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

UNIVERSITÉ YAHIA FARÈS DE MÉDÉA
Département de Génie des Procédés et Environnement
Laboratoire de Biomatériaux et Phénomène de Transport (LBMPT)

THÈSE DE DOCTORAT

Présentée par :

M^lle^ Widad BENMOULOUD

Licencié en Génie des Procédés et Environnement : Université de Médéa, Algérie

Master en Génie chimique : Université de Médéa, Algérie

En vue de l'obtention du titre de Doctorat en Génie chimique

Filière : Génie des Procédés

Thème

Contribution à la modélisation des propriétés physiques et thermodynamiques des liquides ioniques purs et leurs mélanges.

Soutenue devant le jury composé de :

Pr. LAIDI Maamar Professeur, Université Yahia Fares de Médéa	Président
Pr. NASRI Loubna Professeur, Université de Constantine 3	Examinatrice
Dr. AMMI Yamina Maitre de Conférences A, Université Yahia Fares de Médéa	Examinatrice
Pr. HANINI Salah Professeur, Université Yahia Fares de Médéa	Invité
Pr. SI-MOUSSA Cherif Professeur, Université Yahia Fares de Médéa	Directeur de thèse
Pr. BENKORTBI Othmane Professeur, Université Yahia Fares de Médéa	Co Directeur de thèse

Dédicaces

J’ai l’immense plaisir de dédier ce modeste travail à :

Ceux que j’aime le plus au monde, **mes très chers parents** qui m’encouragent et me poussent toujours vers la réussite, que ALLAH les garde et les protège.

Je vous dois ce que je suis aujourd'hui grâce à votre amour, votre patience, vos innombrables sacrifices et toute l'éducation que vous m'avez donnée.

À mon cher frère **Abdelhak**, et sa femme **Romaissaa**

À ma chérie nièce **Hiba**,

À mon grand-père et mes grands-mères,

À toute ma famille **BENMOULOUD**,

À mes copines **Imene** et **Hayet**,

À tous ceux qui m'ont soutenu de près ou de loin.

BENMOULOUD Widad

Remerciement

*A*vant tout je remercie ALLAH le tout puissant de m'avoir donné la volonté, la force, la patience et le courage de mener à bien ce modeste travail car ma réussite n'est due qu'à Allah.

*J*e tiens à adresser mes plus sincères remerciements à mon Directeur de Thèse **Pr. SI-MOUSSA Cherif** qui a cru en moi depuis mes années de master et à mon Co-Directeur de Thèse **Pr. BENKORTBI Othmane** pour le temps précieux qu'ils ont consacrés au cours de mes travaux et recherche et pour leurs suggestions, tant sur le plan scientifique qu'éthique, où ils ont été une source constante de conseils et de réflexions. Et pour la confiance qu'ils m'ont accordée durant ces quatre années.

*J*e tiens également à exprimer mes plus vifs remerciements aux mes chers parents et cher frère pour le suivi quotidien des activités de recherche et pour leurs encouragements.

*J*e tiens également à présenter ma profonde gratitude aux : président du jury le professeur Pr. LAIDI Maamar, et à messieurs les membres de jury Pr. HANINI Salah, Dr. AMMI Yamina à l`université de Médéa et Pr. NASRI Loubna à l'université de Constantine 3, d'avoir accepté de juger ce modeste travail.

*J*e remercie ma copine Imene et tous ceux qui ont contribué de près ou de loin, directement ou indirectement 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.

ملخص

المساهمة في نمذجة الخواص الفيزيائية والديناميكية الحرارية للسوائل الأيونية النقية ومخاليطها

تشكل السوائل الأيونية (LI) فئة جديدة من السوائل تعتبر مذيبات خضراء مبتكرة ذات أهمية كبيرة من المنظور الأكاديمي والصناعي والبيئي، وذلك بسبب خصائصها الجذابة. لتصميم أي عملية تنطوي على السوائل الأيونية على نطاق صناعي، فمن الضروري معرفة مجموعة من الخصائص الفيزيائية والكيميائية والديناميكية الحرارية. ولذلك، فإن النمذجة لها أهمية كبيرة. لذلك، فإن الهدف من هذا العمل هو تطوير نماذج QSPR تساهم في نمذجة خاصيتين للسوائل الأيونية، وهما التوتر السطحي للمخاليط الثنائية وقابلية ذوبان ثاني أكسيد الكربون في السوائل الأيونية باستخدام خوارزميات التعلم الآلي..

ركزت مساهمتنا الأولى على التنبؤ بتوتر 62 خليطًا ثنائيًا (LI+ماء/مذيبات عضوية مختلفة) لقاعدة مكونة من 1623 نقطة بيانات تجريبية تم جمعها من الأدبيات من خلال تطبيق ثلاث تقنيات ذكاء حسابية ،الشبكة العصبية الاصطناعية التنبؤية (ANN)، وآلة ناقلات دعم المربعات الصغرى (LSSVM)، وآلة ناقلات الدعم التي تعمل بتحسين سرب الجسيمات (SVM-PSO). وبشكل عام، فإن النماذج المقترحة تجعل من الممكن ربط البيانات التجريبية التي تم جمعها في الأدبيات بدقة. ومع ذلك، كان التنبؤ من قبل ANN أفضل من النهجين الآخرين. تُقدر معاملات الارتباط والانحرافات المعبر عنها بمتوسط الانحراف النسبي المطلق لنموذج الشبكة العصبية ANN بـ 0.9979 و0.8466% على التوالي.

ركزت مساهمتنا الثانية على التنبؤ بقابلية ذوبان ثاني أكسيد الكربون في 124 سائل أيوني لقاعدة مكونة من 10116 بيانات تجريبية تم جمعها من أدبيات باستخدام الشبكة العصبية الاصطناعية التنبؤية (ANN) وآلة ناقل الدعم المُحسّنة بواسطة خوارزمية اليعسوب (DA-SVM). أظهرت النتائج التي تم الحصول عليها أن كلا النموذجين يمكنهما التنبؤ بشكل موثوق بقابلية ذوبان ثاني أكسيد الكربون في LI مع تفوق طفيف لنموذج ANN مع معامل الارتباط الكلي ومتوسط الخطأ المطلق البالغ 0.9906 و0.0199 على التوالي.

الكلمات المفتاحية: السوائل الأيونية، التوتر السطحي، ذوبان ثاني أكسيد الكربون، QSPR، الذكاء الحسابي، ANN، LSSVM، SVM-PSO، DA-SVM.

Abstract

Contribution to the modeling of the physical and thermodynamic properties of pure ionic liquids and their mixtures.

Ionic liquids (ILs) constitute a new class of liquids, considered innovative green solvents of great importance from academic, industrial and environmental perspectives, due to their attractive properties. To design any process involving ILs on an industrial scale, it is necessary to know a range of physicochemical and thermodynamic properties. Therefore, its modeling is extremely important. The objective of this work is to develop QPSR models which contribute to the modeling of two properties of ionic liquids namely the surface tension of binary mixtures and the solubility of CO_2 in ILs, using algorithms of machine learning.

The first contribution focused on the prediction of the voltage of 62 binary mixtures (IL +Water/different organic solvents) for a base of 1623 experimental data points collected from the literature by the application of three computational intelligence techniques, predictive artificial neural network (ANN), least squares support vector machine (LSSVM), and particle swarm optimization-powered support vector machine (SVM-PSO). Overall, the proposed models allowed to correlate accurately the experimental data. However, the prediction by ANN was better than the other two approaches. The correlation coefficients R and deviations expressed as average absolute relative deviation ERAM for the neural network model are estimated to be 0.9979 and 0.8466%, respectively.

The second contribution focused on the prediction of CO_2 solubility in 124 ILs for a base of 10,116 experimental data using predictive artificial neural network (ANN) and optimized Support Vector Machine by the dragonfly algorithm (DA-SVM). The results obtained showed that both models can reliably predict the solubility of CO_2 in IL with a slight superiority of the ANN model with an overall correlation coefficient R and an average absolute error EAM of 0.9906, and 0.0199, respectively.

Keywords: Ionic liquids, surface tension, CO2 solubility, QSPR, computational intelligence, RNA, LSSVM SVM-PSO, DA-SVM.

Résumé

Contribution à la modélisation des propriétés physique et thermodynamiques des liquides ioniques purs et leurs mélanges.

Les liquides ioniques (LIs) constituent une nouvelle classe de liquides considérés comme des solvants verts innovants revêtant une grande importance du point de vue académique, industriel et environnemental, en raison de leurs propriétés attrayantes. Pour concevoir tout processus impliquant des LIs à l'échelle industrielle, il est nécessaire de connaître une gamme de propriétés physico-chimiques et thermodynamiques. Par conséquent, sa modélisation est extrêmement importante. L'objectif de ce travail était de développer des modèles QPSR contribuant à la modélisation de deux propriétés des liquides ioniques à savoir la tension superficielle de mélanges binaires et la solubilité du CO_2 dans les LIs à l'aide des algorithmes d'apprentissage automatique.

La première contribution a porté sur la prédiction de la tension de 62 mélanges binaires (LI+Eau /différents solvants organiques) pour une base de 1623 points de données expérimentales collectés à partir de la littérature par l'application des trois techniques d'intelligences computationnelles, le réseau neuronal artificiel prédictif (RNA), les machine à vecteurs de support des moindres carrés (LSSVM) et la machine à vecteurs de support optimisée par optimisation par essaim de particules (SVM-PSO). Globalement, les modèles proposés ont permis de corréler avec précision les données expérimentales. Cependant, la prédiction par le RNA était meilleure que les deux autres approches. Les coefficients de corrélation R et les écarts exprimés en écart relatif absolu moyen ERAM% pour le modèle de réseau neuronal sont estimés à 0.9979 et 0.8466%, respectivement.

La deuxième contribution a porté sur la prédiction de la solubilité du CO_2 dans 124 LIs pour une base de 10 116 données expérimentales à l'aide de réseau neuronal artificiel prédictif (RNA) et Machine à vecteurs de support optimisée par l'algorithme de libellule (DA-SVM). Les résultats obtenus ont montré que les deux modèles peuvent prédire de manière fiable la solubilité du CO2 dans les LI avec une légère supériorité du modèle RNA avec un coefficient de corrélation global R et un erreur absolue moyenne EAM de 0,9906, et 0,0199, respectivement.

Mots clés : Liquides ioniques, tension superficielle, la solubilité du CO2, QSPR , intelligence computationnelle, RNA, LSSVM,SVM-PSO, DA-SVM.

Table des matières

Table des matières …..…… i

Liste des figures…….…….....…...v

Liste des tableaux……...……..…..viii

Abréviations ……..……...…....ix

Nomenclature et symbole……..…….....xii

Introduction générale….……...……..……1

Chapitre I : Revue critique sur la prédiction des propriétés des liquides ioniques

I.1 Introduction ……...5

I.2 Revue critique sur le coefficient d'activité à dilution infinie des LIs……9

I.3 Revue critique sur la densité des LIs……...……..16

I.4 Revue critique sur la viscosité des LIs……..21

I.5 Revue critique sur la conductivité thermique des LIs……...30

I.6 Revue critique sur la tension superficielle des LIs……36

I.7 Revue critique sur la solubilité de CO_2 dans les LIs……...……..47

I.8 Conclusion ……72

Chapitre II : Méthodes d'Intelligence Computationnelle

II.1 Introduction ……...73

II.2 Réseaux de neurones artificiels (RNA)……73

II.2.1 Historique……74

II.2.2 Définition……74

II.2.3 Du neurone biologique au neurone artificiel...……75

II.2.4 Concept du réseau de neurones de type feedforward……..77

II.2.5 Fonctions d'activation…….79

II.2.6 Sélection d'un réseau de neurones artificiel……...80

II.2.6.1 Sélection du type de réseau……...80

II.2.6.2 Sélection de la structure du réseau…….81

II.2.6.3 Sélection de la fonction d'activation…….…82

II.2.7 Avantages des réseaux de neurones……83

II.3 Machines à vecteurs de support (Support Vector Machines) (SVM)…….83

II.3.1 C'est quoi un support à vecteur machine ?.....84

II.3.2 SVM : Principe de fonctionnement général……84

II.3.2.1 Notions de base : hyperplan, marge et support vecteur……84

II.3.2.2 Linéarité et non linéarité……....87

II.3.3 SVM pour la régression……....87

II.3.3.1 Fonction de perte en bande ε et insensible à ε……....87

II.3.3.2 ε-SVR linéaire……..88

II.3.3.3 ε-SVR basé sur le noyau (Kernel)……90

II.3.4 Algorithmes d'optimisation du SVM……...…92

II.3.4.1 l'optimisation par essaim de particules (PSO) ……....…92

II.3.4.2 L'algorithme de libellule (DA)…….93

II.4 Les machines à vecteurs avec support des moindres carrés (LSSVM)……..94

II.4.1 Introduction…….94

II.4.2 Définition……95

II.5 Critères statistiques……98

II.6 Conclusion……102

Chapitre III : Résultats et interprétations

III.1 Introduction……103

III.2 Contribution à la modélisation de la tension superficielle de mélanges binaires (LI+ Eau/Solvants organiques) par les méthodes de RNA, LSSVM, SVM-PSO……...103

III.2.1 Elaboration de la base de données……...103

III.2.1.1 Données expérimentales de la tension superficielle……..103

III.2.1.2 Données des descripteurs de profil sigma……104

III.2.2 La modélisation de la tension superficielle par les méthodes d'intelligence computationnelle……106

III.2.2.1 Choix des variables d'entrée et de sortie..…….106

III.2.2.2 Division des données en ensemble d'apprentissage, test et validation…….107

III.2.2.3 Critères statistiques……...107

III.2.2.4 Modélisation par RNA……..……107

III.2.2.5 Modélisation par SVM-PSO……108

III.2.2.5 Modélisation par LSSVM……109

III.2.3 Analyse de sensibilité des paramètres d'entrée……118

III.2.4 Domaine d'applicabilité……...120

III.2.5 Comparaison avec d'autres modèles de la littérature……...122

III.3 Contribution à la modélisation de la solubilité de CO2 dans les liquides ioniques par les méthodes de RNA et DA-SVM……126

III.3.1 Elaboration de la base de données……...126

III.3.1.1 Données expérimentales de la solubilité……...126

III.3.1.2 Les descripteurs QSPR……..126

III.3.1.2.1 Obtention des descripteurs QSPR……..127

III.3.1.2.2 Sélection de descripteurs pertinents……...128

III.3.2 La modélisation de la solubilité de CO2 dans LIs par les méthodes d'intelligence computationnelle……129

III.3.2.1 Choix des variables d'entrée et de sortie..…….129

III.3.2.2 Division des données en ensemble d'apprentissage, test et validation…….129

III.3.2.3 Critères statistiques……...129

III.3.2.4 Modélisation par RNA……..129

III.3.2.5 Modélisation par DA-SVM……...134

III.3.3 Analyse de sensibilité des paramètres d'entrée……139

III.3.4 Domaine d'applicabilité……...140

III.3.5 Validation externe……....142

III.3.6 Comparaison avec d'autres modèles de la littérature……...143

III.4 Conclusion……147

Conclusion générale et perspectives……148

Références bibliographiques……151

Liste des figures

Figure	Titre de figure	Page
Introduction générale		
Figure IG.1	Nombre de publications sur les liquides ioniques sur la période 1996-2020	1
Figure IG.2	Type de catégories versus nombre de publications	2
Chapitre I		
Figure I.1	Exemples de cations et d'anions courants trouvés dans les liquides ioniques, et leurs abréviations préférées	5
Figure I.2	Illustration de quelques applications bien connues des liquides ioniques.	8
Chapitre II		
Figure II.1	Neurone biologique.	75
Figure II. 2	Schéma d'un neurone artificiel.	76
Figure II.3	Mise en correspondance neurone biologique /neurone artificiel.	77
Figure II.4	Structure de Réseau de neurones artificiels de type feedforward multicouches.	77
Figure II.5	Principe des techniques SVM.	84
Figure II.6	Exemple d'un hyperplan séparateur.	85
Figure II.7	Exemple de vecteurs de support.	85
Figure II.8	Exemple de marge maximal (hyperplan valide).	86
Figure II.9	**(a)** L'Hyperplan avec faible marge et (b) le Meilleur hyperplan séparateur	86
Figure II.10	**(a)** cas linéairement séparable et (**b)** cas non linéairement séparable.	87
Figure II.11	Tracé (a) : l'expression graphique de la bande ε de la fonction univariée avec le ε prédéfini. Tracé (b) : comment transformer le problème de régression en problème de classification.	88
Figure II.12	Illustration de la fonction de perte insensible à ε.	89
Figure II.13	Organigramme des étapes pour la création du modèle SVM	91
Figure II.14	Démonstration schématique de l'algorithme LSSVM formé.	97
Chapitre III		
Figure III.1	Criblage de la densité de charge (sigma) et structures des conformères minimaux de 1-butanol, [C2mim] et [NTf2].	105
Figure III.2	Profils sigma des conformères minimum de 1-butanol, [C2mim] et [NTf2].	105
Figure III.3	Descripteur moléculaire basé sur les Profils σ.	106
Figure III.4	L'architecture de RNA	108
Figure III.5	Organigramme du processus d'optimisation par SVM-PSO.	109
Figure III.6	Diagramme de dispersion de la tension superficielle calcule en fonction de la tension superficielle expérimental pour la base globale par les trois topologies du modèle RNA, le modèle LSSVM et le modèle SVM-PSO.	113-114
Figure III.7	Diagramme de dispersion de la tension superficielle calculé	114-115

	en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test par le meilleur modèle de RNA « trainbr, logsig ».	
Figure III.8	Histogramme représentant la différence en termes d'erreur (EQM et ERAM%) entre les modèles étudie.	115
Figure III.9	L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainbr, logsig ».	116
Figure III.10	Diagrammes de distribution d'ER[%] du modèle RNA « trainbr, logsig ».	116
Figure III.11	Les calculs d'interpolation et d'extrapolation de la composition pour les isothermes de variation de tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Diméthylsulfoxyde pour le modèle RNA « trainbr, logsig ».	117
Figure III.12	Les calculs d'interpolation et d'extrapolation de la température pour les isothermes de variation de tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Diméthylsulfoxyde pour le modèle RNA « trainbr, logsig ».	117
Figure III.13	L'importance relative des variables par la méthode Garson du modèle RNA «trainbr, logsig ».	119
Figure III.14	Les Résultats du facteur de pertinence effectué sur le modèle RNA «trainbr, logsig ».	119
Figure III.15	Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.	121
Figure III.16	Comparaison entre les valeurs expérimentales (lignes) et calculées de la tension superficielle par notre modèle (points rouges) et le modèle Valderrama (carrés bleus) pour certains systèmes.	124
Figure III.17	Comparaison entre les valeurs expérimentales (lignes) et calculées de la tension superficielle par notre modèle (points rouges) et par les modèles de Huang et al. (triangle vert) pour son modèle RNA et (étoiles bleues) pour son modèle SM pour certains systèmes.	124
Figure III.18	Comparaison entre les données basées sur les prévisions et les données expérimentales par rapport aux données d'entrée : (a) composition de LI, (b) Température, (c) S_σ1C, (d) S_σ2C, (e) S_σ1A, (f) S_σ2A, (g) S_σ1S, (h) S_σ2S.	125-126
Figure III.19	Diagramme de dispersion de la solubilité prédite du CO2 en fonction de la solubilité expérimentale du CO2 : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test du modèle RNA.	132
Figure III.20	L'erreur relative de la solubilité du CO2 prédite par le modèle de RNA.	133
Figure III.21	L'histogramme de la distribution de l'erreur de prédiction du modèle RNA.	134
Figure III.22	L'organigramme du processus d'optimisation DA-SVM.	135
Figure III.23	Diagramme de dispersion de la solubilité du CO2 prédite en fonction de la solubilité du CO2 expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test	137

	du modèle DA-SVM.	
Figure III.24	L'erreur relative de la solubilité du CO2 prédite par le modèle de DA-SVM.	138
Figure III.25	L'histogramme de la distribution de l'erreur de prédiction du modèle DA-SVM.	138
Figure III.26	L'importance relative des variables d'entrée par la méthode Garson du modèle RNA.	139
Figure III.27	Les Résultats du facteur de pertinence effectué sur le modèle RNA.	140
Figure III.28	Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.	141
Figure III.29	Graphique Insubria de la solubilité du CO2 par le modèle RNA.	142
Figure III.30	Diagramme de dispersion des Valeurs prédites et expérimentales de la solubilité du CO2 de validation externe du modèle RNA.	143
Figure III.31	Comparaisons en termes de distribution d'erreurs.	144
Figure III.32	Comparaison entre les valeurs expérimentales (lignes) et calculées de la solubilité de CO2 dans LI par notre modèle (points mauve) et le modèle Han (carrés rose) pour certains systèmes.	145-146

Liste des tableaux

Tableau	Titre de tableau	Page
Chapitre I		
Tableau I.1	Comparaison des solvants organiques avec les liquides ioniques (LIs), mise à jour à partir des travaux de Plechkova et Seddon. La plage de valeurs a été analysée à partir de la base de données NIST LI Thermo à 25 °C et à pression atmosphérique	6
Chapitre II		
Tableau II.1	Fonctions de transfert usuelles, a=f(n)	79
Chapitre III		
Tableau III.1	Critères statistiques des quatre topologies de modèle ANN, du modèle SVM-PSO et du modèle LSSVM.	110
Tableau III.2	Différentes erreurs des quatre topologies de modèle ANN, le modèle SVM-PSO et le modèle LSSVM.	111
Tableau III.3	Les valeurs de AIC des quatre topologies de modèle RNA, du modèle SVM-PSO et du modèle LSSVM.	112
Tableau III.4	Comparaison entre ce travail et celui de Huang et al. en termes d'ERAM%.	122
Tableau III.5	Comparaison entre ce travail et le travail de Cardona et Valderrama en termes d'ERAM%.	123
Tableau III.6	Caractéristiques des descripteurs sélectionnés dans le modèle optimal et leurs valeurs VIF.	128
Tableau III.7	Critères statistiques du modèle RNA	130
Tableau III.8	Différentes erreurs par le modèles RNA.	131
Tableau III.9	Critères statistiques du modèle DA-SVM.	136
Tableau III.10	Différentes erreurs par le modèle DA-SVM.	136
Tableau III.11	Différentes erreurs pour l'ensemble de validation externe	143
Tableau III.12	Comparaisons des critères statistiques entre notre étude et l'étude précédent.	144

Abréviations

Abréviation	Signification
ACO	Optimisation des colonies de fourmis (Ant Colony Optimization)
AdaBoost	Boost adaptatif (Adaptive Boosting)
AIC	Critère d'information d'Akaike (Akaike Information Criterion)
ANFIS	Système adaptif a inférence neuro-flou (Adaptive Neuro-Fuzzy Inference System)
ASOG-VISCO	Solution analytique des groups- viscosities (Analytical Solution of Groups-VISCO)
BR	Régularisation bayésienne (Bayesian Regularization)
CA	Coefficients d'activité
CAMD	Conception moléculaire assistée par ordinateur (Computer-Aided Molecular Design)
CatBoost	Boosting catégoriel (Categorical Boosting)
Cat-Boost	algorithme de boosting catégoriel (Categorical Boosting Algorithm)
CMIS	Système intelligent de machine de comité (Committee Machine Intelligent System)
CO2	Dioxyde de carbone
COSMO-RS	Modèle en forme d'écran de conducteur pour de vrais solvants (Conductor Screenlike Model for Real Solvents)
COSMO-SAC	Modèle de criblage de type conducteur - Coefficient d'activité de segment (Conductor-like Screening Model - Segment Activity Coefficient)
COV	Composés Organiques Volatils
CPA	ssociation cubique plus (Cubic Plus Association)
CSA	recuit simulé couplé (Coupled Simulated Annealing)
CSGC	Contribution du groupe des États correspondants (Corresponding States Group Contribution)
DA	Algorithme libellule (Dragonfly Algorithme)
DBN	Réseaux de croyance profonde (Deep Belief Network)
DE	Évolution différentielle (differential Evolution)
DNN	Réseau neuronal profond (Deep Neural Network)
DT	Arbre de décision (Decision Tree)
EAM	Erreur absolue moyenne
ELL	Équilibres Liquide-Liquide
ELM	Machine d'apprentissage extrême (Extreme Learning Machine)

EOS	équation d'état cubique (cubic equation of state)
ePC-SAFT	Théorie statistique des fluides associés aux chaînes perturbées des électrolytes (Electrolyte perturbed-chain Statistical Associating Fluid Theory)
EQM	Erreur quadratique moyenne
EQRM	Erreur quadratique relative moyenne
ERAM	Erreur relative absolue moyenne
EVL	Equilibre -Vapeur-Liquide
EVLL	Equilibres triphasés vapeur-liquide-liquide
FFA	Algorithme de luciole (Firefly Algorithm)
FF-RNA	Réseau neuronal artificiel à action directe (Feed-Forward artificial neural network)
FOA	Algorithme d'optimisation mouche des fruits (Fruit-fly Optimization Algorithm)
FVT	Théorie du volume libre (Free Volume Theory)
GA	Algorithme génétique (Genetic Algorithm)
GAMESS	Système général de structure électronique atomique et moléculaire (General Atomic and Molecular Electronic Structure System)
GBR	Régresseur d'amplification de gradient (Gradient Boosting Regressor)
GBT	Arbre d'amélioration du dégradé (Gradient Boosting Tree)
GC	Contribution de groupe (Group Contribution)
GC-ER	Contribution du groupe Esmaeilzadeh-Roshanfekr
GC-NLF EoS	Equation d'état des fluides de réseau non aléatoires à contribution de groupe (group contribution non-random lattice fluid equation of state)
GEP	Programmation d'expression génique (Gene Expression Programming)
GMDH	Méthode de groupe de corrélation de traitement des données (Group Method of Data Handling)
GMDH-RNP	Méthode de groupe de réseau neuronal polynomial de type traitement de données (Group Method of Data Handling type Polynomial Neural Network)
GSA	Algorithme de recherche gravitationnelle (gravitational search algorithm)
H2S	Sulfure d'hydrogène
ICA	Algorithme concurrentiel impérialiste (Imperialist Competitive Algorithm)
IDAC	Coefficient d'activité de dilution infinie (Infinite Dilution Activity Coefficient)
IFC	Contribution de fragments ioniques (Ionic Fragments Contribution)
KNN	Algorithme K-voisin le plus proche (K Nearest Neighbour algorithm)
KNNR	K-Nearest-Neighbour Regression
LI	Liquide ionique

LI-Rev	Liquides Ioniques Réversibles
LIs	Liquides ioniques
LM	Algorithme de Levenberg-Marquardt (Levenberg-Marquardt algorithm)
LSSVM	Machine à vecteur de support aux moindres carrés (Least Squares Support VectorMachines)
MGGP	Programmation génétique multigénique (Multi-Gene Genetic Programming)
ML	Apprentissage automatique (Machine Learning)
MLG	Modèle Linéaire Généralisé
MLP	Modèle de perceptron multicouche (Multi-LayerPerceptron)
MTPS	Source plane transitoire modifiée (Modified Transient Plane Source Method)
MWCNT	Nanotubes de carbone à parois multiples (Multiwalled Carbon Nanotube)
NEILs	Liquides ioniques améliorés par des nanoparticules (Nanoparticle-Enhanced Ionic Liquids)
NLR	Régression non linéaire (Nonlinear Regression)
PC	Propriétés Critiques
PC-SAFT	Théorie des fluides associatifs statistiques à chaîne perturbée (The Perturbed-Chain Statistical Associating Fluid Theory)
PHSC	Équation de chaîne de sphère dure perturbée d'État (Perturbed Hard Sphere Chain Equation of State)
PR	Peng – Robinson
PSO	Optimisation par essaim de particules (Particle swarm optimization)
PSRK	l'approche prédictive-Soave-Redlich-Kwong (Predictive-Soave-Redlich-Kwon)
QM	Modèle quantique (Quantum Model)
QSAR	Relation quantitative structure-activité.
QSPR	Relation quantitative structure-propriété.
RBF	Fonction à base radiale (Radial Basis Function)
RF	Forêt aléatoire (Random Forest)
RL	Régression linéaire
RNA	Réseaux de neurones artificiels
RNG	Réseaux de neurones graphiques
RNN	Réseau neuronal récurrent (Recurrent Neural Network)
RTIL	Liquide ionique à température ambiante (Room Temperature Ionic Liquid)

SA	Recuit simulé (Simulated Annealing)
SCE	Evolution complexe mélangée (Shuffled Complex Evolution)
SCG	Gradient conjugué mis à l'échelle (Scaled Conjugate Gradient)
SPT	Théorie des particules à l'échelle modifiée (scaled particle theory)
SRK	Soave – Redlich – Kwong
SVM	Machine à vecteur de support (Support Vector Machine)
SWMLR	Régression linéaire multiple pas à pas (Stepwise Multiple Linear Regression)
TF	Théorie du frottement
TIs	Indices topologiques généraux (General Topological Indices)
TLBO	Optimisation basée sur l'enseignement-apprentissage (Teaching–Learning-Based Optimization)
tPC-PSAFT	Théorie statistique des fluides associatifs polaires à chaîne perturbée tronquée (Truncated Perturbed Chain-Polar Statistical Associating Fluid Theory)
TPN	Température et Pression Normal
TS	Tension superficielle
UNIFAC	Coefficient d'Activité Fonctionnel UNIversel (UNIversal Functional Activity Coefficient)
UNIQUAC	Universal QUAsiChemical
VaCG	Grains grossiers à variable électrostatique (Electrostatic-Variable Coarse-Grained)
XGB/ XGBoost	Boosting de gradient extrême (Extreme Gradient Boosting)

Nomenclature et symbole

Symbole	**Désignation**	**Unité**
A_f	Facteur de précision	[-]
b	Biais	[-]
B_f	Facteur de biais	[-]
C	Box constraint	[-]
K	Fonction de noyau (kernelScale)	[-]
R	Coefficient de corrélation	[-]
R^2	Coefficient de détermination	[-]
S_σ	Descripteur sur le profil sigma	e/A^2
T	Température	Kelvin
P	Pression	
w	Poids	[-]
X_{LI}	Fraction molaire	[-]
γ	Constante de régularisation	[-]
σ	Tension superficielle	mN/m
σ^2	Paramètre kernel	[-]
η	Viscosité dynamique	cp
λ	Conductivité thermique	W/mK

Introduction Générale

Introduction Générale

Au cours des dernières décennies, Les liquides ioniques (LIs) apparaissent de plus en plus comme de nouveaux solvants verts innovants revêtant une grande importance du point de vue académique, industriel et environnemental. L'intérêt croissant peut être attribué à leurs caractéristiques exceptionnelles, où ils ont été qualifiés de « solvants du futur », en raison du large éventail d'applications dans lesquelles ils pourraient être utilisés, notamment en remplacement des solvants classiques dans les processus industriels qui étaient considérés souvent toxiques, inflammables et hautement volatils. Ainsi que, leur récupération est facile par extraction ou par distillation pour leur recyclage. Cela contribue à réduire leur coût pour une utilisation améliorée dans divers processus d'importance industrielle.[1–3]

La recherche fondamentale et appliquée s'est développée de façon exponentielle et a démontré la valeur des LIs pour la science avec son large éventail d'applications potentielles dans les domaines de la chimie, de la biologie, de la physique, etc. Il n'est pas étonnant qu'actuellement, le nombre de publications sur les liquides ioniques dépasse les 100 000 (104,310 au 23 septembre 2020) et que, chaque année, environ 9 000 nouveaux rapports paraissent (Figure IG.1).[4]

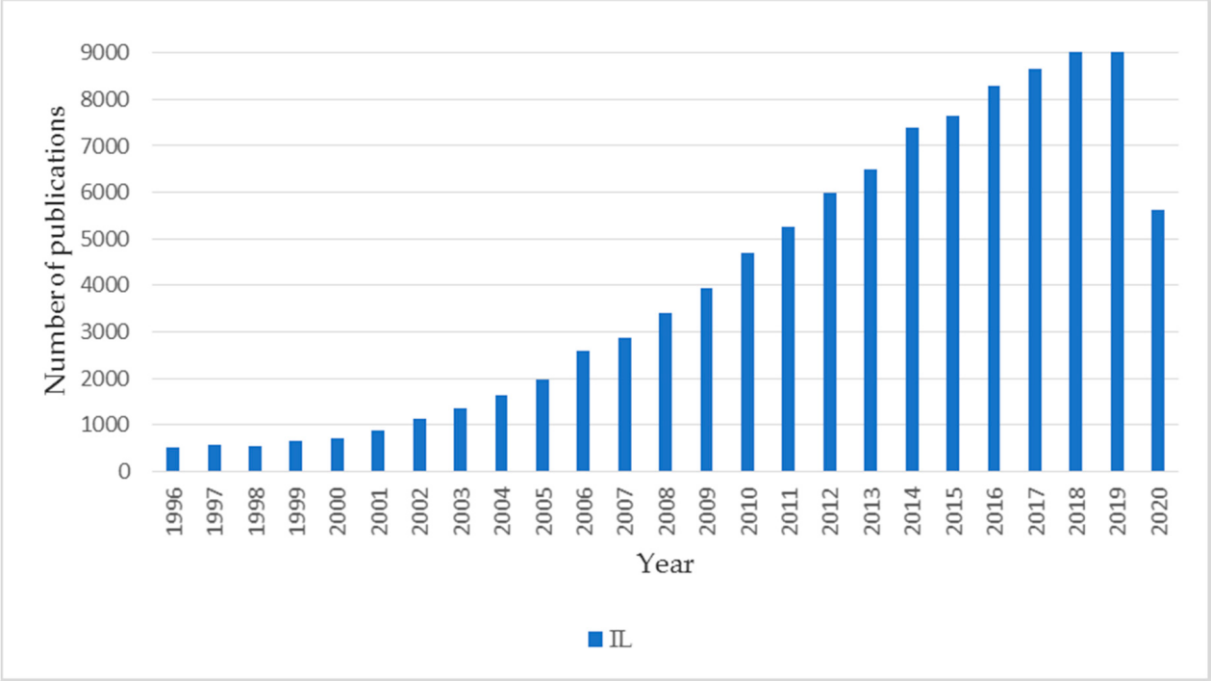


Figure IG.1: Nombre de publications sur les liquides ioniques durant la période 1996-2020.[4]

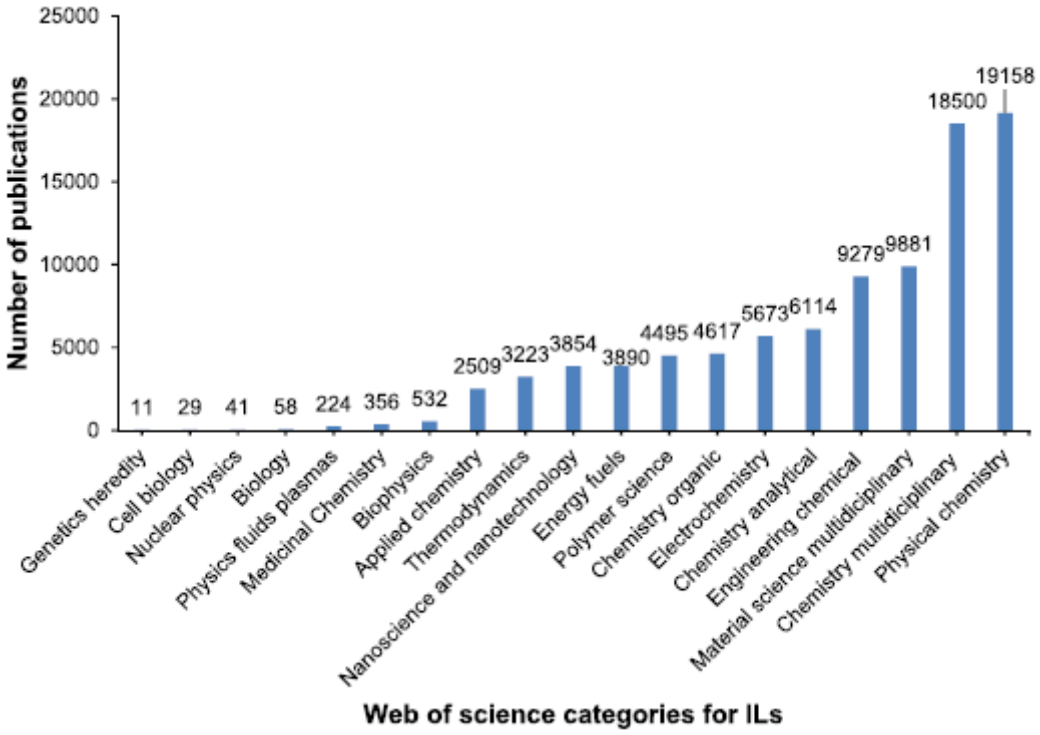


Figure IG.2: Type de catégories en fonction du nombre de publications.[1]

Pour concevoir tout processus impliquant des LIs à l'échelle industrielle, il est nécessaire de connaître une gamme de propriétés physico-chimiques et thermodynamiques de sorte que, la plupart des efforts dans le domaine des LIs se sont concentrés sur l'étude expérimentale et théorique de leurs propriétés, telles que le point de fusion, la viscosité, la densité, La solubilité des gaz, la tension superficielle, etc. Ces données sont nécessaires non seulement pour la conception des processus et des produits, mais également pour l'élaboration de modèles de corrélation et de prévision. La fiabilité des propriétés thermo physiques prévues et des équilibres de phase est un facteur clé dans l'évaluation des programmes de simulation et de conception qui sont malheureusement limitées et parfois contradictoires.[5,6]

Il a été reconnu que la tension superficielle d'un mélange liquide joue un rôle important dans le transfert de masse et de chaleur à l'interface. Il est nécessaire dans de nombreuses applications industrielles telles que l'extraction liquide-liquide, l'absorption des gaz, la distillation, la condensation, etc.[7]. De plus, la tension superficielle révèle des informations précieuses sur la structure et l'énergétique de la région superficielle entre deux phases. La variation de l'énergie de surface dépend de la variation des forces moléculaires et de la densité de tassement ou de la taille moléculaire. Pour cette raison, la détermination de cette propriété pour les mélanges liquides nous permet d'étudier les interactions entre les composants.

