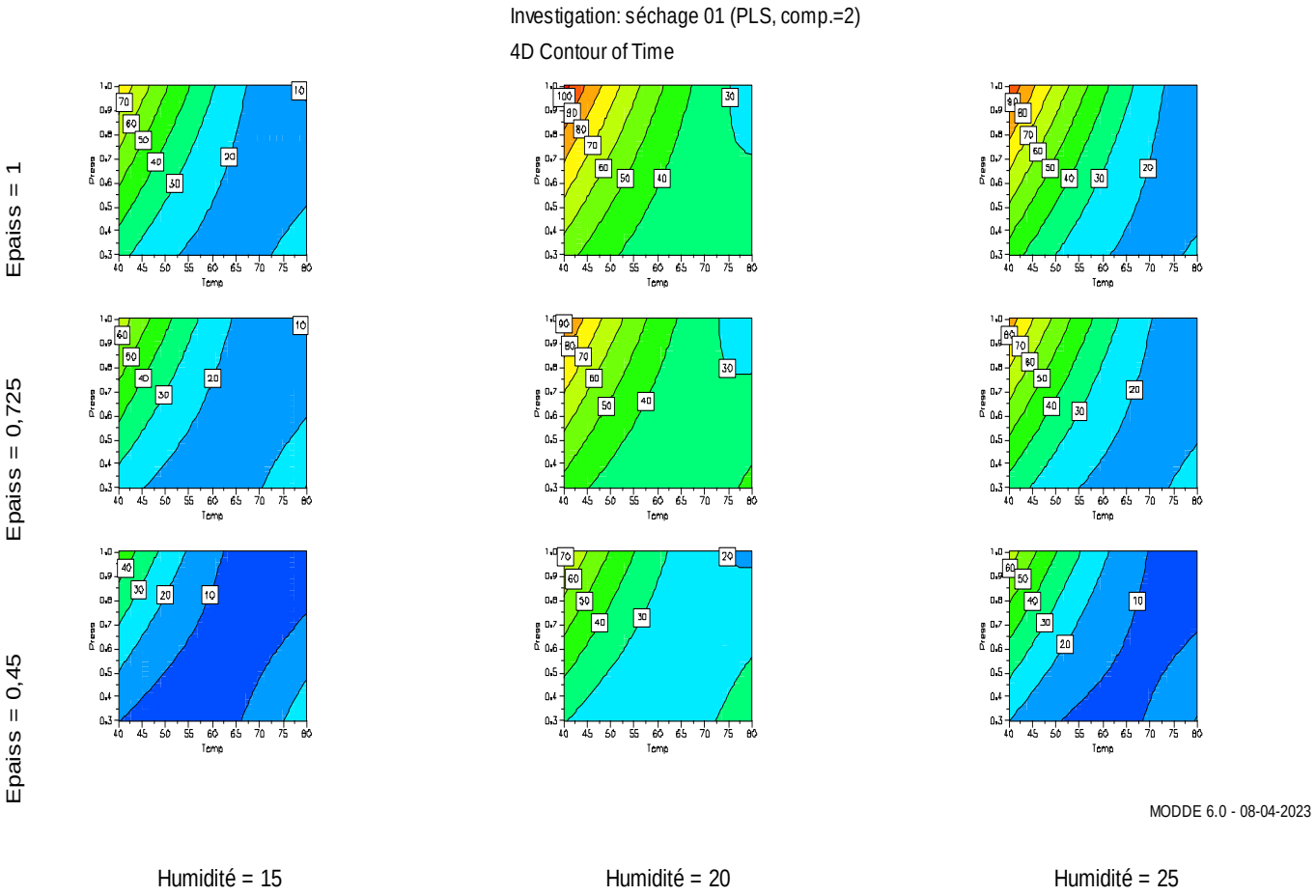


Figure 39. Courbes iso-réponse (paracétamol/eau).

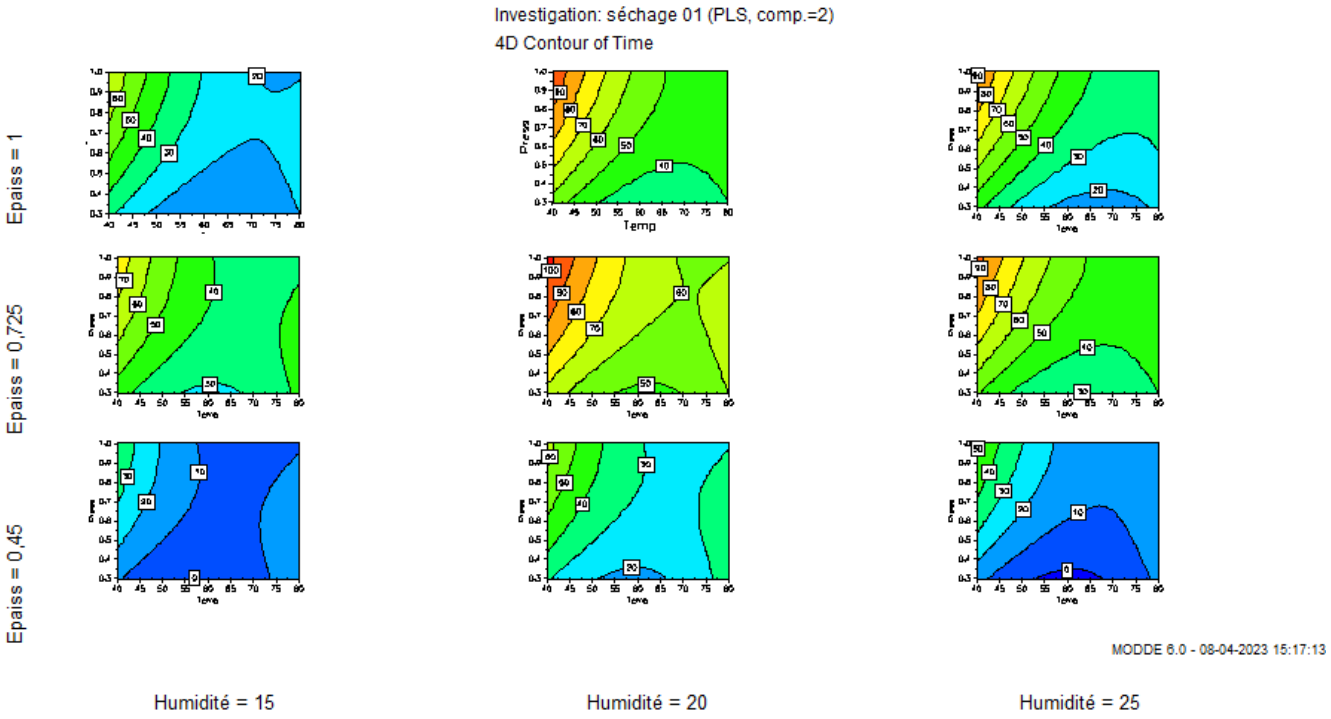


Figure 40. Courbes iso-réponse(paracétamol/éthanol).

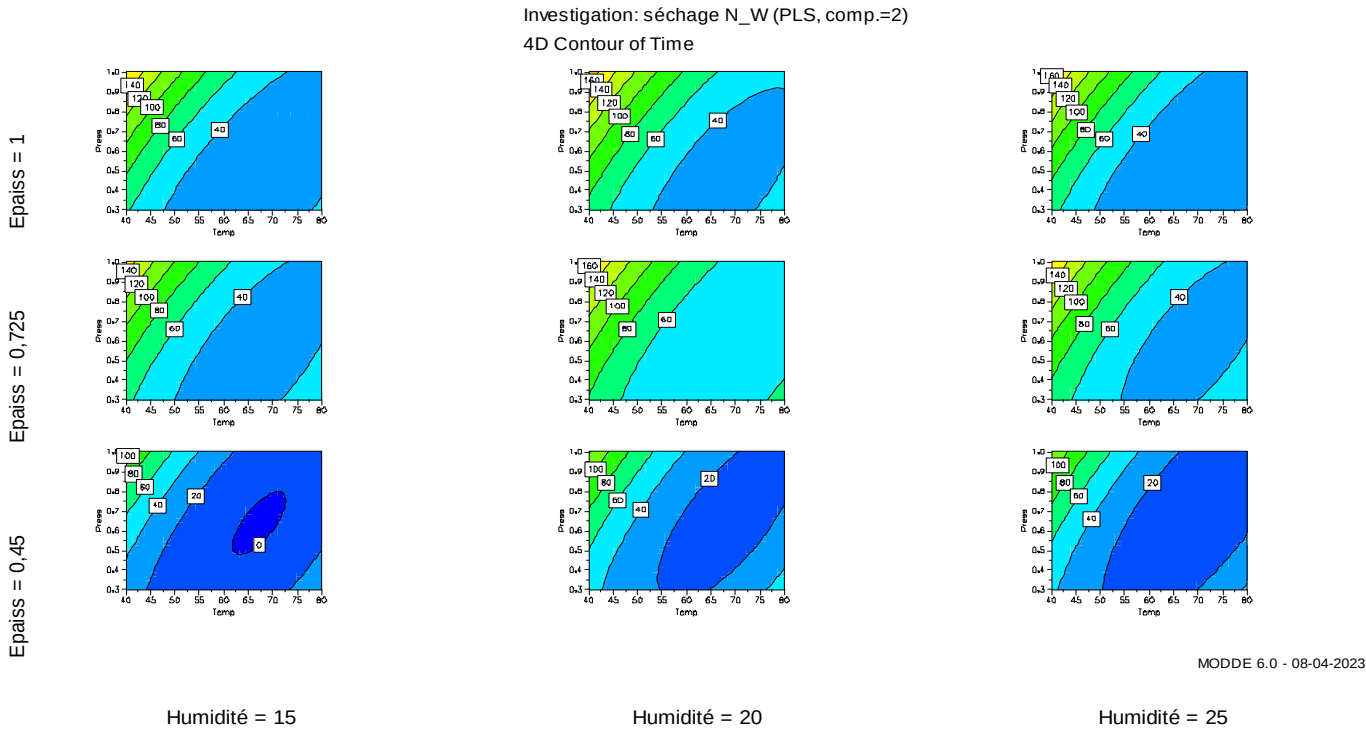


Figure 41. Courbes iso-réponse(nystatine/eau).