En outre, La solubilité des gaz est parmi les propriétés importante des LIs dans le contexte de la séparation des gaz, de l'optimisation des réactions, du captage des gaz à effet de serre, comme antisolvant, du stockage de gaz et de l'absorption des gaz réfrigérants.[8]En particulier, l'étude de la solubilité du dioxyde de carbone dans les liquides ioniques attire l'attention de nombreux chercheurs. Pour surmonter le problème des émissions de CO_2 et réduire leur taux dans l'atmosphère, ce qui est l'un des principaux gaz à effet de serre et qui affecte au changement climatique et au réchauffement climatique.[9]

Cependant, La détermination précise des propriétés thermophysiques des LIs ne peut être réalisée exclusivement par des mesures expérimentales en raison du grand nombre de LIs possibles, où on estime qu'environ (10^{18}) liquides ioniques (LIs) existent car il existe une combinaison illimitée de cations et d'anions pour former des LIs. De plus, la plupart des applications industrielles des LIs ont eu lieu dans des mélanges, provoquant des combinaisons illimitées de mélanges dont les propriétés thermophysiques peuvent être intéressantes. Et comme les mesures expérimentales sont la plupart du temps difficiles, prennent beaucoup du temps et coûteuses avec des incertitudes probables non négligeables, ces mesures ne sont pas toujours possibles. Dans de tels cas, l'élaboration de modèles et de corrélations précis pour la prédiction des propriétés des Lis dans diverses conditions pourrait être considérée comme une alternative appropriée aux mesures expérimentales.[10,11]

L'utilisation des méthodes alternatives et complémentaires à l'expérimentation, telles que les relations quantitatives structure propriété/activité (QSPR/QSAR) sont devenues d'un grand intérêt et sont même recommandées dans les nouvelles règlementations afin d'obtenir les données nécessaires à l'enregistrement des substances.[12]. Le développement des modèles mathématiques QSPR/QSAR liant propriétés physico-chimiques et les activités biologiques à la structure moléculaire permet d'expliquer l'origine de ces activités/propriétés et de les prédire pour des molécules dont les données expérimentales ne sont pas disponibles. [13]

La modélisation des propriétés des liquides ioniques repose principalement sur l'utilisation des équations d'état, mais en raison des limites inhérentes de ces équations d'état, d'autres méthodes de modélisation peuvent être envisagées. Parmi ces alternatives, on cite les corrélations empiriques qui sont simples est très souvent efficaces et les méthodes de l'intelligence computationnelle tels les réseaux de neurones artificiels (RNA), les machines à vecteur de support (SVM) et Machine à vecteurs de support pour les moindres carrés

(LSSVM) sont des méthodes prédictives robustes et précises qui ont récemment été appliquées avec succès pour la prédiction de diverses propriétés des LIs dans des états purs et en mélange.[14]

L'objectif de cette thèse est de développer des modèles QSPR contribuant à la modélisation de deux propriétés des liquides ioniques à savoir la tension superficielle de mélanges binaires (LI-Eau et divers solvants organiques) et la solubilité du CO_2 dans les LIs, à l'aide des algorithmes d'apprentissages automatiques. Le présent travail, verbalisant l'étude, est divisé en trois chapitres.

Le premier chapitre est consacré à une revue critique de la littérature relative à la prédiction des propriétés physicochimiques et thermodynamique des liquides ioniques.

Le deuxième chapitre s'attache à décrire l'apport et le développement des méthodes d'intelligences artificielles basées sur l'apprentissage de la machine pour la modélisation de phénomènes complexes non linéaires. Les méthodes décrites dans ce chapitre sont les réseaux de neurones artificiels, les machines à vecteurs de support des moindres carrés et les machines à vecteurs de support, ainsi que des algorithmes optimisations, tels que l'algorithme libellule (DA) et l'optimisation par essaim de particules (PSO).

Le troisième chapitre est subdivisé en deux parties. La première partie présente les résultats de contributions réalisées par les trois techniques d'intelligences computationnelles, le réseau neuronal artificiel prédictif (RNA), les machines à vecteurs de support des moindres carrées (LSSVM) et la machine à vecteurs de support optimisé par optimisation par essaim de particules (SVM-PSO), pour l'estimation de la tension superficielle des mélanges binaires (LI-Eau et divers solvants organiques), suivi par une étude comparative entre les modèles proposés et d'autres modèles de la littérature. En plus, une étude de sensibilité des paramètres d'entrées et de domaine d'application pour le meilleur modèle proposé est élaborée. Dans la deuxième partie, les résultats de contributions réalisées pour la modélisation de la solubilité du CO_2 dans les LIs à l'aide du réseau neuronal artificiel prédictif (RNA) et machine à vecteurs de support optimisée par l'algorithme de libellule (DA-SVM) sont présentés, ainsi qu'une comparaison entre les modèles proposés et d'autres modèles de la littérature, et une étude de la sensibilité des paramètres d'entrées et du domaine d'application. Actuellement, on procède à la publication de cette modélisation.

L’étude se termine par une conclusion générale tirée de ce travail ainsi que les études qui peuvent être envisagées dans le futur.

Chapitre I : Revue critique sur la prédiction des propriétés des liquides ioniques

Chapitre I : Revue critique sur la prédiction des propriétés des liquides ioniques

I.1 Introduction

Les liquides ioniques (LIs), composés uniquement de cations organiques asymétriques volumineux et d'anions organiques ou inorganiques, sont un type de nouveaux sels qui maintiennent un état liquide stable à ou près de la température ambiante (<100 °C). Les LIs se différencient des autres sels, tels que le chlorure de sodium, qui ne fondent qu'à des températures extrêmement élevées (> 800 °C) en raison de leurs structures serrées et de leurs fortes interactions ioniques. Le point de fusion le plus bas d'un LI est attribué à la structure volumineuse et asymétrique des ions constituants, qui peuvent être associés dans différentes combinaisons pour adapter leurs propriétés thermophysiques. Quelques exemples de cations et d'anions courants sont présentés sur la figure I.1, ainsi que leurs abréviations préférées.[3,11]

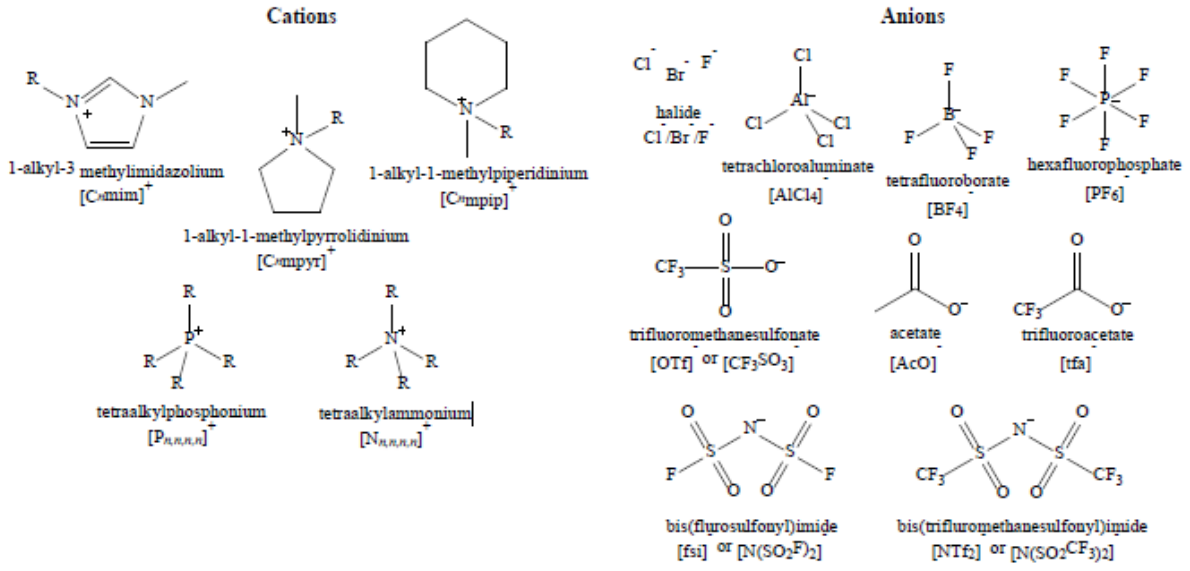


Figure I.1 : Exemples de cations et d'anions courants trouvés dans les liquides ioniques, et leurs abréviations préférées.[3]

Les liquides ioniques (LIs) ont fait l'objet d'une attention accrue grâce à leurs propriétés physiques et chimiques uniques, telles que l'ininflammabilité, la non volatilité, une faible pression de vapeur, une excellente stabilité thermique et chimique, une stabilité électrochimique élevée, une grande capacité de solubilité des substances polaires et non polaires, etc. Ce sont des propriétés intéressantes qui suggèrent que les liquides ioniques pourraient être considérés comme de meilleurs solvants dans des processus plus sûrs et plus

respectueux de l'environnement. Il faut ajouter que les propriétés des liquides ioniques peuvent être ajustées dans une large mesure par la combinaison judicieuse de cations et d'anions et l'adaptation de leurs structures chimiques (par exemple par modification du nombre et/ou de la longueur des substituants alkyles). Cette adaptabilité, en conjonction avec l'ensemble de propriétés mentionnées ci-dessus, a conduit à la création du terme solvant de conception pour souligner la possibilité de « concevoir » un liquide ionique pour qu'il corresponde aux caractéristiques requises par une application spécifique.[14,15]

Cependant, avant de pouvoir substituer tout solvant organique utilisé dans un procédé donné, la compatibilité d'un LI sélectionné doit être vérifiée par rapport à toutes les unités opératoires d'un procédé. Une comparaison des principales propriétés des solvants organiques par rapport à celles des LIs initialement rapportées par Plechkova et Seddon[16]en 2008 a été mise à jour ici, comme indiqué dans le tableau 1, pour mettre en évidence les principales différences entre ces deux classes de matériaux.[3]

Tableau I.1 : Comparaison des solvants organiques avec les liquides ioniques (LIs), mise à jour à partir des travaux de Plechkova et Seddon. La plage de valeurs a été analysée à partir de la base de données NIST LI Thermo à 25 °C et à pression atmosphérique.[3]

Propriétés	**Solvants organiques**	**Liquides ioniques**
Nombre de solvants	>1000	$>10^6$
Applicabilité dans un processus donné	Fonction unique	Multifonction
Coût	Généralement bon marché	2 à 100 fois plus cher que les solvants organiques
Recyclabilité/Toxicité	Impératif vert : l'enquête sur la toxicité des solvants organiques est contrôlée par REACH	Impératif économique : la toxicité et la biodégradabilité sont souvent mal connues
Pression de vapeur	Mesurable et généralement bien connu – plusieurs solvants ont une pression de vapeur > limite utilisée dans la classification des composés organiques volatils (COV)	Pour les LIs aprotiques : pression de vapeur négligeable dans des conditions normales
Inflammabilité	Généralement inflammable	Généralement ininflammable, mais certains LIs sont utilisés comme propulseurs

Accordabilité	Gamme limitée de solvants disponibles	Une portée pratiquement illimitée signifie
Chiralité	Rare	Commun et réglable
Capacité catalytique	Rare	Commun et réglable
Viscosité/mPa.s	0.2-100	20-97,000
Densité/g.cm-3	0.6-1.7	0.8-3.3
Indice de réfraction	1.3-1.6	1.3-2.2
Conductivité électrique/mS.cm-1	Généralement isolant	Jusqu'à 120
Conductivité thermique/W.m-1.K-1	0.1-0.6	0.1-0.3

À la lumière d'une telle comparaison, Les caractéristiques des liquides ioniques les rendent attrayants non seulement pour leur utilisation comme solvants, mais également pour des rôles alternatifs dans une large gamme d'applications variées, comme indiqué sur la figure I.2. Ils sont utilisé comme électrolytes pour le développement de batteries beaucoup plus efficaces et plus sûres, comme solvants dans divers domaines tels que la synthèse organique, la synthèse de nanoparticules, dans divers processus biocatalytiques, dans l'administration de médicaments, dans les industries pharmaceutiques, dans l'adoucissement des gaz, ainsi que divers processus d'extraction d'importance industrielle tel que pour l'élimination des composés soufrés et des contaminants organiques provenant de pétrole brut et d'échantillon de sol et leur utilisation efficace à la place des tensioactifs conventionnels dans divers domaines tels que la récupération assistée du pétrole, les peintures, les détergents, les lubrifiants, etc.[2]

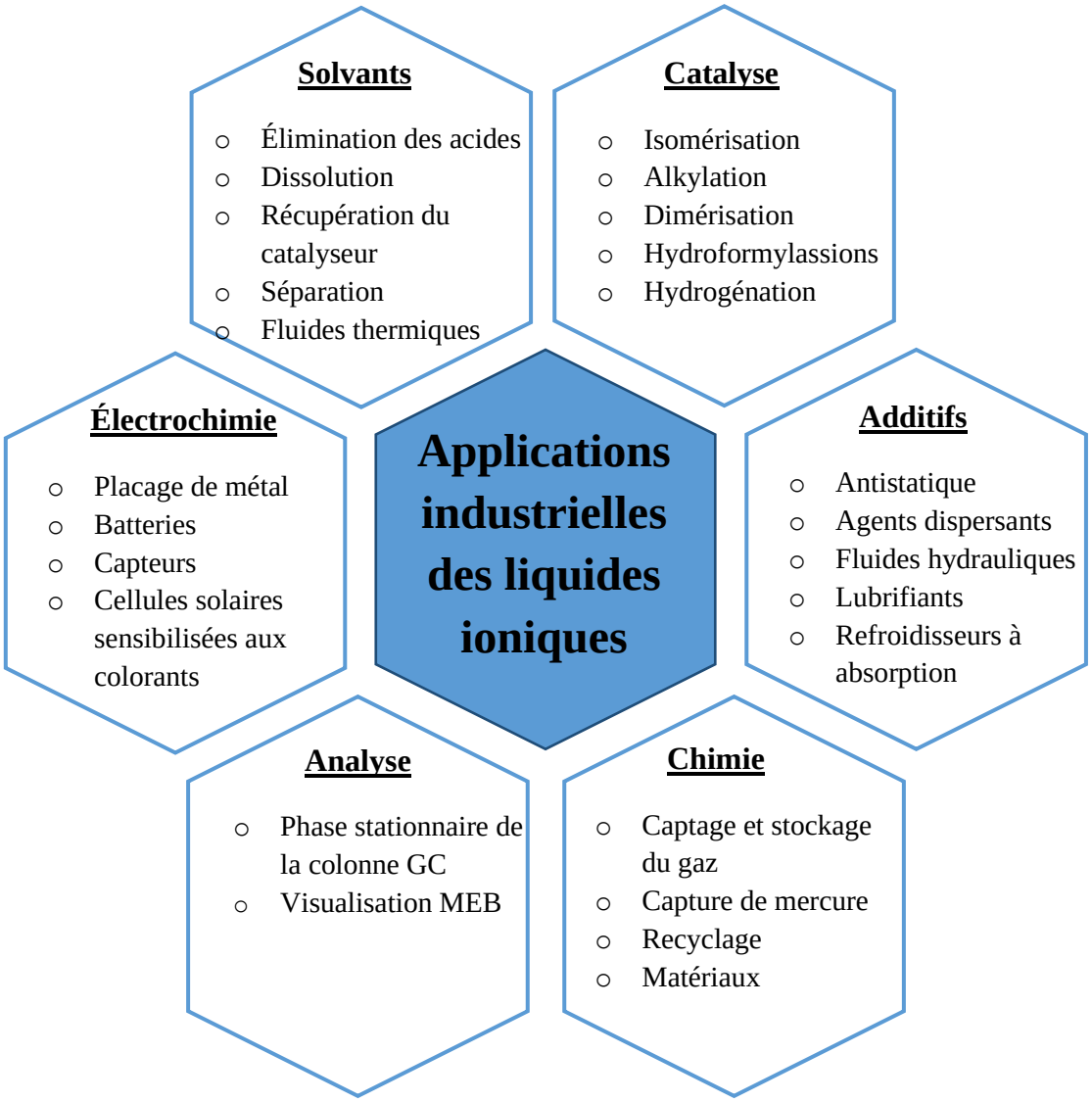


Figure I.2 : Illustration de quelques applications bien connues des liquides ioniques.[3]

En particulier, les LIs ont été utilisés dans un certain nombre de processus industriels pour améliorer la récupération du catalyseur et la séparation des produits grâce à la formation de systèmes biphasiques. Leurs propriétés de solvatation sont fortement influencées par la capacité des ions à agir comme accepteurs/donneurs de liaisons hydrogène et par le degré de délocalisation de la charge anionique. Ils possèdent également une nature non coordonnée et hautement polaire, causée par des ions faiblement coordonnés, ce qui signifie qu'il a été démontré que les Lis améliorent les vitesses de réaction et la sélectivité de nombreuses réactions. Bien qu'il y ait eu de nombreuses découvertes et brevets déposés dans le domaine des LIs, le nombre ayant atteint une application commerciale est relativement limité.[3]

Cependant, l'identification, le développement et la réalisation de conceptions de procédés fiables et rentables pour les applications industrielles des LIs et de leurs mélanges associés reposent en grande partie sur une étude approfondie de leurs propriétés physico-chimiques et thermodynamiques qui sont encore rares, telles que le point de fusion, la tension superficielle, la solubilité des gaz, la densité et la viscosité.[11]. De toute évidence, la mesure expérimentale ne peut pas prendre en compte les propriétés de chaque LI possible et son mélange avec d'autres composés car il existe un nombre infini de systèmes concevables. Par conséquent, il est important de développer des outils informatiques prédictifs pour estimer avec précision ces propriétés de divers systèmes contenant des LIs.[11,17]

Après avoir présenté les liquides ioniques, leurs caractéristiques générales et leurs applications, ce chapitre traitera une revue critique de la littérature relative aux méthodes de prédiction de quelques propriétés physicochimiques et thermodynamiques des liquides ioniques.

I.2 Revue critique sur le coefficient d'activité à dilution infinie des LIs

Le coefficient d'activité à dilution infinie est l'une des propriétés thermodynamiques les plus importantes[18], qui donne un aperçu sur les interactions moléculaires entre le soluté et le solvant ainsi que de l'étendue de la non-idéalité d'un soluté dans un mélange. Il donne des informations clés pour la compréhension de nombreux processus de séparation qui utilisent des liquides ioniques, et il fournit également des informations utiles sur les interactions intermoléculaires soluté-liquide ionique. Ces informations sont utiles pour cribler les solutés pour l'extraction et la distillation extractive et d'autres applications potentielles. Parmi les études théoriques réalisées. [19,20]

Diedenhofen et al. (2003)[21] Pour le calcul des coefficients d'activité à dilution infinie de 38 composés dans les liquides ioniques 1-méthyl-3-éthylimidazolium bis((trifluorométhyl)sulfonyl) imide, 1,2-diméthyl-3-éthylimidazolium bis((trifluorométhyl) sulfonyl)imide et le tétrafluoroborate de 4-méthyl-N-butylpyridinium pour des valeurs de temperature (314 et 344) K, ils ont utilisé la méthode COSMO-RS, une méthode générale et rapide de prédiction des données thermophysiques des liquides ioniques. Ils ont conclu que le modèle COSMO-RS a permis de prédire les coefficients d'activité à dilution infinie dans divers liquides ioniques avec la même précision que celle observée pour les solvants organiques classiques développé, sans aucun ajustement de la théorie ni utilisation de paramètres spécifiques. Ainsi, COSMO-RS et sa mise en œuvre dans le programme

COSMOtherm sont capables de donner des prédictions préalables de la thermodynamique des liquides ioniques, ce qui peut être d'une valeur considérable pour l'exploration des LIs appropriés pour des applications pratiques.

Banerjee et Khanna (2006)[22] ont validé les mesures de coefficients d'activité de dilution infinie sur les liquides ioniques de phosphonium (à base de cations trihexyl-tétradécylphosphonium (thtd-Ph) de haut poids moléculaire avec les anions chlorure [Cl], tétrafluoroborate [BF4] et bis (trifluorométhanesulfonylimide) [(Tf)2N]) pour la séparation des composés aliphatiques et aromatiques dans la gamme de températures (308,15, 318,15 et 328,15) K à l'aide de la méthode a priori COSMO-RS. La prédiction COSMO-RS a d'abord été étalonnée sur les valeurs obtenues avec [thtd-Ph] tris(pentafluoroéthyl) trifluorophosphate, ce qui a donné un écart absolu moyen (EAM) de 15 %. En utilisant COSMO-RS, les valeurs coefficients d'activité ont été obtenues pour les trois LIs de phosphonium, et les EAM correspondants étaient de 9 %, 8 % et 16 %. Les schémas prédictifs de COSMO-RS ont confirmé les résultats expérimentaux que Les LIs à base de phosphonium sont de mauvais solvants pour les procédés de séparation des composés aliphatiques et aromatiques, comme dans le cas de la séparation benzène-hexane, benzène-hexane, benzène-alcool et benzène-cyclohexane. Cependant l'LI [thtd-Ph][BF4] a une sélectivité modérée pour les mélanges benzène-alcool se qui prouve que les LIs de phosphonium sont de nature non polaire.

Lei et al (2006)[23] ont étudié la séparation du 1-hexène et du n-hexane comme représentation des oléfines et des paraffines avec des liquides ioniques, ainsi que de la N-méthyl-2-pyrrolidone criblée par conception moléculaire assistée par ordinateur (CAMD). Sur la base de l'analyse de conformation des solvants et des liquides ioniques, le modèle de criblage de type conducteur pour les solvants réels (COSMO-RS) a été utilisé pour faire une prédiction préalable pour les liquides ioniques appropriés. Ils ont constaté que les liquides ioniques appropriés doivent avoir un petit volume moléculaire, un groupe non ramifié et un effet de protection stérique autour du centre de charge de l'anion. Ces résultats sont prouvés à partir un travail expérimental à dilution finie qui a été fait pour vérifier la fiabilité des résultats calculés à dilution infinie par le modèle COSMO-RS, où les résultats expérimentaux sont qualitativement cohérents avec la conclusion du modèle COSMO-RS. Cela démontre que le modèle COSMORS peut être utilisé comme un outil pour adapter le liquide ionique approprié afin de réduire la quantité de travail expérimental.

Nami et Deyhimi (2010)[24] dans leur travail ont illustré pour la première fois la capacité des réseaux de neurones artificiels (RNA) à prédire les coefficients d'activité à dilution infinie pour les solutés organiques dans les liquides ioniques (LIs). À l'aide d'un réseau d'anticipation multicouche avec algorithme d'optimisation de Levenberg – Marquardt, le modèle RNA résultant a généré un coefficient d'activité à des données de dilution infinie sur une plage de températures de 298 à 363 K. Les données d'entrée non disponibles concernant la douceur des composés organiques (solutés) et le moment dipolaire des liquides ioniques ont été calculés à l'aide des suites GAMESS de programmes de chimie quantique. Le modèle RNA résultant et sa validation sont basés sur l'étude de jusqu'à 24 composés organiques structurellement différents (alcanes, alcènes, alcynes, cycloalcanes, aromatiques et alcools) dans 16 liquides ioniques communs à base d'imidazolium, à différentes températures dans la plage de 298 à 363 K (c'est-à-dire un nombre total de 914 c1 Soluté pour chaque point de données LIs). Les résultats ont montré un accord satisfaisant entre l'RNA prédit et les données expérimentales, où l'erreur quadratique moyenne (EQRM) et le coefficient de détermination (R^2) du réseau neuronal conçus se sont avérés être de 0,103, 0,996 pour les données d'entraînement et de 0,128, 0,994 pour tester les données, respectivement.

Xi et al (2010)[19] ont proposé une nouvelle méthode pour prédire les coefficients d'activité de dilution infinie des LIs à différentes températures. Un modèle de relation quantitative structure-propriété (QSPR) dépendant de la température a été développé pour une série de solutés organiques dans le trihexyl(tétradécyl) phosphonium bis(trifluorométhylsulfonyl) imide liquide ionique. En utilisant des algorithmes génétiques de sélection de sous-ensembles de variables et des méthodes de régression des moindres carrés ordinaires, six variables, dont la température et cinq descripteurs moléculaires significatifs, ont été sélectionnés et utilisés pour construire le modèle de prédiction dépendant de la température. Ils ont trouvé des résultats satisfaisants des validations internes par les approches laisser-un-dehors, bootstrap et brouillage en Y et externes par un ensemble de tests externes ont prouvé la fiabilité, la stabilité et la capacité prédictive du modèle construit.

Anantharaj et Banerjee (2010)[25] ont utilisé les liquides ioniques comme solvants verts pour éliminer ces composés par extraction liquide-liquide (ELL). Environ 168 LIs comprenant des cations comprenant le 1-éthyl-3-méthylimidaozlium [EMIM], le 1-éthylpyridinium [EPY], le 1-éthyl-1-méthylpyrrolidinium [EPYRO], le 1-éthyl-1-méthylpipéridinium [EMPIP], le 4- l'éthyl-4-méthyl morpholinium [EMMOR] et le 1,2,4-triméthylpyrazolium [TMPYZO] combinés à 26 anions ont été étudiés dans ce travail. Le

coefficient d'activité à dilution infinie (IDAC) a été prédit à l'aide du modèle de criblage de type conducteur pour les solvants réels (COSMO-RS) afin de cribler les solvants potentiels. Initialement, le modèle a été étalonné via les prédictions IDAC et équilibres liquide-liquide (ELL). ELL a été prédit pour quatre systèmes ternaires rapportés dans lesquels un hétérocycle azoté était l'un des composés. L'écart moyen quadratique moyen obtenu était de 10 %. Les valeurs IDAC ont été prédites pour la pyridine dans deux liquides ioniques, à savoir [BMIM][BF4] et [EMIM][TOS], avec une erreur quadratique moyenne de 8 %. Ensuite, pour évaluer les performances ils ont calculé la sélectivité, la capacité et l'indice de performance à dilution infinie. Ils ont constaté que les espèces azotées à cinq chaînons ayant une densité électronique délocalisée élevée possédaient une sélectivité supérieure de 3 ordres de grandeur à celle des espèces azotées à six chaînons. Pour les structures cycliques à cinq chaînons, la sélectivité s'est avérée suivre l'ordre [EPYRO] > [EMPIP] > [EPY] > [EMMOR] > [EMIM] > [TMPYZO]. Pour l'hétérocycle à six chaînons, il suivait l'ordre [EPY] > [EMMOR] > [EPYRO] > [EMPIP] > [EMIM] > [TMPYZO]. Indépendamment de l'hétérocycle azoté, les anions tels que le thiocyanate [SCN] et l'acétate [Ac] ont donné des valeurs élevées de sélectivité. En général, les cations sans cycles aromatiques tels que [EPYRO], [EMPIP] et [EMMOR] ont donné une sélectivité et une capacité plus élevées quelque soit l'hétérocycle azoté.

Passos et al(2014)[26] ont utilisé la théorie des fluides associatifs statistiques à chaîne perturbée (PC-SAFT) pour décrire les données expérimentales qui ils ont obtenues. Les LIs ont été traités comme des espèces associatives moléculaires avec deux sites d'association par LIs. Les paramètres du modèle pour les fluides purs et le paramètre d'interaction binaire kij entre l'eau et les LIs ont été déterminés par un ajustement simultané aux densités des LIs pur, aux coefficients d'activité de l'eau à 298,15 K et aux données d'EVL à 0,1 MPa. Les paramètres PC-SAFT ont été appliqués pour prédire les coefficients d'activité de l'eau à dilution infinie en LIs. Ils ont constaté que PC-SAFT est capable de prédire de manière satisfaisante les coefficients d'activité de l'eau à dilution infinie pour les systèmes où les paramètres du modèle ont été ajustés aux coefficients d'activité de l'eau à de faibles molalités LIs

Gonfa et al (2016)[18] ont développé des modèles de relation quantitative structure-propriété/activité (QSPR/QSAR) pour la prédiction du coefficient d'activité de l'eau à dilution infinie à différentes températures, utile pour la sélection préliminaire des liquides ioniques appropriés pour la déshydratation du gaz naturel. Pour la construction de ces modèles, ils ont

utilisé des descripteurs moléculaires basés sur le profil sigma COSMO-RS et des régressions linéaires multiples. Deux approches ont été appliquées pour développer les relations QSPR/QSAR ; y compris la température comme variable indépendante (dans le modèle I) et la génération de QSPR/QSAR pour chaque température sans le terme de température (modèle II). La deuxième approche (modèle II) fournit un modèle légèrement plus précis que le modèle I. Cela pourrait être dû au fait que l'effet de la température sur le coefficient d'activité de l'eau dépend de la nature des LIs. La fiabilité des modèles a été confirmée par une analyse statistique des ensembles de données d'entraînement et de test, ce qui indique que les modèles développés peuvent être utilisés pour la prédiction des coefficients d'activité dans les LIs. Les modèles donnent également un aperçu clair de l'effet des variations structurelles des cations et des anions sur les valeurs du coefficient d'activité et peuvent aider à concevoir des LIs pour la déshydratation du gaz naturel.

Benimam et al.(2019)[27] pour modéliser le coefficient d'activité à dilution infinie de l'eau dans 53 liquides ioniques, ils ont appliqué Trois modèles basés sur les réseaux de neurones artificiels, de la machine à vecteurs de support et de la machine à vecteurs de support des moindres carrés en utilisant six variables d'entrées indépendantes à savoir la température du système eau/liquide ionique, la température critique, la pression critique, le volume critique, le facteur acentrique et le poids moléculaire du liquide ionique, tandis que le coefficient d'activité à dilution infinie de l'eau dans les liquides ioniques est défini comme la variable dépendante de la sortie. Ils ont constaté que dans l'ensemble, les modèles proposés sont capables de corréler avec précision 318 points de données expérimentales recueillies de la littérature. Selon les résultats, le RNA est une machine d'apprentissage informatique plus puissante et plus efficace que les deux autres. Les coefficients de corrélation et les écarts exprimés en écart relatif absolu moyen pour le modèle de réseau de neurones sont estimés à 0,99997 et 0,56 %, respectivement. En outre, ils ont démontré les capacités d'interpolation et d'extrapolation du modèle de réseau neuronal et ils ont comparé sa précision avec d'autres modèles proposés dans la littérature basés sur la régression multilinéaire, les moindres carrés prennent en charge la machine vectorielle et un autre réseau neuronal à anticipation.

Thangarajoo et al.(2020)[10] ont appliqué le modèle de van't Hoff via la méthode GC pour développer un modèle permettant de prédire le coefficient d'activité à dilution infinie de l'éthanol et du propanol dans plusieurs types de LIs. Le modèle développé utilisant la régression linéaire multiple où il est ajusté par un algorithme non linéaire à gradient réduit généralisé prédit la valeur IDAC des solutés d'éthanol et de propanol dans les 46 LIs à

différentes températures. Une valeur de ERAM globale de 7,49 % et 9,91 % a est obtenue à partir du soluté d'éthanol et de propanol, respectivement, qui a été formé par la combinaison de l'ensemble de formation et de validation. L'efficacité du modèle avec un écart inférieur à 10 % a été validée par la comparaison entre les valeurs IDAC prédites et expérimentales basées sur l'anion, le groupe de tête cationique et la longueur de la chaîne alkyle.

Benimam et al.(2020)[28] ont développé deux modèles l'un basé sur l'algorithme d'apprentissage de la machine à vecteurs de support pour la régression SVMr et l'autre basé sur l'algorithme Dragonfly (DA) associé à SVMr pour l'estimation du coefficient d'activité de dilution infinie (IDAC) de 17 solutés dans 44 LIs d'imidazolium à l'aide de 2666 points de données expérimentales recueillies à partir de la littérature. Où les deux modèles prennent en compte le même ensemble de variables prédictives qui sont la température, le poids moléculaire du soluté et du solvant, et cinq modèles de criblage de type conducteur pour les descripteurs de profil σ de solvants réels (COSMO-RS) liés au soluté et à l'IL. Les résultats trouvés ont montré la supériorité du modèle DA-SVMr démontrée par son coefficient de corrélation et ses valeurs d'erreur quadratique moyenne de 0,996 et 0,170, respectivement.

Malik et al.(2023)[20] ont effectué le criblage de vingt-cinq (25) LIs à l'aide de COSMO-RS pour sélectionner les LIs appropriés comme entraîneurs verts en fonction du coefficient d'activité, de la capacité et de la sélectivité. Les résultats obtenus ont montré que le chlorure de tétraméthylammonium ([TMAm][Cl]), en raison de sa forte capacité de liaison hydrogène, s'est avéré être le meilleur entraîneur de LIs.

Rather et al.(2023)[29] ont développé une corrélation permettant de prédire le coefficient d'activité du CO_2 dans les liquides ioniques par analyse de régression multi-non linéaire dans une plage de pression de 1 à 50 bars et de température de 298,15 à 33,15 K pour des fractions molaires de 0,3, 0,5 et 0,7. En outre, ils ont utilisé la méthode du boxplot pour analyser les valeurs aberrantes et évaluer si les plages de paramètres d'entrée sélectionnées sont appropriées. À cet égard, ils ont utilisé les données expérimentales pour vérifier les résultats de COSMO-RS et celles prédites par la corrélation. Les valeurs de l'erreur absolue moyenne et l'erreur quadratique moyenne de corrélation prédite obtenus, sont de l'ordre de 19,53 % et 0,465 respectivement. À cet égard, ils ont conclu que la corrélation pourra être utilisée à l'avenir pour prédire les valeurs du coefficient d'activité du CO_2 dans les liquides ioniques afin de faciliter les analyses qualitatives de leur efficacité de capture du CO_2.

Rittig et al.(2023)[30] ont utilisé le modèle réseaux de neurones graphiques (RNG) pour la prédiction des coefficients d'activité CA à dilution infinie dépendant de la température des solutés dans les LIs. Ils ont formé le RNG sur une base de données comprenant plus de 40 000 valeurs de coefficients d'activité CA et en fonction directe des graphiques moléculaires de LIs et du soluté ainsi que de la température. Les résultats trouvés ont montré que le modèle RNG permet d'obtenir des prédictions CA de haute précision, supérieures aux modèles classiques tels que COSMO-RS et UNIFAC-LI et compétitives avec les méthodes de complétion matricielle pour les solutions LIs, et que le RNG permettant en outre des prédictions de haute qualité pour les CA des solutions contenant des LIs et des solutés non pris en compte lors de la formation. Ce qui en fait un constituant très prometteur de la conception assistée par ordinateur des LIs.

Li et al. (2023)[31] ont établi un modèle de calcul du coefficient d'activité de dilution infinie ($\gamma\infty$) du dichlorométhane (DCM) dans les LIs sur la base de l'algorithme de contribution de fragments ioniques (IFC) et de régresseur d'amplification de gradient (GBR). Où, le IFC est utilisé pour obtenir la zone de distribution de densité de charge de surface des IL (profil Sσ) qui est l'entrée du GBR. Le GBR a été utilisé pour apprendre la relation de mappage entre la caractéristique d'entrée et $\gamma\infty$ du DCM dans les LIs. À cet égard, pour établir le modèle IFC -GBR et prédire le $\gamma\infty$ du DCM dans les LIs ils ont utilisé une base de données composée de 29 cations et 22 anions comprenant 72 données expérimentales et 421 données de calcul COSMO. Les résultats ont indiqué que le modèle proposé peut prédire avec précision le $\gamma\infty$ du DCM dans les LIs, et fournir des données importantes pour développer un nouveau processus d'adsorption et de désorption du DCM par des technologies basées sur les LIs. Où le coefficient de détermination et l'erreur absolue moyenne de l'ensemble de test du modèle IFC-GBR sont respectivement de 0,9703 et 0,0519.

Kabane et al.(2023)[32] à l'aide d'une étude expérimentale, ils ont calculés les coefficients d'activités à dilution infinie à partir des données de rétention acquises par une chromatographie gaz-liquide opérée à différentes températures, T= (313,15 – 373,15) K et à pression atmosphérique pour les solutés organiques volatils (alcanes, alcènes, alcynes, alcools, cétones, hydrocarbures aromatiques, tétrahydrofuranne, thiophène et acétonitrile) dans du 1-éthyl-3-méthylimidazolium tétrachloroaluminate et un solvant eutectique (1-éthyl-3-méthylimidazolium tétrachloroaluminate et diéthylèneglycol). Ensuite l'étude théorique effectuée a montré que les prédictions de COSMO-SAC étaient qualitativement en très bon accord avec les données expérimentales. Le signal de l'effet de la température sur les

coefficients d'activité à des valeurs de dilution infinies a été prédit avec précision pour les solutés dans le liquide ionique et le solvant eutectique.

I.3 Revue critique sur la densité des LIs

La densité est l'une des propriétés physico-chimiques importante d'un liquide ionique, dont une connaissance claire de sa valeur est essentielle dans diverses applications scientifiques et techniques telles que la conception d'équipements, les calculs de dosage de liquides et de nombreuses autres applications.[33,34]. De sorte que, elle a attiré l'attention de nombreux scientifiques et chercheurs tel que :

Gardas et Coutinho (2007)[35] ont étendu la méthode de contribution du groupe Ye et Shreeve pour l'estimation des densités des liquides ioniques avec succès pour prédire les densités des LIs pour une large plage de température (273,15 à 393,15 K) et de pression (0,10 à 100 MPa). En utilisant 1521 points de données, de liquides ioniques à base de cations imidazolium, pyridinium, pyrrolidinium et phosphonium. Ils ont constaté que les densités prévues de mélanges binaires purs et homogènes des LIs ayant le même cation ou anion sont en bon accord avec les valeurs de la littérature expérimentale disponible. Où ils ont trouvé que Pour 12 LIs à base d'imidazolium, l'erreur relative absolue moyenne (ERAM) est de 0,45 % et pour les LIs à base de phosphonium, le MPD est de 1,49 %. Ils ont également observé le faible ERAM allant de 0,41 % à 1,57 % pour les LIs à base de pyridinium et de pyrrolidinium.

Valderramaa et Zarricueta (2008)[36] ont proposé un modèle basé sur une corrélation généralisée convenablement modifiée et en utilisant des données de la littérature expérimentale pour ajuster les cinq paramètres du modèle, pour finalement proposer une équation permettant de prédire les densités de tout liquide ionique. A cet égard, un ensemble de 602 données de densité pour 146 liquides ioniques pour une plage de températures (270 à 360 K) a été utilisé dans cette étude. Les résultats ont montré que la nouvelle corrélation simple donne de faibles écarts et peut être utilisée en toute confiance dans les calculs thermodynamiques et techniques.

Valderrama et al.(2009)[37] ont utilisé une méthode hybride cohérente avec succès, la méthode du réseau neuronal et de la contribution de groupe (GC + RNA) pour corréler et prédire la densité de103 liquides ioniques avec 399 points de données. Les résultats ont démontré que le réseau choisi et la méthode de contribution de groupe utilisée sont capables

d'estimer la densité des liquides ioniques avec une précision acceptable pour les calculs techniques.

Lazzu´(2010)[38] a utilisé une méthode de contribution de groupe simple pour prédire la densité de 210 liquides ioniques. En outre il a utilisé aussi les données de densité de 76 liquides ioniques (3 530 points de données) pour obtenir le modèle linéaire pour de larges plages de températures et de pressions : 258 à 393 K et 0,09 à 207 MPa. Les résultats ont montré que la nouvelle méthode de contribution de groupe (R^2=0,9687 ; ERAM =1,93 %) et le modèle linéaire (R^2=0,9989 ; ERAM =0,73 %) sont suffisamment précis et donnent des écarts plus faibles que les autres modèles disponibles dans la littérature.

Paduszynski et Domanska (2011)[39] ont proposé une nouvelle méthode basée sur la corrélation empirique généralisée et les contributions de groupe (GC). Elle a été développée sur une base de données expérimentales contenant plus de 18 500 points pour une grande variété de 1 028 LI, dans des plages de température et de pression de 253 473 K et 0,1 300 MPa, respectivement. Alors que, ils ont adopté une équation de type-Tait avec quatre paramètres réglables pour décrire la dépendance température-pression de la densité. Les résultats ont montré que le GC proposé permet de prédire avec précision les densités de haute pression pour une variété des Lis avec un ERAM de prédiction de 0,45 %, ce qui est l'une des valeurs les plus basses par rapport aux corrélations similaires rapportées dans la littérature. De plus, ils ont montré que la méthode présentée est capable de capturer avec précision d'autres propriétés volumétriques des LIs purs, telles que le volume molaire et les propriétés dérivées (coefficient de dilatation thermique et compressibilité isotherme), ainsi que leurs dépendances en matière de température et de pression.

Patela et Joshipurab (2013)[40] ont généré un paramètre réglable pour les liquides ioniques purs en conjonction avec l'approche prédictive-Soave-Redlich-Kwong (PSRK) pour la prédiction de la densité d'IL pure. En utilisant un ensemble de 735 points de données de 47 liquides ioniques purs et ils ont utilisé 4 fonctions de cohésion différentes dans l'étude. Ils ont observé que les modèles spécifiques aux composés ont donné des résultats très précis. En outre, ils ont comparé les résultats trouvés au modèle linéaire généralisé (MLG) développé par Valderrama pour vérifier la pertinence et l'exactitude de l'expression présentée. Ils ont constaté que l'approche PSRK spécifique au composé donne des prédictions très précises et que l'expression généralisée est comparable au modèle MLG.

Huang et al.(2015)[41] ont utilisé des modèles semi-empiriques et des modèles de réseaux de neurones artificiels (RNA) pour un ensemble de 659 points de données, ils ont constaté que le modèle semi-empirique de densité peut obtenir des résultats de corrélation très précis. Où la prévision globale ERAM% n'était que de 1,1 %. En outre, le modèle RNA a également donné des résultats de prédiction très précis avec un ERAM% global de 0,42 %.

Ji a et Held (2015)[42] ont utilisé Théorie statistique des fluides associés aux chaînes perturbées des électrolytes (ePC-SAFT) pour modéliser les densités de liquides ioniques (LIs) pour des pressions et des températures élevées. Dans le cadre ePC-SAFT, ils ont appliqué des paramètres spécifiques aux ions LIs et sont valides indépendamment de LI dont ils font partie. Ils ont modélisé chaque ion de LI comme une espèce non sphérique exerçant des forces répulsives, dispersives et coulombiennes. Les résultats obtenus ont montré que le modèle ePC-SAFT a permis de représenter/prédire de manière fiable la densité des LIs purs et des LIs mixtes jusqu'à des pressions élevées.

Xu et al. (2016)[33] ont utilisé une équation d'état de contribution de groupe (EOS), déjà développée dans leurs travaux antérieurs, pour calculer la densité des liquides ioniques à base d'imidazolium avec une bonne précision pour corrélées/prédites la densité des LIs purs. Pour cela, sur la base des nouveaux paramètres de diamètre du noyau mou et d'énergie dispersive qui ont été obtenu, ils ont prédit la densité de 908 points de données de 42 LIs purs de 278,15 à 393,15 K et jusqu'à 65,01 MPa. Les résultats ont indiqué que l'EOS pourrait bien capturer les effets de la température, de la pression et des types d'alkyles sur la densité. En outre, ils ont évalué l'EOS plus en détail pour prédire les données de densité pour les mélanges binaires des LIs sans introduire de nouveaux paramètres, et a témoigné de l'applicabilité étendue de l'EOS avec un écart relatif moyen de 0,70 %.

Keshavarz et al.(2016)[43] pour estimer la densité d'une large gamme de différents types de liquides ioniques ont utilisé une méthode reposée sur deux corrélations inédites dans lesquelles la taille, la structure et le type des cations et des anions sont des paramètres essentiels. Où la première corrélation est dérivée de la composition élémentaire, modifiée sur la base de fonctions de correction et dans la seconde afin de prendre en compte les interactions inter ioniques et d'obtenir des résultats plus fiables. Ils ont constaté que les résultats prédits de deux corrélations pour 249 types différents de composés mentionnés ont donné des résultats plus fiables par rapport aux résultats d'une méthode complexe.

Najafi-Marghmaleki et al. (2016)[44] ont développé deux modèles d'un réseau neuronal artificiel (RNA), à savoir le modèle de perceptron multicouche MLP-RNA et le modèle de fonction à base radiale RBF-RNA pour prédire les valeurs de densité dans un mélange binaire d'eau et de liquides ioniques à différentes températures pour différents liquides ioniques à base d'imidazolium où les paramètres d'entrées étaient la température, le rapport eau/liquide ionique, le poids moléculaire, la température critique, la pression critique et la température d'ébullition des liquides ioniques. Le paramètre de sortie était la densité. Les résultats obtenus ont montré que les deux modèles sont précis et efficaces pour prédire la densité du mélange liquide ionique + eau. Cependant, le modèle MLP-RNA a présenté de meilleures performances par rapport au modèle RBF-RNA en raison d'un coefficient de corrélation plus élevé (R^2 MLP-RNA = 0,9992, R^2 RBF-RNA = 0,9671) et des valeurs inférieures des valeurs AARD% et RMSE.

Barati-Harooni et al.(2016)[45] ont développé un modèle basé sur des algorithmes du système d'inférence neuro-flou adaptatif (ANFIS) pour prédire la densité de liquides ioniques à différentes températures en utilisant une base de données comportant 602 valeurs de densité de 146 liquides ioniques. Les résultats ont montré que le modèle développé prédit avec précision les données expérimentales avec des valeurs globales de R^2 et ERAM% de 0,985 et 0,657 %, respectivement. De plus, le modèle développé surpasse efficacement les corrélations de la littérature et présente des prédictions plus précises et plus fiables.

Farzi et Esmaeilzadeh (2016)[46] ont utilisé la contribution du groupe Esmaeilzadeh-Roshanfekr (GC-ER) pour calculer la densité de 294 liquides ioniques à différentes températures et pressions ambiantes sur la base de 1775 données expérimentales. Les résultats ont montré que l'erreur du modèle était inférieure à celle des autres équations testées dans ce travail. L'écart relatif absolu moyen (ERAM%) du modèle proposé était de l'ordre de 11,49 %.

Najafi-Marghmaleki et al.(2016)[34] ont développé un modèle basé sur la technique de programmation d'expression génique (GEP) pour la prédiction des valeurs de densité expérimentales de 146 liquides ioniques à différentes températures en utilisant une base de 602 données. Où ils ont utilisé la température, le poids moléculaire, la température normale du point d'ébullition, la pression critique, le volume critique et le facteur acentrique comme paramètres d'entrées du modèle. Les résultats obtenus ont indiqué que le modèle développé est capable d'estimer les données de densité cible avec une précision acceptable. De plus,

après une comparaison aux prédictions des corrélations de la littérature ils ont constaté que le modèle GEP est largement supérieur par rapport aux autres corrélations et présente une précision et une fiabilité supérieures.

Barati-Harooni et al. (2017)[47] ont développé un modèle sur la base du réseau neuronal à fonction de base radiale (RBF-RN) pour l'estimation des valeurs de densité de 146 liquides ioniques à différentes températures. Les résultats de cette étude ont révélé que le nouveau modèle est capable de prédire les données de densité avec une précision acceptable. Et après une comparaison aux corrélations de la littérature ils ont observé que le modèle surpasse efficacement les autres corrélations et présente une précision et une fiabilité supérieures.

Shekari et al.(2019)[48] ont utilisé l'équation d'état du fluide d'association statistique de chaîne perturbée (PC-SAFT) pour prédire la pression de vapeur et la densité des solutions non aqueuses du liquide ionique dans certains solvants organiques polaires dans la région diluée du liquide ionique. Les résultats obtenus à partir de l'équation d'état PC-SAFT ont fourni des prédictions fiables et précises avec un écart moyen absolu de 0,01.

Zimmermann et Mattedi (2020)[49] ont développé un réseau de neurones artificiels pour la prédiction de la densité d'un mélange binaire d'eau et de liquides ioniques à base d'ammonium, en utilisant la température, la fraction massique du liquide ionique et les groupes structurels des réactifs utilisés pour synthétiser le liquide ionique comme variable d'entrée. Le réseau final proposé était un réseau en cascade 12-10-10-1 avec deux couches cachées entraînées avec une régularisation bayésienne a atteint un EQM, ERAM% et R^2 pour l'ensemble de tests de densité de 0,0000, 0,0107 % et 1,0000, respectivement.

Amara et al.(2020)[50] ont mis en œuvre une technique intelligente rigoureuse, à savoir la programmation d'expression génique (GEP) pour modéliser la densité de mélanges binaires et ternaires de liquides ioniques et de solvants moléculaires à l'aide d'une base de données expérimentale complète comprenant 1985 mesures réelles dans diverses conditions opérationnelles. Les résultats obtenus ont révélé que les corrélations basées sur le GEP nouvellement établies peuvent prédire la densité de mélanges binaires et ternaires de LIs et de solvants moléculaires avec un degré élevé d'intégrité. Où les corrélations basées sur le GEP ont présenté des valeurs d'erreur relative absolue moyenne globale (ERAM%) de 0,5621 % et 0,2128 % pour les cas binaires et ternaires, respectivement. En outre, ils ont constaté que les corrélations explicites proposées suivaient la tendance attendue par rapport aux variables

considérées. Ainsi que, la supériorité et la fiabilité des corrélations basées sur le GEP ont été démontrées par rapport aux meilleures approches existantes dans la littérature.

Bagheri et al. (2021)[51] ont étudié six équations semi-empiriques simples avec différentes règles de mélange et deux cubiques (CEoS)(Soave-Redlich Kwong EoS et Patel-TejaEoS) pour prédire la densité d'un mélange comprenant un liquide ionique. Pour chaque équation semi-empirique, les coefficients ont été optimisés à l'aide d'un algorithme d'optimisation de baleine Laplacienne basé sur la famille de cations. Les résultats obtenus basés sur des équations semi-empiriques ont indiqué que la modification effectuée fournit une faible déviation des données expérimentales et peut être appliquée en toute confiance dans les calculs techniques et thermodynamiques. Cependant, les résultats basés sur deux CEoS ont montré une précision inacceptable.

Ashraf Mansouri (2021)[52] a utilisé la technique de contribution SAFT-γgroup (GC) pour estimer la densité et les propriétés dérivées thermodynamiques du second ordre de certaines LIs à base d'imidazolium et de cyano (avec [SCN], [DCA], [TCM] ou [TCB]) dans une plage de température de 262 à 412 K. Les résultats ont démontré que l'équation SAFT-γ pouvait estimer la densité des LIs étudiés avec une bonne précision.

I.4 Revue critique sur la viscosité des LIs

La viscosité est l'une des propriétés les plus importantes, largement utilisée pour étudier les IL et leurs puretés. La viscosité est une propriété de transport provoquée par le frottement entre les particules voisines dans un fluide qui se déplacent à des vitesses différentes. En d'autres termes, cette propriété fournit une mesure de la résistance du fluide à une déformation progressive par contrainte de cisaillement ou contrainte de traction.[53]

Ainsi, la viscosité est la propriété physique qui a une influence directe sur les applications industrielles. Par exemple, il est clair que lorsque les LIs sont utilisés comme lubrifiants, une viscosité élevée est préférée, en particulier pour les applications exigeantes et les températures élevées. En revanche, une faible viscosité est souhaitable lorsque les LIs sont utilisés comme solvants verts, catalyseurs ou fluides caloporteurs. Par conséquent, la connaissance de la valeur de la viscosité est d'une grande importance.[54]

Étant donné que les données expérimentales actuelles sur cette propriété sont encore rares et limitées aux familles LIs connues. Plusieurs études théoriques ont été réalisées capables de donner des prévisions fiables sur les viscosités des LIs.

Gharagheizi et al.(2012)[55] ont développé une méthode fiable de contribution de groupe pour prédire la viscosité de 443 liquides ioniques en utilisant un ensemble de données composé de 1672 valeurs et un total de 46 sous-structures en plus de la température pour prédire la viscosité des LIs. Ils ont constaté que le modèle peut être utilisé pour prédire la viscosité de divers liquides ioniques avec un faible écart relatif moyen inférieur à 6,4 %.

Chena et al.(2013)[56] ont développé la viscosité des LIs à base d'imidazolium à l'aide modèle QSPR couplé aux descripteurs de contribution de groupe où ils ont formulé une relation mathématique entre la viscosité des LIs à base d'imidazolium et les paramètres descriptifs des anions et des cations. Ils ont constaté que la forte corrélation entre la valeur calculée de QSPR et la viscosité expérimentale offre une possibilité de prédiction de la viscosité pour la conception de nouveaux liquides ioniques. Le coefficient de corrélation était de $R^2 = 0,9888$.

Bajić et al (2014)[57] ont calculé la viscosité de dix mélanges liquides binaires constitués de liquides ioniques et de composés organiques à pression atmosphérique et dans une large plage de températures à l'aide de deux approches différentes, prédictive par modèles de contribution de groupe (UNIFAC-VISCO et ASOG-VISCO) et corrélative (Seddon, Grunberg – Nissan, McAlister, Eyring-UNIQUAC et Eyring-NRTL). Aussi, une étude comparative avec les données de la littérature a été effectuée. Ils ont constaté que ces modèles ont donné de très bons résultats, qui sont proches des modèles corrélatifs.

Hosseinzadeh et Hemmati-Sarapardeh (2014)[53] pour prédire la viscosité des mélanges ternaires ont utilisé une approche de modélisation rigoureuse, à savoir la machine à vecteurs de support des moindres carrés (LSSVM) où ils ont optimisé ces paramètres à l'aide de l'outil d'optimisation du recuit simulé couplé. À cette fin, ils ont collecté 729 points de données des mélanges de LIs pour l'apprentissage et le test du modèle en utilisant la température du point d'ébullition, le poids moléculaire et la composition des composés du mélange comme variables d'entrées. Ils ont constaté à partir des résultats des analyses d'erreurs statistiques et graphiques que le modèle développé peut prédire la viscosité du mélange ternaire avec une précision significative.

Fatehi et al.(2014)[58] ont construit un réseau neuronal à action directe pour estimer les viscosités de mélanges binaires de cinq types différents de liquides ioniques avec divers solvants polaires et non polaires dans une large plage de températures en utilisant le poids moléculaire et la structure de groupe du liquide ionique, le poids moléculaire et la température

d'ébullition réduite du solvant, ainsi que la composition molaire, la température et la pression du système comme variables d'entrées indépendantes, pour une base de données contenant 1996 points de données expérimentales. Les résultats obtenus en terme de l'erreur relative moyenne de 0,6 %. Le réseau RNA obtenu peut être considéré comme un outil efficace et précis pour estimer la viscosité de mélanges binaires de liquides ioniques et de solvants et que peut être utilisée pour l'estimation de systèmes similaires.

Paduszyński et Domańska (2014)[59] ont collecté plus de 13000 points de données sur la viscosité dépendant de la température et de la pression pour la prédiction de la viscosité de 1484 liquides ioniques en utilisant une stratégie de réseau neuronal artificiel à action directe à deux couches pour représenter la relation entre la viscosité et les variables d'entrées: la température, la pression et la contributions de groupe (GC). Ils ont obtenu pour les processus d'apprentissage, de validation et de test, les erreurs quadratiques moyennes de 0,0334, 0,0595 et 0,0603 unités log, correspondant à des coefficients de corrélation quadratique de 0,986, 0,973, et 0,972 et des écarts relatifs globaux au niveau de 11,1, 13,8 et 14,7 %, respectivement. À cet égard, ils ont révélé que les résultats calculés dans ce travail se sont plus précis que ceux obtenus avec le meilleur modèle GC actuel pour la viscosité des LIs décrit dans la littérature.

Haghbakhsh et Raeissi (2014)[60] ont étudié les viscosités de huit liquides ioniques différents et leurs mélanges binaires avec l'eau en appliquant deux corrélations empiriques, l'une pour l'estimation des viscosités de liquides ioniques purs, et l'autre pour leurs mélanges aqueux à différentes températures à pression atmosphérique et, dans le cas de la corrélation de mélange, à diverses compositions molaires. Ils ont constaté que les deux corrélations pour les viscosités pures et celles des mélanges étaient précises, ils ont obtenus pour la corrélation des composants purs, un écart relatif moyen absolu de 2,97 % pour un total de 245 données sur des liquides ioniques purs (et de l'eau pure), tandis qu'ils ont obtenus pour la corrélation des mélanges, un écart relatif moyen absolu de 6,72 % pour un total de 245 données sur les liquides ioniques purs (et l'eau pure).

Zhao et al.(2015)[13] ont développé deux nouveaux modèles QSPR pour prédire la viscosité des liquides ioniques à l'aide des algorithmes de régression linéaire multiple (MLR) et de machine à vecteurs de support (SVM) basés sur les descripteurs moléculaires du modèle de dépistage de type conducteur pour les solvants réels (COSMO-RS) (Sσ -profil). En utilisant un ensemble de données total de 1502 points de données expérimentales sur la

viscosité dans une large plage de température et de pression pour 89 liquides ioniques. Ils ont trouvé un écart relatif absolu moyen de l'ensemble total de données du MLR et du SVM de 10,68% et 6,58%. Respectivement. Ces résultats ont montré que les modèles MLR et SVM peuvent prédire la viscosité des LIs et que les performances du modèle non linéaire développé par le SVM sont supérieures à celles du modèle linéaire (MLR). En outre, ils ont conclu que les modèles dérivés peuvent également éclairer les caractéristiques structurelles liées à la viscosité des LIs.

Lazzús et Pulgar-Villarroel (2015)[61] ont estimé la viscosité des liquides ioniques à différentes températures à l'aide d'une méthode simple de contribution de groupe en utilisant une approche de type-Andrade. Tout d'abord ils ont utilisé les données expérimentales de 183 liquides ioniques (1 110 points de données) avec une viscosité comprise entre 6 et 20 883 (cP) à des températures comprises entre 253 K et 395 k pour obtenir les contributions des groupes cation-anion dans un ensemble de corrélations. Ensuite, ils ont prédit les viscosités de 143 autres liquides ioniques (335 points de données). Les résultats ont montré que cette méthode de contribution de groupe représente une excellente alternative pour l'estimation de la viscosité de divers liquides ioniques à différentes températures à partir de la connaissance de leur structure moléculaire avec un écart moyen de 4,5% et un coefficient de corrélation de 0,94.

Sattaria et al.(2016)[62] ont rassemblé deux ensembles de données complets de liquides ioniques contenant du fluor (F-LI) l'un avec des sources fiables constituées de 863 viscosités expérimentales pour 85 F-LI, et l'autre comprenant des sources peu fiables constituées de 1160 points de données expérimentales pour 332 F-LI pour estimer les viscosités des F-LI en fonction de la température à l'aide d'un modèle efficace de contribution de groupe (GC) qui utilise un total de 35 groupes fonctionnels. Ils ont trouvé que le modèle GC proposé a produit un faible écart relatif absolu moyen de 3,23 %, une erreur quadratique moyenne de 0,19 et un coefficient de détermination de 0,977.

Zhao et al. (2016)[63] ont développé une méthode de contribution de groupe UNIFAC−VISCO modifiée pour la corrélation et la prédiction de la viscosité des liquides ioniques en fonction de la température à 0,1 MPa. Ils ont déterminé les paramètres d'interactions binaires pour des cations et des anions sélectionnés en ajustant les données expérimentales de viscosité disponibles dans la littérature pour des liquides ioniques sélectionnés. Ainsi que, ils ont adapté la dépendance en température de la viscosité des

cations et des anions à un comportement de Vogel−Fulcher−Tamman. Les résultats de toutes les données de viscosité calculées ont montré un excellent accord avec les données expérimentales avec un écart moyen absolu relatif inférieur à 1,7 %.

Zhaoa et al.(2016)[64] ont intégré la méthode des états correspondants à la contribution des fragments avec deux algorithmes, notamment l'algorithme de régression linéaire multiple (MLR) et l'algorithme de machine à vecteurs de support (SVM) pour prédire la viscosité des liquides ioniques à base d'imidazolium. En utilisant un vaste ensemble de données de 1 079 points de données expérimentales de 45 LIs à base d'imidazolium couvrant une large plage de pression et de température. Sur la base de des résultats qu'ils ont obtenus ils ont constaté que Le modèle non linéaire développé par l'algorithme SVM est bien meilleur que le modèle linéaire construit par le MLR, ce qui indique que l'algorithme SVM est plus fiable dans la prédiction de la viscosité des LIs à base d'imidazolium. Où ils ont trouvé que L'écart relatif absolu moyen de l'ensemble des données du MLR et du SVM est respectivement de 24,2 % et 3,95 %.

Fatehi et al. (2017)[65] ont construit un réseau neuronal artificiel à trois couches, avec 24 neurones, pour estimer les viscosités d'un large éventail de familles de liquides ioniques, y compris celles basées sur l'imidazolium, l'ammonium, le pyridinium, le pyrrolidinium, le phosphonium et l'isoquinolinium dans une plage de température de (273,15 à 393,15) K et une plage de pression de (0,1 à 150) MPa. Ils ont constaté que le modèle RNA proposé peut-être considérer comme un outil capable d'estimer les viscosités des liquides ioniques purs et que ce modèle peut également être appliqué à des systèmes similaires selon les résultats obtenus tel que, l'erreur quadratique moyenne et le coefficient de corrélation sur les ensembles de données d'apprentissage et de validation sont de 4,36. 10^{-8}, 1,625.10^{-6}, 1,00000 et 0,99955 respectivement. Ainsi que, l'erreur absolue moyenne sur l'ensemble de données de test était de 1,310 %.

Baghban et al.(2017)[66] pour estimer la viscosité de 443 liquides ioniques différentes ils ont développé un modèle basé sur la combinaison de la machine hybride à vecteurs de support des moindres carrés (LSSVM) et de la méthode de contribution de groupe. De plus, ils ont comparé le modèle LSSVM proposé avec un autre modèle de contribution de groupe développé par Gharagheizi et al (2012). Les résultats obtenus à partir des analyses statistiques ont confirmé que les approximations du modèle LSSVM est en bon accord avec les viscosités réelles. Ainsi que, les résultats de l'erreur quadratique moyenne (EQM) et le R-carré (R^2) qu'ils

ont obtenus sont de 0,007, 0,979, respectivement ils ont confirmé la meilleure capacité prédictive du modèle.

Kang et al.(2017)[67] ont développé deux nouveaux modèles en utilisant un algorithme d'intelligence machine d'apprentissage extrême (ELM) pour prédire la viscosité des liquides ioniques basés sur de larges plages de pression (1 à 3 000 bars) et de température (258,15 à 395,32 K). Ils ont utilisé la température, la pression et un nombre de descripteurs de profils Sσ calculés à l'aide de la méthode de chimie quantique comme paramètres d'entrées. Ils ont constaté que les modèles ELM ont de bonnes performances basées sur des plages de température et de pression plus larges, et le pourcentage de ERAM de l'ensemble total des modèles ELM1 et ELM2 était de 2,21 et 4,10 %, respectivement. En comparant avec d'autres modèles, ils ont vérifié que les modèles établis par ELM ont une meilleure corrélation et stabilité pour l'ensemble des données.

Valderrama et al.(2019)[68] ils ont proposé en premier lieu une équation d'état de viscosité généralisée incluant un seul paramètre réglable pour corréler et estimer la viscosité des liquides ioniques à différentes températures et pressions en utilisant une base de 3857 données de viscosité pour 187 liquides ioniques dans la plage de température de 253 à 573 K et des pressions de 1,0 à 1 500 bar. Un certain nombre de paramètres a été déterminé en ajustant le modèle à l'aide des données expérimentales de viscosité sélectionnées de différents types de liquides ioniques. Les résultats ont montré que le nouveau modèle d'équation d'état cubique de viscosité a fourni des résultats précis et cohérents en tenant compte de la simplicité des expressions généralisées qui ne contiennent qu'un seul paramètre réglable pour chaque liquide ionique.

Valderrama et al.(2019)[69] dans une deuxième étude, ils ont étendu l'équation d'état générale utilisée sous forme directe dans la première partie pour corréler et prédire la viscosité des mélanges binaires de liquides ioniques dans une large plage de températures à l'aide d'un ensemble de 2520 points de données où l'extension considère également les règles classiques de mélange et de combinaison couramment utilisées dans les études sur les équations d'états pression-température-volume, qui a compris deux cas : (i) avec paramètre de non-interaction (modèle prédictif) ; et (ii) avec un paramètre d'interaction réglable (modèle de corrélation). Les résultats obtenus ont montré que même sans paramètre d'interaction, le modèle prédictif fournit des valeurs adéquates et cohérentes pour les premières estimations dans la conception

et la simulation de processus. En résumé, le modèle est plus simple et plus précis que les modèles multiparamétriques présentés dans la littérature.

Paduszyński (2019)[70] a proposé de nouvelles relations quantitatives structures-propriétés (QSPR) pour estimer la viscosité dynamique (η) de liquides ioniques purs (LIs) en fonction de la température et des contributions de groupe (GC) à l'aide de trois algorithmes d'apprentissage automatique courants (régression linéaire multiple pas à pas "SWMLR", réseau neuronal artificiel à action directe "FF-RNA " et machine à vecteurs de support des moindres carrés "LSSVM") sur la base de la plus grande base de données comprenant les données de 2068 Lis distinctes (3236 ensembles de données et 22 268 points de données). Il a constaté que la modélisation de la viscosité des LIs en termes de GC n'est pas un problème facile, quelle que soit la méthode de calcul appliquée. En particulier, une simple règle d'additivité, qui constitue la base du SWMLR, a donné de mauvais résultats. Les méthodes d'apprentissage automatique plus sophistiquées (FF-RNA et LSSVM) ont permis d'obtenir des performances de modélisation nettement meilleures, mais l'interprétation physique ou mécaniste des paramètres des modèles résultants est fondamentalement irréalisable. Ainsi, ces modèles sont plus précieux du point de vue purement utilitaire que pour la connaissance fondamentale ou la compréhension physique du comportement visqueux des LIs.

Nabi et al.(2019)[71] ont développé un système d'inférence neuro-flou adaptatif (ANFIS) pour prédire la viscosité des mélanges ternaires contenus dans les liquides ioniques en fonction de 11 paramètres d'entrées et une base de données contenant 800 points de données de 12 mélanges ternaires. Ils ont constaté que le modèle ANFIS proposé a donné une grande capacité à prédire la viscosité des mélanges avec un ERAM % de 0,945.

Darabi et Zare (2020)[72] ont proposé une nouvelle équation linéaire simple pour représenter la dépendance à la pression de la viscosité des liquides ioniques. A cet égard, ils ont utilisé et corrélé les données expérimentales de viscosité de la littérature en fonction de la pression. Ils ont montré que cette équation a permis de fournir des descriptions précises de la viscosité des liquides ioniques en fonction de la pression. De plus, ils ont proposé une nouvelle équation pour intégrer à la fois la dépendance de la viscosité (η – T – p) à la température et à la pression. Les résultats ont indiqué que le modèle η – T – p présente une grande précision avec un écart relatif absolu moyen global de 2,69 % dans la plage de température de 283,15 à 363,15 K et la plage de pression de 0,06 à 300 MPa. De plus, Ils ont

constaté que le coefficient pression-viscosité calculé augmente avec la longueur de la chaîne alkyle du cation.

Lopez-Cortés et al (2020)[73] ont développé une équation pour corréler les viscosités cinématiques de mélanges binaires de liquides ioniques avec des solvants organiques et avec de l'eau dans une large plage de température à pression atmosphérique en utilisant une base de 5746 points de données expérimentales de 111 mélanges binaires et la fraction de surface moléculaire comme variable de composition. Ensuite, ils ont également comparé les performances de la nouvelle équation avec des équations récemment développées pour la viscosité cinématique telles que les équations de Nava-Ríos d'Eyring Modified Two Suffix Margules, de Fang-He et d'Eyring-Wilson. Ils ont trouvé que la nouvelle équation corrèle la viscosité cinématique avec un pourcentage absolu moyen d'écart de 2,64 % pour les mélanges binaires de liquides ioniques + eau et de 1,44 % pour les mélanges binaires de liquides ioniques + composés organiques.

Boldoo et al. (2021)[74] ont développé un modèle de réseau neuronal artificiel (RNA) pour prédire la viscosité dynamique de divers cation 1-hexyl-3-méthylimidazolium [HMIM] avec des nanoparticules de nanotubes de carbone à parois multiples (MWCNT). Ils ont conclu que le modèle RNA proposé peut prédire avec précision la viscosité dynamique des MWCNT NEIL avec des agents de base RTIL [HMIM]-cation pour toute application où ils ont trouvé que les valeurs prédites et expérimentales présentaient un ajustement exceptionnel, comme le prouve les valeurs de l'erreur quadratique moyenne et la valeur R du modèle prédictif RNA proposé sont respectivement de 0,001291 et 0,9985. La marge d'écart maximale du modèle est de 9,63 %.

Chen et al.(2022)[75] pour prédire la viscosité des mélanges liquide ionique (LI)-eau ont proposé un modèle non linéaire intégrant la méthode de contribution de groupe (GC) à un algorithme d'apprentissage automatique bien connu, c'est-à-dire un réseau neuronal artificiel (RNA), en utilisant un ensemble de données couvrant 8 523 points de données de viscosité de mélanges LI-H2O à différentes températures (272,10 K - 373,15 K). Les résultats ont montré que ce modèle RNA-GC avec 4 ou 5 neurones dans la couche cachée était capable de fournir des prédictions fiables sur les viscosités des mélanges LI-H2O avec 4 neurones dans la couche cachée, le modèle RNA-GC a donné une erreur absolue moyenne (EAM) de 0,0091 et un coefficient de corrélation (R^2) de 0,9962 pour les 6586 points de données d'entraînement, et pour les 1937 points de données de test, ils sont de 0,0095 et 0,9952, respectivement.

Lorsque ce modèle non linéaire comporte 5 neurones dans la couche cachée, il a donné un EAM de 0,0098 et un R^2 de 0,9958 pour l'ensemble de données d'entraînement, et pour l'ensemble de données de test, ils sont respectivement de 0,0092 et 0,9990.

Boualem et al.(2022)[54] ont développé deux nouveaux modèles de contribution de groupe GC complets sur la base des approches régression non linéaire traditionnelle (NLR) la régression de machine à vecteurs de support (SVM) pour prédire la viscosité dynamique des LIs en fonction de la température, à l'aide d'une grande base de données expérimentale contenant 15251 points de données expérimentales couvrant 1654 LIs distinctes. Les résultats obtenus des deux modèles ont montré des prédictions précises et fiables avec une supériorité prédictive du modèle SVM par rapport au modèle NLR et d'autres modèles existants dans la littérature.

Zimmermann et Mattedi (2022)[76] ont appliqué des réseaux de neurones (les réseaux feed forward et cascade-forward) pour la prédiction de la viscosité de mélanges binaires de liquides ioniques à base d'ammonium et d'eau en utilisant une approche de contribution de groupe. Ils ont trouvé des écarts relatifs absolus moyens pour l'ensemble de tests variaient de près de 200 % à 0,1999 % en fonction des paramètres sélectionnés. Les résultats ont montré que les réseaux neuronaux peuvent être utilisés avec succès pour prédire les données de viscosité, mais une sélection minutieuse de l'architecture et de l'algorithme de formation est nécessaire.

Huang et al. (2023)[77] ont développé des modèles d'apprentissage automatique interprétables en intégrant des informations physiques pour relier les modèles universels quasi-chimiques de coefficient d'activité du groupe fonctionnel (UNIFAC) en tant que nouveaux outils prédictifs de la viscosité des mélanges LI-eau. Après la comparaison des algorithmes d'apprentissage automatique, notamment Adaptive Boosting (AdaBoost), CategoricalBoosting (CatBoost), Extreme Gradient Boosting (XGBoost), K-Nearest-NeighbourRegression (KNNR) et Random Forest (RF), ils ont trouvé que le modèle Catboost a montré de meilleures précisions (EQRM = 0,0084, EQM = 0,0001, EAM = 0,0020, R2 = 0,9941). En plus, ils ont démontré que le modèle UNIFAC et la relation de Stokes-Einstein ont contribué à réduire la dimensionnalité des caractéristiques et à améliorer le pouvoir prédictif de la viscosité des mélanges LI-eau. De sorte que, le modèle UNIFAC contient des informations physiques préalables pour optimiser les caractéristiques de l'expérience et combler le vide entre les caractéristiques d'entrées et la viscosité des mélanges LI-eau.

Moslehi et al.(2023)[78] ont utilisé la théorie du frottement (TF) pour estimer les viscosités dynamiques de liquides ioniques purs dans la plage 273-373K et des pressions allant jusqu'à 298,9 MPa à l'aide d'une équation d'état à chaîne dure perturbée. Le nombre total de points de données de viscosité utilisés dans ce travail était de 3 074. Cependant, ils ont exclué 167 points de données pour l'obtention des coefficients de corrélation. Ensuite, ils ont testé ce modèle TF par rapport aux 2 907 points de données restants en utilisant trois paramètres moléculaires ainsi que 5 coefficients de corrélation dans les termes de frottement. Ils ont révélé que leur modèle surpassait des autres modèles existants dans la littérature en termes de précision avec un écart absolu moyen (ERAM) de 2,46 %. De plus, ils ont formé un réseau neuronal sur 3074 points de données pour représenter la viscosité des LIs purs qu'ils ont étudiés avec un ERAM global de 1,44 %.

I.5. Revue critique sur la conductivité thermique des LIs

La conductivité thermique (λ) est un indicateur extrêmement crucial de la capacité de transfert de chaleur des liquides ioniques et joue un rôle essentiel dans leurs applications industrielles. Par exemple, lors de la sélection du catalyseur, la conductivité thermique est l'un des paramètres pertinents nécessaires aux propriétés de transport. Il est également nécessaire d'obtenir le coefficient de transfert thermique des matériaux, crucial pour la conception d'un agent caloporteur et d'un équipement.[79,80] Parmi les méthodes alternatives importante qui ont été étudiées pour compenser le manque de données expérimentales on trouve

Hezave et al.(2012)[81] ont proposé un modèle de réseau de neurones artificiels (RNA) pour prédire la conductivité thermique des liquides ioniques purs en utilisant une base de 209 points de données provenant de 21 liquides ioniques différents où le nombre optimal de couches cachées a été déterminé comme étant 1, avec 13 neurones dans la couche cachée et des fonctions logarithmique-sigmoïde et pureline comme fonctions de transfert dans les couches cachées et de sortie, respectivement. Les résultats obtenus ont révélé que le réseau proposé est capable de corréler et de prédire la conductivité thermique de tous les liquides ioniques purs avec un pourcentage d'écart relatif moyen absolu global (ERAM %) de 0,5 % et une erreur quadratique moyenne (EQM) de $1,2 \times 10^{-6}$.