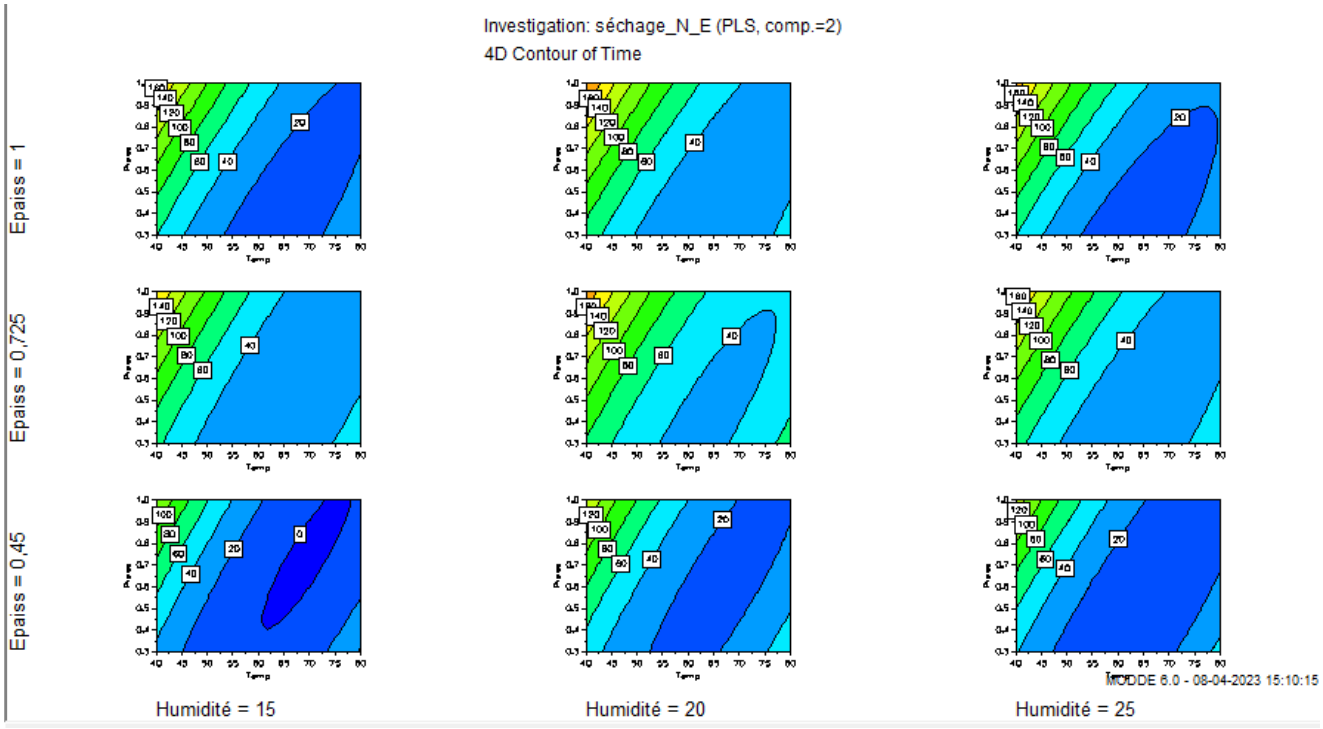


Figure 42. Courbes iso-réponse (nystatine/éthanol).

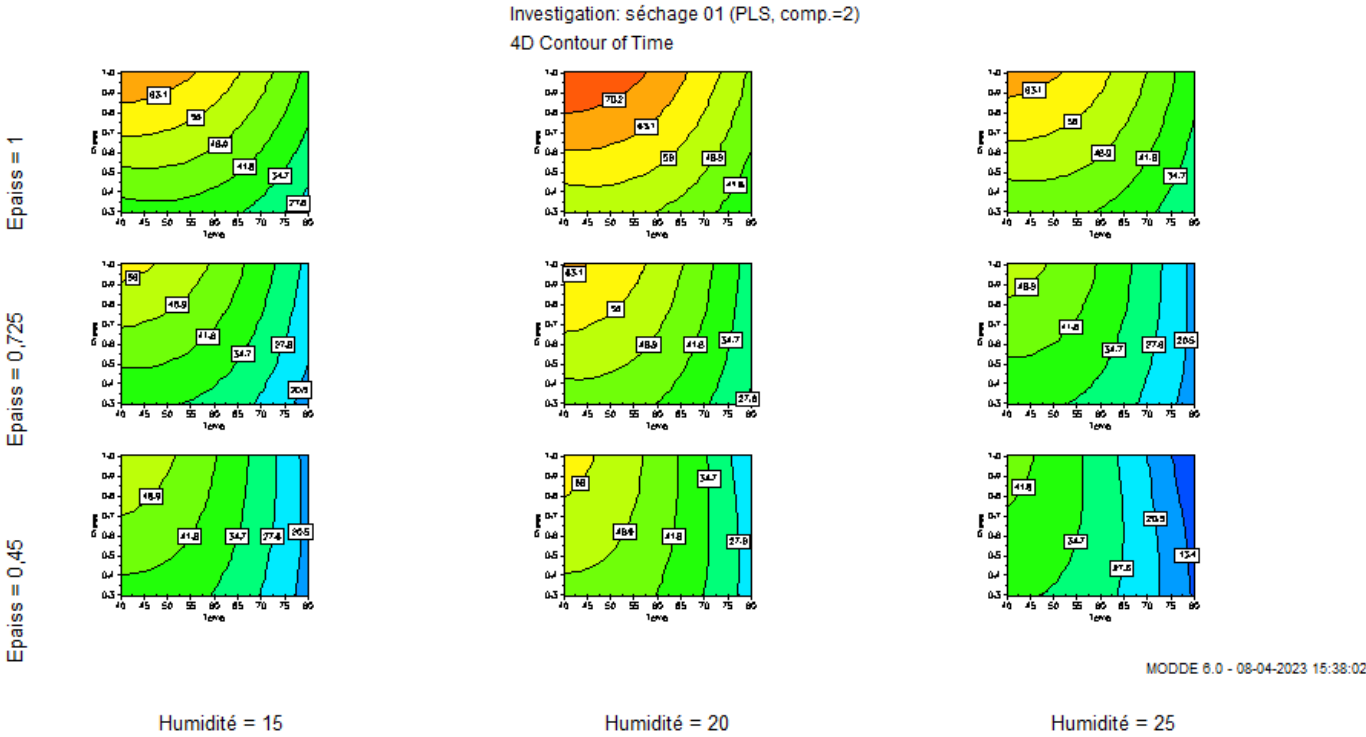


Figure 43. Courbes iso-réponse (PVP/eau).

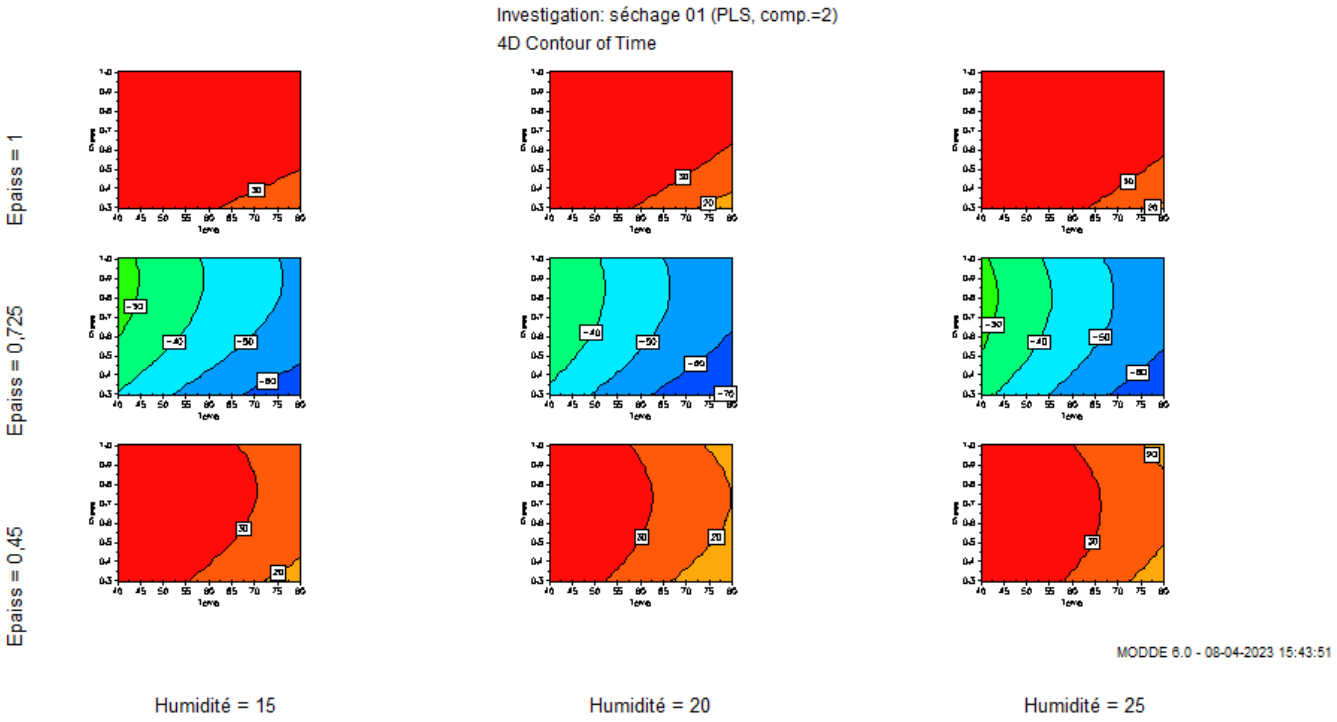


Figure 44. Courbes iso-réponse (PVP/éthanol).

Ces iso-réponses permettent de visualiser l'évolution des poudres en fonction des quatre facteurs à la fois. On peut ainsi tirer directement à partir de ces graphiques, la valeur estimée du paramètre désiré, à n'importe quel point du domaine d'étude.

Conclusion

En conclusion, cette contribution vise à étudier expérimentalement l'effet de certains paramètres opératoires sur le processus de séchage sous vide de trois poudres pharmaceutiques, à savoir le paracétamol, la nystatine et le PVP (polyplasdone). Les expériences et la modélisation seront réalisées au laboratoire de contrôle qualité SAIDAL de Médéa.

L'objectif est d'analyser l'effet des paramètres opératoires tels que la température, l'épaisseur, la pression, la masse et le taux d'humidité sur le temps de séchage. La modélisation du temps de séchage sera basée sur l'utilisation de la méthode des plans d'expériences avec un minimum d'expérimentations, grâce à l'utilisation du logiciel MODDE®.

En étudiant ces paramètres et en établissant un modèle empirique quadratique, nous pourrons mieux comprendre l'influence de chaque paramètre sur le processus de séchage. Les résultats obtenus contribueront à améliorer les connaissances dans le domaine du séchage des poudres pharmaceutiques et pourront être utilisés pour optimiser les conditions opératoires afin d'obtenir des temps de séchage plus efficaces et reproductibles.

Contribution 2 : Modélisation par l'intelligences artificielle

1. Modélisation des données expérimentale par l'intelligences artificielle

L'utilisation de l'intelligence artificielle pour modéliser nos données expérimentales est une approche novatrice qui permet d'analyser et d'interpréter les résultats de nos expérimentations. Cette méthode exploite les capacités d'apprentissage automatique des systèmes d'intelligence artificielle afin de tirer des enseignements précieux à partir de nos données expérimentales.

2. Préparation de la base de données

Dans cette section, nous avons exploité les données recueillies lors de la première étape de notre recherche afin de développer des modèles mathématiques basés sur l'intelligence artificielle. Nous avons identifié cinq facteurs clés, à savoir la température (T), la pression (P), l'humidité relative (H%), et le temps (t), ainsi que trois types de poudres pharmaceutiques (Paracétamol, Nystatine et PVP) et deux solvants (eau, eau-éthanol) utilisés pour mouiller les poudres. Ces facteurs ont été pris en compte dans la construction de modèle de l'intelligence artificielle.

Pour distinguer les différentes poudres séchées, nous avons utilisé l'empreinte de molécule. Une empreinte moléculaire est une représentation numérique ou binaire d'une molécule qui capture certaines de ses caractéristiques structurelles, fonctionnelles ou physico-chimiques. Elle permet de résumer les informations contenues dans la structure moléculaire sous une forme utilisable pour diverses applications, la même chose pour différencier entre les deux solvants de mouillage.

2.1. Calcul de l'empreinte digitale (fingerprint)

L'empreinte Morgan, également connue sous le nom de calcul ECFP-4 (Extended Connectivity Fingerprint-4), est une méthode largement utilisée en chimio-informatique et en découverte de médicaments pour la création de empreintes moléculaires. Elle permet de capturer les caractéristiques topologiques et les environnements chimiques d'une molécule. Le calcul de l'empreinte Morgan implique plusieurs étapes. Tout d'abord, la molécule est transformée en une représentation graphique, où les atomes sont des nœuds et les liaisons sont des arêtes. Ensuite, à partir de chaque atome, une série d'expansions itératives est effectuée pour générer des sous-structures circulaires de rayon croissant autour de chaque atome. Ces sous-structures circulaires sont appelées fragments centrés sur l'atome.

Pendant le processus d'expansion, les atomes et les liaisons voisines sont pris en compte, créant un environnement local autour de chaque atome. Cette expansion se poursuit pendant un nombre d'itérations spécifié, généralement quatre, ce qui donne l'empreinte ECFP-4. Le nombre quatre indique que le rayon maximal des sous-structures circulaires est de quatre liaisons à partir de l'atome central. L'empreinte ECFP-4 finale est une représentation sous forme de vecteur de bits, où chaque bit encode la présence ou l'absence d'un fragment spécifique centré sur l'atome dans la molécule. La taille de l'empreinte est déterminée par le nombre de fragments uniques dans l'ensemble de données ou la bibliothèque chimique analysée, nous avons choisi la taille de vecteur de 1024 données.

2.2. Simplified Molecular Input Line Entry System (SMILES)

La spécification Simplified Molecular Input Line Entry System (SMILES) est une notation chimique linéaire utilisée pour représenter les structures moléculaires de manière textuelle. Elle est largement utilisée en chimie théorique et informatique pour stocker, transmettre et générer des informations sur les molécules. Le SMILES représente les atomes, les liaisons et les informations stéréochimiques d'une molécule en utilisant une chaîne de caractères. Les atomes sont représentés par leurs symboles chimiques, les liaisons sont indiquées par des tirets (-) et des chiffres pour représenter la connectivité entre les atomes, et les informations stéréochimiques sont généralement représentées par des caractères spéciaux tels que "@", "@@", "/", et "\\". Cette notation permet de décrire de manière compacte et univoque la structure d'une molécule, ce qui facilite son utilisation dans les bases de données chimiques, les logiciels de modélisation moléculaire et d'autres applications liées à la chimie. le SMILES est une spécification qui permet de représenter les structures moléculaires sous forme de chaînes de caractères, offrant ainsi une méthode pratique pour stocker et manipuler des informations chimiques. Les figures suivent représente les SMILES des molécules suivent Nystatine, Paracétamol, PVP XL10.

```
CC1OC(OC2/C=C/C=C/C=C/C=C/CC/C=C/C=C/[C@H](C)[C@@H](O)[C@@H](C)[C
@H](C)OC(=O)CC(O)CC(O)CC(O)CCC(O)C(O)C[C@]3(O)CC(O)C(C(=O)O)C(C2)O3)
C(O)C(N)C1O
```

Figure 45: SMILES de Nystatine

```
CC(=O)Nc1ccc(O)cc1
```

Figure 46: SMILES Paracétamol

O=C1CCCN1C(P)CP

Figure 47: SMILES de Molécule Crospovidone PVP XL10

2.3. Analyse statistique descriptive des données

L'analyse statistique descriptive des données est une étape fondamentale de l'analyse de données qui vise à décrire, résumer et présenter les caractéristiques essentielles d'un ensemble de données. Elle fournit des informations sur la distribution, la tendance centrale, la dispersion et les relations entre les variables étudiées, le tableau suivent représente un résumé de l'analyse statistique des données expérimentale.

Variable category	Factors	A. B	Unit	Min	Max	Standard deviation	Variance	Kurtosis	Skewness	Harm mean	size
Input	Température	T	°C	40	80	13,54	183,27	-1,20	0,234	53,01	(3658,1)
	Pression	P	bar	0,3	1	0,29	0,08	-1,55	-0,158	0,55	
	Humidité	H	%	15	25	4,99	24,95	-1,99	-0,097	18,98	
	Epaisseur	e	mm	0,45	1	0,27	0,07	-1,75	-0,503	0,68	
	Temps	t	Min	0	132	19,45	378,47	4,03	1,613	-	
	Fingerprint solvant	Fs	-	0	1	-	-	-	-	-	(3658,1024)
	Fingerprint poudres	Fp				-	-	-	-	-	
Output	Teneur en Humidité	Mr				0,32	0,10	-0,45	0,865	-	(3658,1)

2.4. Analyse de multicolinéarité et la corrélation des entrées

L'analyse de la multicolinéarité et la corrélation des entrées font partie des techniques utilisées pour évaluer les relations entre les variables dans un modèle statistique, La multicolinéarité se produit lorsqu'il existe une forte corrélation entre deux ou plusieurs variables indépendantes (entrées) dans un modèle de régression ou d'analyse. Cela signifie que ces variables fournissent des informations similaires ou redondantes, ce qui peut entraîner des problèmes dans l'estimation des coefficients de régression et la pertinence des prédictions. Lorsque la multicolinéarité est présente, il devient difficile de déterminer l'effet spécifique de chaque variable indépendante sur la variable dépendante.