Wu et al.(2013)[82] ont développé une méthode de contribution de groupe (GC) pour estimer la conductivité thermique des liquides ioniques couvrant une large plage de température (273,15–353,15 K) et de conductivité thermique (0,1050–0,2137 W m−1 K^{-1}) sur

la base de données expérimentales collectée à partir de la littérature. Ils ont obtenu un bon accord entre les données calculées et celles de la littérature pour un total de 286 points de données de 36 liquides ioniques avec un écart absolu moyen (ERAM) de 1,66 %. A partir de ces résultats ils ont conclu que cette méthode de contribution de groupe a permis d'offrir une approche simple et fiable pour estimer la conductivité thermique de nouveaux liquides ioniques dans une large plage de températures à pression atmosphérique.

Shojaee et al.(2013)[83] ont proposé une nouvelle corrélation pour prédire la conductivité thermique des LIs pour un ensemble de 209 points de données provenant de 21 liquides ioniques différents à partir des publications, en utilisant cinq paramètres de corrélation réglable obtenus à l'aide de l'algorithme génétique. Ils ont trouvé un pourcentage d'écart relatif absolu moyen (ERAM%) au cours de la phase de formation était de 5,22 %, tandis que les résultats ont révélé un ERAM de 10,76 % pour la phase de test. En général, ils ont conclu que les résultats globaux ont révélé une assez bonne précision de la corrélation proposée, surtout compte tenu de sa capacité extrapolative.

Albert et Müller (2014)[84] ont développé un modèle QSPR pour la détermination de la conductivité thermique de 39 liquides ioniques en utilisant 375 points de données couvrant 15 cations et 18 anions. Ils ont trouvé pour les données des substances non utilisées pour l'ajustement ont pu être estimées avec une erreur absolue moyenne de 4,99 %, où l'erreur pour les substances utilisées pour l'ajustement des paramètres n'était que de 1,21 %.

Albert et Müller (2014)[85] dans une autre étude, ils ont développé une méthode incrémentale pour l'estimation des propriétés thermiques des LIs en utilisant des contributions de groupes de premier ordre pour décrire la structure moléculaire des ions. À cet égard, 372 points de données pour 39 liquides ioniques ont été traités de manière analogue pour la prédiction de la conductivité thermique. Ils ont trouvé que les données pour les liquides ioniques du sous-ensemble non utilisés pour le raccord ont pu être reproduites avec une erreur absolue moyenne de 8,1 %.

Chen et al(2014)[86] ont développé une méthode d'indice topologique pour estimer la conductivité thermique des LIs sur la base des données collectées qui ont été divisées en ensembles de formation (299 points de données, 38 LIs) et de tests (60 points de données, 7 LIs). Ils ont constaté que la méthode de l'indice topologique peut garantir une précision satisfaisante avec une forme simple et peut être facilement étendue à de nouveaux types de

LIs, avec un écart absolu moyen total de 2,3 %, dont 2,1 % pour l'ensemble de formation et 2,9 % pour l'ensemble de tests.

Wu et al.(2014)[87] ont déterminé les conductivités thermiques des liquides ioniques sur la base d'une équation de Bridgman modifiée dans laquelle seules les données de densité et de vitesse du son sont nécessaires. À cet égard, ils ont développé une méthode CSGC de second ordre avec une bonne précision pour estimer la vitesse du son des liquides ioniques, qui constitue une propriété thermodynamique importante des liquides ioniques et est toujours choisie comme source pour déterminer d'autres propriétés. Ils ont trouvé que pour une base de données de 276 points de données sur 34 liquides ioniques, les conductivités thermiques calculées concordent bien avec les données expérimentales de la littérature avec un ERAM de 3,91 %.

Lazzús et Pulgar-Villarroel (2015)[80] ont estimé la conductivité thermique à différentes températures de liquides ioniques à l'aide d'un modèle de relation quantitative structure-propriété (QSPR) impliquant 11 descripteurs moléculaires basés sur des calculs semi-empiriques pour 195 points de données expérimentales dans une large plage de températures (273–390 K). Les résultats ont montré que le modèle QSPR pour estimer la conductivité thermique était suffisamment précis et a donné des écarts inférieurs à tous les autres modèles QSPR existants avec ERAM = 1,97 et R^2=0,9919.

Lazzús (2015)[88] afin d'estimer la conductivité thermique de liquides ioniques à différentes températures et pressions a optimisé un réseau neuronal artificiel à l'aide d'un algorithme génétique en utilisant un ensemble de 400 points de données expérimentales de conductivité thermique de 41 liquides ioniques dans la plage de 0,10 à 0,22 $W.m^{-1}.K^{-1}$ pour une plage de température et de pression de (273-390 K) et de (100- 20, 000kPa) respectivement. Les résultats ont montré que la méthode proposée pour estimer la conductivité thermique des liquides ioniques à différentes températures et pressions présentait une bonne précision avec des écarts plus faibles tels que ERAM inférieur à 0,91 % et R^2 de 0,9969 pour l'ensemble d'entraînement, et ERAM inférieur à 0,84 % avec R^2 de 0,9963 pour l'ensemble de prédiction.

Lazzús (2015)[89] a réalisé une autre étude pour la prédiction de la conductivité thermique des liquides ioniques pour la même base de données et les mêmes plages de température et de pression utilisé dans son étude précédente. Les résultats ont montré que la méthode de contribution de groupe proposée pour prédire la conductivité thermique des

liquides ioniques en fonction de la température et de la pression a présenté une bonne précision avec des écarts plus faibles tels que ERAM inférieur à 1,90 % et R^2 de 0,9879 pour l'ensemble de corrélation, et ERAM inférieur à 2,33. % avec R^2 de 0,9754 pour l'ensemble de prediction.

Atashrouz et al.(2015)[90] ont proposé une méthode de groupe de modèle de traitement des données afin d'estimer la conductivité thermique des liquides ioniques purs, basée sur 23 liquides ioniques correspondant à 216 points de données expérimentales. L'écart relatif absolu moyen pour tous les systèmes étudiés était de 1,81 %, ce qui constitue un degré de précision satisfaisant pour le modèle proposé.

Atashrouz et al.(2016)[91] afin d'estimer la conductivité thermique des LIs purs ils ont proposé une machine à vecteurs de support aux moindres carrés, basée sur 22 liquides ioniques. Ils ont trouvé que le pourcentage d'erreur relative absolu moyen pour tous les systèmes étudiés est de 1,03 %, ce qui constitue un degré de précision satisfaisant pour leur modèle proposé. En outre, le modèle proposé a présenté une précision supérieure à celle des autres modèles disponibles dans la littérature.

He et al.(2017)[79] ont présenté un modèle linéaire basé sur les indices-normes pour décrire la relation entre la conductivité thermique des LIs et leur structure ainsi que la température et la pression. À cet égard, ils ont utilisé un ensemble total de données de 475 points de données expérimentales de conductivité thermique dans une large plage de températures (273,15 à 355,07 K) et de pressions (0,1 à 20,0 MPa) pour cinq types de liquides ioniques contenant de l'ammonium, du phosphonium, de l'imidazolium, du pyridinium et du pyrrolidinium. Les résultats ont révélé que cette méthode proposée présentait une bonne précision avec un F élevé de 1 540, un R^2 élevé de 0,984 et un ERAM faible de 1,45 %.

Soman et al.(2019)[92] pour la prédiction de la conductivité thermique du liquide ionique aqueux déterminées à différentes températures allant de 302,52 à 337,92 K et à des concentrations de 0,1 à 0,6 % sous une pression ambiante. Ils ont tout d'abord développé un modèle quadratique en utilisant la méthodologie de surface de réponse avec R^2 = 0,99 et EQM = 2,212. Ensuite un réseau neuronal à propagation rétroactive à deux couches 2-10-1 a été également modélisé avec une erreur quadratique moyenne (EQM), une racine moyenne carrée (EQRM) et un pourcentage d'erreur moyen absolu (ERAM) de 0,005, 0,071 et 2,684, respectivement. Ils ont constaté que les deux modèles sont explicites dans la prédiction des

conductivités thermiques expérimentales des liquides ioniques aqueux. Cependant, le modèle RNA a démontré de meilleures performances par rapport à l'autre modèle.

Cardona et Valderrama (2020)[93] dans la première partie de leur étude, ils ont proposé un modèle basé sur le concept de similitude géométrique pour calculer la conductivité thermique et la vitesse du son des liquides ioniques en utilisant 320 données de conductivité thermique (40 LIs) dans la plage de température de 263,15 K à 390,15 K et de pression de 0,86 à 2 000 bar. De sorte que, le modèle d'équation d'état empirique proposé ne contenait qu'un seul paramètre réglable pour chaque liquide ionique, paramètre qui peut être déterminé à l'aide d'une seule donnée expérimentale ou qui peut être généralisé en termes d'autres propriétés disponibles, comme proposé dans ce travail. Les résultats ont montré que le modèle d'équation d'état a fourni des résultats acceptables pour la corrélation et la prédiction de la conductivité thermique pour les liquides ioniques purs étudiés.

Cardona et Valderrama (2020)[94] dans la deuxième partie de leur étude, ils ont utilisé l'équation d'état générale basée sur le concept de similarité géométrique pour corréler et prédire la conductivité thermique et la vitesse du son des mélanges binaires eau + liquide ionique. De sorte que, l'extension considère les règles classiques de mélange et de combinaison couramment utilisées dans les études et les applications des équations d'état pression-température-volume. Pour cela, ils ont recueilli un total de 77 données sur la conductivité thermique de mélanges binaires eau + liquides ioniques dans la plage de température de 278 K à 343 K à pression atmosphérique. Les résultats de l'extension aux mélanges ont montré que le modèle empirique d'équation d'état a fourni des valeurs adéquates et cohérentes en utilisant uniquement des données de composants purs, prouvant ainsi sa bonne capacité prédictive.

Mousavi et al.(2020)[95] afin de prédire la conductivité thermique de la variété des liquides ioniques, ils ont développé trois modèles à savoir, (I) une méthode simple de contribution de groupe basée sur la température, la pression, le poids moléculaire, la température d'ébullition et le facteur acentrique ; (II) un modèle basé sur les propriétés thermodynamiques, la pression et la température des LIs ; et (III) un modèle basé sur la structure chimique, la pression et la température en utilisant un ensemble de 504 données expérimentales provenant de 50 LIs sur une large plage de pression (100 à 20 000 KPa) et de température (273 à 390 K). De plus, pour développer le modèle (I) ils ont utilisé une simple corrélation et pour le développement des modèles (II) et (III) ils ont appliqué des approches

intelligentes comprenant des réseaux neuronaux à fonction de base radiale (RBF) et à perceptron multicouche (MLP). En outre, pour l'optimisation des réseaux neuronaux afin de prédire avec précision la conductivité thermique, ils ont utilisé différentes techniques d'optimisation, notamment l'algorithme génétique (GA), l'algorithme de recherche gravitationnelle (GSA), le gradient conjugué mis à l'échelle (SCG), la régularisation bayésienne (BR) et l'algorithme de Levenberg-Marquardt (LM). Les résultats ont révélés que bien que la corrélation empirique ne prédise pas avec précision la conductivité thermique des LIs (erreur relative moyenne absolue en pourcentage (ERAM%) = 12,26 %), sa simplicité intrinsèque est pourtant précieuse et supérieure aux modèles de la littérature. Le modèle (II) (RBF-GSA avec un MAPRE% de 1,051) a montré de meilleures performances avec des prédictions plus précises. Le modèle (III) (RBF-GSA) a montré également une précision acceptable (ERAM % = 1,901). En outre, les résultats de comparaison avec les modèles/corrélations/contribution de groupe précédentes ont révélé que les modèles proposés sont plus précis et plus performants que les autres modèles de la littérature.

Zhang et al.(2021)[96] ont établi des modèles QSPR basés sur des descripteurs de normes pour prédire les propriétés des liquides ioniques à des températures et des pressions variables. À cet égard, ils ont développé un modèle f-T-P (variation de la propriété f avec la température T et la pression P) avec 608 points de données de 87 LIs pour la conductivité thermique. Ils ont constaté que le modèle QSPR présente une bonne stabilité et prévisibilité, avec des résultats statistiques satisfaisants.

Wan et al.(2022)[97] ont proposé un modèle de relation quantitative structure-propriété (QSPR) en utilisant l'approche de régression linéaire multiple (MLR) pour décrire la conductivité thermique(λ) de 44 liquides ioniques pour un ensemble de 606 points de données expérimentales où ils ont utilisé deux descripteurs pour la construction du modèle, à savoir la zone de distribution de la densité de charge des ions à un intervalle spécifique (Sσi) obtenue à l'aide du modèle de blindage de type conducteur pour le coefficient d'activité du segment (COSMO-SAC) et le volume de la cavité des LIs (Vcosmo). Ils ont constaté que pour le modèle construit à partir des données totales, ses coefficients de détermination (R^2), son erreur quadratique moyenne (EQRM) et son écart relatif absolu moyen (ERAM) sont respectivement de 0,9713, 0,004304 et 2,18 % ; ainsi, le modèle λ-QSPR développé offre une prédiction relativement bonne de conductivité thermique pour les LIs.

Mital et al.(2022)[98] aprés une étude expérimentale réalisé à l'aide de la méthode de source plane transitoire modifiée (MTPS) pour mesurer la conductivités thermiques de neuf liquides ioniques différents sur la plage de température de 298 à 348 K à pression atmosphérique. Ils ont étendu deux modèles de contribution de groupe (GCM) existants avec des philosophies de modélisation différentes par rapport à une base de données composée de ces nouvelles données combinées à toutes les données de conductivité thermique des LIs existantes dans la littérature où Les GCM étendus ont donné des écarts relatifs absolus moyens de 5,5 % et 4,9 %, respectivement, pour l'ensemble de l'ensemble de données.

I.6 Revue critique sur la tension superficielle des LIs

La tension superficielle est l'une des caractéristiques physicochimiques les plus importantes à prendre en compte dans les systèmes hétérogènes, la conception de procédés et de nombreuses applications scientifiques et industrielles. Dans les processus chimiques, la tension superficielle est une propriété appropriée car elle fait référence à un phénomène dans lequel il existe une interface vapeur-liquide. Dans les sciences de l'ingénieur, la tension superficielle affecte non seulement la production de certains processus tels que l'extraction, l'adsorption et la distillation, mais domine également les caractéristiques de nombreux produits industriels, tels que les produits agrochimiques, les cosmétiques, les peintures, les revêtements et les détergents. Étant donné que la tension superficielle est une propriété physique qui dépend de la structure et de la composition des molécules, l'étude de la tension superficielle des LIs peut jouer un rôle central dans l'explication des énergies intrinsèques impliquées dans les interactions ioniques existants dans diverses applications.[99]

Gardas et Coutinho(2008)[100] ont montré dans leur étude comment une corrélation quantitative structure-propriété (QSPR) pour les parachors peut être utilisée avec une méthode d'estimation des densités pour prédire les tensions superficielles des LIs dans de larges plages de température, 268,15–393,15 K. La validité de cette corrélation a été évaluée en comparant les prédictions aux tensions superficielles expérimentales recueillies de la littérature. Les résultats obtenus ont montré un bon accord avec les données de la littérature. Pour une base de données de 361 points de données, pour environ 40 liquides ioniques, un pourcentage d'écart moyen de 5,75 % avec un écart maximum inférieur à 16 % a été observé. Pour l'estimation des tensions de surface à température ambiante, ils ont développé une corrélation des tensions de surface avec les volumes moléculaires des LIs. L'écart observé est de 4,50 % pour 47

points de données de tension superficielle disponibles dans la littérature à 298,15 K pour un total de 22 LIs à base d'imidazolium

Gharagheizi et al.(2012)[101] ont développé une méthode de contribution de groupe fiable pour l'estimation de la tension superficielle des liquides ioniques. Un ensemble de données comprenant 920 données expérimentales pour 51 liquides ioniques collectées à partir de diverses références a été utilisé pour développer et évaluer la capacité prédictive du modèle. L'ensemble de données couvre une plage de température de 268,3 à 743,7 K et une plage de tension superficielle de 0,02499 à 0,0647 N m^{-1}. Où Ils ont utilisé un total de 19 sous-structures plus la température pour prédire la tension superficielle. Pour mieux distinguer les effets de l'anion et du cation sur la tension superficielle des liquides ioniques, 12 sous-structures liées à la structure chimique de l'anion et 7 sous-structures liées à la structure chimique du cation ont été mises en œuvre. Les résultats de cette méthode ont montré un écart relatif absolu moyen (ERAM) de 3,6 % par rapport aux données expérimentales.

Wu et al. (2012)[102] ont proposé une méthode de contribution de groupe d'état correspondant pour estimer la tension superficielle des liquides ioniques couvrant une large plage de température (268,29–533,20 K) et de tension superficielle (15,50–76,20 mN/m) sur la base de données expérimentales recueillies de la littérature. Ils ont obtenu un bon accord entre les données prédites et expérimentales de la littérature pour un total de 1224 points de données de 105 liquides ioniques avec un écart absolu moyen (ERAM) de 4,95 %.

Shang et al. (2012)[103] ont appliqué les indices topologiques généraux (TIs) pour prédire les tensions de surface des liquides ioniques par la méthode relation quantitative structure-propriété (QSPR). Dans cette étude, outre les deux ensembles de TIs générés à partir du cation et de l'anion, un troisième TIs a été utilisé pour décrire l'interaction de l'anion et du cation. Contrairement aux méthodes d'estimation précédentes pour les tensions de surface des LIs, la dépendance à la température a été également corrélée par les paramètres de structure. Pour développer le modèle QSPR 930 points de données de tension superficielle pour 115 LIs ont été appliqués où le coefficient de corrélation (R^2) et l'écart absolu moyen global (ERAM) sont respectivement de 0,9516 et 3,43 %.

Mirkhani et al.(2013)[104] ont développé des modèles QSPR (Quantitative Structure Property Relationship) pour estimer la tension superficielle des liquides ioniques en utilisant un modèle basé sur la machine à vecteurs pris en charge par les moindres carrés (LSSVM) pour capturer la nature non linéaire de la tension superficielle. Des LIs via des descripteurs

moléculaires théoriques obtenus uniquement à partir de leurs structures. Un grand ensemble de données de 930 points de données de tensions superficielles expérimentales pour 48 liquides ioniques a été appliqué pour dériver le modèle. Sept descripteurs ont été sélectionnés par approximation des fonctions génétiques pour relier la tension superficielle des liquides ioniques à leurs anions et structures cationiques correspondantes où les résultats des techniques de validation ont montré que le modèle est stable et précis et qu'il est exempt de corrélation fortuite tel que le coefficient de corrélation (R^2) et l'écart relatif absolu (ERAM %) qui sont 0.989 et 1.03 respectivement.

Ghasemian et Zobeydi (2013)[5] ont utilisé une équation simple pour prédire les tensions superficielles de liquides ioniques purs sur la base de l'enthalpie de vaporisation. Le facteur acentrique, la température critique et la pression critique des liquides ioniques sont des entrées nécessaires pour cette prédiction. L'erreur relative moyenne obtenue à partir de la comparaison des valeurs de tension superficielle expérimentales et calculées pour les liquides ioniques étudiés a été inférieure à 2 % démontré que le modèle présente une bonne précision par rapport à d'autres équations prédictives.

Lemraski et Zobeydi (2014)[105] ont appliqué la méthode Parachor basée sur la contribution de groupe modifiée pour estimer la tension superficielle des liquides ioniques à base d'imidazolium couvrant une large plage de température et de longueur de chaîne, sur la base de données expérimentales collectée de la littérature. À partir de la comparaison des valeurs de tensions superficielles expérimentales et calculées pour les liquides ioniques étudiés ils ont obtenu une erreur relative moyenne (moins de 8,5 %) et a montré que le modèle a une bonne précision par rapport à d'autres équations prédictives.

Lemraski et pouyanfar (2014)[106] ont prédit la tension superficielle des systèmes binaires liquides ioniques à base d'imidazolium en fonction de la concentration en utilisant l'équation de l'équation de Sprow et Prausnitz à 298,15 K. la fraction molaire en vrac, la surface molaire de surface et la tension superficielle du composant pur sont des entrées nécessaires pour cette équation et la méthode de contribution du groupe UNIFAC a été également appliquée pour tenir compte de la non-idéalité de la masse et de la surface de la solution. Les erreurs relatives moyennes obtenues à partir de la comparaison des valeurs de tensions superficielles expérimentales et calculées pour 16 systèmes binaires sont inférieures à 4 %. Par conséquent, ce modèle présente une bonne précision par rapport à d'autres équations prédictives.

Huang et al. (2015)[41] ont utilisé des modèles semi-empiriques et des modèles de réseaux de neurones artificiels (RNA) pour un ensemble de 545 points de données, ils ont constaté que les deux modèles ont donné des résultats de corrélation très précis avec des écarts absolus moyens globaux (ERAM) des modèles semi-empiriques et RNA inférieurs à 2 %.

Hashemkhani et al. (2015)[11] ont examiné la capacité et la faisabilité de trois algorithmes intelligents pour prédire la tension superficielle des systèmes binaires contenant des LIs. Pour construire et tester les modèles, 748 points de données correspondants aux valeurs expérimentales de tensions superficielles des mélanges binaires contenant des LIs ont été extraits de la littérature. La tension superficielle était comprise entre 0,0157 et 0,07185 $N \cdot m^{-1}$. La température absolue, la fraction molaire et le poids moléculaire des composants LIs et la densité des composants LIs ainsi que le point d'ébullition et le poids moléculaire des composants non-LIs ont été considérés comme variables d'entrées du modèle pour différencier les différents composés impliqués dans les systèmes binaires. Une comparaison entre les données expérimentales et les valeurs prédites utilisant les trois méthodes (en termes de paramètres statistiques) a montré un bon accord. Cependant, la prédiction CSA-LSSVM était meilleure que les deux autres approches avec un écart relatif moyen absolu (ERAM %) et un coefficient de corrélation (R) de 1.3785 %, 0.987044 respectivement.

Atashrouz et al(2016)[107] ont développé des approches informatiques souples. Ce sont des modèles basés sur la machine à vecteurs de support (SVM), les machines à vecteurs de support des moindres carrés (LSSVM) et la méthode de groupe des systèmes de réseau neuronal polynomial de type traitement de données (GMDH-RNP) pour la prédiction des propriétés physicochimiques des mélanges des LIs telle que la tension superficielle. Ils ont trouvé que le modèle LSSVM est plus robuste et fiable pour la prédiction des propriétés physico-chimiques des mélanges LIs avec des écarts relatifs absolus moyens de1,18 % pour la tension superficielle.

Soleimani et al. (2017)[108] ont comparé le modèle RNA entraîné avec l'algorithme trainbr par d'autres modèles tels que SVM, GA-SVM, GA-LSSVM, CSA-LSSVM, GMDH-RNP et RNA entraînés avec l'algorithme trainlm pour la prédiction de la tension superficielle, ils ont conclu que le modèle proposé était meilleur en termes de précision avec un ERAM % moyen de 0,44.

Atashrouz et al.(2017)[109] ont développé un modèle mathématique basé sur la machine à vecteurs de support des moindres carrés pour prédire la tension superficielle des liquides ioniques. Le poids moléculaire, la température réduite, la pression réduite, le facteur de compressibilité critique et le facteur acentrique ont été sélectionnés comme paramètres d'entrées et la tension superficielle comme paramètre de sortie. Une base de données étendue comprenant 868 points de données expérimentales pour la tension superficielle de 61 LIs a été prise en compte pour développer le modèle LSSVM dont les paramètres réglables du modèle ont été calculés par une programmation d'algorithme génétique. Ils ont conclu que le modèle proposé a présenté une bonne précision avec un écart relatif absolu moyen de 1,41, 1,74 et 1,54 % pour l'ensemble d'apprentissage, l'ensemble de test et les données totales, respectivement. La majorité des points de données sont situés dans un domaine d'erreur standard démontrant l'acceptabilité statistique du modèle et que seulement 3,6 % des points de données sont considérés comme des données suspectes. Et partir d'une étude d'importance relative ils ont trouvé que le facteur de compressibilité critique des LIs a le plus grand effet sur la tension superficielle des liquides ioniques, avec une importance relative de 22,97 %.

Lazzús et al.(2017)[110] pour estimer la tension superficielle des liquides ioniques en fonction de la température absolue ont utilisé une méthode de contribution de groupe. Un total de 2 286 points de données expérimentales provenant de 226 ensembles de données de 154 liquides ioniques a été collectés de la littérature spécialisée dans une plage de température de 263 à 533 K et une plage de tension superficielle de 0,015 à 0,062 N•m−1, pour un ensemble hétérogène de types de liquides ioniques tels que l'imidazolium, l'ammonium, le phosphonium, le pyridinium, le pyrrolidinium et le pipéridinium. Un ensemble de corrélation contenant 75 % de la base de données globale a été utilisé afin de calculer les valeurs de contribution de 10 groupes de cations structuraux et de 30 groupes d'anions structuraux. Ensuite, un ensemble de prédiction contenant les 25 % restants de la base de données globale avec des ensembles de données non utilisées dans la phase de corrélation a été utilisé afin de tester les capacités et l'exactitude de la méthode proposée. Les résultats ont montré que la méthode proposée a permis d'estimer la tension superficielle de plusieurs liquides ioniques avec une meilleure précision que les autres méthodes disponibles dans la littérature, avec un écart relatif absolu moyen de 2,8 % et un coefficient de corrélation de 0,98.

Shahsavari et al(2018)[111] ont développé un modèle de contribution atomique simple général pour prédire la tension superficielle des liquides ioniques purs (LIs) à pression atmosphérique en fonction de la température et du nombre d'atomes dans le LI (carbone

anionique, carbone cationique, oxygène, azote, etc.). La corrélation a été développée avec 1246 points de données de 61 LIs, dont plusieurs contenaient différents anions et cations. Afin de trouver les constantes de la relation de fonctionnalité proposée grâce à l'algorithme de recuit simulé (SA), ils ont divisé la base de données au hasard en deux sous-ensembles de données de formation et de test avec des ratios de 70 et 30 %, respectivement. Les résultats d'évaluation statistique approfondie ont indiqué que la corrélation proposée avec l'erreur quadratique moyenne relative (EQRM) de 2,232, le pourcentage d'écart relatif absolu moyen (ERAM %) de 4,16 et le coefficient de corrélation (R2) de 0,9190 est très robuste et présente une tendance d'erreur acceptable. Ainsi que les valeurs de validation externe de 15 LI non inclus dans le processus d'optimisation étaient en accord acceptable avec les données expérimentales correspondantes bien que certaines de ces LIs comprenaient des cations et des anions non utilisés dans le développement de la corrélation.

Esmaeili et Hashemipour (2018)[112] pour générer un réseau et obtenir une corrélation simple et précise pour prédire la tension superficielle des LIs ont utilisé une des techniques les plus puissantes du soft computing, la programmation génétique multigénique où les valeurs de température réduite, de pression réduite, de facteur de compatibilité critique et de facteur acentrique ont été sélectionnées comme paramètres d'entrées du réseau. La corrélation obtenue a une forme mathématique simple, fonction de la température réduite avec une bonne précision (R^2=0,99). Cette corrélation comporte trois coefficients, qui peuvent être déterminés à l'aide de GA ou d'un simple ajustement de courbe, ou peuvent être trouvés dans cette étude pour certains des liquides ioniques importants. L'autre méthode proposée pour déterminer les coefficients consiste à utiliser six corrélations présentées dans ce travail.

Jarvas et al. (2018)[112] afin de construire un modèle QSPR précis et fiable pour l'estimation de la tension superficielle dépendant de la température des liquides ioniques, ont utilisé les moments sigma de base COSMO-RS comme descripteurs moléculaires et le réseau neuronal artificiel comme modèle de régression. Ils ont confirmé la précision et la robustesse du modèle par l'analyse statistique des ensembles de données d'apprentissage, de validation et de test, qui ont montré clairement le pouvoir prédictif du modèle proposé sur une large plage de valeurs de tension superficielle. Ils ont aussi validé le pouvoir de prédiction du modèle avec un ensemble de données externes, avec un coefficient de corrélation R^2 et une erreur absolue moyenne EAM de l'ordre de 0,97 et 1 mNm^{-1} respectivement où ils ont conclu que le modèle non linéaire proposé a constitué un outil optimal pour prédire la tension superficielle des liquides ioniques.

Shojaeian et Asadizadehb (2019)[113] ont étudié l'effet des paramètres d'entrées sur les performances du réseau neuronal artificiel (RNA) pour l'estimation de la tension superficielle du mélange binaire de 33 liquides non ioniques + ioniques. Pour cela ils ont collecté un ensemble de données de 1537 points de données binaires de tension superficielle à la pression atmosphérique ou ils ont considéré 8 approches différentes pour les paramètres d'entrées en utilisant les propriétés physiques telles que la température, la température réduite, la température critique, la pression critique, le volume critique, le poids moléculaire, le facteur acentrique et le facteur de compressibilité critique ainsi que deux règles de mélange différentes. L'erreur quadratique moyenne (EQRM) de l'approche la plus précise, obtenue est de 7,88 × 10−3. En outre, la comparaison entre les résultats de RNA et les résultats de cinq autres méthodes d'intelligences différentes telles que ANFIS, ANFIS-ACO, ANFIS-DE, ANFIS-GA et ANFIS-PSO appliqué pour estimer la tension superficielle des mélanges binaires sur la base des meilleurs paramètres d'entrées a montré que la méthode RNA a donné une meilleure précision.

Cardona et Valderrama(2020)[114] ont développé un modèle basé sur une équation d'état cubique et le concept de similarité géométrique pour prédire la tension superficielle des substances pures et des mélanges contenants des substances organiques, de l'eau et des liquides ioniques de 90 mélanges binaires (2660 données) et 12 mélanges ternaires (467 données) considérés dans une large gamme de température de 278,15 à 348,15 K. L'étude a conclu que le modèle est précis pour l'estimation primaire dans la conception et la simulation de processus.

Esmaeili and Hashemi (2020)[115] ont utilisé la programmation génétique multigène (MGGP), une des techniques les plus puissantes de soft computing pour la prédiction de la tension superficielle du mélange en utilisant la fraction molaire, la température, le poids moléculaire de deux composants et le point d'ébullition comme paramètres d'entrées du réseau. Les résultats de cette étude ont révélés que le modèle proposé a atteint un écart relatif absolu moyen acceptable (ERAM) et le coefficient de corrélation (R^2) qui sont de 0,9567 % et 0,994 respectivement.

Wang et al.(2020)[116] ont prédit la tension superficielle des liquides ioniques de 1-Ethyl-3-methylimidazolium éthyl sulfate (EMIES) aqueux en utilisant l'équation empirique améliorée d'Eötvös sur la base de la définition de l'énergie de Gibbs de la surface molaire où les valeurs prédites de la tension superficielle des EMIES aqueux étaient très cohérentes avec

les valeurs expérimentales de la tension superficielle. Le coefficient de corrélation trouvé était de 0,998.

Paduszyński(2021)[117] a proposé des relations quantitatives structure-propriété (QSPR) pour calculer la dépendance à la température de la tension superficielle (σ) des liquides ioniques en termes de contributions de groupe (GC). Il a appliqué une méthode d'apprentissage statistique comprenant une régression linéaire multiple pas à pas et deux méthodes d'apprentissages automatiques, notamment un réseau neuronal artificiel à action directe et une machine à vecteurs de support des moindres carrés pour exprimer la tension superficielle σ en fonction des contributions de groupe GC de 570 LIs. L'étude réalisée a démontré également que la modélisation de la tension superficielle en termes de parachor semble être la voie la plus efficace, quelques soit la complexité de la méthode informatique adoptée pour calibrer le modèle.

Mousavi et al.(2021)[99] ont utilisé deux modèles mathématiques basés sur la fonction de base radiale (RBF) et la machine vectorielle des moindres carrés (LSSVM), couplés à deux algorithmes d'optimisation, nommément algorithme de luciole (FFA) et d'évolution différentielle (DE) pour prédire la tension superficielle des LIs en fonction de leur structure chimique et de leur température, à l'aide d'un large ensemble de données comprenant 1042 données expérimentales provenant de 69 LIs dans une plage de tension superficielle et la température de 18,5 à 70,3 mN/m et de 268,29 à 532,4 K, respectivement, où la pression est constante (0,101 MPa). Afin d'évaluer les performances et l'exactitude des modèles proposés ils ont utilisé l'analyse de sensibilité, les erreurs statistiques et graphiques. Les résultats ont montré que le modèle LSSVM-FFA prédit avec précision la tension superficielle des LIs avec un pourcentage d'écart relatif absolu moyen de 1,8440 % et un coefficient de détermination de 0,9828.

Setiawan et al.(2021)[17] ont appliqué un modèle de réseaux de neurones artificiels RNA avec quatre algorithmes d'optimisation à savoir l'optimisation basée sur l'enseignement-apprentissage (TLBO), l'optimisation par essaim de particules (PSO), l'algorithme génétique (GA) et l'algorithme de concurrence impérialiste (ICA) pour la prédiction de la tension superficielle (TS) dans des mélanges binaires impliquant 31 liquides ioniques différents en fonction de la température de fonctionnement et des spécifications des LIs et des non- LI. En utilisant 748 points de données collectés à partir de différentes ressources documentaires comme ensembles de formation et de test. Les analyses statistiques ont indiqué que le TLBO-

RNA était le plus précis parmi les modèles proposés, comme le confirme l'approche mathématique à effet de levier. Basé sur une analyse statistique, ce modèle a la capacité de prédire les valeurs de laboratoire avec R^2 = 0,998 et EQM = 0,0000002. Ils ont réalisé ensuite une analyse de sensibilité qui a illustré que la fraction molaire de l'IL et le poids moléculaire de la non- LI comme étant respectivement le facteur le plus efficace et le moins efficace sur la ST.

Ge et al.(2022)[118] ont proposé un modèle à grains grossiers à variable électrostatique (VaCG) pour prédire les propriétés physiques de liquide ionique aqueux de l'eau/[BMIM]+[BF4]- des mélanges binaires et du [BMIM]+[SCN] pur- en introduisant une concentration ou une variable dépendante de l'environnement. Ce modèle gère l'effet de polarisation sur la plage de fraction molaire d'eau x = 0,100 % en mole avec les charges partielles explicites et le traitement implicite du milieu continu à longue portée sans augmenter le coût de calcul. Ils ont constaté que le modèle VaCG a de bonnes performances pour prédire les propriétés thermodynamiques et les propriétés diélectriques des solutions aqueuses.

Baskin et al.(2022)[119] ont réalisé une étude comparative à grande échelle des modèles de relations quantitatives structure-propriété (QSPR) construites en combinant trois méthodes traditionnelles de ML et des réseaux de neurones avec sept architectures différentes avec cinq types de représentations moléculaires (sous la forme de descripteurs moléculaires numériques ou de chaînes de texte SMILES) pour prédire six propriétés physiques importantes des LIs parmi eux tension superficielle. Les ensembles de données constitués de 407 à 1204 LI divers composés de divers ions organiques et inorganiques. Ils ont conclu que :(i) les méthodes de ML non linéaires fonctionnent bien mieux que les méthodes linéaires, (ii) les réseaux de neurones fonctionnent mieux que les méthodes de ML traditionnelles, (iii) les transformateurs, qui sont activement utilisés dans le traitement du langage naturel, fonctionnent mieux que les autres types de réseaux neuronaux en raison de leur capacité avancée à analyser les structures chimiques des LIs codées dans les chaînes de texte SMILES.

Abourehab et al. (2022)[120] ont développé des modèles prédictifs basés sur l'apprentissage automatique pour calculer la tension superficielle de deux LIs couramment utilisés, notamment le 1-octyl-3-méthylimidazolium bis[(trifluorométhyl) sulfonyl] imide et le 1-butyl-2,3-diméthylimidazolium bis[(trifluorométhyl) sulfonyl] imide. Pour cela, ils ont fait un ensemble de données comportant de nombreuses fonctionnalités d'entrées ou ils ont utilisé

trois algorithmes d'optimisation distincts pour la sélection des fonctionnalités avec l'algorithme KNN, notamment l'optimisation GA, PSO et FOA. Ils ont trouvé que GENETIC-KNN, FOA-KNN et PSO-KNN avaient des scores R2 de 0,962, 0,941 et 0,827 respectivement. Avec la métrique MAPE, leurs taux d'erreur sont de 1,67E-02, 2,13E-02 et 4,26E-02 dans le même ordre. Au final, GENETIC-KNN a été déterminé comme étant le modèle le plus efficace dans cette enquête.

Alshahrani et al.(2022)[121] ont développé trois modèles prédictifs basés sur l'intelligence artificielle (IA) pour mesurer la tension superficielle de deux liquides ioniques LIs répandues, notamment le méthylsulfate de 1,3-diméthylimidazolium et l'éthylsulfate de 1-éthyl-3-méthylimidazolium. Pour créer des modèles de prédiction, ils ont utilisé Lasso, K-Nearest-Neighbour (KNN) et la régression linéaire (RL) comme modèles de base d'un ensemble de régression par bagging. En utilisant un ensemble de données comportant 45 caractéristiques d'entrées et une caractéristique de sortie (tension superficielle) pour découvrir la meilleure combinaison de fonctionnalités. A cet effet, ils ont utilisé l'optimisation des colonies de fourmis comme stratégie de recherche pour découvrir la meilleure combinaison de fonctionnalités. Les modèles ont été testés à l'aide de diverses mesures standards. Bagging RL, Bagging Lasso et Bagging KNN ont illustré des valeurs de R^2 de 0,934, 0,915 et 0,977 respectivement ou ils ont sélectionné Bagging KNN comme modèle principal car il comportait moins d'erreurs et un R2 plus élevé.