L'analyse de la multicolinéarité et la corrélation des entrées sont importantes pour plusieurs raisons :

2.4.1. Impact sur les estimations des coefficients :

Lorsque la multicolinéarité est présente, les estimations des coefficients de régression peuvent devenir instables et peu fiables. De plus, il devient difficile d'interpréter l'impact spécifique de chaque variable indépendante.

2.4.2. Robustesse des prédictions :

La multicolinéarité peut affecter la précision des prédictions du modèle. Des variables indépendantes redondantes peuvent introduire du bruit et rendre les prédictions moins fiables.

2.4.3. Interprétation des résultats :

Lorsque des variables indépendantes sont fortement corrélées, il peut être difficile de déterminer quelle variable a un impact plus important sur la variable dépendante. Cela peut rendre l'interprétation des résultats plus complexe.

La figure suivent représente l'analyse de multicolinéarité des entrées

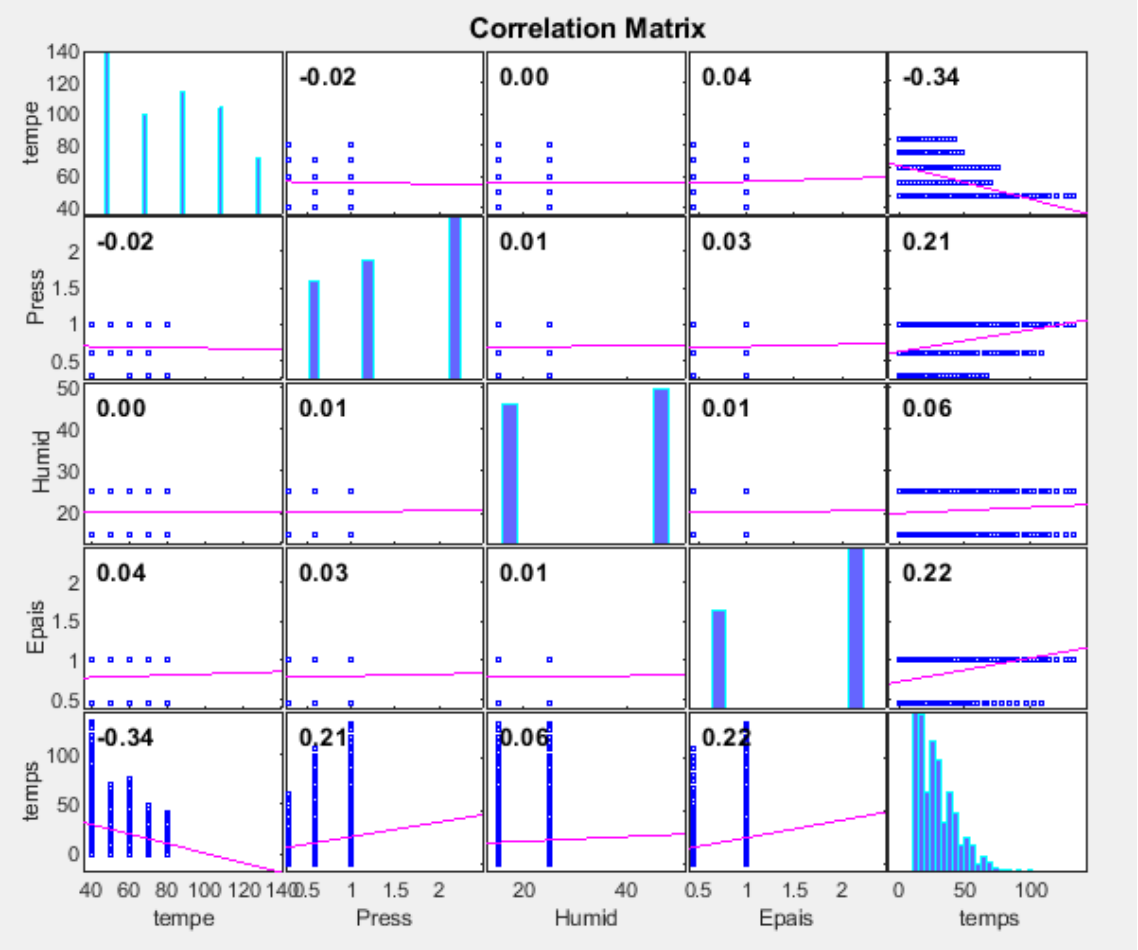


Figure 48.Multicolinéarité des entrées.

Tableau 11. La matrice de valeurs de P Value des entré

	Température	Press	Humidité	Epaisseur	Temps
Température	1,000	0,242	0,967	0,021	0,000
Press	0,242	1,000	0,409	0,116	0,000
Humidité	0,967	0,409	1,000	0,532	0,001
Epaisseur	0,021	0,116	0,532	1,000	0,000
temps	0,000	0,000	0,001	0,000	1,000

La valeur p de sortie présente des valeurs p par paires toutes inférieures au niveau de signification par défaut de 0,05, ce qui indique que toutes les paires de variables présentent une corrélation significativement supérieure à zéro.

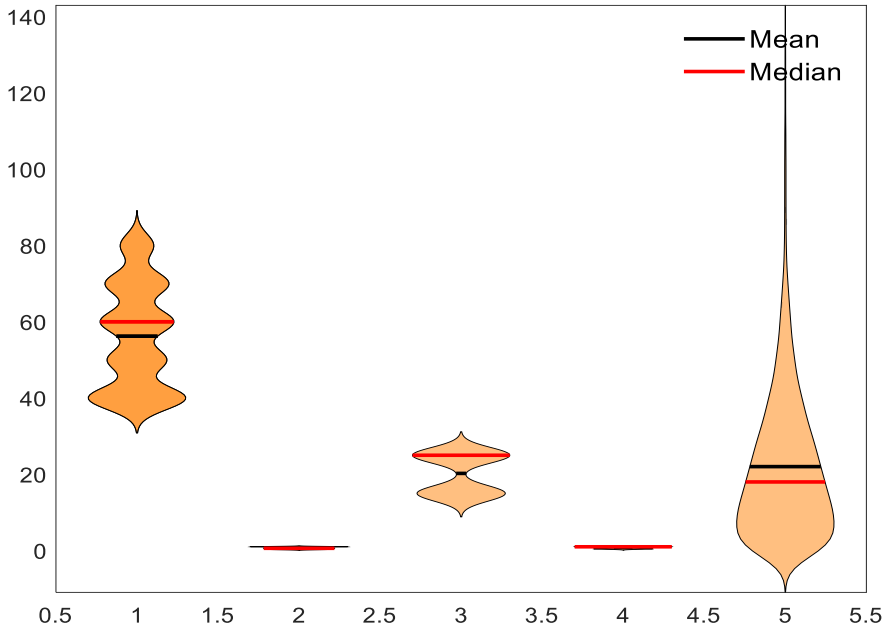


Figure 49. Diagramme en violon des entrés

Un diagramme en violon est un type de visualisation de données qui combine les aspects d'un diagramme en boîte et d'un diagramme de densité de noyau. Il est utilisé pour afficher la distribution et les statistiques récapitulatives des données numériques dans différentes catégories ou groupes. L'intrigue tire son nom de sa ressemblance avec un violon ou une forme de violon musical.

Dans un tracé de violon, chaque catégorie ou groupe est représenté par une forme de "violon". La largeur du violon à un point donné représente la densité ou la fréquence des valeurs de données à ce point. Les parties plus larges du violon indiquent une densité de données plus élevée, tandis que les parties plus étroites indiquent une densité plus faible. Le graphique

comprend également une ligne centrale, qui représente la médiane des données, et comprend souvent des statistiques récapitulatives supplémentaires telles que des quartiles ou des intervalles de confiance.

Avant de commencer la modélisation, nous avons divisé la base de données en deux en utilisant la première partie comme ensemble d'apprentissage et la deuxième partie comme ensemble de validation (70%,30%) respectivment.

Cela signifie que la base de données d'origine a été séparée en deux ensembles distincts. L'ensemble de test, également appelé ensemble de validation, est utilisé pour évaluer les performances du modèle final. Il est généralement utilisé pour estimer à quel point le modèle sera performant lorsqu'il sera utilisé sur de nouvelles données.

2.5. Modélisation AE-MLP

Nous avons choisi d'utilisé la méthode Auto-encoder avec MLP de deep learning pour résumé les donné de fingerprint des solvants et poudre utilisé, après l'AE alimenté les réseaux des neurones de type MLP, la figure suivent représente le modèle choisi pour prédire le taux d'humidité relative Mr.

AE-MLP signifie Autoencoder Multilayer Perceptron. Il fait référence à une architecture de réseau de neurones qui combine les concepts d'un auto-encodeur et d'un perceptron multicouche (MLP).

Un auto-encodeur est un type de réseau de neurones conçu pour apprendre des représentations efficaces des données d'entrée en les encodant dans un espace latent de dimension inférieure, puis en les décodant à l'entrée d'origine. L'objectif d'un auto-encodeur est de reconstruire les données d'entrée aussi précisément que possible, ce qui aide à capturer les caractéristiques importantes et à réduire le bruit ou les informations non pertinentes.

D'autre part, un perceptron multicouche (MLP) est un type de réseau de neurones à anticipation avec plusieurs couches de nœuds (neurones). Les MLP sont couramment utilisés pour des tâches telles que la classification et la régression, où ils apprennent à mapper des entités d'entrée sur des étiquettes de sortie.

La combinaison d'un auto-encodeur et d'un MLP, appelée AE-MLP, intègre les points forts des deux architectures. Il exploite la capacité des auto-encodeurs à apprendre des représentations significatives des données par codage et décodage, tout en utilisant également les puissantes capacités de mappage des MLP pour des tâches telles que la classification ou la régression.

Dans une architecture AE-MLP, les données d'entrée sont d'abord codées à l'aide du réseau de codeurs d'un auto-encodeur, ce qui réduit la dimensionnalité des données et capture les caractéristiques importantes. La représentation codée est ensuite introduite dans un MLP, qui effectue d'autres transformations et apprend à mapper les caractéristiques codées sur les

étiquettes de sortie souhaitées. Cette combinaison permet un apprentissage plus efficace des fonctionnalités et peut améliorer les performances du réseau sur diverses tâches.

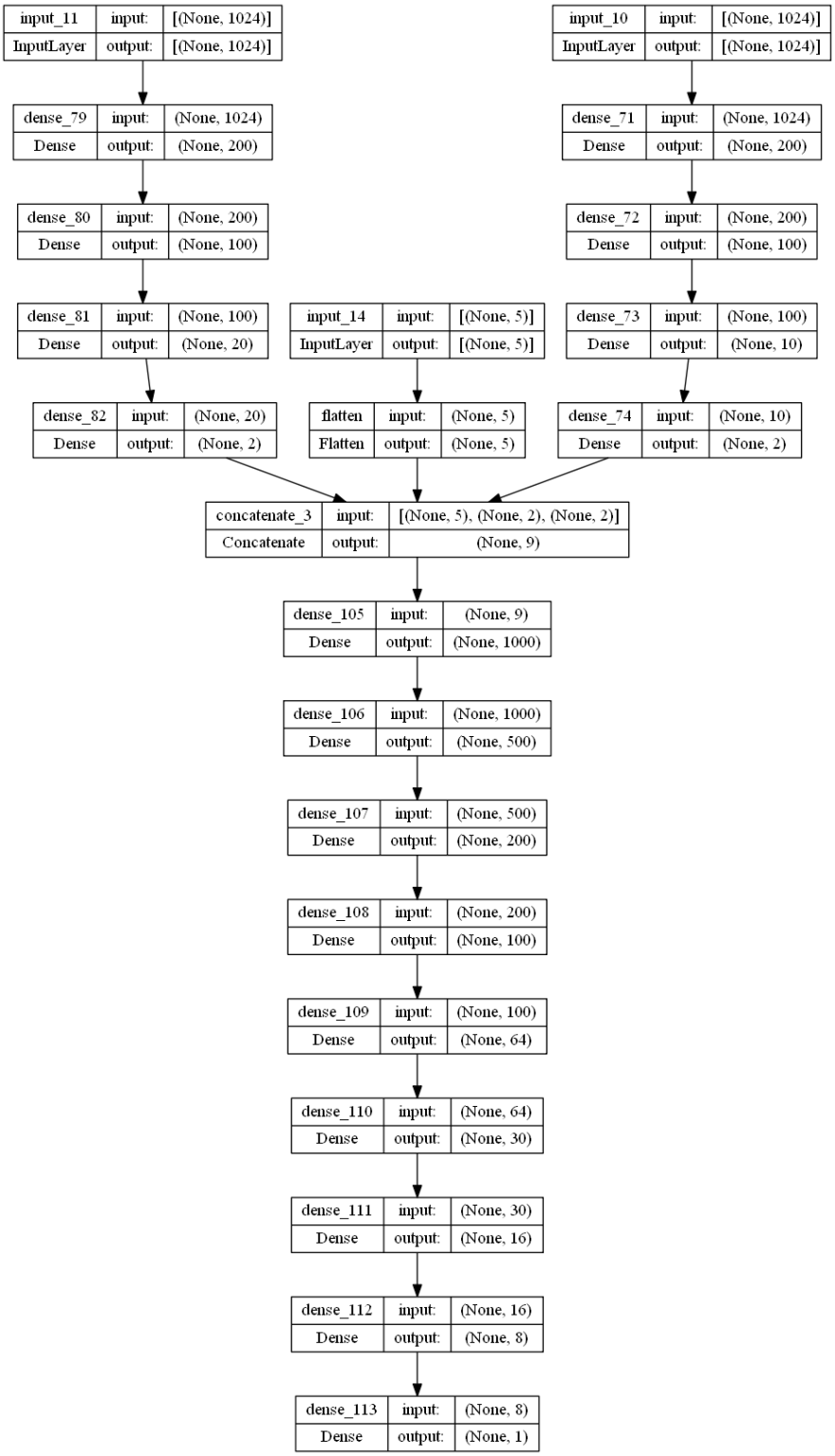


Figure 50. Modèle AE-MLP.

2.6. Résultats de la modélisation

Une fois que nous avons entraîné le modèle sélectionné, nous avons obtenu des résultats très fiables qui sont représentés par le coefficient de détermination et le RMSE.

Les mesures du coefficient de détermination et du RMSE sont couramment utilisées pour évaluer la performance d'un modèle de régression. Le coefficient de détermination, également appelé R^2, mesure la proportion de la variance des données de sortie qui est expliquée par le modèle. Une valeur de R^2 proche de 1 indique que le modèle s'ajuste bien aux données et explique une grande partie de la variabilité. D'autre part, le RMSE (Root Mean Squared Error) mesure la moyenne des écarts entre les valeurs prédites par le modèle et les valeurs réelles des données. Un RMSE plus faible indique que le modèle a une meilleure capacité à prédire avec précision les valeurs cibles.

Tableau 12. Valeurs des critères d'évaluation du modèle optimale à deux phases apprentissage et validation

	Apprentissage	Validation	All
MSE	0.0016	0.0030	0.0022
MAE	0.0293	0.0407	0.0346
RMSE	0.0400	0.0543	0.0467
R2	0.9899	0.9728	0.9811

Dans notre cas, les résultats du coefficient de détermination et du RMSE indiquent que le modèle choisi a une bonne capacité à expliquer la variance des données et à prédire avec précision les valeurs cibles. Cela suggère que le modèle est fiable et peut être utilisé pour faire des prédictions sur de nouvelles données.

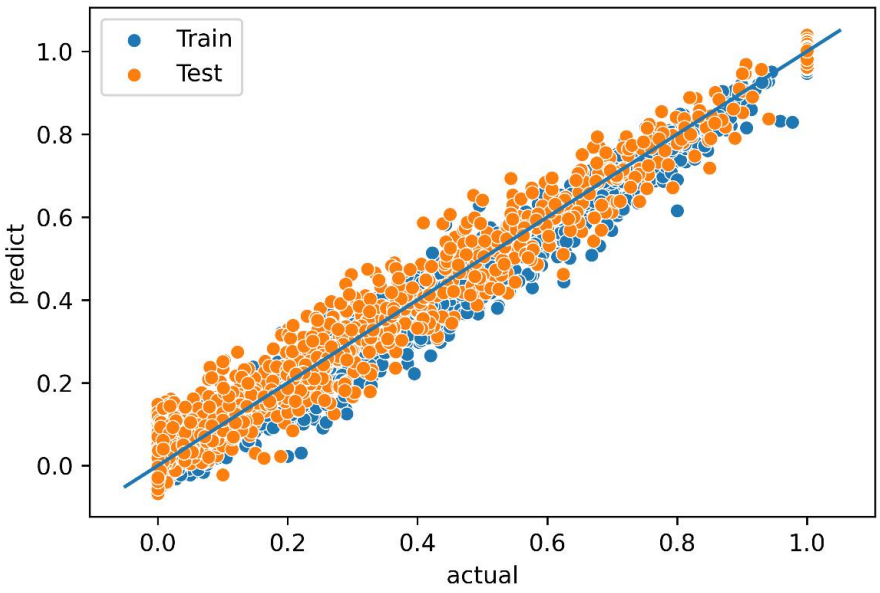


Figure 51. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase Apprentissage.

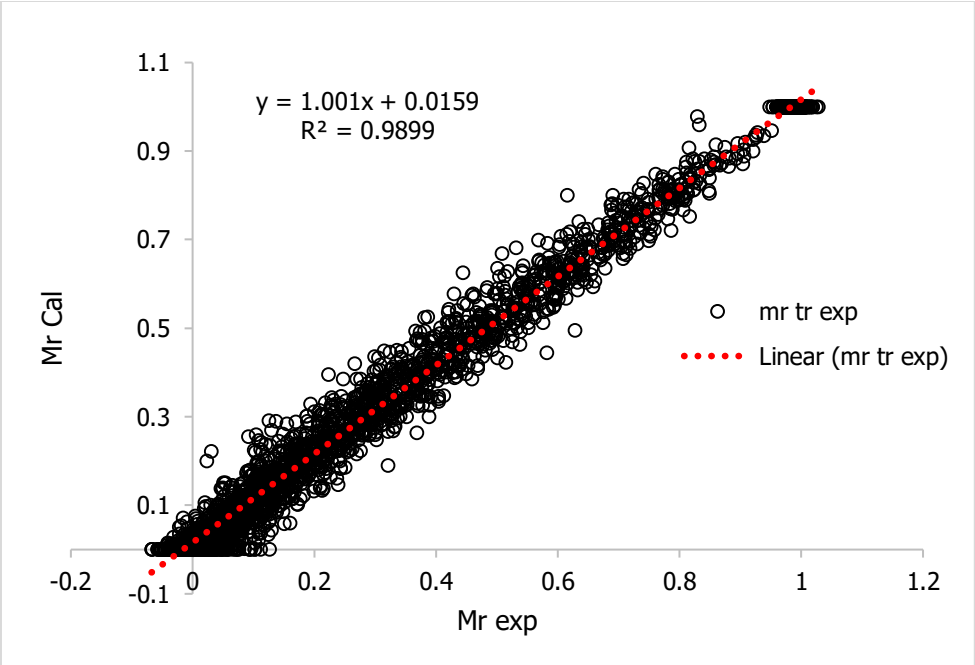


Figure 52. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase Apprentissage.