Fu et al.(2022)[122] ont développé un modèle RNA-GC pour prédire la tension superficielle des systèmes mixte LI-H2O à différentes températures en combinant la méthode de contribution de groupe (GC) et l'algorithme de réseau neuronal artificiel (RNA). Pour ce faire, ils ont collecté une base de données pertinentes sur la tension superficielle de 4 578 points de données pour 138 systèmes hybrides liquides ioniques (LI)-H2O. Les résultats ont montré que les modèles avec 4, 5 et 7 neurones dans la couche cachée fournissaient tous des prédictions fiables pour la tension superficielle du système mixte LI-H2O avec des coefficients de corrélation (R^2) de l'ensemble d'apprentissage de 0,92, 0,93 et 0,93 ; et des MAE sont respectivement de 0,0021, 0,0020 et 0,0020. Alors que les coefficients de corrélation (R^2) de l'ensemble de test sont de 0,94, 0,91, 0,91 et que le EAM est de 0,0023, 0,0029 et 0,0027, respectivement. De plus, le modèle RNA-GC développé pour le système hybride a également été vérifié et extrapolé aux systèmes LIs purs sur la base de ces 172 points de données. Ils ont trouvé que le modèle RNA-GC développé pour le système LIs pur a atteint un EAM inférieur à celui du système hybride. Néanmoins, le modèle RNA-GC qui ils

ont développé pour les systèmes hybrides avec 4 et 5 neurones dans la couche cachée peut toujours être appliqué pour la prédiction de la tension superficielle au système LIs pur.

Mohan et al. (2023)[123] ont utilisé un modèle d'apprentissage automatique (ML) basé sur un arbre de décision (algorithme GBT) pour la prédiction de de la tension superficielle, sur une large plage de températures et de pressions en utilisant des descripteurs moléculaires dérivés du COSMO-RS, un modèle basé sur la chimie quantique. Le modèle ML a montré un excellent accord avec les observations expérimentales, avec une valeur de R^2 de 0,96 et un RMSE de 1,71 mN/m. Dans l'ensemble, ces résultats ont souligné l'utilité des profils σ COSMO-RS en tant que descripteurs moléculaires utiles des LIs.

Mohan et al.(2023)[124] dans une autre étude ont examiné comment les différentes architectures de ML, allant des approches arborescentes aux réseaux de neurones artificiels à action directe, fonctionnent pour générer des modèles quantitatifs multivariés non linéaires de relation structure-propriété pour la prédiction de la tension superficielle dépendante de la température et de la pression sur une large gamme de (16,9 à 76,2 mN/m). Ensuite, ils ont interrogé les modèles ML à l'aide de la puissante méthode d'interprétation, les explications additives Shapley où ils ont constaté que plusieurs modèles de ML différents ont offert une grande précision, selon les mesures statistiques traditionnelles. Les approches basées sur des arbres de décision semblent être les plus précises, les arbres à forte accélération du gradient et les arbres à accélération du gradient étant les plus performants.

Afsharpour(2023)[125] a utilisé l'équation d'état (EoS) de l'association cubique plus (CPA), seule et en combinaison avec la théorie du volume libre (FVT) et les modèles Pelofsky modifiés pour estimer certaines propriétés thermophysiques parmi elles; la tension superficielle de sept liquides ioniques. Selon les résultats obtenus, les approches de corrélation et de prédiction ont présenté un écart moyen absolu (ERAM %) égal à 0,1632 et 1,3820 pour la tension superficielle. Comme le montre les résultats, les approches corrélatives et prédictives ont démontré une bonne précision. L'inter/extrapolation des résultats pourraient donc être réalisées pour d'autres LIs de cette famille.

I.7 Revue critique sur la solubilité de CO_2 dans les LIs

La solubilité des gaz purs constitue également un sujet prometteur pour étudier les performances de solvatation des composants dans les LIs, qui dépend des conditions de pression, de température, d'interactions attractives/répulsives et du volume libre au sein de LI.

Par conséquent, une prévision précise de la solubilité des gaz dans les LIs est cruciale pour leur application dans l'industrie[126].

Kim et al.(2005)[127] après une étude expérimentale de la solubilité du dioxyde de carbone dans les liquides ioniques, [bmim][PF6], [C6mim][PF6], [emim][BF4], [C6mim][BF4], [emim][Tf2N] et [C6mim][Tf2N] à 298,15 K et jusqu'à 1,0 MPa, ils ont appliqué une forme de contribution de groupe d'un modèle réseau-fluide non aléatoire pour prédire les solubilités du dioxyde de carbone dans les liquides ioniques. Ils ont constaté que les accords entre les données et les solubilités calculées se sont révélés généralement bons pour les liquides ioniques contenant des cations 1-alkyl-3-méthylimidazolium et [PF6], [BF4] et les anions [Tf2N].

Shiflett et Yokozeki (2005)[128] ont corrélé avec succès pour la première fois à l'aide d'une simple équation d'état cubique (EOS) Les données expérimentales sur la solubilité du CO_2 dans les liquides ioniques déterminées à partir des données de saturation d'absorption (équilibre) à chaque température et pression fixes après l'étude expérimentale réalisée où ils ont trouvé que certain comportement de phase des mélanges de CO_2 et de liquides ioniques ainsi que les constantes de la loi de Henry et le changement de volume des solutions ont été bien prédits avec le modèle EOS actuel. Ils ont obtenu pour la première fois en utilisant un modèle de diffusion simple à partir de données d'absorption dépendant du temps, des coefficients de diffusion binaires du CO_2 dans les liquides ioniques.

Kim et al. (2006)[129] ont appliqué une équation d'état de fluide de réseau à contribution de groupe pour prédire la solubilité du dioxyde de carbone pur et du propane et celles des gaz mixtes dioxyde de carbone + azote et dioxyde de carbone + propane dans le liquide ionique 1-hexyl-3-méthylimidazolium tri(bisfluorométhylsulfonyl)imide, Ces composés ont été étudié expérimentalement à 298,15K et à 1MPa. Ils ont trouvé que les résultats prédits étaient en assez bon accord avec les données expérimentales.

Karakatsani et al.(2007)[130] ont utilisé les paramètres tPC-PSAFT pour les LIs à base d'imidazolium ajustés aux données de densité liquide et d'ordre de grandeur de la pression de vapeur pour la corrélation de la solubilité du dioxyde de carbone dans diverses LIs. De plus, les solubilités du monoxyde de carbone, de l'oxygène et du trifluorométhane dans l'hexafluorophosphate de 1-butyl-3-méthylimidazolium à faible et haute pression. Ils ont constaté que dans tous les cas, un très bon accord entre la corrélation tPC-PSAFT et les données expérimentales des mélanges.

Andreu et Vega (2008)[131] ont utilisé un nouveau modèle simple pour les LIs d'alkylimidazolium-[Tf2N] dans le contexte de soft-SAFT pour vérifier la capacité de soft-SAFT à capturer la solubilité de différents gaz dans ces LIs. Pour imiter les principales interactions entre l'anion et le cation pour les LIsalkylimidazolium-[Tf2N] ils ont utilisé un site d'association pour représenter une force d'attraction à courte portée et hautement directionnelle entre les molécules. Ils ont estimé les paramètres du modèle pour les liquides ioniques à l'aide de données de densité expérimentale provenant de différents auteurs et ils ont utilisé le nouvel ensemble de paramètres pour étudier le comportement de solubilité de l'hydrogène, du dioxyde de carbone et du xénon dans ces LIs sur une large plage de températures et de pressions. Les résultats ont montré que soft-SAFT a pu capturé le comportement des LIs complexes avec une approche simple, ce qui a permis de développer des modèles plus raffinés et détaillés pour des LIs plus complexes, de manière systématique tel que la solubilité de l'hydrogène, du dioxyde de carbone et du xénon dans la famille des LIsimidazolium-[Tf2N] a été décrite avec précision sans la nécessité d'interactions faibles explicites. En plus, le modèle a pu reproduire avec précision la solubilité de l'hydrogène à différentes températures de manière purement prédictive.

Carlisle et al. (2008)[132] ont utilisé une méthode de contribution de groupe pour prédire les paramètres de solubilité des RTIL fonctionnalisés, et ces valeurs ont été utilisées avec la théorie des solutions régulières pour prédire la solubilité et la sélectivité des trois gaz du CO_2, du N_2 et du CH_4 à TPN. Ils ont trouvé que ces valeurs prédites de solubilité des gaz se sont révélées être en bon accord avec celles mesurées expérimentalement. De plus, l'inclusion des groupes fonctionnels nitrile et alcyne augmentaient le paramètre de solubilité par rapport aux RTIL analogues substitués par n-alkyle. Ces résultats ont montré que la méthode de contribution de groupe a permis d'offrir un guide précieux pour la conception systématique de RTIL fonctionnalisés avec des performances spécifiques de solubilité des gaz et de sélectivité.

Zhang et al. (2008)[133] ont proposé une méthode de criblage pour la conception moléculaire de liquides ioniques afin de capturer le dioxyde de carbone (CO_2), basée sur la prédiction de la constante de la loi d'Henry en utilisant la méthode COSMO-RS sans aucun ajustement de paramètre spécifique où ils ont criblé 408 types de LIs, avec 24 cations et 17 anions. Ils ont constaté que les LIs contenant l'anion tris (pentafluoroéthyl) trifluorophosphate ([FEP]) ont présenté une capacité améliorée à capturer le CO_2 par rapport aux autres LIs

rapportées dans la littérature. Les résultats de la prédiction sont en bon accord avec ceux expérimentaux, indiquant ainsi que la méthode proposée est fiable.

Shimoyama et Ito (2010)[134] ont prédit les solubilités de CO_2, N_2 et CH_4 dans le liquide ionique à l'aide d'un modèle de coefficient d'activité basé sur COSMO, la méthode COSMO-SAC dans lequel ils ont ciblé les 1-alkyl-3-méthylimidazolium comme cations, avec les espèces d'anions comprenant le tétrafluoroborate, l'hexafluorophosphate, le triflate, le dicyanamideet le bis(trifluorométhane)-sulfonimide. Ils ont trouvé que les résultats prédits des solubilités du CO_2 dans les liquides ioniques par la méthode COSMO-SAC sont en accord avec les données expérimentales dans la limite d'un écart moyen de 0,0017 en fraction molaire.

Yokozekia et Shiflett(2010)[135] ont appliqué l'équations d'état volumétrique original de van der Waals pour modéliser et prédire la solubilité des gaz ((équilibre -vapeur-liquide(EVL), équilibres triphasés vapeur-liquide-liquide (EVLL) et séparation liquide-liquide (ELL)) dans des liquides ioniques à température ambiante. Lesrésultats de modèle EOS ont montré que les données expérimentales sur la solubilité des gaz (CO_2, $CF_3 – CFH_2$, SO_2 et NH_3) dans les liquides ioniques à température ambiante sont bien corrélées, et également les comportements de phase tels que ELL (séparations liquide-liquide) sont prédits de manière satisfaisante.

Eslamimanesh et al. (2011)[136] ont appliqué avec succès un réseau neuronal artificiel FeedForward pour l'estimation de la solubilité du dioxyde de carbone supercritique dans divers liquides ioniques. Pour cela, ils ont développé un réseau neuronal optimisé à trois couches en utilisant les propriétés critiques des liquides ioniques ainsi que la température et la pression de fonctionnement comme paramètres d'entrées ou l'application de ce modèle pour 1128 points de données de 24 liquides ioniques a montré des coefficients de corrélation quadratiques de 0,993 et un écart absolu moyen de 3,6 % par rapport aux valeurs expérimentales pour les solubilités calculées montrant ainsi une capacité de prédiction.

Kima et al.(2011)[137] ont utilisé l'équation d'état des fluides de réseau non aléatoires à contribution de groupe (GC-NLF EoS) proposée par Lee et ses collaborateurs , pour corréler les données de la solubilité du dioxyde de carbone dans les liquides ioniques 1-butyl-3-méthylimidazolium hexafluorophosphate et 1-nonyl-3-méthylimidazolium hexafluorophosphate mesurées à des températures de 293,15 et 298,15 K et à une pression allant jusqu'à 4 MP. Ils ont légèrement modifié les paramètres de groupe pour CG-NLF EoS

sur une plage limitée pour tenir compte des données expérimentales récentes et d'une meilleure prévision à haute pression et à longue chaîne alkyle. Les résultats ont montré que la matrice des paramètres d'interaction de groupe peut être améliorée si davantage de données expérimentales sont utilisées dans le processus d'ajustement et qu'elle peut être applicable pour la prédiction de la solubilité avec des structures similaires.

Bermejo et al.(2012)[138] ont utilisé l'équation d'état de contribution de groupe développée par Skjold-Jørgensen pour décrire la solubilité de différents gaz (dioxyde de carbone, méthane, éthane, monoxyde de carbone et hydrogène) dans des liquides ioniques avec un anion alkylsulfate. Ils ont constaté que l'équation d'état de contribution de groupe peut être utilisée pour décrire la solubilité des gaz dans les liquides ioniques de sulfate d'alkyle, avec des écarts moyens entre les valeurs expérimentales et prédites inférieurs à 10 %.

Chen et al.(2012)[139] ont estimé la solubilité du dioxyde de carbone (CO2) dans les liquides ioniques à l'aide de l'équation PC-SAFT. Les résultats obtenus démontrent la capacité du modèle à décrire avec précision la solubilité du dioxyde de carbone dans les liquides ioniques à base d'imidazolium.

Gonzalez-Miquel et al.(2012)[140] ont utilisé l'approche de chimie quantique COSMO-RS pour développer des relations structures-propriétés des solvants liquides ioniques réversibles (LI-Rev) pour la sequestration du CO_2. Cette méthode a été appliquée sur une gamme des structures, y compris des silylamines avec différentes longueurs de chaîne alkyle attachées à la fonctionnalité silicium et amine, des silylamines avec des chaînes alkyles fluorées, des silylamines à encombrement stérique et des analogues à base de carbone. Ils ont conclu que les approches informatiques simples COSMO-RS peuvent servir de guide pour la conception de nouveaux LI-Rev et que COSMO-RS permet de déterminer les propriétés thermodynamiques relatives du CO_2 dans ces systèmes et dans des systèmes associés.

Macías-Salinas et al. (2013)[141] ont utilisé une équation d'état cubique (Soave ou Peng−Robinson) couplée à des règles de mélange modernes de type Wong−Sandler pour représenter la solubilité du CO_2 dans certains LIs. Ils ont constaté que l'approche de modélisation qui en résulte s'avère capable de corréler et/ou de prédire la solubilité du CO_2 dans les liquides ioniques dans diverses conditions de température, de pression et de composition du soluté.

Haghbakhsh et al.(2013)[142] ont proposé une nouvelle corrélation basée sur une méthode de régression multiple pour calculer la solubilité du CO_2 dans 27 liquides ioniques

courants basé sur une banque de données de 3073 données de solubilité, en fonction de la température et de la pression. Les résultats de ce modèle ont montré un total de 4,9 % d'écart relatif moyen absolu (ERAM %) pour tous les liquides ioniques considérés. Ils ont constaté que la corrélation actuelle est plus facile à utiliser en la comparant aux modèles de Peng–Robinson (PR) et Soave–Redlich – Kwong (SRK). Les valeurs obtenues des ERAM % sont de l'ordre de 43,5%, 44,3% et 4,9% pour PR, SRK et le modèle proposé pour un total de 3073 données lorsque les coefficients d'interaction binaires de PR et SRK sont optimisés pour les données expérimentales (valeurs ERAM % de 17,2%, 16,9% et 4,9% pour PR, SRK et le modèle proposé, respectivement).

Miraraba et al.(2014)[143] ont utilisé le réseaux de neurone artificiel (RNA) comme nouvelle approche pour prédire la solubilité du CO_2 dans les mélanges des liquides ioniques éthanol-[EMIM][Tf2N]. Les valeurs expérimentales du taux de capture du CO_2 ont été mesurées à 0, 20, 50 et 80 et 100 % en poids du mélange de liquide ionique [EMIM [Tf2N], à des températures de 313,2 et 333,2 °C et dans une plage de pression de 0 à 7 MPa. Sur la base d'une procédure d'optimisation effectuée par un algorithme génétique, ils ont développé un RNA à réaction anticipée à trois couches avec 15 neurones dans la couche cachée, une fonction de transfert log-sigmoïde dans la couche cachée et une fonction de transfert linéaire dans la sortie ultérieure. Les résultats ont également prouvé que le modèle RNA développé surpasse le modèle Peng-Robinson avec une très petite différence entre les données prédites et expérimentales du taux de capture de CO_2, donnant une très faible valeur de l'écart absolu moyen (ERAM) et une valeur élevée des moindres carrés (R^2) très proche de l'unité, indiquant une grande précision de ce modèle pour prédire la variable de sortie.

Sedghamiz et al.(2015)[144] ont appliqué les réseaux de neurones artificiels (RNA) pour étudier la solubilité du dioxyde de carbone et du sulfure d'hydrogène dans différents liquides ioniques. En utilisant 2930 et 664 points de données de solubilité du CO_2 et du H_2S respectivement.de sorte que le réseau a été formé, validé et testé par 70, 15 et 15 % du total des données avec une couche cachée via la fonction de transfert sigmoïde tangent hyperbolique. Les résultats ont montré que ce réseau à action directe a une couche avec fonction sigmoïde tangente hyperbolique peut prédire les solubilités avec des écarts absolus moyens (ERAM %) et le coefficient de détermination (R^2) sont de 3,58 % et 0,9947% pour le CO_2 et de 2,07% et 0,9987% pour le système H_2S. De plus, les résultats de comparaison en terme d'écart entre le modèle RNA développé et l'EoS de Peng – Robinson avec et sans kij optimisé et une corrélation empirique ont montré que le modèle RNA peut corréler la

solubilité des gaz acides dans les LIs avec une grande précision avec une erreur minimale parmi les trois approches.

Baghban et al.(2015)[145] ont développé deux model pour estimer la solubilité du dioxyde de carbone en présence de diverses ILs sur une large plage de pression, de température et de concentration, à savoir : le réseau neuronal artificiel de perception (MLP-RNA) et le système d'interférence neuro-flou adaptatif (ANFIS). Le poids moléculaire, le facteur acentrique, la pression critique et la température critique des LIs ainsi que la pression (P) et la température ont été utilisés comme variables d'entrée et la concentration de dioxyde de carbone a été utilisée comme variable de sortie. À cette fin, ils ont utilisé 728 points de données expérimentales collectées de la littérature pour la mise en œuvre de ces modèles. De plus, ils ont réalisé une comparaison entre les solubilités expérimentales et prédites du dioxyde de carbone par des modèles thermodynamiques basés sur l'équation d'états de Peng-Robinson (PR) et de Soave – Redlich – Kwong (SRK). Les résultats obtenus ont démontré la capacité et une meilleure performance du MLP-RNA en tant que corrélation précise permettant d'estimer la solubilité du dioxyde de carbone dans les LIs par rapport à d'autres modèles alternatifs avec un coefficient de détermination (R^2) et une erreur quadratique moyenne (EQM) de l'ordre de 0,9972 et 0,000133 respectivement.

Hizaddin et al. (2015)[146] ont appliqué le modèle prédictif Soave – Redlich – Kwong (SRK) pour prédire la solubilité du CO_2 dans des liquides ioniques à base d'imidazolium contenant des anions bis (trifluorométhylsulfonyl) imide, tétrafluoroborate et hexafluorophosphate. Ils ont obtenu les paramètres d'interactions binaires en régressant les données expérimentales d'équilibre vapeur-liquide disponibles dans la littérature. Seules les données cohérentes ont été prises en compte lors de la régression. Ils ont conclu que le modèle PSRK a fourni une prédiction raisonnable de la solubilité du CO_2 dans les liquides ioniques à des pressions loin des pressions critiques du CO_2 et des liquides ioniques. A des pressions allant jusqu'à 5 MPa, l'écart quadratique moyen (EQRM) global était de 0,31. De plus, la capacité de prédiction du modèle a été testée pour les liquides ioniques contenant de longues chaînes alkyle, présentant un EQRM global de 0,40.

Hamzehie et al. (2015)[147] ont établi un modèle de réseau neuronal artificiel (RNA) pour prédire la solubilité des gaz acides (H_2S et CO_2) dans 32 solutions d'amines liquides ioniques. Après optimisation des meilleurs paramètres (nombre de neurones, nombre de couches cachées et fonction de transfert), le modèle RNA développé s'est montré capable

d'estimer avec précision la solubilité des gaz acides dans différents liquides ioniques avec une valeur d'erreur relative moyenne de 3,104 et un coefficient de détermination de 0,997.

Bazargani et Sabzi (2016)[148] ont appliqué l'équation d'état de la chaîne de sphère dure perturbée (PHSC) pour modéliser l'absorption du CO_2 dans différents liquides ioniques à longue chaîne alkyle constituée de tétrafluoroborate de 1-hexyl-3-méthylimidazolium, le 1-hexyl-3-méthylimidazolium. hexafluorophosphate, 1-hexyl-3-méthylimidazolium bis(trifluorométhylsulfonyl)imide, tétrafluoroborate de 1-octyl-3-méthylimidazolium, 1-octyl-3-méthylimidazolium hexafluorophosphate et 1-octyl-3-méthylimidazolium bis(trifluorométhylsulfonyl)imide. Ils ont constaté que l'utilisation de trois paramètres moléculaires ainsi que la régression d'un seul paramètre d'interaction binaire étaient suffisants pour que PHSC EoS puisse atteindre des objectifs de modélisation notables.

Mehraein et Riahi (2016)[149] ont développé deux modèles pour prédire la solubilité du dioxyde de carbone (CO_2) dans 21 liquides ioniques, dont une approche basée sur la combinaison d'un algorithme génétique de régression multilinéaire (GA-MLR) et de machines vectorielles de support des moindres carrés (LS-SVM) a été utilisée pour construire un modèle QSPR non linéaire. Ils ont trouvé que le modèle non linéaire peut donner des résultats de prédiction très satisfaisants : Le coefficient de corrélation (R^2) et l'erreur quadratique moyenne (EQRM) étaient respectivement de 0,962 et 0,015 pour l'ensemble des données. De plus, un autre modèle QSPR, la régression multilinéaire (MLR), a également été mise en œuvre. R^2 et EQRM étaient respectivement de 0,876 et 0,027. Les résultats ont démontré que le modèle LS-SVM est supérieur à celui MLR et a permis d'améliorer considérablement la capacité de prédiction.

Kamgar et Rahimpour (2016)[150] ont calculé la solubilité du CO_2 dans les liquides ioniques la base de deux modèles thermodynamiques, à savoir le modèle UNIQUAC combiné avec Peng-Robinson EOS et le modèle quantique QM, basé sur la théorie COSMORS de la charge de surface moléculaire en interaction. Ils ont constaté que la précision de COSMO-RS est bonne, mais seulement pour certains LIs, alors qu'UNIQUAC est la méthode la plus efficace pour prédire la solubilité des gaz.

Tatar et al.(2016)[151] ont appliqué quatre méthodes différentes basées sur l'intelligence artificielle, à savoir le réseau neuronal à fonction de base radiale (RBF), machine à vecteurs de support des moindres carrés (LSSVM), perceptron multicouche (MLP) et système intelligent de machine de comité (CMIS) pour prédire avec précision la solubilité du CO_2

dans 14 liquides ioniques différents. Ils ont utilisé 728 points de données, parmi lesquels 582 points de données ont été utilisés pour construire les modèles et les 146 points de données restants ont été utilisés pour valider les modèles développés. Ils ont constaté que les valeurs prédites sont en grand accord avec les données expérimentales et que l'écart d'erreur absolu maximum pour le meilleur urédicteur ne dépassait pas 3,5 %. De plus, Une comparaison entre les modèles développés et ceux publiés ont révèlé la supériorité des modèles proposés.

Ali et al.(2016)[152] ont utilisé le modèle COSMO-RS et l'équation d'état (EoS) de Peng-Robinson (PR) pour prédire la solubilité du dioxyde de carbone (CO_2) dans huit des liquides ioniques spécifiques à base d'imidazolium résultant de la combinaison de quatre cations: éthyle, butyle, hexyle et octyl-imidazolium avec deux anions : bis(trifluorométhylsulfonyl)imide et Trifluorométhanesulfonate, dans un profil de température basse (313–343 K) et une plage de pression de 7–450 atm. L'étude a démontré une cohérence qualitative relativement acceptable entre les valeurs expérimentales et prédites. Ils ont conclu que le PR EoS pouvait estimer avec précision la solubilité du CO_2 à des valeurs de pression inférieures à la pression critique du CO_2.

Al-fnaish et Lue(2017)[153] ont utilisé la théorie des fluides d'association statistique de chaîne perturbée (PC-SAFT) pour étudier la solubilité du dioxyde de carbone (CO_2) et du sulfure d'hydrogène (H_2S) dans plusieurs liquides ioniques. Ils ont estimé les paramètres des composants purs des LIs en ajustant les données expérimentales de densité et les données binaires de solubilité des gaz acides dans les LIs rapportées dans la littérature. Ainsi ils ont examiné deux stratégies pour modéliser les LIs. Dans la première stratégie, les LIs ont été traités comme des molécules neutres. Dans la deuxième stratégie, les LIs ont été modélisés comme deux ions chargés : le cation imidazolium $[Cnmim]^+$ et l'anion bis (trifluorométhylsulfonyl) imide $[NTf2]^-$ et pour chaque stratégie, ils ont examiné quatre schémas d'auto-association différents : schémas non associatifs, à 2 sites, à 3 sites et à 4 sites. Ils ont constaté que le schéma d'association à 4 sites avec deux donneurs et deux accepteurs a fourni les meilleurs résultats pour presque tous les systèmes binaires gaz acides-LI étudiés, avec un ERAM de 2,76 % à 6,62 % pour les systèmes H2S-LI et de 1,54 % à 4,98 % pour les systèmes CO_2-LI.

Mehraein et Riahi (2017)[154] ont construit un modèle QSPR non linéaire pour prédire la solubilité de CO_2 en utilisant une approche efficace basée sur la combinaison de l'algorithme génétique-régression multi-linéaire (GA-MLR) et des machines vectorielles de

support des moindres carrés (LS-SVM). Ils ont trouvé des résultats très satisfaisants avec un coefficient de corrélation (R^2) et une erreur quadratique moyenne (EQRM) de 0,962 et 0,015 respectivement, ce qui a démontré que le modèle LS-SVM est supérieur au MLR et a amélioré significativement la capacité prédictive.

Li et al.(2017)[155] ont utilisés plusieurs modèles thermodynamiques (modèle Kent-Eisenberg, modèle Austgen, modèle Hu-Chakma, modèle Liu et al.) et des modèles RNA pour corréler et prédire la solubilité à l'équilibre du CO_2 de la solution aqueuse 3-diméthylamino-1 propanol. Ils ont constaté que les modèles RNA pouvaient très bien prédire les valeurs expérimentales en comparant avec des modèles thermodynamiques.

Venkatraman et Alsberg (2017)[156] ont étudié plus de 10 000 données de solubilité LI-CO_2 mesurées à différentes températures et pressions de fonctionnement, en utilisant des méthodes d'apprentissage automatique (ML) avec des descripteurs moléculaires géométriques et basés sur la charge dérivée de manière semi-empirique. Ils ont découvert que les méthodes d'apprentissage automatique étaient plus précises que COSMOtherm, basé sur la chimie quantique, pour estimer la solubilité du CO_2 dans différents liquides ioniques.

Baghbana et al. (2017)[157] ont évalué la capacité prédictive de la machine à vecteurs de support LeastSquare (LSSVM), du système d'inférence neuro-fuzzy adaptatif (ANFIS), du réseau neuronal artificiel Multi-LayerPerceptron (MLP-RNA) et du réseau neuronal artificiel à fonction de base radiale (RBF-RNA) pour estimer la solubilité du dioxyde de carbone dans 67 LIs différents en fonction de la température de fonctionnement, de la pression accompagnée des propriétés des LIs, notamment la température critique, la pression critique et le facteur acentrique . À cet égard, ils ont utilisé une vaste banque de données contenant 5 368 données recueillies de la littérature. Les résultats obtenus à partir de l'approche LSSVM ont indiqué ses prédictions satisfaisantes par rapport aux autres stratégies avec un R^2 et une erreur quadratique moyenne (EQM) de 0,9942 et 0,00035 respectivement.

Panah (2017)[158] a étudié la solubilité de deux gaz uniques, le sulfure d'hydrogène et le dioxyde de carbone, dans quatorze liquides ioniques à l'aide du d'équation d'état CPA EoS via une nouvelle approche. De plus, il a modélisé les systèmes proposés par le SRK EoS et les résultats obtenus ont été comparés avec ceux obtenus par le CPA EoS. Il a trouvé que pour tous les systèmes étudiés, le pourcentage d'écart absolu moyen (ERAM %) était inférieur à 10 %, ce qui a montré un bon accord entre les valeurs expérimentales et calculées.

Akbari et Rahimpour (2018)[159] ont prédit la solubilité du dioxyde de carbone dans huit liquides ioniques à base d'imidazolium à différentes températures et pressions en utilisant la théorie des particules à l'échelle modifiée (SPT). Ils ont constaté que les résultats sont en meilleur accord avec les solubilités expérimentales. De plus, les résultats prédits par le SPT original et modifié ont été comparé avec ceux du modèle UNIFAC-Lei. Les écarts relatifs moyens absolus totaux (ERAM%) pour tous les systèmes ce sont trouvés être respectivement de 2,49 %, 8,45 % et 14,6 % pour le SPT modifié, le SPT original et le modèle UNIFAC-Lei.

Dashti et al. (2018)[160] ont estimé la solubilité du CH_4 et du CO_2 dans 11 liquides ioniques à l'aide de deux méthodes intelligentes, à savoir CSA-LSSVM et PSO-ANFIS en utilisant la fraction molaire de CH_4, et du CO_2, le pourcentage du CH_4 et du CO_2 dans la charge, la température, la pression critique, le facteur acentrique et la température critique des liquides ioniques comme des entrées. Sur la base des résultats obtenus, il est devenu évident que les modèles améliorés étaient très précis et montraient plus de précision que d'autres algorithmes intelligents dans la prédiction des données réelles de solubilité des gaz. En fait, le modèle LSSVM a démontré de meilleures performances et plus de précision par rapport aux autres modèles développés.

Mesbah et al. (2018)[161] ont prédit la solubilité du CO_2 et du CO_2 supercritique (SC) dans 20 liquides ioniques sur une large gamme de pression (0.25–100.12 MPa) et de température (278.15–450.49 K) en utilisant une méthode d'apprentissage automatique robuste du réseau de neurones perceptrons multicouches (MLP-RN). Ils ont également développé une approche de corrélation empirique pour estimer la solubilité du CO_2. Les résultats ont montré que l'approche MLP-RN est de loin supérieure aux autres approches en précision et en généralisation avec un R^2 de 0,9987, EQM de 0,6293 et ERAM% de 1,8416.

Hadj-Kali et al.(2019)[162] ont utilisé le modèle de dépistage de type conducteur pour solvants réels (COSMO-RS) pour prédire la solubilité du CO_2, CH_4, C_2H_6 et C_3H_8 dans huit LIs sélectionnés avec diverses combinaisons de cations et d'anions. Les résultats obtenus ont permis de démontrer des tendances similaires et un accord modéré pour chacun des LIs sélectionnés avec les données expérimentales. Les écarts observés étaient respectivement de 8 à 62 %, 8 à 40 %, 1 à 37 % et 29 à 35 % pour le CO_2, le CH_4, le C_2H_6 et le C_3H_8.

Deng et al.(2019)[163] ont combiné des propriétés critiques des liquides ioniques avec des modèles d'apprentissage en profondeur (CP-DNN, CP-CNN, CP-RNN) pour établir des modèles théoriques de prédiction de la solubilité du CO_2 dans les liquides ioniques. Ils ont

trouvé pour CP-RNN, le coefficient de détermination (R^2) entre les valeurs expérimentales et prédites est de 0,988, CP-CNN est de 0,999 et CP-DNN est de 0,984. Ce qui leur ont permis de conclure que les performances prédictives de ces cadres étaient capables d'atteindre ou de dépasser les effets prédits de la méthode basée sur des modèles thermodynamiques (PR, SRK) et la méthode d'apprentissage automatique (XGBoost).

Song et al.(2020)[164] ont développé deux algorithmes d'apprentissage automatique différents : les réseaux de neurones artificiels (RNA) et la machine à vecteurs de support (SVM) à l'aide de méthodes de contribution de groupe (GC) pour la prédiction de solubilité du CO2 dans les liquides ioniques en utilisant une base de données complète contenant 10 116 données de solubilité du CO_2 mesurées dans divers liquides ioniques à différentes températures et pressions. Les résultats obtenus ont démontré que les deux modèles pouvaient donner des prédictions fiables sur la solubilité du CO_2 dans les LIs ainsi que le modèle RNA-GC a montré une légère supériorité par rapport au modèle basé sur SVM avec un erreur absolue moyenne (EAM) et un coefficient de détermination (R^2) de 0,0202 et 0,9836 respectivement pour les données de l'ensemble de test.

Aghaie et Zendehboudi (2020)[165] ont utilisé quatre méthodes d'apprentissage automatique ML, notamment la machine à vecteurs de support des moindres carrés (LSSVM), l'arbre de décision (DT), la forêt aléatoire (RF) et la régression multilinéaire (MLR) pour obtenir la solubilité du CO_2 dans un ensemble structurellement diversifié de LIs basé sur les propriétés thermodynamiques et le modèle de relation quantitative structure-activité (QSPR). En outre, ils ont utilisé deux ensembles de données sur la solubilité du CO_2 dans diverses conditions de fonctionnement ; un ensemble de données considérant les propriétés critiques, le poids moléculaire et le facteur acentrique des LIs purs comme données d'entrée, et le second incluant des descripteurs structurels de cations et d'anions comme paramètres d'entrée. Ils ont constaté que pour les deux ensembles de données, les modèles DT et RF ont montré de meilleures performances en termes de précision et de fiabilité pour prédire la solubilité du CO_2 dans les LIs, par rapport aux approches LSSVM et MLR. Aussi, ils ont trouvé que les erreurs résultants des modèles basés sur QSPR étaient légèrement inférieures à celles des modèles basés sur la thermodynamique qui pourrait être dû à la forte relation entre les propriétés structurelles des LIs et la solubilité du CO_2 dans les LIs. Par conséquent, Les modèles basés sur des descripteurs structurels sont plus fiables, précis et robustes que ceux construits sur les propriétés critiques.

Mesbah et al.(2020)[166] ont proposé trois méthodes différentes : à savoir, la méthode de groupe de corrélation de traitement des données (GMDH)(modèle I), le réseau de fonction de base radiale ((RBF)-RNA) (modèle II) et une relation empirique (modèle III) pour décrire la solubilité du CO_2 dans les liquides ioniques sur un intervalle de pression (99,8 à 85 800,0 kPa) et de température (283,15 à 373,35 K) en utilisant un ensemble de données composé de 1155 points. Les résultats ont indiqué que tous les modèles proposés étaient vigoureux et possédaient des tendances d'erreur satisfaisantes basées sur une évaluation statistique complète. Les valeurs de R^2, EQRM et ERAM% sont de l'ordre de 0,9650, 4,2727 et 8,46 pour le modèle I, de 0,9942, 1,7172 et 4,23 pour le modèle II et de 0,9886, 2,4248 et 4,25 pour le modèle III.

Moosanezhad-Kermani et al.(2020)[167] ont utilisé la méthode de groupe de traitement des données (GMDH) en tant qu'approche informatique puissante pour prédire la solubilité du CO_2 dans différents liquides ioniques avec une température inférieure et supérieure à 324 K. À cet égard, 4726 points de données, y compris la solubilité du CO_2 dans 60 LIs, ont été utilisés. Les résultats des paramètres statistiques calculés ont montré une précision acceptable pour le réseau neuronal hybride GMDH avec un R^2 et une erreur quadratique moyenne (EQRM) de l'ordre de 0,90435 et 0.0765 respectivement.

Zhang et Xu (2020)[168] ont développé le modèle d'amplification des moindres carrés pour prédire la solubilité du dioxyde de carbone supercritique dans 24 liquides ioniques en fonction des propriétés critiques des LIs et des conditions du système. Ils ont constaté que le modèle est très précis, stable et prometteur en tant qu'outil rapide, robuste et peu coûteux pour les estimations de solubilité.

Valeh-e-Sheyda et al. (2021)[169] ont développé deux modèles différents de réseaux neuronaux à rétro-propagation avec trois et cinq variables d'entrée à l'aide d'une collection de 430 mesures expérimentales pour prédire la solubilité du CO_2 à l'équilibre dans les LIs à base de pyridinium. En premier lieu, les entrées du modèle considéré sont le poids moléculaire, la température d'absorption et la pression d'équilibre. En deuxième lieu, présentation d'un nouveau modèle chimioinformatique basé sur des descripteurs avec les données d'entrée des informations structurelles. L'analyse statistique des résultats obtenus a montré que le modèle développé basé sur un descripteur moléculaire, avec l'erreur quadratique moyenne (EQRM) et un R^2 de 9,06-E03 et 0,995 pour les données d'essai, peut être utilisé pour l'évaluation détaillée de la capacité de chargement en CO_2 des solutions LIs à base de pyridinium.

Daryayehsalameh et al.(2021)[170] ont estimé la solubilité du CO_2 dans le tétrafluoroborate de 1-n-butyl-3-méthylimidazolium ([Bmim][BF4]) à l'aide de six techniques d'intelligence artificielle (IA) différentes, dont quatre réseaux de neurones artificiels (RNA), prenant en charge les machines vectorielles (LS-SVM), système d'interface neuro-fuzzy adaptatif (ANFIS). Ils ont trouvé que le réseau neuronal à réaction en cascade avec dix neurones cachés constitue le paradigme basé sur l'IA le plus précis pour les analyses thermodynamiques de l'absorption du CO2 par [Bmim] [BF4] dans une large gamme de conditions de fonctionnement avec une excellente précision de ERAM = 6,88 %, EQM = 8.10^{-4} et R^2 = 0,98808.