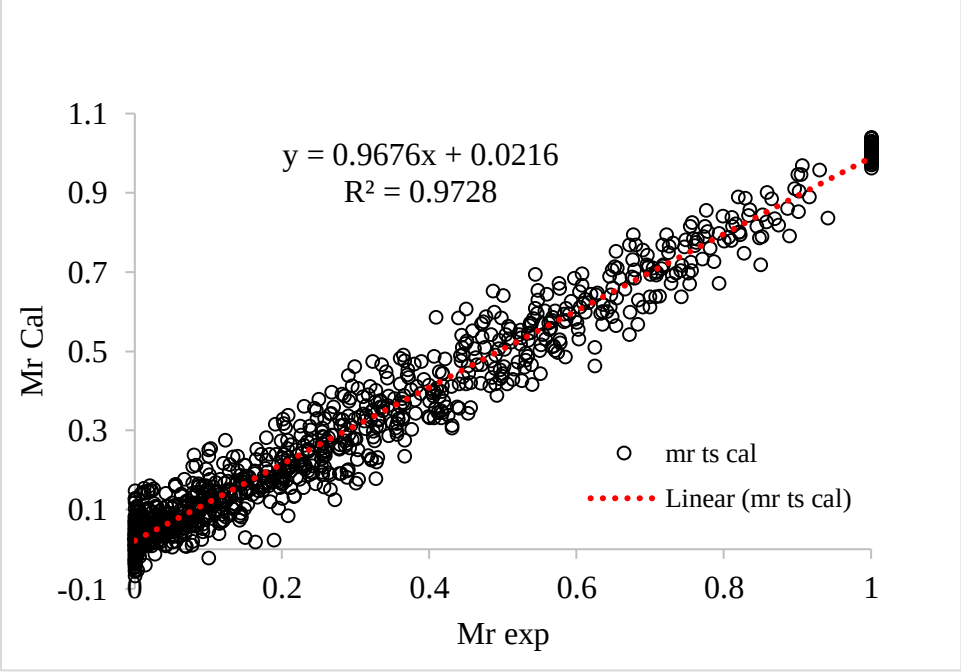


Figure 53. Corrélation linière entre les valeurs prédit et expérimentale de taux d'humidité relative de la phase validation.

2.7. Analyse de la variance des valeurs prédit et les valeurs de l'expérimentale

2.7.1. Hypothèse H0

Puisque p-value > α, H0 est accepté. Les moyennes de tous les groupes sont supposées égales. En d'autres termes, la différence entre les moyennes de tous les groupes n'est pas assez grande pour être statistiquement significative.

2.7.2. Valeur P

la valeur de p est égale à 0,0208334, [p($x \leq F$) = 0,979167]. Cela signifie que si nous rejetions H0, le risque d'erreur de type 1 (rejeter un H0 correct) serait trop élevé : 0,02083 (2,08 %). Plus la valeur de p est élevée, plus elle supporte H0.

3. Les statistiques

La statistique de test F est égale à 5,342991, ce qui se situe dans la région d'acceptation à 99 % : [-∞ : 6,6384].

Tableau 13 . Analyse de la variance

Source	DF	Sum of Square	Mean Square	F Statistic	P-value
Groups (between groups)	1	0.5295	0.5295	5.343	0.02083
Error (within groups)	7314	724.7647	0.09909		
Total	7315	725.2942	0.09915		

2.7.3. Taille de l'effet

La taille d'effet observée f est faible (0,027). Cela indique que l'ampleur de la différence entre les moyennes est faible. Le η2 est égal à 0,00073. Cela signifie que le groupe explique 0,07 % de la variance par rapport à la moyenne (similaire à R2 dans la régression linéaire)

2.7.4. Tukey HSD / Tukey Kramer

Il n'y a pas de différence significative entre les moyennes d'une paire.

Tableau 14 . Analyse de teste Tukey HSD / Tukey Kramer

Pair	Difference	SE	Q	Lower CI	Upper CI	Critical Mean	p-value
x1-x2	0.02	0.005	3.3	-0.002	0.04	0.02	0.02

2.8. Validation par Egalité des variances

L'outil a utilisé le test de Levene pour évaluer l'égalité des variances. Les variances de la population sont considérées comme égales. (valeur de p = 0,207). La puissance de test de Levene est considérée comme forte (1). La taille des groupes est considérée comme similaire. (Le rapport entre le plus grand groupe et le plus petit groupe est : 1) Le test ANOVA est considéré comme robuste à l'hypothèse d'homogénéité des variances lorsque les tailles des groupes sont similaires.

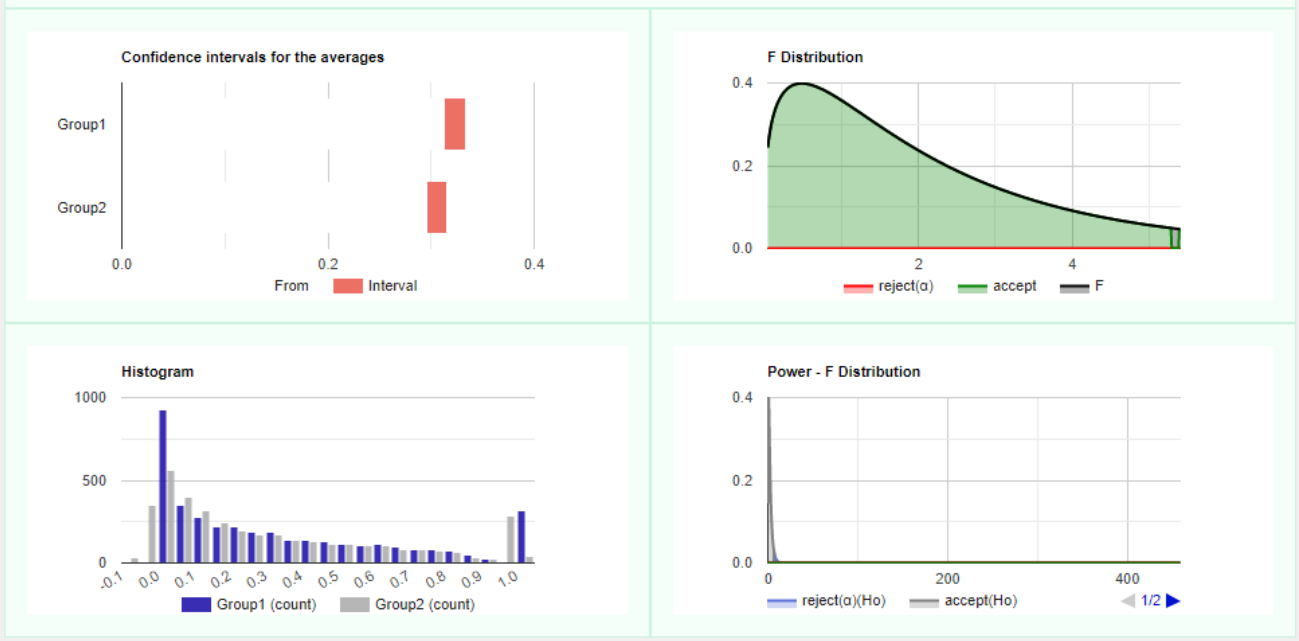


Figure 54. Description statistique des données expérimentale et calculée par le modèle.

Conclusion

Les résultats expérimentaux obtenus en utilisant un auto encodeur et un perceptron multicouche (MLP) ont montré des performances prometteuses. Tout d'abord, l'auto encodeur a été utilisé pour résumer les empreintes digitales générées à partir des données moléculaire des poudres utilisées. Ensuite, nous avons modélisé le taux d'humidité (MR) en fonction des paramètres expérimentaux tels que la température, la pression, l'épaisseur, le taux d'humidité initial et le temps, ainsi que le résumé des empreintes digitales.

Les résultats de nos modèles ont été évalués en utilisant P-valeur, la non signification de P-valeur indique que nos modèles ont réussi à représenter de manière significative les relations entre les paramètres expérimentaux et les Mr. De plus, le coefficient de détermination (R2) a montré des résultats proches de 1, ce qui suggère que nos modèles capturent bien les variations des taux de réponse en fonction des paramètres expérimentaux.

En outre, l'erreur quadratique moyenne (RMSE) est également proche de 0, ce qui confirme la précision de nos modèles dans la prédiction des taux de réponse. Ces résultats indiquent que

nos modèles basés sur l'auto encodeur et le MLP sont performants et fournissent des prédictions précises pour les taux de réponse en fonction des paramètres expérimentaux étudiés.

Nos résultats expérimentaux et les performances de nos modèles suggèrent que l'utilisation de l'auto encodeur pour résumer les empreintes digitales et le modèle MLP pour modéliser les taux de réponse en fonction des paramètres expérimentaux est une approche efficace. Ces résultats ouvrent des perspectives intéressantes pour l'application de ces modèles dans d'autres domaines de recherche où la prédiction des résultats expérimentaux est cruciale.

Conclusion Générale

Ce travail est basé sur une étude expérimentale du phénomène de séchage des poudres pharmaceutiques paracétamol, nystatine, PVP. La méthodologie du manuscrit est divisée en deux. Une partie théorique et une partie expérimentale.

Dans l'aspect théorique, nous avons présenté différentes méthodes de modélisation numérique utilisée dans les domaines de la chimie, des sciences pharmaceutiques et de l'apprentissage automatique. Les plans d'expérience sont examinés comme une approche statistique pour concevoir des expériences systématiques et optimiser les processus. Le deep learning est également abordé, en mettant en évidence les réseaux neuronaux profonds tels que le MLP et les CNN, qui sont utilisés dans des domaines tels que la reconnaissance d'images et le traitement du langage naturel.

Dans l'aspect expérimentale, nous avons étudié le séchage des poudres pharmaceutiques par la méthode de plan expérimental et nous avons modélisé certains des facteurs qui en effets sur le temps de séchage final.

Ensuite nous avons étudié Les résultats obtenus en utilisant un autoencodeur et un perceptron multicouche (MLP) ont montré des performances prometteuses. En modélisant le taux d'humidité (MR) en fonction de différents paramètres, ainsi que le résumé des empreintes digitales, les modèles ont démontré leur capacité à représenter de manière significative les relations entre ces paramètres et les taux d'humidité.

L'évaluation des modèles a été réalisée en utilisant la p-valeur, qui a confirmé la significativité des relations modélisées. De plus, le coefficient de détermination (R2) a montré des résultats proches de 1 tel que R2=0.9811, indiquant que les modèles sont capables de capturer efficacement les variations des taux de réponse en fonction des paramètres expérimentaux. L'erreur quadratique moyenne (RMSE) proche de 0 tel que RMSE=0.0467 atteste également de la précision des modèles dans la prédiction des taux de réponse. Ces résultats démontrent ainsi que les modèles basés sur l'autoencodeur et le MLP sont performants et permettent de fournir des prédictions précises pour les taux d'humidité en fonction des paramètres expérimentaux étudiés.

Référence Bibliographique

[1] J.T. Carstensen, P. Toure, Compression cycles in tableting, Powder Technology. 26 (1980) 199–204. https://doi.org/10.1016/0032-5910(80)85062-5.

[2] J.T. Carstensen, P. Touré, D. Duchĕne, F. Puisieux, Correlation Between Hysteresis Loop Areas of Lower Punch and of Die Pressures Versus Upper Punch Pressures, Drug Development and Industrial Pharmacy. 7 (1981) 645–648. https://doi.org/10.3109/03639048109057736.

[3] B. Lassudrie-Duchene, J.-L. Mucchielli, Les échanges intra-branche et la hiérarchisation des avantages comparés dans le commerce international, Revue Économique. 30 (1979) 442. https://doi.org/10.2307/3501231.

[4] P. Arlabosse, Séchage industriel, Techniques de l'ingénieur. (2008).

[5] E. Koller, Aide-mémoire Génie chimique, 3rd ed., Paris, 2009.

[6] J.H. Ferrasse, Développement d'outils expérimenataux pour le dimensionnement du procédé de séchage conductif avec agitation : application à des boues de station d'épuration urbaines, Thèse Univ. Toulouse II. (2000).

[7] D. Farges, Contribution à l'étude du séchage par contact sous atmosphère inerte et sous vide: étude expérimentale et théorique, Thèse de Doctorat. (1996).

[8] A. Seghier, S. Rouas, Étude du retrait au cours du séchage des fruits Soutenu, These de Master. (2016).

[9] W.L. McCabe, J.C. Smith, P. Harriott, Unit operation of chemical engineering, Paris, 1985.

[10] R. Cavaillé, A. Charreau, Séchage: théorie et calculs, Techniques de l'ingénieur. (1995) 1–23.

[11] S. Trichaiyaporn, Contribution à l'étude du séchage en discontinu de sucre cristalisé en lit fluidisé, These de Doctorat. (1991).

[12] J.C. Charpentier, génie des procédés, Technique de'Ingénieur. (1996).

[13] A. Charreau, R. Cavaille, Séchage théorie et pratique, Technique de'Ingénieur. (1991).

[14] L. Bennamoun, A. Belhamri, Contribution à l'étude de faisabilité de séchoirs solaires pour les produits agricoles Revue des Energies Renouvelables, CER. (2007).

[15] S.J. Friedman, W.R. Marshall, STUDIES IN ROTARY DRYING. 2. HEAT AND MASS TRANSFER, Chemical Engineering Progress. 45 (1949) 573–588.

[16] L. Busin, F. Mortazavi, J.J. Bimbenet, Étude expérimentale du séchage par atomisation de solutions sucrées, Industries Alimentaires et Agricoles. 112 (1995).

[17] A. Rankelli, H. Lieberman, R. Schiffman, Drying, the theory and practice of industrial pharmacy, Philadelphia, 1996.

[18] J. Claude, Contribution à l'étude du séchage en discontinue d'acide adipique en lit fluidisé, Thèse de Doctorat. (1993).

[19] Marcel Loncin, Livres sur Google Play Les opérations unitaires du génie chimique, 1961.

[20] H. Bozkir, Y. Tekgül, E.S. Erten, Effects of tray drying, vacuum infrared drying, and vacuum microwave drying techniques on quality characteristics and aroma profile of orange peels, Journal of Food Process Engineering. 44 (2021). https://doi.org/10.1111/jfpe.13611.

[21] F. Dahmoune, B. Nayak, K. Moussi, H. Remini, K. Madani, Optimization of microwave-assisted extraction of polyphenols from Myrtus communis L. leaves, Food Chemistry. 166 (2015) 585–595. https://doi.org/10.1016/j.foodchem.2014.06.066.

[22] R. Jeantet, T. Croguennec, P. Schuck, G. Brulé, Sciences des aliments. 1 Stabilisation biologique et physicochimique., 1st ed., Paris, 2006.

[23] G.D. Saravacos, Mass transfer properties of foods. In Engineering properties of foods, 1995.

[24] J.J. Bimbenet, Les échanges dans l'industrie agricole et alimentaire, Cahiers Du Génie Industriel Alimentaire (GIA). (1984).

[25] D. Djerroud, Modélisation markovienne du séchage continu par contact avec agitation, Doctoral Dissertation. (2010) 1–181.

[26] G. V Barbosa-Cánovas, H. Vega-Mercado, Dehydration of foods, Springer Science & Business Media, 1996.

[27] J.W.-C. Y.H. Hui, C. Clary, M.M. Farid, O.O. Fasina, A. Noomhorm, Food Drying Science and Technology Microbiology, Chemistry, Applications, 2008.

[28] J. Crank, The mathematics of diffusion 2nd Edition, Oxford Science Publications,. (1975) 32.

[29] P. et al, Grain drying: Theory and practice, John Wiley & Sons, 1998.

[30] K. Kahveci, A. Cihan, Drying of food materials: transport phenomena, Nova Science Publishers, 2008.

[31] C.W.H. Donald B. Brooker, Fred W. Bakker-Arkema, Drying Cereal Grains. AVI Publishing Company, 1974.

[32] S. Pabis, D.S. Jayas, S. Cenkowski, Grain drying: theory and practice, John Wiley & Sons, 1998.

[33] P.S. Madamba, Thin Layer Drying Models for Osmotically Pre-dried Young Coconut, Drying Technology. 21 (2003) 1759–1780. https://doi.org/10.1081/DRT-120025507.

[34] W.K. Lewis, The Rate of Drying of Solid Materials., Journal of Industrial & Engineering Chemistry. 13 (1921) 427–432. https://doi.org/10.1021/ie50137a021.

[35] B. Vijayaraj, R. Saravanan, S. Renganarayanan, Studies on thin layer drying of bagasse, International Journal of Energy Research. 31 (2007) 422–437. https://doi.org/10.1002/er.1237.

[36] A. Ghazanfari, S. Emami, L.G. Tabil, S. Panigrahi, Thin-Layer Drying of Flax Fiber: II. Modeling Drying Process Using Semi-Theoretical and Empirical Models, Drying Technology. 24 (2006) 1637–1642. https://doi.org/10.1080/07373930601031463.

[37] G.E. Page, Factors Influencing the Maximum Rates of Air Drying Shelled Corn in Thin layers., (1949).

[38] I. Doymaz, Thin-layer drying behaviour of mint leaves, Journal of Food Engineering. 74 (2006) 370–375.

[39] D.G. Overhults, G.M. White, H.E. Hamilton, I.J. Ross, Drying soybeans with heated air, Transactions of the ASAE. 16 (1973) 112–113.

[40] G.M. White, I.J. Ross, C.G. Poneleit, Fully-exposed drying of popcorn, Transactions of the ASAE. 24 (1981) 466–468.

[41] C.Y. Wang, R.P. Singh, A single layer drying equation for rough rice, ASAE paper, 1978.

[42] M.H. Jazini, M.S. Hatamipour, A new physical pretreatment of plum for drying, Food and Bioproducts Processing. 88 (2010) 133–137.