Feng et al.(2021)[171] ont utilisé 12 algorithmes d'apprentissage automatique différents pour estimer la solubilité du CO_2 et du H_2S dans les liquides ioniques, dont 4107 points de données du CO_2 dans 17 liquides ioniques différents et 549 points de données de H_2S dans 10 liquides ioniques différents, en utilisant la température, la pression, le poids moléculaire, la température critique et la pression critique des solutions comme variables d'entrée ; tandis que la solubilité comme variable de sortie. Les résultats ont indiqué que l'algorithme d'empilement possède le modèle le plus précis du système LI-CO2, avec une erreur quadratique moyenne (EQM), l'erreur quadratique moyenne relative (EQRM), l'erreur absolue moyenne (EAM), le coefficient de corrélation (R^2) de 0,001, 0,025, 0,018 et 0,969 respectivement en moyenne.

Valeh-e-Sheyda et al.(2022)[172] ont utilisé trois algorithmes d'apprentissage automatique intelligents distincts, à savoir les modèles FFNN, RBFNN et SVM pour prédire la solubilité à l'équilibre du CO2 dans 43 LIs différentes à base d'imidazolium. À l'aide des descripteurs basés sur la chimioinformatique qui identifient pleinement la divergence des IL à base d'imidazolium, y compris les groupes constitutifs des noyaux de cations et d'anions, le poids moléculaire, la densité, la teneur en eau, la liaison et les informations structurelles, le nombre de carbone, d'hydrogène, d'oxygène, les atomes de fluor et les informations sur les liaisons sont considérées comme les entrées du modèle proposé. Les résultats des trois méthodes d'apprentissage automatique démontrent que les performances de prédiction des modèles proposés sont assez fiables. De plus, la comparaison des résultats indique que le FFNN avec les valeurs correspondantes de l'erreur quadratique moyenne RMSE = 0,071, le coefficient de détermination R2 = 0,952 et l'erreur quadratique moyenne moyenne EQRM = 0,544 est le meilleur paradigme pour prédire la solubilité à l'équilibre du CO2 dans les LIs à base d'imidazolium.

Nakhaei-Kohani et al.(2022)[173] ont estimé la solubilité du N2O dans 25 LIs (pour un ensemble de 801 points de données) en développant quatre modèles intelligents puissants, notamment le réseau de croyance profonde (DBN), l'algorithme d'amplification catégorielle (Cat-Boost), un machine d'apprentissage extrême(ELM) et l'amplification de gradient extrême (XGB) sur la base deux méthodes différentes, la première basée sur la structure chimique et l'autre basée sur les propriétés thermodynamiques. Les résultats obtenus ont indiqué que les nouvelles approches estiment de manière appropriée la solubilité du N2O dans les LIs. De plus, l'approche XGB s'est avérée être la technique de prévision supérieure dans les deux méthodes ((I) : R2 = 0,9999 et EQRM = 0,0016, (II) : R2 = 0,9998 et EQRM = 0,0025).

Balchandani et Dey (2022)[174] ont développé cinq modèles de régression statistique non rigoureux pour prédire la solubilité du dioxyde de carbone (CO2) dans différents mélanges aqueux de liquides ioniques et des mélanges de LIs avec des alcanolamines, de l'éthanol, etc. Où ils ont constaté qu'il existe un bon accord entre les résultats des modèles proposés et les données expérimentales rapportées sur la solubilité du CO2. En outre, ils ont testé 22 systèmes de liquides ioniques avec des amines sur trois types de RNA. Dans l'ensemble, ils ont trouvé que la fonction de formation par rétro-propagation de Levenberg-Marquardt s'est avérée bien adaptée à la plupart des 22 systèmes.

Fierro et al. (2022)[175] ont prédit la solubilité du CO_2 dans 33 types différents liquides ioniques dans une large plage de température et de pression pour 2099 données expérimentales à l'aide d'un perceptron multicouche (MLP) composé de quatre étapes pour déterminer les paramètres du modèle de l'architecture de type. Où ils ont étudié huit algorithmes d'apprentissage différents avec des architectures à deux couches cachées de type (4,m,n,1), et ils ont entraîné le réseau neuronal artificiel avec trois combinaisons d'entrée avec trois combinaisons de variables thermodynamiques telles que la température (T), la pression (P), la température critique (Tc), la pression critique(Pc), le facteur de compressibilité critique(Zc) et le facteur acentrique (ω). Les résultats ont montré que l'architecture 4-6-8-1 avec la combinaison d'entrée T-P-Tc-Pc et l'algorithme d'apprentissage de Levenberg Marquard donne la meilleure prédiction avec un écart absolu moyen de 2,29 %.

Yazdani et al.(2023)[176] ont étudié la modélisation thermodynamique de la solubilité du CO_2 et du H_2S dans différents liquides ioniques à l'aide de la méthode GE/EOS en utilisant L'équation d'état modifiée de Peng-Robinson (MPR2) avec les règles de mélange LCVM. De plus, ils ont également appliqué les modèles d'énergie libre Excess Gibbs de Wilson, NRTL et

UNIQUAC pour être combinés avec les modèles GE/EOS. Sur la base des résultats de la modélisation thermodynamique, ils ont trouvé que les modèles MPR2-LCVM-Wilson, MPR2-LCVM-NRTL et MPR2-LCVM-UNIQUAC offraient une bonne corrélation avec les données expérimentales. Ainsi, ils ont constaté que le modèle MPR2-LCVM-Wilson peut présenter la corrélation la plus élevée avec le moins d'erreur par rapport aux données expérimentales.

Wang et al.(2023)[177] ont appliqué l'équation d'état cubique plus association (CPA EoS) pour la modélisation intégrée des propriétés (la densité, la capacité thermique, la solubilité du CO2 et l'enthalpie d'absorption) dans l'absorption physique et chimique du CO_2 par 22 LIs à base d'imidazolium dans une plage de température (281 à 403 K) et de pression (0,1 à 23 MPa). Les résultats obtenus ont montré que le modèle a présenté une grande précision à basse et haute pression et que les valeurs calculées de solubilité du CO_2 dans 22 LIs sont en bon accord avec les données expérimentales avec un ERAM inférieur à 3,73 %.

Tian et al.(2023)[178] pour prédire la solubilité du CO_2 et du N_2 dans différents types de LIs avec différentes plages de température et de pression ils ont construit un modèle de relation quantitative structure-propriété (QSPR) en combinant la contribution de fragments ioniques (IFC) avec une machine à vecteurs de support (SVM) et un réseau neuronal artificiel (RNA).en prenant le nombre de chaque IF, la température et la pression comme descripteurs d'entrée et la solubilité du gaz comme valeurs de sortie. Pour cela, ils ont collecté deux bases de données à savoir 13055 points de données de solubilité du CO_2 dans 164 types de LIs et 415 données de solubilité du N_2 dans 38 types de LIs. Les résultats ont indiqué que les modèles IFC-SVM et IFC-RNA ont permis de prédire avec précision et fiabilité la solubilité du CO_2 et du N_2 dans les Lis. Les coefficients du détermination obtenus pour la prévision de la solubilité du CO2 étaient de 0,9855 et 0,9732 pour les ensembles d'apprentissage par IFC-SVM et IFC-RNA, respectivement. Tandis que pour la prédiction de la solubilité du N_2, le R^2 obtenu est de l'ordre de 0,9966 et 0,9909 pour les ensembles d'apprentissage par IFC-SVM et IFC-RNA, respectivement.

Heidari et al. (2023)[179] ont utilisé un système d'inférence neuro-flou adaptatif (ANFIS)pour prédire la solubilité du CO_2 dans 26 liquides ioniques différents avec différentes structures ioniques dans des conditions du CO_2 sous-critiques et supercritiques sur une plage de pression (0,34 à 14 596 kPa) et de température (278,1 à 368,4 K) à l'aide de deux modèles. Dans le modèle n°1, les variables d'entrée sont : la température critique (Tc), la température

de fusion (Tm) et le poids moléculaire (Mw) des liquides ioniques, Dans le modèle n°2: les variables d'entrée sont : la température critique (Tc), la pression critique (Pc), et le facteur acentrique (ω) des liquides ioniques.. A la lumière des résultats obtenus, le facteur statistique pour le modèle n°1 était de 0,95, légèrement inférieur à celui du modèle n°2, qui est de 0,9950. Ils ont constaté que malgré que les deux modèles de point de vu mathématique aient une grande précision, le modèle n°2 prédit mieux la solubilité du CO_2 dans les liquides ioniques que le modèle n°1.

Bazooyar et al.(2023)[180] ont prédit la solubilité du CO_2 dans les liquides ioniques contenant de l'eau à l'aide de trois modèles d'intelligence: le réseau neuronal artificiel (RNA), les machines à vecteurs de support (SVM) et la machine à vecteurs de support des moindres carrés (LSSVM). Ils ont utilisé l'évolution complexe mélangée (SCE) pour optimiser les hyperparamètres des modèles intelligents SVM et LSSVM (σ2 et γ), tandis que les essais et erreurs ont été utilisés pour déterminer le nombre optimal de neurones et de couches pour le RNA. Les résultats des trois modèles proposés ont démontré que les performances de prédiction étaient assez fiables. Alors que le RNA s'est avéré modérément plus précis que les SVM ou LSSVM pour prédire solubilité avec un coefficient de détermination (R^2) et une erreur quadratique moyenne (EQRM) de l'ordre de 0,9965 et 0,0104 respectivement, pour les points de données de test.

Loreno et al.(2023)[181]ont utilisé l'équation d'état (EoS) GC-sPC-SAFT (Group Contribution Simplified Perturbed Chain Statistical Associating Fluid Theory) pour comparer deux stratégies permettant de corréler les propriétés thermophysiques des liquides ioniques et de prédire la solubilité du CO2 dans ces LIs. Pour ce faire, ils ont comparé la méthodologie utilisant la densité de la pression atmosphérique, la vitesse du son et la capacité thermique comme données d'entrée avec ceux qui utilisaient une densité de pression élevée comme donnée d'entrée. Àu moyen des familles [Cn-mim][BF4], [Cn-mim][PF6] et [Cn-mim][NTf2] utilisant le schéma d'association 2B, ils ont conclu que la solubilité du CO2 a été mieux prédite à pression atmosphérique en utilisant la stratégie 1. Ceux-ci peut être dû au poids implicite de la compressibilité isotherme dans cette fonction.

Rather et al.(2023)[29] ont développé une corrélation en utilisant une analyse de régression multi-non linéaire entre la plage de pression de 1 à 50 bars et la plage de température de 298,15 à 33,15 K pour des fractions molaires de 0,3, 0,5 et 0,7 pour prédire le coefficient d'activité du CO2 dans les liquides ioniques. Les résultats de la régression ont

révélé une erreur absolue de 19,53 % et une valeur d'erreur quadratique moyenne de 0,465. Ils ont constaté que le développement de la corrélation était bénéfique pour une analyse qualitative robuste du degré de solubilité du CO2 dans différents liquides ioniques

I.8 Conclusion

A la lumière de la littérature scientifique consultée, il ressort que la recherche basée sur l'étude des propriétés physicochimiques et thermodynamique des liquides ioniques développées au cours de ces dernières années pour évaluer la possibilité de leurs applications dans les divers domaines et d'être une alternative aux solvants organique volatils. L'intérêt croissant peut être attribué à leurs caractéristiques exceptionnelles, telles que les capacités de solvant spécifiques, des pressions de vapeur négligeables et de larges plages de température des liquides.

En outre, vue la diversité des liquides ioniques et leur très grand nombre, une caractérisation expérimentale de tous ces composés est pratiquement impossible, d'où alors une modélisation computationnelle basée sur des algorithmes d'apprentissage de la machine semble capable d'accomplir la tâche.

Chapitre II : Méthodes d'Intelligence Computationnelle

Chapitre II : Généralités sur les méthodes d'intelligence computationnelle :

II.1 Introduction

L'intelligence computationnelle (IC) traite principalement des méthodes informatiques « inspirées de la nature » pour faciliter un comportement intelligent dans des environnements complexes et potentiellement changeants. En outre, une approche algorithmique nécessite l'écriture du processus à suivre pour résoudre le problème. Lorsque le problème est complexe, il peut être une étape coûteuse ou impossible. D'autre part, l'intelligence computationnelle est un successeur de l'intelligence artificielle (IA). Elles sont des nouvelles techniques discriminantes dans la théorie de l'apprentissage statistique. Elles permettent d'aborder plusieurs problèmes divers et variés comme la régression, la classification, la fusion etc., comme les systèmes flous, les réseaux de neurones et les algorithmes évolutionnistes.

Par ailleurs, l'intelligence computationnelle embrasse aussi les techniques utilisant l'Intelligence collective, la Fractale, la Théorie du chaos, le Système immunitaire artificiel, les ondelettes, etc...

L'intelligence computationnelle combine divers domaines : l'adaptation, l'apprentissage, l'évolution et la logique floue pour créer des programmes qui sont, dans un certain sens, intelligent. La recherche dans le domaine de l'intelligence computationnelle ne rejette pas les méthodes statistiques, mais offre souvent une vue complémentaire. Les réseaux de neurones artificiels sont une branche de l'intelligence computationnelle qui est très proche du domaine de l'apprentissage automatique.[182]

Dans ce chapitre, une attention particulière sera donnée aux méthodes d'intelligence computationnelle telles que : les réseaux de neurones artificiels (RNA), machine à vecteur de support (SVM) et machine à vecteur de support aux moindres carrés (LSSVM), ainsi que les algorithmes d'optimisation, tels que l'algorithme libellule (DA) et l'optimisation par essaim de particules (PSO).

II.2 Réseaux de neurones artificiels

Les réseaux de neurones, fabriqués de structures cellulaires artificielles, constituent une approche permettant d'aborder sous des angles nouveaux les problèmes de perception, de mémoire, d'apprentissage et de raisonnement. Ils s'avèrent aussi des alternatives très

prometteuses pour contourner certaines des limitations des ordinateurs classiques. Grâce à leur traitement parallèle de l'information et à leurs mécanismes inspirés des cellules nerveuses (neurones)qui sont identifiés comme éléments cellulaires responsables du traitement de l'information dans le cerveau, ils infèrent des propriétés émergentes permettant de solutionner des problèmes jadis qualifiés de complexes.

II.2.1 Historique

Le domaine des réseaux de neurones débutera avec le modèle simple des neurones biologiques proposé par les deux américains de Chicago, le neurophysiologiste Warren Mc cullers et le logicien Walter Pitt en 1943.

- En 1949, Hebb a proposé une formulation du mécanisme d'apprentissage, sous la forme d'une règle de modification des connexions synaptiques. Cette règle repose sur le fait que si deux neurones sont connectés entre eux et activés en même temps, les connexions qui les relient doivent être renforcées, dans le cas contraire, elles ne sont pas modifiées.
- En 1958, Rosenblatt, a développé le modèle du perceptron et a prouvé son théorème de convergence.
- 1960, Widrow et Hoff ont proposé le modèle Adaline (élément linéaire adaptatif).
- En 1969, Minsky et Papert ont fait des commentaires sur la nature du perceptron.
- En 1972, Kohonen a proposé sa mémoire associative automatique.
- En 1982, Hopfield a démontré l'intérêt d'utiliser les réseaux récurrents ('feed-back'). En parallèle des travaux de Hopfield, a conçu son algorithme de rétropropagation de l'erreur, qui offre un mécanisme d'apprentissage pour les réseaux multicouches de typeperceptron (MLP), fournissant ainsi un moyen d'entraîner les neurones des couches cachées. Cet algorithme de 'back- propagation n'a été popularisé qu'en 1986 par Rumelhart.[183]

II.2.2 Définition

Les RNA sont des programmes informatiques qui tentent de simuler quelques aspects de fonctionnement du cerveau humain, notamment l'apprentissage et la généralisation. Ces systèmes sont capables de discerner des relations entre des couples de variables indépendantes et de variables dépendantes durant le processus d'apprentissage. La capacité de généralisation

permet aux RNA de produire un résultat correct lorsqu'ils sont appliqués sur des exemples de données différents de ceux utilisés dans la phase d'apprentissage.

Les RNA sont constitués d'unités de calcul, les neurones, organisées dans une structure spécifique pour former un réseau.[184]

II.2.3 Du neurone biologique au neurone artificiel

L'idée de réseaux neuraux artificiels a été inspirée de la manière par laquelle les neurones biologiques traitent les informations. Ce concept est utilisé pour mettre en œuvre des logiciels de simulations des processus massivement parallèles qui impliquent des éléments de traitement connectés dans une architecture de réseau. L'apprentissage dans le cerveau humain se fait à travers un réseau des neurones qui sont connectés par l'axone, les synapses et des dendrites(Figure II.1). [184]

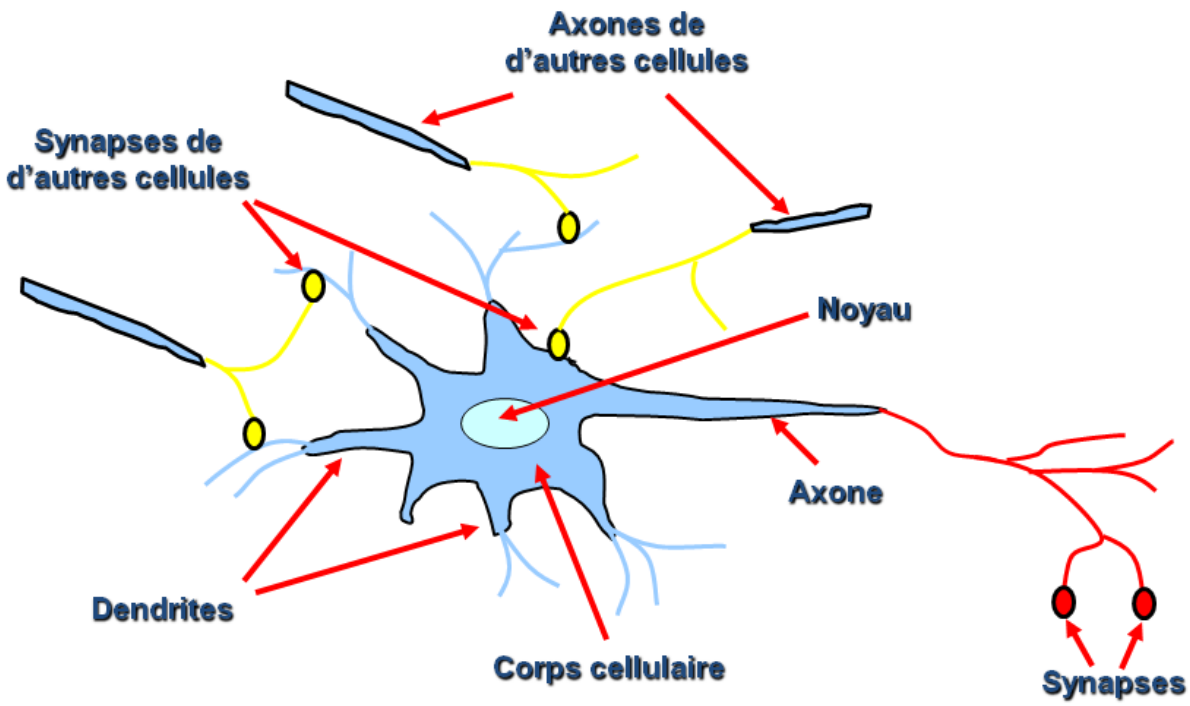


Figure II.1 : Neurone biologique.

Le neurone biologique est donc une cellule composée d'un corps cellulaire et d'un noyau. Le corps cellulaire se ramifie pour former ce que l'on nomme les dendrites. C'est par les dendrites que l'information est acheminée de l'extérieur vers le soma, corps du neurone. L'information traitée par le neurone chemine ensuite le long de l'axone (unique) pour être transmise aux autres neurones. La jonction entre deux neurones est appelée la synapse. Une

résistance synaptique variable affecte le flux d'informations entre deux neurones biologiques. Les dendrites, recueillent, via les synapses, les signaux électrochimiques ou les potentiels d'action venant d'autres cellules ou de l'extérieur et les achemine vers le corps cellulaire ou le soma. Le soma, outre son rôle concernant le métabolisme de la cellule, collecte et concentre les informations et en fait une sommation dite "spatiotemporelle".

Si le potentiel somatique dépasse un certain seuil, il y a émission d'un potentiel d'action, qui correspond à une oscillation électrique très brève. L'axone est un prolongement cellulaire chargé de faire passer l'information nerveuse à d'autres cellules à travers les synapses grâce à son arborisation terminale. Les synapses se trouvent au point de contact entre neurones, elles convertissent les impulsions nerveuses en signaux chimiques c'est-à-dire en molécules de neurotransmetteur. En effet, l'arrivée d'un potentiel d'action à l'extrémité d'un axone entraîne le passage d'une quantité de neurotransmetteurs dans l'intervalle synaptique. Ils permettent alors un passage pondéré de l'influx nerveux. Le nombre de synapses par neurone varie considérablement de plusieurs centaines à une dizaine de milliers.

Le neurone artificiel (Figure II. 2) reçoit les entrées qui sont analogues aux impulsions électrochimiques que les dendrites de neurones biologiques reçoivent d'autres neurones. Il est donc un opérateur mathématique très simple qui traduit la représentation mathématique d'un neurone biologique. Il possède des entrées qui peuvent être les sorties d'autres neurones, ou des entrées de signaux extérieurs, et une sortie. La valeur de la sortie résulte du calcul de la somme des entrées, pondérées par des coefficients (dits poids synaptiques) et du calcul d'une fonction non linéaire bornée de cette somme pondérée. Cette fonction (dite fonction d'activation) est le plus souvent une fonction sigmoïde, mais peut également être une gaussienne, un échelon. [185]

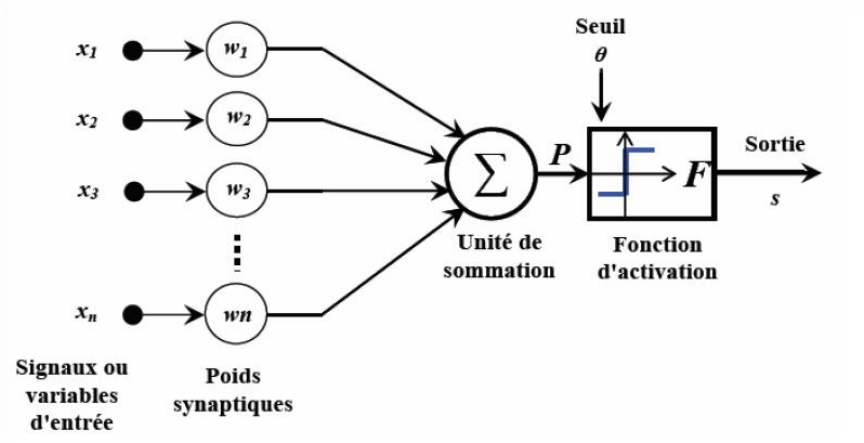


Figure II. 2: Schéma d'un neurone artificiel.[185]

La figure II. 3 résume la correspondance entre le neurone artificiel et le neurone biologique.

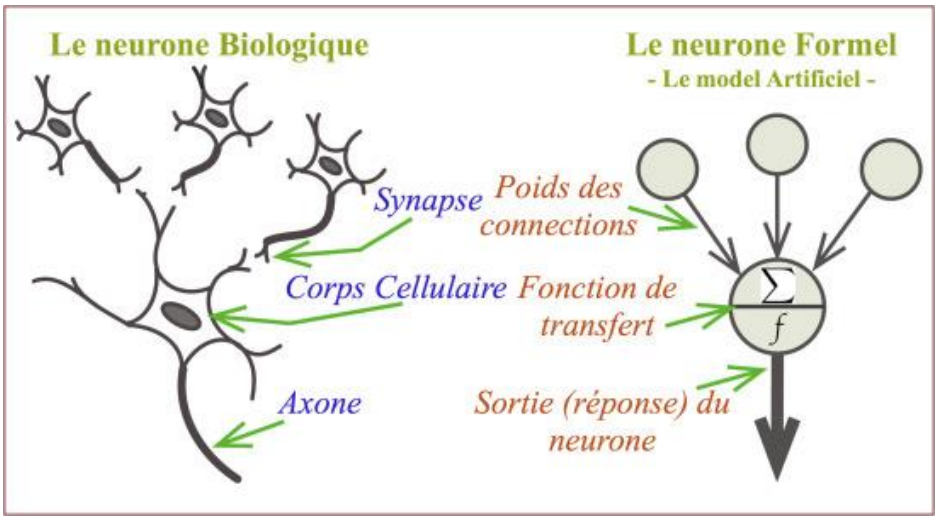


Figure II.3: Mise en correspondance neurone biologique /neurone artificiel.[186]

II.2.4 Concept du réseau de neurones de type feedforward

Les neurones artificiels sont arrangés en couches (figure II.4), où la couche d'entrée reçoit des signaux d'entrée (*ai*) du monde réel et chaque couche successive reçoit les sorties pondérées (*wij ai*) de la couche précédente comme son entrée résultant ainsi en un RNA detype feedforward (cas de perceptrons multicouches), dans lequel chaque sortie d'une coucheest alimentée en avant à la couche qui la succède où elle est traitée. Les sorties de la dernièrecouche constituent les sorties au monde réel.[185]

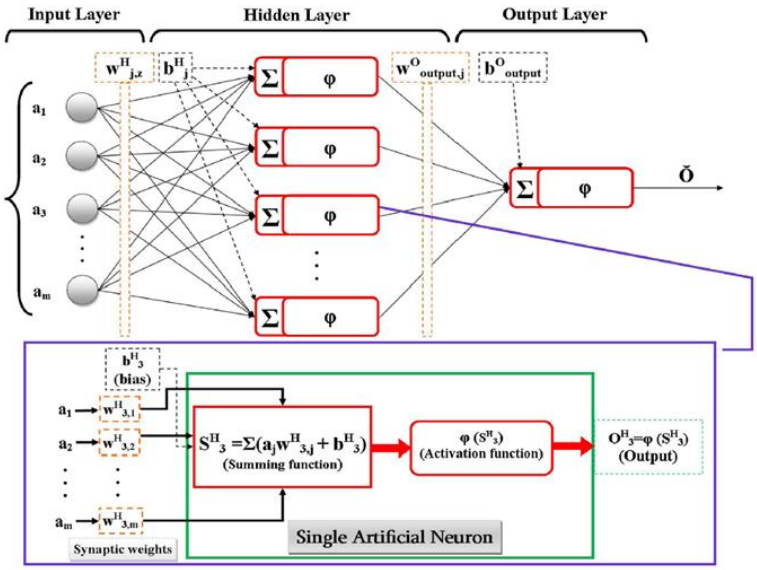


Figure II.4: Structure de Réseau de neurones artificiels de type feedforward multicouches.[145]

Dans un tel RNA de type feedforward un neurone, dans une couche cachée ou une couche de sortie, a deux tâches : il additionne le biais aux entrées pondérés de plusieurs connexions et applique ensuite une fonction d'activation (dite aussi fonction de transfert ou encore fonction de seuillage) à cettesomme comme donnée par l'équation II.1 (pour le neurone *j* de la couche cachée de la Figure II. 4) :

$$Z_j = f_h\left(\sum_{i=1}^{n} {w_{ji}}^I u_i + b_{hj}\right); \quad j = 1,2,\dots,m \qquad \text{II.1}$$

Il propage la valeur résultante par les connexions partantes aux neurones de la couche suivante où elle subit le même processus tel que donné par l'équation II.2 (par exemple les sorties de la couche cachée alimentée au neurone *k* de la couche de sortie donnent la sortie v_k):

$$v_k = f_o\left(\sum_{j=1}^{m} {w_{kj}}^h z_j + b_{ok}\right); \quad k = 1,2,\dots,l \qquad \text{II.2}$$

En combinant des équations II.1 et II.2 on obtient la relation entre la sortie $v_k$et les entrées $u_i$du RN:

$$v_k = f_o\left(\sum_{j=1}^{m} {w_{kj}}^h f_h\left(\sum_{i=1}^{n} {w_{i=1}}^I u_i + b_{hj}\right) + b_{ok}\right); \quad k = 1,2,\dots,l \qquad \text{II.3}$$

Le nombre de neurones dans la couche d'entrée et la couche de sortie est fixé par le nombre de variables indépendantes et dépendantes respectivement. Le modélisateur définit le nombre de couches cachées et le nombre de neurones dans chaque couche cachée. Le développement du modèle est réalisé par un processus d'apprentissage dans laquelle un ensemble de données expérimentales des variables indépendantes est présenté à la couche d'entrée du réseau. Les sorties de la couche de sortie comprennent une prédiction des variables dépendantes. Le réseau apprend les relations entre les variables indépendantes et dépendantes par la comparaison itérative des sorties calculées et des sorties expérimentales et l'ajustement postérieur de la matrice de poids et le vecteur des biais de chaque couche par un algorithme d'apprentissage de rétro propagation du gradient de l'erreur. Ainsi le réseau développe un modèle neuronal capable de prédire avec précision acceptable les variables desortie comprises dans le domaine des données d'apprentissage. Par conséquent, l'objectif de lamodélisation neuronale est de réduire au minimum les erreurs de prédiction de l'ensemble desdonnées de validation présentées au réseau à la fin de l'étape d'apprentissage.[185]

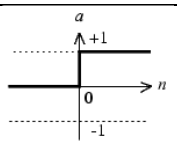
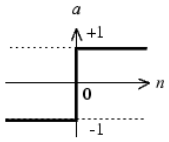
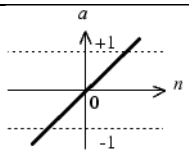
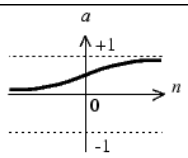
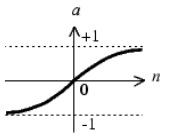
II.2.5 Fonctions d'activation

La fonction d'activation calcule la valeur de l'état du neurone qui est transmise aux neurones avals. La fonction d'activation sert aussi à introduire une non-linéarité dans le

fonctionnement du neurone. A la différence des neurones biologiques dont l'état est binaire, la plupart des fonctions de transfert sont continués, offrant une infinité de valeurs possibles comprises dans l'intervalle [0, +1] ou [-1, +1]. Il existe de nombreuses formes possibles pour la fonction de transfert. Le tableauII.1 résume quelques fonctions de transfert utilisées dans le programme de calcul techniques MATLAB®

En outre, du fait que la continuité et la différentiabilité des fonctions d'activation à tous les points sont les caractéristiques requises par les algorithmes d'optimisation actuels, les fonctions d'activation typiques qui accomplissent ces exigences sont la fonction linéaire, la fonction tangente hyperbolique et la fonction sigmoïde logarithmique. La fonction linéaire est plus appropriée pour les neurones de sorties alors que les fonctions sigmoïdes sont plus appropriées pour les neurones des couches cachées. [185]

Tableau II.1 : Fonctions de transfert usuelles, a=f(n) [185]

Nom de la fonction	**Relation entré /sortie**	**Graphe de la fonction**	**Fonction MATLAB correspondante**
Seuil	a=0 Si n < 0 a=1 Si n ≥ 0		***hardlim***
Seuil symétrique	$a = -1$ Si $n < 0$ $a = 1$ Si $n \geq 0$		***hardlims***
Linéaire	$a = n$		***purelin***
Logarithmique sigmoïde (Logistique)	$a=\frac{1}{1+e^{-n}}$		***logsig***
Tangente Hyperbolique	$a=\frac{e^{+n}-e^{-n}}{e^{+n}+e^{-n}}$		***tansig***

II.2.6 Sélection d'un réseau de neurones artificiel

Bien que la stratégie de sélection d'un modèle neuronal reste encore un sujet de débat, en raison peut être que le développement d'un modèle neuronal relève plus de l'art que de la science, il est clair que l'application réussie des RNA pour la modélisation des problèmes d'ingénierie est fortement affectée par quatre facteurs majeurs :

- Type de réseau.
- Structure de réseau (nombre de couches cachées, nombre de neurones par couche cachée).
- Fonctions d'activation.
- L'algorithme d'apprentissage.

II.2.6.1 Sélection du type de réseau

Parmi les réseaux de neurones de type feedforward qui possèdent la propriété d'approximateur universel de fonction, on trouve les MLP et les RBF. Bien que ces deux types de réseaux aient cette propriété, ils présentent, cependant, un certain nombre de différences qui sont essentiellement :

- Un RBF n'a qu'une seule couche cachée, alors qu'un MLP peut avoir une ou plusieurs couches cachées.
- Pour un MLP, les neurones des couches cachées ou de la couche de sortie ont le même modèle, alors que les neurones de la couche cachée du RBF n'ont pas le même modèle et en plus différent, au niveau de leurs fonctions d'activation, des neurones de la couche de sortie.
- La couche cachée d'un RBF est non linéaire et la couche de sortie est linéaire alors que, les couches cachées et la couche de sortie d'un MLP, sont tous non linéaires. Cependant, en cas de régression non linéaire la couche de sortie est de préférencelinéaire.

Pour une approximation non linéaire d'une fonction et pour un même degré de précision, le MLP nécessite un nombre plus réduit de paramètres (poids synaptiques) que le RBF. L'utilisation de fonctions radiales implique une modification de l'algorithme de rétro propagation. L'apprentissage est souvent réalisé en deux étapes distinctes : une phase de détermination des centres des fonctions radiales suivie d'une phase de détermination des

poids synaptiques. Cependant, cette approche ne garantit pas de conduire à un optimum de la fonction d'erreur.

Toutefois, pour traiter des problèmes multidimensionnels, les RBF nécessitent un nombre de neurones de type RBF (et, donc, un nombre de paramètres), bien plus important, que le nombre de neurones d'un MLP utilisant des fonctions sigmoïdes C'est ce qui fait, en pratique, la plupart des applications des réseaux de neurones en génie chimique ont utilisé des MLP. C'est, particulièrement, le cas des applications relatives à la modélisation des équilibres de phases par les RNA. [185]

II.2.6.2 Sélection de la structure du réseau

Il est bien établi que la variation du nombre des neurones de(s) couche(s) cachée(s) a un effet significatif sur la capacité prédiction du réseau. Le nombre adéquat de neurones cachées dépend d'une façon complexe de :

- Les nombres de neurones de la couche d'entrées et de sorties.
- La taille ou la dimension de l'échantillon d'apprentissage.
- La quantité de bruit dans les cibles (sorties du réseau).
- La complexité de la fonction ou de la classification à faire apprendre au réseau.
- L'architecture du réseau.
- Le type de fonction d'activation des neurones cachés.
- L'algorithme d'apprentissage.
- La régularisation.

Dans la plupart des situations, il n'y a aucune façon de déterminer le nombre adéquat de neurones cachés sans faire l'apprentissage de plusieurs réseaux et d'évaluer ensuite l'erreur de généralisation de chaque réseau. Très peu de neurones cachés donnent des erreurs d'apprentissage et de généralisation top élevées en raison du sous apprentissage. Par contre un nombre est trop élevé de neurones cachés, donne une erreur d'apprentissage très faible mais une erreur de généralisation top élevée en raison du sur apprentissage.

La façon la plus commune d'optimiser la performance d'un RNA est en variant le nombre de neurones dans la couche (s) cachée et choisissant l'architecture avec la meilleure capacité degénéralisation. Les règles heuristiques n'ont aucun sens parce qu'elles ignorent la taille de l'échantillon d'apprentissage, la quantité de bruit des cibles et la complexité de la fonction.

Même si on limite la considération à la minimisation de l'erreur d'apprentissage sur des données non bruités, il est très facile de construire les contre-exemples qui réfutent ces règles. La plus simple, *raisonnable* et probablement la plus populaire encore est la stratégie "du praticien" qui consiste à se concentrer sur la capacité de généralisation du RN et sur la minimisation de la racine de l'erreur quadratique moyenne. La difficulté évidente pour cette stratégie est le fardeau de calcul excessif provoqué par le processus d'expérimentation de modèles différents : la cause est "l'explosion combinatoire" du nombre des modèles neuronales différents qui doivent être exécutés [185]

II.2.6.3 Sélection de la fonction d'activation

Les fonctions d'activation pour les unités cachées sont nécessaires pour introduire le non linéarité dans le réseau. Sans non linéarité, les unités cachées ne seraient pas de réseaux plus puissants que juste des perceptrons purs (aucun neurone caché, juste des neurones d'entrée et de sortie). La raison est qu'une fonction linéaire de fonctions linéaires est de nouveau une fonction linéaire. Cependant, c'est la non linéarité (la capacité de représenter des fonctions non-linéaires) qui fait des réseaux à plusieurs couches, si puissants. Presque n'importe quelle fonction non-linéaire fait le travail, à part des polynômes. Pour l'apprentissage par rétropropagation, la fonction d'activation doit être différentiable et il est préférable qu'elle soit bornée ; les fonctions sigmoïdes comme ***logsig*** et ***tansig*** et la fonction de Gaussienne sont les choix les plus habituels. La fonction ***tansig*** qui produit des valeurs positives et négatives a tendance à donner un apprentissage rapide que des fonctions qui produisent seulement des valeurs positives comme la fonction ***logsig***, à cause du meilleur conditionnement numérique.

Pour des neurones cachés, des fonctions d'activation sigmoïdes sont d'habitude préférables aux fonctions d'activation de seuil. Les réseaux avec des unités de seuil sont difficiles à faire apprendre parce que la fonction d'erreur est constante, de là le gradient n'existe pas ou égal à zéro, écartant ainsi la possibilité d'utiliser les algorithmes d'apprentissage de rétro propagation du gradient de l'erreur qui sont plus efficaces. Même l'apprentissage par les méthodes qui n'utilisent pas les gradients " comme le recuit simulé (simulated annealing) et les algorithmes génétiques" les fonctions sigmoïdes sont plus faciles pour l'apprentissage que les fonctions de seuil. Avec des neurones sigmoïdes, une petite variation des poids synaptiques entraîne d'habitude une variation des sorties, qui permet de dire si la variation des poids est bonne ou mauvaise. Avec des unités de seuil, une petite variation des poids ne produira souvent aucune variation des sorties.

Pour les neurones de la couche de sortie, le choix d'une fonction d'activation doit obéir à la convenance de la distribution des valeurs cibles pour :

- Des cibles binaires (0/1), la fonction logistique est un choix excellent.
- Des cibles continuent et bornées, les fonctions ***logsig*** de ***tansig*** peuvent être utilisées, pourvu que les valeurs cibles soient normalisées.
- Les valeurs cibles positives, mais n'ont aucune borne supérieure, la fonction d'activation exponentielle peut être utilisée, mais faire attention au problème de débordement.
- Des cibles continuent sans bornes connues, la fonction d'activation linéaire ***(purelin)*** est préférable.

II.2.7 Avantages des réseaux de neurones

Les avantages des réseaux de neurones peuvent être résumés par les points suivants :

- Possibilité de traiter les problèmes complexes
- Possibilité d'analyser des données incomplètes
- Capacité d'apprentissage et d'adaptation
- Capacité de généralisation
- Efficacité et rapidité
- Flexibilité

II.3 Machines à vecteurs de support (Support Vector Machines) (SVM)

Parmi les méthodes à noyaux, inspirées de la théorie statistique de l'apprentissage de Vladimir Vapnik, les machines à vecteurs de support (SVM) constituent la forme la plus connue. SVM est une méthode de classifications binaire par apprentissage supervisé, elle fut introduite par Vapnik en 1995.[187] Cette méthode est donc une alternative récente pour la classification. Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprie. Puisque c'est un problème de classification à deux classes, cette méthode fait appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elle est basée sur l'utilisation de fonction dites noyau (kernel) qui permettent une séparation optimale des données. [64,188]

Les Support Vector Machines souvent traduit par l'appellation de Séparateur à Vaste Marge (SVM) sont une classe d'algorithmes d'apprentissage initialement définis pour la

discrimination c'est-à-dire la prévision d'une variable qualitative binaire. Ils ont été ensuite généralisés à la prévision d'une variable quantitative.

II.3.1 C'est quoi un support à vecteur machine ?

L'idée principale des SVM consiste à projeter les données dans un espace de plus grande dimension appelé, espace de caractéristiques, afin que les données non linéairement séparables dans l'espace d'entrée deviennent linéairement séparables dans l'espace de caractéristiques. En appliquant dans cet espace la technique de construction d'un hyperplan optimal séparant les deux classes, on obtient une fonction de classification qui dépend d'un produit scalaire des images des données de l'espace d'entrée dans l'espace des caractéristiques. Ce produit scalaire peut être exprimé, sous certaines conditions, par des fonctions définies dans l'espace d'entrée, qu'on appelle les noyaux.

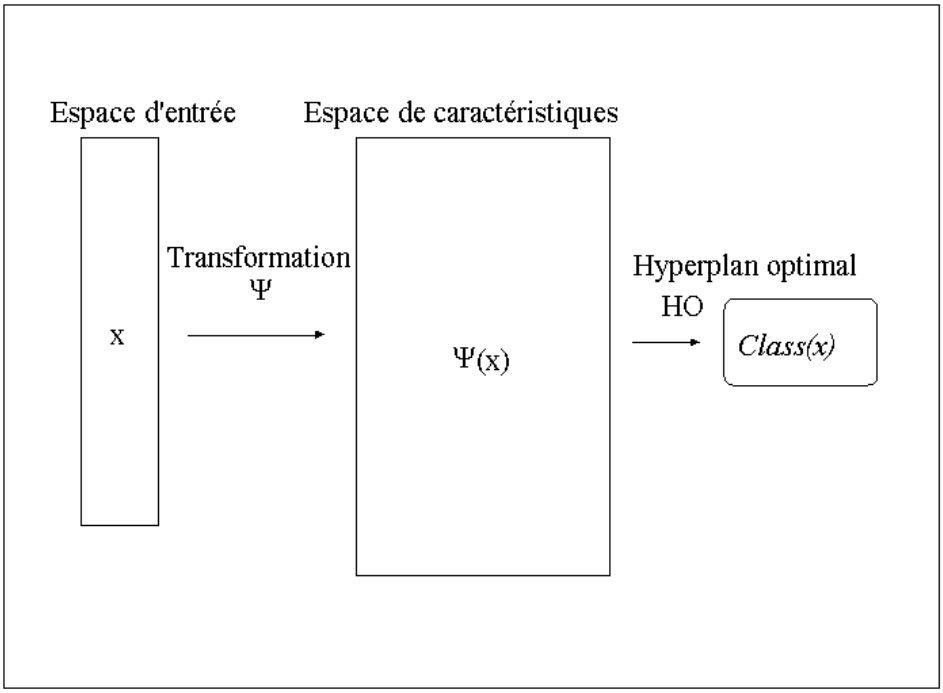


Figure II.5 : Principe des techniques SVM.[189]

II.3.2 SVM : Principe de fonctionnement général

II.3.2.1 Notions de base : hyperplan, marge et support vecteur

Pour deux classes d'exemples donnes, le but de SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec SVM, ce classificateur est un classificateur linéaire appelé hyperplan. Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points.[190]

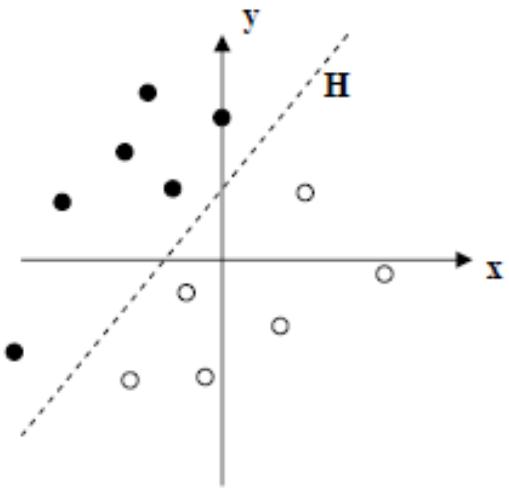


Figure II.6: Exemple d'un hyperplan séparateur[190]

Les points les plus proches, qui seuls sont utilisés pour la détermination de l'hyperplan, sont appelés vecteurs de support.

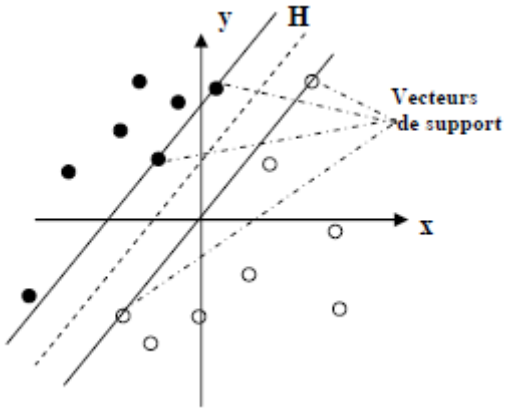


Figure II.7: Exemple de vecteurs de support[190]

Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe " au milieu " des points des deux classes d'exemples. Intuitivement, cela revient à chercher l'hyperplan le " plus sur ". En effet, supposons qu'un exemple n'ait pas été décrit parfaitement, une petite variation ne modifiera pas sa classification si sa distance a l'hyperplan est grande. Formellement, cela revient à chercher un hyperplan dont la distance minimale aux exemples d'apprentissage est maximale.

On appelle cette distance " marge " entre l'hyperplan et les exemples. L'hyperplan séparateur optimal est celui qui maximise la marge. Comme on cherche à maximiser cette marge, on parlera de séparateurs à vaste marge.[191]

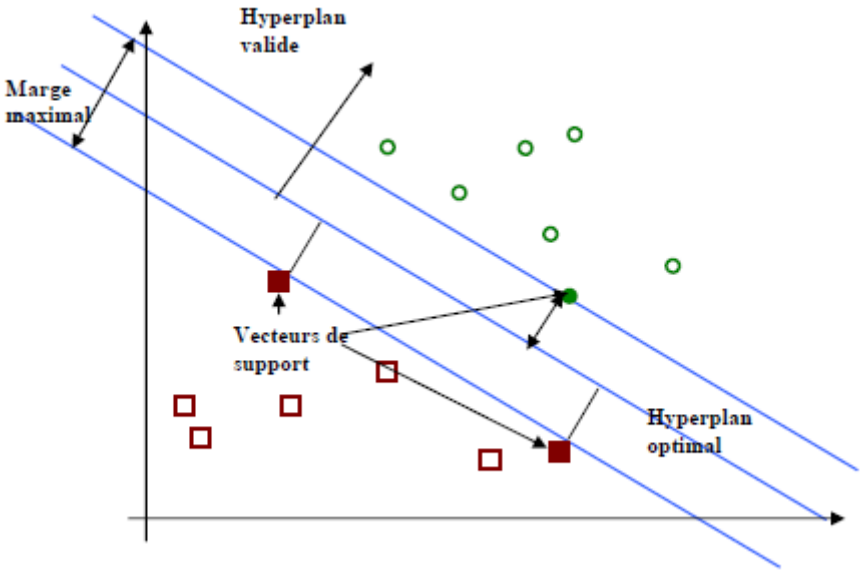


Figure II.8 : Exemple de marge maximal (hyperplan valide)

De plus, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé.

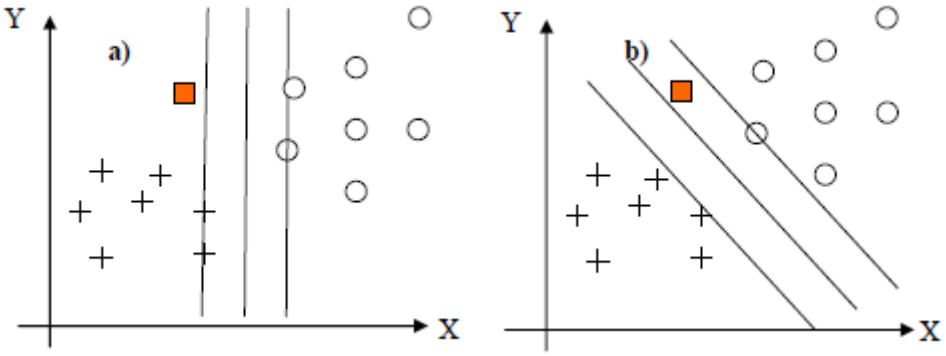


Figure II.9 : (a) L'Hyperplan avec faible marge et (b) le Meilleur hyperplan séparateur.

En général, la classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma suivant, le nouvel élément sera classe dans la catégorie des ≪ + ≫.

II.3.2.2 Linéarité et non linéarité

Parmi les modèles des SVM, on constate les cas linéairement séparables et les cas non linéairement séparable. Les premiers sont les plus simples de SVM car ils permettent de trouver facilement le classificateur linéaire. Dans la plupart des problèmes réels il n'y a pas de séparation linéaire possible entre les données, le classificateur de marge maximale ne peut pas être utilise car il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables.[190]

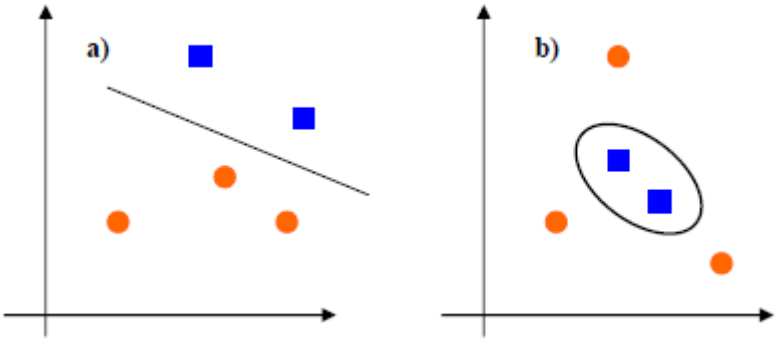


Figure II.10 : (a) cas linéairement séparable et **(b)** cas non linéairement séparable.

II.3.3 SVM pour la régression

Les machines à vecteurs de support ont été étendues pour résoudre le problème de régression avec l'ensemble de données donné $D = \{(x_i, y_i)\}_{i=1}^{N}$ obtenu à partir d'une fonction latente où x_i désigne le vecteur échantillon et y_i la réponse correspondante. N est le nombre total d'échantillons. Deux concepts fondamentaux importants doivent d'abord être pris en compte dans SVR. L'une est la bande ε et l'autre est la fonction de perte ε-insensible. Comme pour la classification des vecteurs de support (SVC), les données originales sont d'abord mappées de manière non linéaire dans un espace de caractéristiques de grande dimension, puis une fonction linéaire est ajustée pour se rapprocher de la fonction latente entre x et y.[192]

II.3.3.1 Fonction de perte en bande ε et insensible à ε

Prenons par exemple une fonction à variable unique, comme le montre la figure II.11 La bande ε fait référence à la région située entre les deux lignes pointillées. Cette région peut être trouvée en déplaçant la ligne continue selon un déplacement de ε de haut en bas. Ici, ε est un nombre positif prédéfini.

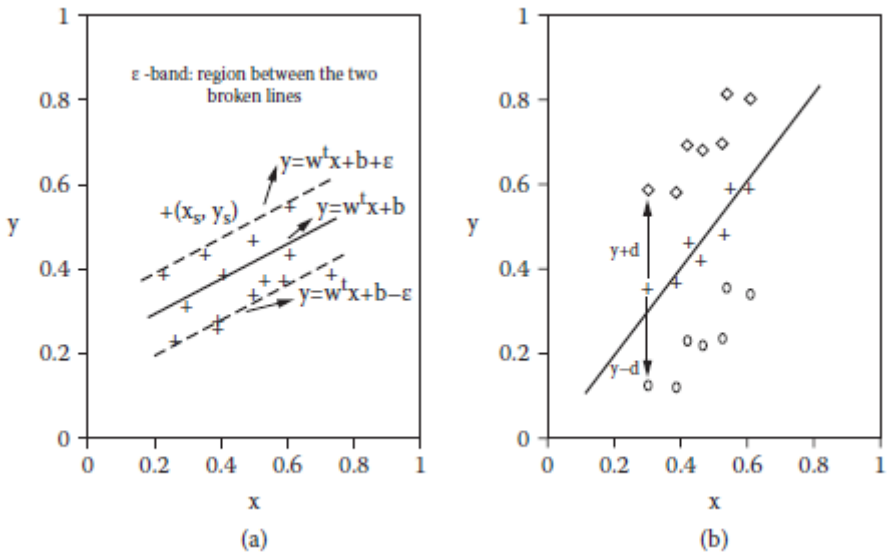


Figure II.11 : Tracé (a) : l'expression graphique de la bande ε de la fonction univariée avec le ε prédéfini. Tracé (b) : comment transformer le problème de régression en problème de classification.

Pour chaque échantillon xi dans les données d'apprentissage (plus), le y_i correspondant est ajouté par un nombre positif d pour produire un nouvel échantillon (x_i, y_i^1) appartenant à la classe 1. De même, le y_i peut également être soustrait du même d à produire un autre nouvel échantillon (x_i, y_i^{-1}) appartenant à la classe –1. En répétant cette procédure, les N échantillons de régression sont doublés et classés en deux classes. Ensuite, le problème de régression se transforme en problème de classification binaire. Ainsi, l'algorithme de SVC peut être appliqué ici pour résoudre le problème de régression.

La fonction de perte ε-insensible à la forme par morceaux :

$$L(y - f(X), \varepsilon) = \begin{cases} |y - f(X)| - \varepsilon, & |y - f(X)| \geq \varepsilon \\ 0, & autrement \end{cases} \quad \text{II.4}$$

Autrement dit, seuls les points de données en dehors de la bande ε (par exemple (x_s, y_s) sur la figure II.11a) provoquent une perte. Ce type de fonction de perte est illustré dans la figure II.12.

II.3.3.2 ε-SVR linéaire

Afin de résoudre le problème de régression, Cortes et al. A intelligemment transformé le problème de régression en problème de classification. Ensuite, la fonction de régression peut être calculée en utilisant le même algorithme que celui décrit dans SVC. La première étape

pour résoudre le problème de régression consiste à le convertir en un problème de classification, illustré à la figure II.11 (tracé (b)).

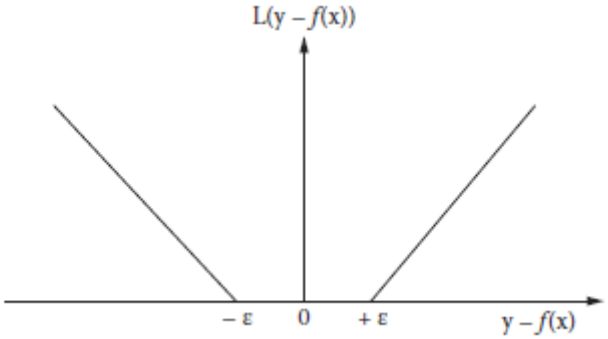


Figure II.12 : Illustration de la fonction de perte insensible à ε.

Étant donné les données d'apprentissage $D = \{(x_i, y_i)\}_{i=1}^{N}$, l'algorithme ε-SVR linéaire vise théoriquement à résoudre le problème d'optimisation, qui peut s'écrire sous la forme suivante avec un terme de perte ε-insensible :

$$minimiser : \frac{1}{2}\|w\|^2 + \frac{C}{N}\sum_{i=1}^{N} L(y_i - f(x_i),\varepsilon) \quad \text{II.5}$$

Où C'est un paramètre de régularisation prédéfini. Le problème de minimisation ci-dessus peut en outre être exprimé sous la forme suivante avec la variable de mou $\xi^{(*)}_i$ introduite :

$$minimiser : L(w, b, \xi^{(*)} = \frac{1}{2}\|w\|^2 + \frac{C}{N}\sum_{i=1}^{N}(\xi_i + \xi^*{}_i)$$

$$sujet\ \text{à}: (w^t x_i + b) - y_i \leq \varepsilon + \xi_i\ , i = 1,2, \dots, N \quad \text{II.6}$$

$$y_i - (w^t x_i + b) \leq \varepsilon + \xi^*{}_i\ , i = 1,2, \dots, N$$

$$\xi^{(*)}{}_i \geq 0, i = 1,2, \dots, N$$

À l'aide de la méthode du multiplicateur de Lagrange et de l'algorithme de programmation quadratique (voir l'annexe B pour plus de détails), la fonction de régression peut être dérivée comme suit :

$$f(x) = \sum_{i=1}^{N}(\alpha^*{}_{i,f} - \alpha_{i,f})(x_i{}^t x) + b_f \quad \text{II.7}$$

$b_f = y_i - \sum_{i=1}^{N}(\alpha^*_{i,f} - \alpha_{i,f})(x_i{}^t x) + \varepsilon$ II.8

Où $\alpha^*_{i,f}$ et $\alpha_{i,f}$ sont respectivement les multiplicateurs de Lagrange optimisés.

II.3.3.2 ε-SVR basé sur le noyau (Kernel)

Il est bien connu que la non-linéarité est le cas le plus courant dans les ensembles de données du monde réel. Par conséquent, il est nécessaire d'étendre le ε-SVR linéaire en régression non linéaire. En introduisant la fonction noyau (Kernel), l'entrée d'origine a d'abord été mappée de manière non linéaire dans l'espace des caractéristiques, et le ε-SVR résultant devient si flexible qu'il peut être utilisé pour résoudre le problème complexe de régression non linéaire en chimie.

Comme la procédure de dérivation de la fonction de décision finale est assez similaire à celle du cas linéaire, nous donnons ici uniquement la forme mathématique ultime :

$f(x) = \sum_{i=1}^{N}(\alpha^*_{i,f} - \alpha_{i,f})K(x_i, x) + b_f$ II.9

$b_f = y_i - \sum_{i=1}^{N}(\alpha^*_{i,f} - \alpha_{i,f})K(x_i, x_j) + \varepsilon$ II.10

Où $\alpha^*_{i,f}$ et $\alpha_{i,f}$ sont les multiplicateurs de Lagrange optimisés.

De plus, il convient de souligner que ν-SVR (ν ∈ [0, 1]) pourrait être traité comme une version modifiée du ε-SVR original basé sur la fonction du noyau. Dans ν-SVR, ν est un paramètre de parcimonie. Après avoir choisi ν, la valeur ε est automatiquement ajustée par l'algorithme. Ici, ν est la limite inférieure de la proportion de vecteurs de support par rapport au total des échantillons et la limite supérieure de la fraction d'erreurs en même temps.[192]

Exemples de noyaux :

- Linéaire $k(x, x') = x.\, x'$
- Polynomial $k(x, x') = (x.\, x')^d$ *ou* $(c + x.\, x')^d$
- Gaussien $k(x, x') = e^{-\|x-x'\|2/\sigma}$
- Laplacien $k(x, x') = e^{-\|x-x'\|1/\sigma}$

On remarque en général que le noyau gaussien donne de meilleurs résultats et groupe les données dans des paquets nets. En pratique, on combine des noyaux simples pour en obtenir de plus complexe.

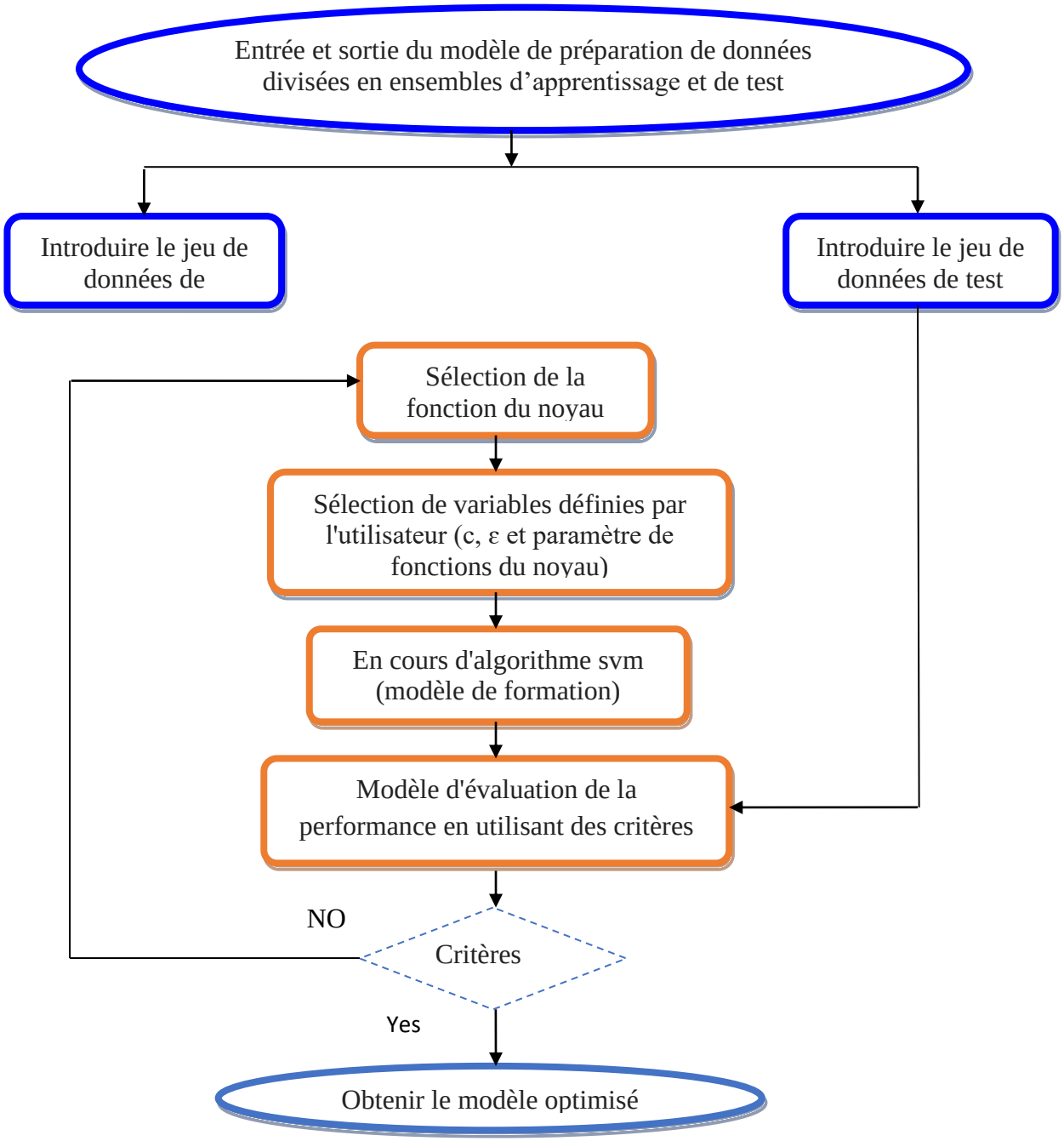


Figure II.13 : Organigramme des étapes pour la création du modèle SVM.[11]

II.3.4 Algorithmes d'optimisation du SVM

Le processus de modélisation prédictive des machines à vecteurs de support, le facteur de pénalité C et le paramètre du noyau σ doivent être sélectionnés pour spécifier une SVM. Ils sont généralement sélectionnés par essai et erreur, ce qui entraîne des prédictions difficiles et des erreurs importantes. Par conséquent, les paramètres du SVM ont été optimisés à l'aide de différents algorithmes méta-heuristiques ces dernières années, tels que l'optimisation par essaim de particules (PSO), l'algorithme de libellule (DA) et l'algorithme génétique (GA).[193]

II.3.4.1 Optimisation par essaim de particules (PSO)

En 1995, le PSO a été introduit par Kennedy et Eberhart[194] sur la base d'un modèle de simulation sociale connu sous le nom d'algorithme d'optimisation stochastique, inspiré du comportement social de volées d'oiseaux ou de bancs de poissons. PSO est similaire aux algorithmes génétiques (GA), en termes d'initialisation de la population avec des solutions aléatoires et de recherche de l'optimum en mettant à jour les générations. Cependant, contrairement à GA, la PSO ne subit pas de croisement ou de mutation, car les particules se déplacent dans l'espace problématique en suivant les particules optimales actuelles. Puisqu'il n'y a que quelques paramètres à définir dans PSO, il est facile à mettre en œuvre. Le concept sous-jacent est que ces particules se déplacent sur la zone de recherche à une vitesse flexible et conservent la meilleure position qu'elles ont découverte dans l'espace de recherche. Chaque particule peut réviser son vecteur vitesse pour explorer la meilleure position grâce à son expertise de vol et à celle des autres particules dans l'espace de recherche.[195,196]

Mathématiquement, un essaim de particules est initialisé de manière aléatoire sur l'espace de recherche et se déplace dans l'espace D-dimensionnel pour rechercher de nouvelles solutions. Soit x^i_k et v^i_k respectivement la position et la vitesse de la $i^{ème}$ particule dans l'espace de recherche à la k^{ith} itération, puis sa vitesse et la position de cette particule à (k+1) l'itération sont mises à jour à l'aide des équations suivantes : [195]

$$v^i{}_{k+1} = w.v^i{}_k + c_1.r_1.\left(p^i{}_k - x^i{}_k\right) + c_2.r_2.(p^g{}_k - x^i{}_k) \quad \text{II.11}$$

$$x^i{}_{k+1} = x^i{}_k + v^i{}_{k+1} \quad \text{II.12}$$

Où r_1 et r_2 représentent des nombres aléatoires entre 0 et 1, c_1 et c_2 sont des constantes, p^i_k représente la meilleure position jamais atteinte de la $i^{ème}$ particule et p^g_k correspond à la meilleure position globale dans l'essaim jusqu'à la k^{ith} itération.

II.3.4.2 Algorithme de libellule (DA)

L'algorithme de la libellule (Dragonfly) est un nouvel algorithme métaheuristique basé sur le comportement des libellules proposé par Seyedali Mirjalili en 2015.[197] où L'étude vise à comprendre les différentes interactions entre les créatures vivantes individuelles qui peuvent aboutir à l'intelligence sociale. Il existe cinq comportements dans la population de libellules, à savoir la séparation, l'alignement, la cohésion, le comportement de recherche de nourriture et d'évitement des ennemis.[198]

Chacun de ces comportements a été modélisé mathématiquement comme suit :

(1) Les comportements d'évitement des collisions :

$$S_i = -\sum_{j=1}^{N} X - X_j \qquad \text{II.13}$$

X est la position de l'individu libellule actuel ; Xj indique la position du j-ième individu libellule voisin ; N est le nombre d'individus libellules voisins et.

(2) Le comportement consistant à maintenir le vol coordonné avec le groupe libellule :

$$A_i = \frac{\sum_{j=1}^{N} V_j}{N} \qquad \text{II.14}$$

Vj indique la vitesse du j-ième individu libellule voisin.

(3) Le comportement de rapprochement pour chaque individu :

$$C_i = \frac{\sum_{j=1}^{N} X_j}{N} - X \qquad \text{II.15}$$

(4) Le comportement de recherche de nourriture :

$$F_i = X^{+} - X \qquad \text{II.16}$$

X est la position de l'individu libellule actuel ; X+ indique la position de la source de nourriture.

(5) Le comportement des ennemis qui échappent :

$$E_i = X + X^{-} \qquad \text{II.17}$$

X est la position de l'individu libellule actuel, X- indique la position de l'ennemi.

Afin de mettre à jour l'emplacement des libellules et de simuler le comportement de vol dans l'espace de recherche, deux vecteurs : la taille du pas (ΔX) et la position (X) sont introduits. Le vecteur de pas est calculé comme suit :

$$\Delta X^{t+1} = w\Delta X^{t} + (sS_i + aA_i + cC_i + fF_i + eE_i) \quad \text{II.18}$$

où s, a, c, f , e indiquent les poids de cinq comportements, u est le poids d'inertie, t est l'itération courante.

Le vecteur de position est mis à jour comme suit :

$$X^{t+1} = X^{t} + \Delta X^{t+1} \quad \text{II.19}$$

Lorsqu'il n'y a pas d'individus adjacents, une stratégie de marche aléatoire est introduite pour améliorer le caractère aléatoire de la recherche. Dans ce cas, l'équation du vecteur de position est indiquée dans l'équation. (II.20):

$$X^{t+1} = Levy(d) \times X^{t} + X^{t} \quad \text{II.20}$$

où d représente la dimension de l'individu libellule.

La stratégie de vol Levy est décrite comme dans les équations. (II.21 etII.22):

$$Levy(X) = 0.01 \times \frac{r_1 \times \sigma}{|r_2|^{\frac{1}{\beta}}} \quad \text{II.21}$$

$$\sigma = \left(\frac{r(1+\beta) \times sin\left(\frac{\pi\beta}{2}\right)}{r\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\frac{\beta-1}{2}}} \right)^{\frac{1}{\beta}} \quad \text{II.22}$$

où r1, r2 sont deux nombres stochastiques dans [0,1], β est une constante, qui est pris égal à 1,5, Γ(X) = (X-1)!

II.4 Machines à vecteurs avec support des moindres carrés (LSSVM)

II.4.1 Introduction

Suykens et Vandewalle (1999) [199,200]ont proposé une simplification de la SVM en utilisant des machines à vecteurs avec support des moindres carrés (LSSVM). Les LSSVM ont été appliqués avec succès dans divers domaines. La LSSVM présente des avantages similaires à celle du SVM, mais un avantage supplémentaire est qu'elle ne nécessite que résoudre un ensemble d'équations linéaires ce qui est beaucoup plus facile et plus simple en calcul, La méthode utilise des contraintes d'égalité au lieu de contraintes d'inégalité, et adopte le système linéaire des moindres carrés comme fonction de perte, le rendant attractif du point de vue du calcul. Le LSSVM a également une bonne convergence et une grande précision; par conséquent, cette méthode est plus facile à utiliser par rapport aux solveurs de programmation quadratiques de la méthode SVM.[201]

Des études empiriques approfondies (par exemple, Wang et Hu, 2005) ont montré que le LSSVM est comparable à un SVM en termes de performances de généralisation. Le principal avantage du LSSVM est qu'il est peu coûteux en calcul, tout en maintenant les propriétés importantes associées à un SVM. Dans le domaine des ressources en eau. La méthode LSSVM a reçu très peu d'attention et seulement quelques applications de LSSVM dans la modélisation des systèmes environnementaux et écologiques, tels que la prévision de la qualité de l'eau, ont été effectués. Les auteurs doivent encore découvrir une étude qui a été réalisée par utilisation des capacités du modèle LSSVM en prévision de flux.[201]

II.4.2 Définition

Le LSSVM, c'est une modification d'un SVM, a été introduit par Suykens (2000). Le LSSVM fournit un avantage de calcul sur le SVM standard en convertissant un problème d'optimisation quadratique en un système d'équations linéaires. Cette nouvelle version de SVM simplifie le problème et converge rapidement vers une solution. Le prédicateur LSSVM est formé à l'aide d'un ensemble de valeurs historiques de séries chronologiques en tant qu'entrées et d'une sortie unique en tant que valeur cible. La LSSVM a été développée pour trouver la fonction de régression non linéaire optimale.[201]

L'approche LS SVM améliore les performances de l'algorithme SVM en résolvant la programmation linéaire au lieu de la programmation quadratique, alors LSSVM semble plus simple que SVM L'ensemble de données $\{(x_1, y_1), (x_2, y_2),\ldots, (x_N, y_N)\}$ est supposé et la fonction non linéaire suivante est appliquée pour l'approximation :

$$f(x) = b + (\text{ø}(x), \omega) \qquad \text{II.23}$$

Ici, *f (x)* exprime la sortie, *x* indique la variable d'entrée, *φ* est connue comme fonction d'approximation non linéaire, *ω* et *b* indiquent les termes de pondération et de biais, respectivement et *(,)* Exprime le produit scalaire.

Le LSSVM peut être exprimé comme un problème d'optimisation :

$$\text{Min}_{w,b,e}\, j(w,e)=\frac{1}{2}\|w\|^2+\frac{1}{2}\gamma\sum_{k=1}^{n} e_k^2 \qquad \text{II.24}$$

$$y_k = (\omega, \varphi(x_k)) + b + e_k \quad k = 1, 2, \ldots, N$$

Dans les équations susmentionnées, e_k et y indiquent l'erreur et coefficients de régularisation, respectivement. L'optimisation expliquée est construite à partir de l'équation de Lagrange :

$$L_{lssvm} = \frac{1}{2}\|w\|^2+\frac{1}{2}\gamma\sum_{k=1}^{n} e_k^2-\sum_{k=1}^{n} a_k\{(\omega.\, \text{ø}(x_k)) +b+e_k-y_k\} \qquad \text{II.25}$$

Ici, indique le multiplicateur de la fonction lagrangienne. Pour résoudre Cette équation utilise le concept de différenciation partielle :

$\frac{\partial Llssvm}{\partial w}$=0 → w=$\sum_{k=1}^{n} a_k$ø ($x_k$)

$\frac{\partial Llssvm}{\partial w}$=0 → $\sum_{k=1}^{n} a_k$=0

$\frac{\partial Llssvm}{\partial w}$=0 →$a_k$ =γek=1,……, N

$\frac{\partial Llssvm}{\partial w}$=0 → (ω. ø ($x_k$)) +b+ e_k-y_k =0 k=1,…..N

En considérant α = [α1,…, αN], lv = [1,… l], Y = [y1,… yN] et élimination de w et e_i, le problème peut être défini comme suit:

$$\begin{bmatrix} 0 & \text{ITN} \\ \text{IN}\ \Omega + & \gamma - 1\text{LN} \end{bmatrix}$$

Dans l'algorithme LSSVM, les paramètres correspondants ont un grand impact sur la capacité et la performance du modèle, ainsi, le choix de ces paramètres doit être soigneusement déterminé.[202]

La détermination des hyperparamètres optimaux est réalisée à l'aide d'une méthode d'optimisation par algorithme génétique (AG) basée sur l'organigramme présenté à la Figure II.14

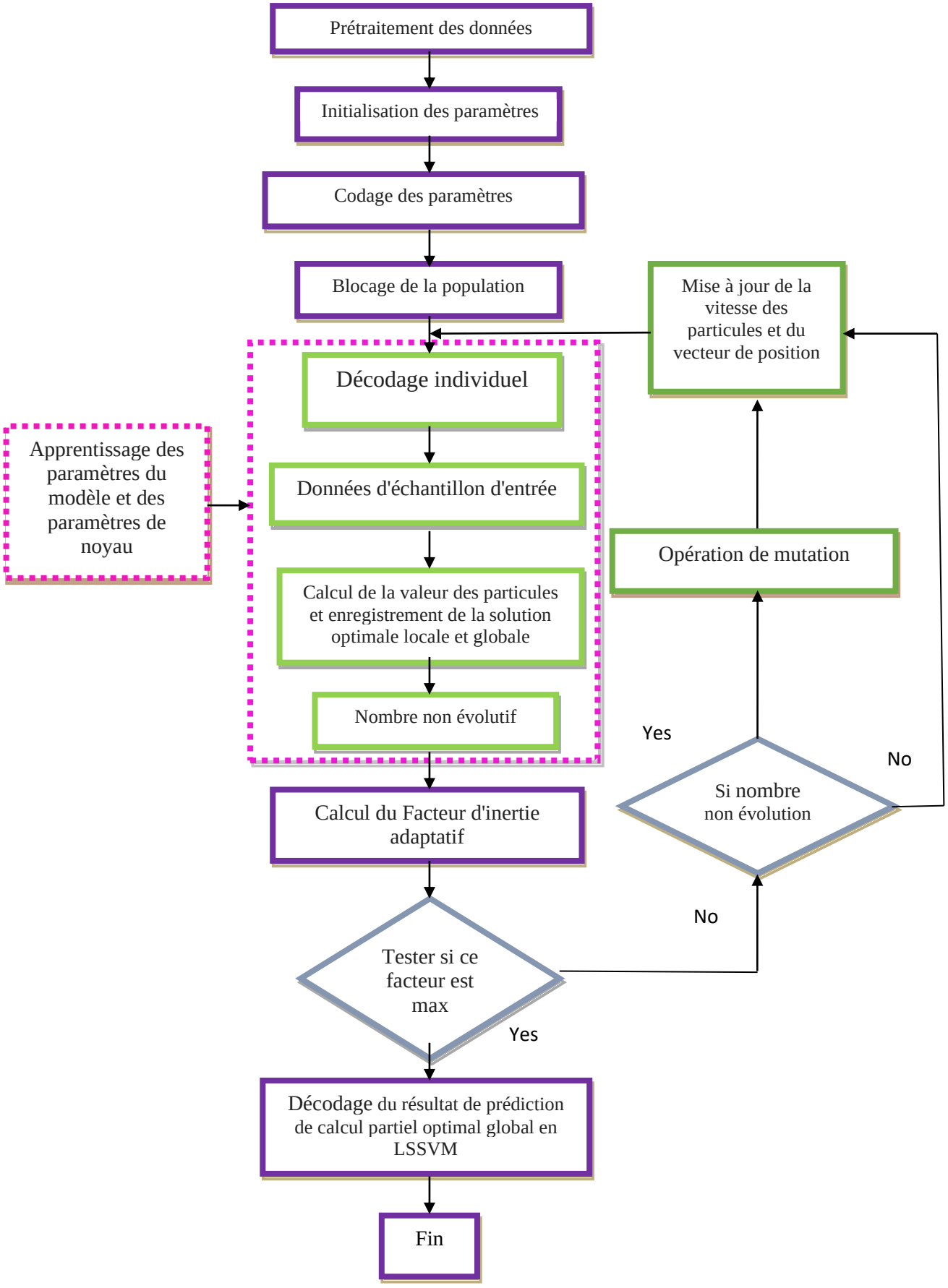


Figure II.14 : Démonstration schématique de l'algorithme LSSVM formé.