[43] C. Ertekin, M.Z. Firat, A comprehensive review of thin-layer drying models used in agricultural products, Critical Reviews in Food Science and Nutrition. 57 (2017) 701–717. https://doi.org/10.1080/10408398.2014.910493.

[44] L.M. Diamante, P.A. Munro, Mathematical modelling of hot air drying of sweet potato slices, International Journal of Food Science & Technology. 26 (1991) 99–109.

[45] D.G.P. Kumar, H.U. Hebbar, M.N. Ramesh, Suitability of thin layer models for infrared–hot air-drying of onion slices, LWT-Food Science and Technology. 39 (2006) 700–705.

[46] C. Chen, P.C. Wu, Thin-layer drying model for rough rice with high moisture content, Journal of Agricultural and Engineering Research. 80 (2001) 45–52. https://doi.org/10.1006/jaer.2000.0677.

[47] S.M. Hendreson, S. Pabis, Grain drying theory. I. Temperature effect on drying coefficients, Journal of Agricultural Engineering Research. 6 (1961) 169–174.

[48] V.T. Karathanos, Determination of water content of dried fruits by drying kinetics, Journal of Food Engineering. 39 (1999) 337–344.

[49] O. Corzo, N. Bracho, C. Alvarez, Determination of suitable thin layer model for air drying of mango slices (Mangifera indica L.) at different air temperatures and velocities, Journal of Food Process Engineering. 34 (2011) 332–350.

[50] I.T. Togrul, D. Pehlivan, Mathematical modelling of solar drying of apricots in thin layers, Journal of Food Engineering. 55 (2002) 209–216. https://doi.org/10.1016/S0260-8774(02)00065-1.

[51] S.M. Henderson, Progress in developing the thin layer drying equation, Transactions of the ASAE. 17 (1974) 1167–1172.

[52] E. Demirhan, B. ÖZBEK, Microwave-drying characteristics of basil, Journal of Food Processing and Preservation. 34 (2010) 476–494.

[53] G. Dadali, E. Demirhan, B. Özbek, Color change kinetics of spinach undergoing microwave drying, Drying Technology. 25 (2007) 1713–1723.

[54] R. Baini, T.A.G. Langrish, Choosing an appropriate drying model for intermittent and continuous drying of bananas, Journal of Food Engineering. 79 (2007) 330–343.

[55] J. Chen, C. Wu, S. Yoon, Y. You, A stabilized conforming nodal integration for Galerkin mesh-free methods, International Journal for Numerical Methods in Engineering. 50 (2001) 435–466.

[56] D. Hutchinson, L. Otten, Thin-layer air drying of soybeans and white beans, International Journal of Food Science & Technology. 18 (2007) 507–522. https://doi.org/10.1111/j.1365-2621.1983.tb00292.x.

[57] L.R. Verma, R.A. Bucklin, J.B. Endan, F.T. Wratten, Effects of drying air parameters on rice drying models, Transactions of the ASAE. 28 (1985) 296–301.

[58] A.S. Kassem, Comparative studies on thin layer drying models for wheat, in: 13th International Congress on Agricultural Engineering, 1998: pp. 2–6.

[59] A. Midilli, H. Kucuk, Z. Yapar, A new model for single-layer drying, Drying Technology. 20 (2002) 1503–1513.

[60] I. Doymaz, Mathematical modelling of thin-layer drying of kiwifruit slices, Journal of Food Processing and Preservation. 33 (2009) 145–160.

[61] S.J. Babalis, E. Papanicolaou, N. Kyriakis, V.G. Belessiotis, Evaluation of thin-layer drying models for describing drying kinetics of figs (Ficus carica), Journal of Food Engineering. 75 (2006) 205–214.

[62] M. Mundada, B.S. Hathan, S. Maske, Convective dehydration kinetics of osmotically pretreated pomegranate arils, Biosystems Engineering. 107 (2010) 307–316.

[63] O. Corzo, N. Bracho, A. Pereira, A. Vásquez, Weibull distribution for modeling air drying of coroba slices, LWT - Food Science and Technology. 41 (2008) 2023–2028. https://doi.org/10.1016/j.lwt.2008.01.002.

[64] A. Vega-Gálvez, A. Andrés, E. Gonzalez, E. Notte-Cuello, M. Chacana, R. Lemus-Mondaca, Mathematical modelling on the drying process of yellow squat lobster (Cervimunida jhoni) fishery waste for animal feed, Animal Feed Science and Technology. 151 (2009) 268–279. https://doi.org/10.1016/j.anifeedsci.2009.01.003.

[65] S. Jena, H. Das, Modelling for vacuum drying characteristics of coconut presscake, Journal of Food Engineering. 79 (2007) 92–99. https://doi.org/10.1016/j.jfoodeng.2006.01.032.

[66] J. Neter, M.H. Kutner, C.J. Nachtsheim, W. Wasserman, Applied linear statistical models, (1996).

[67] L. Kong, X. Yang, Z. Hou, J. Dong, Mathematical Modeling of Drying Kinetics for Pulp Sheet Based on Fick's Second Law of Diffusion, Journal of Korea Technical Association of the Pulp and Paper Industry. 52 (2020) 23–31. https://doi.org/10.7584/JKTAPPI.2020.04.52.2.23.

[68] K. Sacilik, A.K. Elicin, The thin layer drying characteristics of organic apple slices, Journal of Food Engineering. 73 (2006) 281–289.

[69] H. Toğrul, N. Arslan, Moisture sorption isotherms and thermodynamic properties of walnut kernels, Journal of Stored Products Research. 43 (2007) 252–264.

[70] V. Demir, T. Gunhan, A.K. Yagcioglu, Mathematical modelling of convection drying of green table olives, Biosystems Engineering. 98 (2007) 47–53. https://doi.org/10.1016/j.biosystemseng.2007.06.011.

[71] M. Hentabli, A.-E. Belhadj, H. Benimam, F. Dahmoune, S. Keskes, Vacuum drying of the Terbinafine HCl powder: A kinetics study and mathematical modeling, Powder Technology. 383 (2021). https://doi.org/10.1016/j.powtec.2021.01.038.

[72] L. David, A. Thakkar, R. Mercado, and O. Engkvist, "Molecular representations in AIdriven drug discovery: a review and practical guide," J. Cheminform., vol. 12, no. 1, pp. 1–22, 2020, doi: 10.1186/s13321-020-00460-5.

[73] M. Krenn et al., "SELFIES and the future of molecular string representations," pp. 1– 34, 2022, [Online]. Available: http://arxiv.org/abs/2204.00056. [15] R. Sayle and J. Mayfield, "Advances in (standardization of) organometallic and inorganic structure representations," no. March, 2022.

[74] R. Sayle and J. Mayfield, "Advances in (standardization of) organometallic and inorganic structure representations," no. March, 2022.

[75] Accelrys, "CTfile Formats," CTfile Formats, no. October, 2012, [Online]. Available: http://www.symyx.com/downloads/public/ctfile/ctfile.jsp. [17] A. Dalby et al., "Description of Several Chemical Structure File Formats Used by Computer Programs Developed at Molecular Design Limited," J. Chem. Inf. Comput. Sci., vol. 32, no. 3, pp. 244–255, 1992, doi: 10.1021/ci00007a012.

[76] A. Dalby et al., "Description of Several Chemical Structure File Formats Used by Computer Programs Developed at Molecular Design Limited," J. Chem. Inf. Comput. Sci., vol. 32, no. 3, pp. 244–255, 1992, doi: 10.1021/ci00007a012.

[77] D. S. Wigh, J. M. Goodman, and A. A. Lapkin, "A review of molecular representation in the age of machine learning," Wiley Interdiscip. Rev. Comput. Mol. Sci., no. August 2021, pp. 1–19, 2022, doi: 10.1002/wcms.1603.

[78] A. C. K. To and R. For, "S TEPPING B ACK TO SMILES T RANSFORMERS FOR," pp. 1–14, 2022.

[79] N. M. O'Boyle, "Towards a Universal SMILES representation - A standard method to generate canonical SMILES based on the InChI," J. Cheminform., vol. 4, no. 9, pp. 1– 14, 2012, doi: 10.1186/1758-2946-4-22.

[80] D. Weininger, "J. Chem. Inj Comput.," J. Chem. Inf. Comput. Sci., vol. 28, pp. 31–36, 1988.

[81] S. Honda, S. Shi, and H. R. Ueda, "SMILES Transformer : Pre-trained Molecular Fingerprint for Low Data Drug Discovery," 2018.

[82] G. Landrum, "RDKit: Open-source Cheminformatics," Http://Www.Rdkit.Org/, vol. 3, no. 04, p. 2012, 2006, [Online]. Available: http://www.rdkit.org.

[83] H. Gao, T. J. Struble, C. W. Coley, Y. Wang, W. H. Green, and K. F. Jensen, "Using Machine Learning to Predict Suitable Conditions for Organic Reactions," ACS Cent. Sci., vol. 4, no. 11, pp. 1465–1476, 2018, doi: 10.1021/acscentsci.8b00357.

[84] A. C. Mater and M. L. Coote, "Deep Learning in Chemistry," J. Chem. Inf. Model., 2019, doi: 10.1021/acs.jcim.9b00266.

[85] P. J. Chun, T. Yamane, and Y. Maemura, “A deep learning-based image captioning method to automatically generate comprehensive explanations of bridge damage,” Comput. Civ. Infrastruct. Eng., pp. 1–15, 2021, doi: 10.1111/mice.12793.