Ce model permet d'estimer les propriétés physico-chimiques du liquide ionique.

II.5 Critères statistiques :

La fiabilité et la précision des modèles, développés à partir des méthodes d'intelligence computationnelle, reposent sur diverses quantités statistiques : erreur quadratique moyenne (EQM); Erreur absolue moyenne (EAM); Erreur quadratique relative moyenne (EQRM); Erreur relative absolue moyenne (ERAM); Coefficient de détermination (R^2); Coefficient de corrélation (R); Facteur de précision (A_f) ; facteur de biais (B_f) et Critère d'information d'Akaike (AIC) qui sont données par les formules suivantes :[108]

$$EQM = \frac{1}{N}\sum_{i=1}^{N}({y_i}^{exp} - {y_i}^{cal})^2 \quad \text{II.26}$$

$$EAM = \frac{1}{N}\sum_{i=1}^{N}|{y_i}^{exp} - {y_i}^{cal}| \quad \text{II.27}$$

$$EQMR = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{{y_i}^{exp}-{y_i}^{cal}}{{y_i}^{exp}}\right)^2 \quad \text{II.28}$$

$$ERAM\% = 100 * \frac{1}{N}\sum_{i=1}^{N}\left|\frac{{y_i}^{exp}-{y_i}^{cal}}{{y_i}^{exp}}\right| \quad \text{II.29}$$

$$ER\% = 100 * \left|\frac{{y_i}^{exp}-{y_i}^{cal}}{{y_i}^{exp}}\right| \quad \text{II.30}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left({y_i}^{exp}-{y_i}^{cal}\right)^2}{\sum_{i=1}^{N}\left({y_i}^{exp}-\overline{y^{exp}}\right)^2} \quad \text{II.31}$$

$$R = \frac{\sum_{i=1}^{N}\left({y_i}^{exp}-\overline{y^{exp}}\right)\times\left({y_i}^{cal}-\overline{y^{cal}}\right)}{\sqrt{\sum_{i=1}^{N}\left({y_i}^{exp}-\overline{y^{exp}}\right)^2}\times\sqrt{\sum_{i=1}^{N}\left({y_i}^{cal}-\overline{y^{cal}}\right)^2}} \quad \text{II.32}$$

- **Facteur de biais Bf et facteur de précision Af:**

Bon modèle	B_f =0.9 – 1.05
Modèle acceptable	B_f = 0.7 – 0.9
Utiliser avec précaution	B_f = 1.06 – 1.15
Modèle inacceptable	B_f< 0.7 et B_f> 1.15

Ross (1996) a proposé deux indices différents pour la détermination de la performance du modèle. Le facteur de biais (B_f) est une mesure de la concordance globale entre les valeurs prédites et les valeurs observées. Il indiquera si, et dans quelle mesure, les prévisions se situent au-dessus ou au-dessous de la ligne d'équivalence. Une pondération égale est donnée à

la sur-prédiction, où la valeur prédite est supérieure à la valeur observée, et à la sous-prédiction. Le B_f est montré dans l'équation II.24

$$B_f = 10^{\left(\frac{\sum_{i=1}^{N} log\frac{y_i^{cal}}{y_i^{exp}}}{N}\right)} \quad \text{II.33}$$

Un accord parfait entre les valeurs observées et les valeurs prédites donnerait un B_f de 1. Une valeur supérieure à 1 indiquerait que, en moyenne, les prédictions étaient plus grandes que les valeurs observées et représenteraient un modèle à sécurité intrinsèque ; par exemple un B_f de 1,2 indiquerait que les prévisions étaient 20% plus grandes que les valeurs observées.

Une valeur inférieure à 1 indiquerait un modèle à sécurité intrinsèque où, en moyenne, les prévisions étaient plus courtes que celles observées. Le B_f est maintenant largement utilisé comme mesure de la performance du modèle et constitue un critère commun pour la comparaison de différents modèles. Dans le calcul du B_f, les prédictions sur (+) et sous (-) peuvent s'annuler. Ils ne montrent donc pas les erreurs absolues dans le modèle. Ceci peut être réalisé en utilisant le facteur de précision (A_f). Ceci est basé sur une équation similaire à II.24 mais ignore si la différence entre la valeur prédite et la valeur observée est positive ou négative:

$$A_f = 10^{\left(\frac{\sum_{i=1}^{N} \left|log\frac{y_i^{cal}}{y_i^{exp}}\right|}{N}\right)} \quad \text{II.34}$$

Toutes les différences étant désormais positives, la valeur de Af sera toujours égale à (si l'accord est parfait) ou supérieure à un. Une précision de 1,5 signifie qu'en moyenne, la valeur prédite est différente de 50% (plus petite ou plus grande) de la valeur observée.

Où N est le nombre de points de données (paires d'entrée et de sortie), y^{exp} est la valeur des ensembles de données expérimentales au point d'échantillonnage i, yi^{cal} est la ième valeur des jeux de données prédites par le modèle et c'est la moyenne des données expérimentales et prédites utilisant le modèle, sont définies respectivement comme suit:

$$\overline{y^{exp}} = \frac{1}{N}\sum_{i=1}^{N} {y_i}^{exp} \quad \text{II.35}$$

$$\overline{y^{cal}} = \frac{1}{N}\sum_{i=1}^{N} {y_i}^{cal} \quad \text{II.36}$$

En outre, les critères suggérés par les chercheurs ont été utilisés pour assurer une plus grande validité des modèles développés. Ces critères statistiques sont donnés comme suit:

$$k = \frac{\sum {y_i}^{exp} {y_i}^{cal}}{\sum ({y_i}^{cal})^2} \quad \text{II.37}$$

$$k' = \frac{\sum {y_i}^{exp} {y_i}^{cal}}{\sum ({y_i}^{exp})^2} \quad \text{II.38}$$

Avec : $0.85 \leq k \leq 1.15$; $0.85 \leq k' \leq 1.15$ comme critères d'acceptabilité.

$${R_o}^2 = 1 - \frac{\sum ({y_i}^{cal} - k{y_i}^{cal})^2}{\sum ({y_i}^{cak} - \overline{{y_i}^{cal}})^2} \quad \text{II.39}$$

$${R_o}'^2 = 1 - \frac{\sum ({y_i}^{exp} - k'{y_i}^{exp})^2}{\sum ({y_i}^{exp} - \overline{y^{exp}})^2} \quad \text{II.40}$$

$$M = \frac{R^2 - {R_0}^2}{R^2} \leq 0.1 \quad \text{II.41}$$

$$N = \frac{R^2 - {R_o}'^2}{R^2} \leq 0.1 \quad \text{II.42}$$

$${R_m}^2 = R^2 \times \left(1 - \sqrt{\left|R^2 - {R_o}^2\right|}\right) \geq 0.5 \quad \text{II.43}$$

- **Critère d'information d'Akaike (AIC).**

Le critère d'information d'Akaike (AIC) est une approche statistique couramment utilisée pour comparer différents modèles afin d'identifier le meilleur modèle qui représente les résultats expérimentaux. Le modèle le mieux adapté aux données expérimentales a la valeur d'AIC la plus faible. Pour un échantillon de petite taille, l'AIC d'un modèle peut être calculé à l'aide de l'équation suivante :[203]

$$AIC = N ln\left(\frac{SSE}{N}\right) + 2N_p + \frac{2N_p(N_p+1)}{N-(N_p+1)} \quad \text{II.44}$$

où N est le nombre d'échantillons expérimentaux, Np est le nombre de paramètres du modèle et SSE est la somme des erreurs quadratiques qui peuvent être calculées par l'équation suivante :

$$SSE = \sum_{i=0}^{N} \left(y_i^{exp} - y_i^{cal}\right)^2 \quad \text{II.45}$$

- **Définitions des métriques Q^2 :**

Dans ce paragraphe, les métriques Q^2 analysées sont brièvement présentées. Les trois premiers indicateurs (Q_{F1}^2 et Q_{F2}^2 et Q_{F3}^2) sont définis comme suit:[204]

$$Q_{F1}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{TR})^2} \quad \text{II.46}$$

$$Q_{F2}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{OUT})^2} \quad \text{II.47}$$

$$Q_{F3}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2 / n_{OUT}}{\sum_{i=1}^{n_{TR}} (y_i - \bar{y}_{TR})^2 / n_{TR}} \quad \text{II.48}$$

- **Cefficient de corrélation de concordance (CCC) :**

Le coefficient de corrélation de concordance (CCC) a été proposé pour évaluer la concordance de deux essais d'un test ou d'un instrument en termes de validité et de répétabilité de la méthode. Plus tard, le coefficient de corrélation de concordance CCC a été introduit pour évaluer la capacité de prédiction des modèles de régression. En particulier, le coefficient de corrélation de concordance, appelé ici Q^2_{ccc}, est défini comme suit:

$$Q^2{}_{ccc} = \frac{2.\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{EXP}).(\hat{y}_{i/i} - \bar{y}_{CAL})}{\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{EXP})^2 + \sum_{i=1}^{n_{OUT}} (\hat{y}_{i/i} - \bar{y}_{CAL})^2 + n_{OUT}.(\bar{y}_{EXP} - \bar{y}_{CAL})^2} \quad \text{II.49}$$

En supposant que les réponses expérimentales de l'ensemble externe sont désignées par X et que les réponses prédites par Y, Q^2_{ccc} peuvent être exprimées comme suit:

$$Q^2{}_{ccc} = \frac{2.S_{XY}}{S_{XX} + S_{YY} + (\bar{X} - \bar{Y})^2} \quad \text{II.50}$$

Où S_{XY}, S_{XX} et S_{YY} sont la covariance entre X et Y et leurs variances, respectivement

La métrique R_m, appelée ici Q^2_{Rm}, a été proposée pour évaluer la capacité prédictive des modèles QSAR par régression. En supposant toujours que les réponses expérimentales et prédites de l'ensemble externe sont respectivement notées X et Y, Q^2_{Rm} est défini comme suit:

$$Q^2{}_{Rm} = \frac{1}{2}.\left[R^2.\left(1 - \sqrt{R^2 - R^2{}_{0.XY}}\right) + R^2.\left(1 - \sqrt{R^2 - R^2{}_{0.YX}}\right)\right] = R^2.\left[1 - \frac{\sqrt{R^2 - R^2{}_{0.XY}} + \sqrt{R^2 - R^2{}_{0.YX}}}{2}\right] \quad \text{II.51}$$

Où R^2 est le coefficient de détermination calculé entre X et Y, $R^2_{0,XY}$ est le coefficient de détermination des valeurs prédites par rapport aux valeurs expérimentales et $R^2_{0,YX}$ est le

coefficient de détermination des valeurs expérimentales par rapport aux valeurs prédites, tous deux calculés pour une droite de régression sur l'origine.

II.6 Conclusion

Dans ce chapitre, le concept de système d'apprentissage a été présenté. Une vision générale et purement mathématique a été donnée sur les méthodes d'intelligence computationnelle qui sont des méthodes alternatives et complémentaires à l'expérimental. Ces dernières ont permis de faciliter la résolution d'un problème mathématique.

Les méthodes décrites dans ce chapitre sont les réseaux de neurones artificiels, les machines à vecteurs de support des moindres carrés et les machine à vecteurs de support, ainsi que des algorithmes optimisations, tels que l'optimisation par essaim de particules (PSO) et l'algorithme libellule (DA).

Également, les critères et métriques statistiques ont été présentés pour être utilisé dans le test et la validité des modèles développés.

Chapitre III : Résultats et interprétations

Chapitre III : Résultats et Discussion

III.1 Introduction

Le présent chapitre est scindé en deux parties :

La première partie représente les résultats de contribution à la modélisation de la tension superficielle de mélanges binaires (LI-Eau et divers solvants organiques) par les méthodes de réseau neuronal artificiel prédictif (RNA), les machine à vecteurs de support des moindres carrés (LSSVM) et la machine à vecteurs de support optimisée par optimisation par essaim de particules (SVM-PSO)

La deuxième partie représente les résultats de contribution à la modélisation de la solubilité du CO_2 dans les LIs à l'aide de réseau neuronal artificiel prédictif (RNA) et Machine à vecteurs de support optimisée par l'algorithme de libellule (DA-SVM)

III.2 Contribution à la modélisation de la tension superficielle de mélanges binaires.

III.2.1 Elaboration de la base de données

III.2.1.1 Données expérimentales de la tension superficielle

Les données expérimentales de la tension superficielle des mélanges binaires ont été publiées à différentes températures et compositions du mélange. Les cations et les anions constituants les LIs considérés dans cette étude sont représentés dans les tableaux III.1 et III.2 respectivement.

Les données expérimentales utilisées pour la modélisation consistaient en 1623 points de données de tension superficielle (σ) rassemblés à partir de 62 mélanges binaires d'LI et de solvant moléculaire à différentes températures et les compositions ont été collectées à partir de nombreuses publications. Les cations impliqués sont : l'imidazolium [Im], le pyridinium [Py], l'ammonium [N] et le phosphonium [P], tandis que les anions sont : le tétrafluoroborate [BF4], l'hexafluorophosphate [PF6], le bis(trifluorométhylsulfonyl)imide [BTI], le bromure. [Br], chlorure [CL], acétate [AC], sulfate d'alkyle [RSO4], phosphate de diméthyle [DMP], trifluorométhylsulfonate [TfO], nitrate [NO3] et dicyanamide [Dca]. Composants moléculaires destinés à couvrir les solvants courants tels que l'eau, les alcools, le diméthylsulfoxyde, l'acétonitrile et le tétrahydrofurane. Les source et les domaines de données expérimentales des mélanges binaires étudiés sont indiquées dans le tableau 1 (voir AnnexeA)

III.2.1.2 Données des descripteurs de profil sigma

Suite aux résultats obtenus à partir des travaux antérieurs [205] les descripteurs moléculaires basés sur les profiles σ ont été adoptés. Ces descripteurs dérivent de la théorie COSMO-RS considérée comme modèle de solvatation continue qui combine la théorie de chimique quantique, les modèles de continuum diélectrique, les interactions de surface et la thermodynamique statistique. Il calcule les interactions moléculaires à partir des densités de charge de blindage (polarisation) sur les segments de surface moléculaires. Les densités de charge de polarisation ont été calculées par des méthodes de chimie quantique. Les calculs de chimie quantique permettent de fournir un segment de surface discret autour d'une seule molécule intégrée dans un conducteur virtuel. La surface de chaque segment est caractérisée par sa surface et la densité de charge de blindage du segment, qui tient compte de l'interaction électrostatique de la molécule de soluté par son entourage (qui dans un conducteur virtuel est un criblage parfait) et de la rétro-polarisation de la molécule de soluté. La densité de charge de criblage est habituellement donnée dans le diagramme de probabilité de la distribution de la densité de charge statistique à la surface d'une molécule, et également connue sous le nom de profil σ. Le profil σ d'une molécule montre la probabilité de la quantité relative de surface avec la polarité σ sur la surface de la molécule, il est considéré comme les propriétés caractéristiques de la molécule. Les figures III.1et III.2 montrent respectivement la densité de charge en 3D du cation [C2Mim], de l'anion [NTf2] ([BTI]) et du solvant 1-butanol et les profils sigma correspondants.[206]

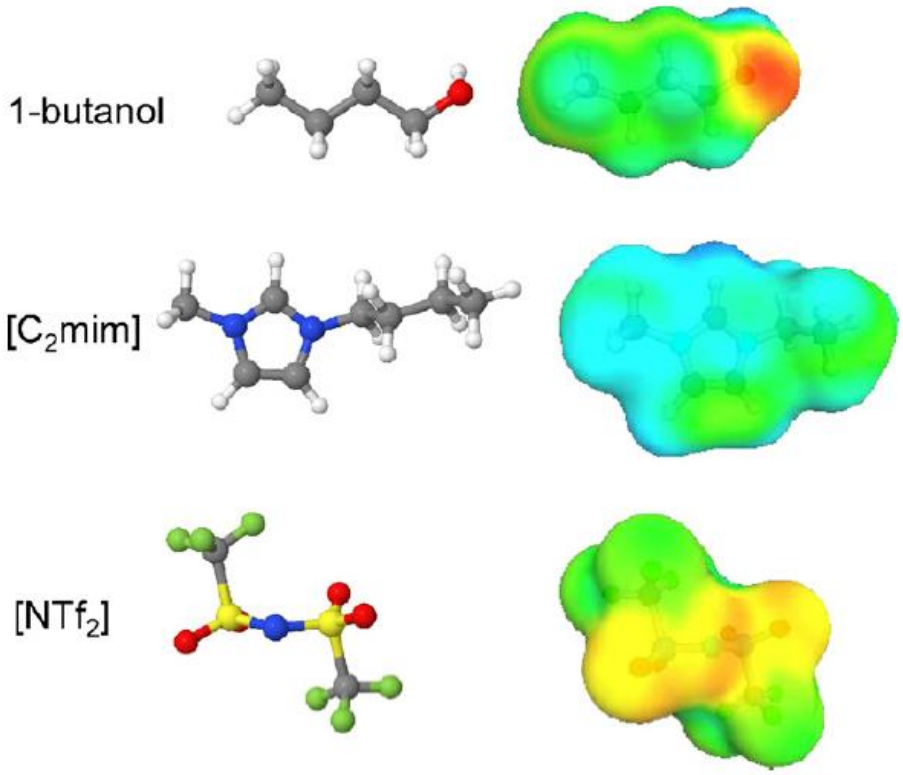


Figure III.1 : Criblage de la densité de charge (sigma) et structures des conformères minimaux de 1-butanol, [C2mim] et [NTf2].

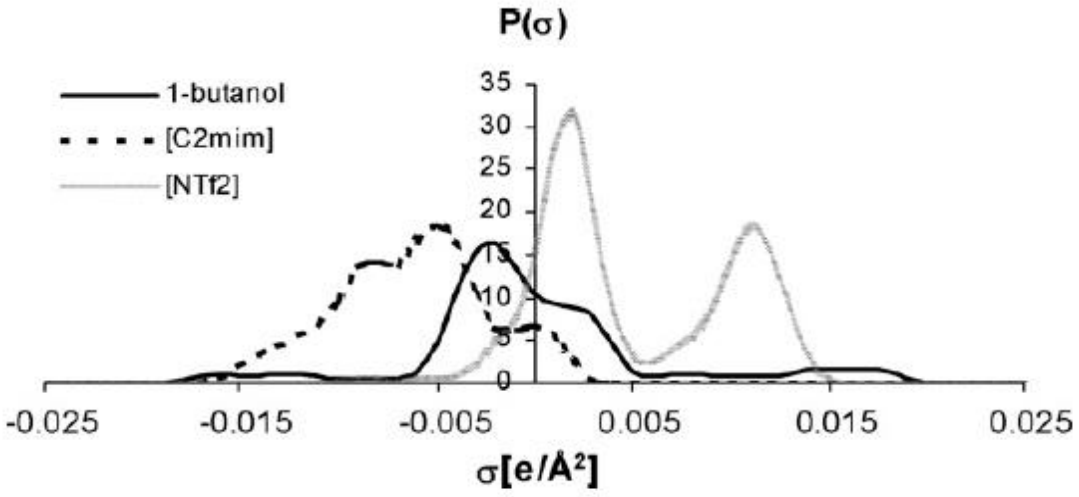


Figure III.2 : Profils sigma des conformères minimum de 1-butanol, [C2mim] et [NTf2].

En d'autres termes, le profil σ d'une molécule comprend les informations chimiques principales nécessaires pour prédire les interactions électrostatiques, de liaison hydrogène et de dispersion possible de la molécule dans un fluide. Les profils σ se sont révélés être un descripteur moléculaire efficace pour établir des modèles de QSPR capables de prédire les propriétés physiques, chimiques et toxicologiques des LIs.

La première étape du développement du modèle consistait en la construction de descripteurs moléculaires appropriés à partir des profils σ des cations, des anions et des solvants des mélanges binaires étudiés. Le profil σ et les descripteurs (S_σ) des cations, des anions et des solvants utilisés dans cette étude ainsi que les divisions des aires des profils σ sont illustrés dans les Figure III.3

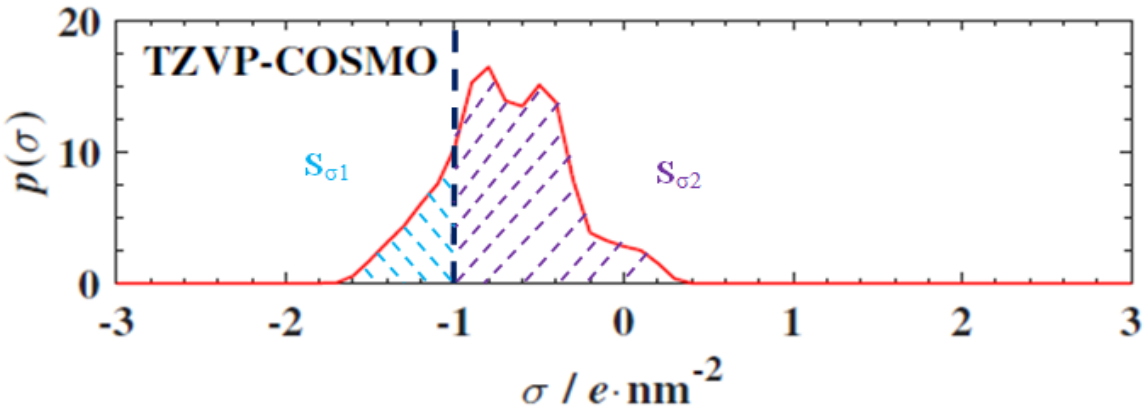


Figure III.3 : descripteurs moléculaire basés sur les Profils σ.

Ainsi, $S_{\sigma 1}$ représente le caractère donneur de liaisons hydrogène, $S_{\sigma 2}$ représente le comportement accepteur de liaison hydrogène.

Dans ce travail un programme MATLAB a été élaboré pour calculer les aires $S_{\sigma i}$=1,..,6 des cations, des anions et des solvants considérés après numérisation des graphes de profiles σ obtenus à partir de la référence[207]et les profiles σ numérisés pour l'eau et les solvants organiques obtenus à partir de bases VT-2005 et VT-2006 des références. [208,209]Les descripteurs fondés sur le profil σ (Sσi) des cations, des anions et des solvants étudiés sont regroupés dans le tableau 1 (Annexe A)

III.2.2 Modélisation de la tension superficielle par les méthodes d'intelligence computationnelle

III.2.2.1 Choix des variables d'entrée et de sortie

Pour le développement des modèles d'intelligence computationnelle, la tension superficielle du mélange a été exprimée comme fonction de la Température (K) la fraction molaire X_{LI} du LI et les descripteurs moléculaires $S_{\sigma 1C}$ et $S_{\sigma 2C}$ du cation, $S_{\sigma 1A}$ et $S_{\sigma 2A}$ d'Anion et $S_{\sigma 1S}$ et $S_{\sigma 2S}$ du solvant :

$$\sigma_m = f(T, X_{LI}, S_{\sigma 1C}, S_{\sigma 2C}, S_{\sigma 1A}, S_{\sigma 2A}, S_{\sigma 1S}, S_{\sigma 2S}) \qquad \text{III.1}$$

III.2.2.2 Division des données en un ensemble d'apprentissage, test et validation

L'ensemble de données d'apprentissage se compose de 70 % de données choisies au hasard à l'aide de la fonction ***dividerand*** de MATLAB 2018. Les 30 % restants ont été utilisés comme un ensemble de test pour LSSVM et SVM-PSO. Dans le cas de la modélisation RNA, ces 30 % ont été divisés en deux ensembles de 15 % pour le test et 15 % pour la validation. Cette division a pour but d'éviter le sur-apprentissage.

III.2.2.3 Critères statistiques :

Pour juger la performance des modèles, une fonction MATALAB a été élaborée pour calculer 15 valeurs statistiques et 3 types d'erreurs pour les ensembles d'apprentissage de test et global. Les formules et significations de ces critères de validité du modèle ont été mentionnées dans le chapitre II.

III.2.2.4 Modélisation par RNA

Pour ce faire, l'algorithme d'apprentissage de retro-propagation de Levenberg-Marqhardt (fonction ***trainlm***de MATLAB) et l'algorithme de la régularisation Bayisienne (fonction ***trainbr*** de MATLAB) ont été utilisés. Ce choix est justifié par le fait que ces deux algorithmes sont considérés comme les meilleures dans plusieurs études antérieures.

Pour la fonction d'activation, les fonctions logsig et tansig ont été utilisés pour les couches cachées et la fonction purelin pour la couche de sortie.

Un programme sous MATLAB 2018 a été élaboré en faisant varier le nombre de neurone dans la couche cachée, l'algorithme d'apprentissage et la fonction d'activation de la couche cachée. Afin d'arriver au modèle qui permet de donner la meilleure erreur quadratique moyenne, le programme a été exécuté 20 fois pour la même topologie. L'évaluation du modèle a été réalisée à l'aide d'outils statistiques et graphiques. La figure III.4 illustre l'architecture de RNA utilisée dans cette étude.

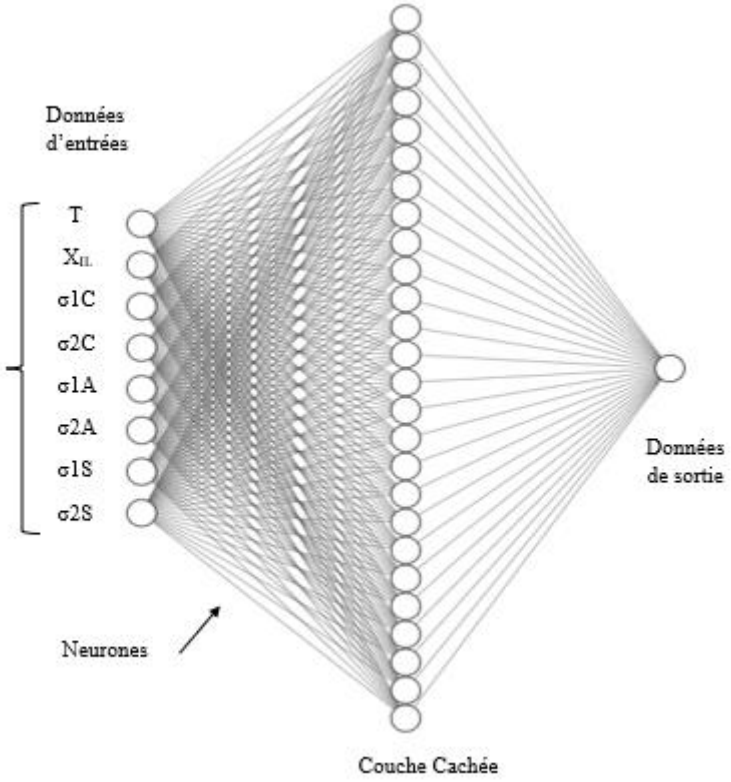


Figure III.4 : L'architecture du RNA

III.2.2.5 Modélisation par SVM-PSO

Le développement du modèle SVM-PSO a été réalisé par la fonction ***fitrsvm*** de l'environnement MATLAB. Cette fonction dispose de plusieurs options de calcul relatives au choix de la méthode de validation croisée, de la fonction noyau (gaussienne, rbf et polynomiale), des valeurs optimales de la constante C (box Constraint), de l'epsilon et du paramètre de la fonction noyau (KernelScale) du modèle optimisé au moyen de l'algorithme PSO (100 itérations). Dans cette étude, la méthode de validation croisée « holdout » fixée à 0,3 a été choisie, ce qui signifie que 70% des données sont utilisées pour l'apprentissage et 30% pour la validation.

Les valeurs optimales sont 200, 1 et 0,006 pour c, k et ep respectivement avec gaussienne comme fonction Kernel. Les étapes conduisant au développement du modèle optimal SVM-PSO hybride sont illustrées dans l'organigramme présenté sur la figure III.5.

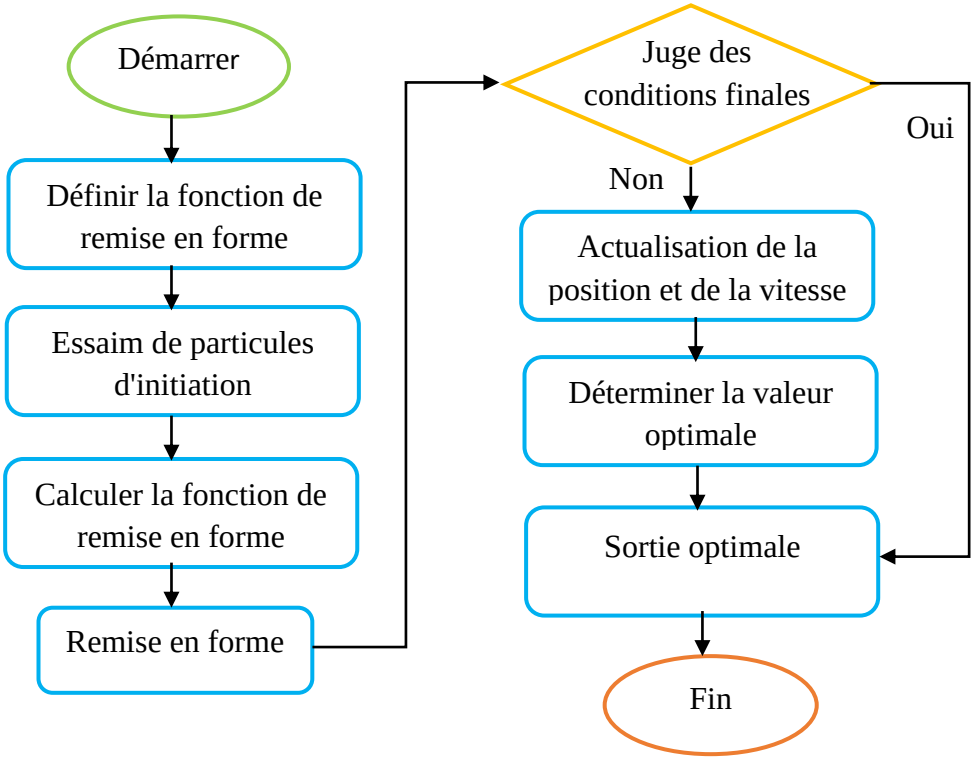


Figure III.5 : Organigramme du processus d'optimisation par SVM-PSO

III.2.2.5 Modélisation par LSSVM

La modélisation LSSVM a été réalisée à l'aide de la boîte à outils LS-SVMlab. Pour ce faire, un algorithme approprié implémenté dans LS-SVMlab détecte et met à l'échelle les variables continues, catégorielles et binaires. Une étape appelée réglage consiste à déterminer les paramètres LSSVM (constante de régularisation γ et paramètre de noyau σ^2) en minimisant une mesure de performance sélectionnée. Les constantes α et b du modèle LSSVM ont été déterminées à partir de l'étape d'apprentissage qui utilise la constante de régularisation γ et le paramètre de noyau σ^2 ajustés dans la première étape. La dernière étape a consisté à tester la capacité de généralisation du modèle sur l'ensemble des données réservées au test. Les paramètres optimisés étaient respectivement de 3 548,43 et 1,6510 pour γ et σ^2.

Le tableau III.1 résume le calcul des différentes métriques statistiques pour les quatre topologies du modèle RNA (« trainbr, logsig » ; « trainbr, tansig » ; « trainlm, logsig » ; « trainlm, tansig »), pour le modèle SVM-PSO et pour le modèle LSSVM.

Le tableau III.2 résume respectivement le calcul des différentes erreurs pour les quatre topologies du modèle RNA, pour le modèle SVM-PSO et pour le modèle LSSVM.

Tableau III.1 : Critères statistiques des quatre topologies de modèle RNA, du modèle SVM-PSO et du modèle LSSVM.

(a): Trainbr, logsig			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9955	0.9964	0.9958
Q2F2	0.9956	0.9964	0.9958
Q2F3	0.9955	0.9965	0.9958
Q2CCC	0.9978	0.9982	0.9979
R^2	0.9955	0.9964	0.9958
R	0.9978	0.9982	0.9979
K	1.0000	0.9994	0.9999
k’	0.9996	1.0003	0.9998
R_0^2	1.0000	1.0000	1.0000
$R_0'^2$	1.0000	1.0000	1.0000
M	-0.0045	-0.0036	-0.0043
N	-0.0045	-0.0036	-0.0043
R_m	0.9291	0.9365	0.9309
A_f	1.0080	1.0098	1.0085
B_f	1.0001	1.0003	1.0002

(b): Trainbr, tansig			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9951	0.9962	0.9955
Q2F2	0.9952	0.9962	0.9955
Q2F3	0.9951	0.9958	0.9953
Q2CCC	0.9976	0.9981	0.9977
R^2	0.9951	0.9962	0.9955
R	0.9976	0.9982	0.9977
K	1.0000	1.0030	1.0008
k’	0.9996	0.9967	0.9988
R_0^2	1.0000	0.9999	1.0000
$R_0'^2$	1.0000	0.9999	1.0000
M	-0.0049	-0.0037	-0.0046
N	-0.0049	-0.0037	-0.0046
R_m	0.9258	0.9358	0.9284
A_f	1.0083	1.0116	1.0091
B_f	1.0001	0.9983	0.9996

(c): Trainlm, logsig			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9946	0.9907	0.9934
Q2F2	0.9947	0.9906	0.9934
Q2F3	0.9946	0.9907	0.9935
Q2CCC	0.9953	0.9973	0.9967
R^2	0.9946	0.9906	0.9934
R	0.9973	0.9953	0.9967
K	0.9999	1.0002	1.0000
k’	0.9997	0.9990	0.9995
R_0^2	1.0000	1.0000	1.0000
$R_0'^2$	1.0000	1.0000	1.0000
M	-0.0054	-0.0095	-0.0066
N	-0.0054	-0.0095	-0.0066
R_m	0.9217	0.8945	0.9129
A_f	1.0101	1.0150	1.0116
B_f	1.0003	0.9990	0.9999

(d): Trainlm, tansig			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9953	0.9941	0.9949
Q2F2	0.9953	0.9941	0.9949
Q2F3	0.9953	0.9938	0.9949
Q2CCC	0.9977	0.9971	0.9975
R^2	0.9953	0.9941	0.9949
R	0.9977	0.9971	0.9975
K	1.0001	1.0013	1.0005
k’	0.9995	0.9982	0.9991
R_0^2	1.0000	1.0000	1.0000
$R_0'^2$	1.0000	1.0000	1.0000
M	-0.0047	-0.0059	-0.0051
N	-0.0047	-0.0059	-0.0051
R_m	0.9272	0.9182	0.9242
A_f	1.0091	1.0128	1.0102
B_f	0.9999	0.9981	0.9993

(e): Modèle LSSVM			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9906	0.9408	0.9754
Q2F2	0.9907	0.9406	0.9754
Q2F3	0.9394	0.9906	0.9752
Q2CCC	0.9953	0.9694	0.9875
R^2	0.9906	0.9406	0.9754
R	0.9953	0.9699	0.9876
K	1.0002	1.0011	1.0005
k’	0.9991	0.9942	0.9976
R_0^2	1.0000	1.0000	1.0000
$R_0'^2$	1.0000	0.9996	0.9999
M	-0.0095	-0.0631	-0.0252
N	-0.0095	-0.0627	-0.0251
R_m	0.8947	0.7114	0.8225
A_f	1.0082	1.0219	1.0123
B_f	1.0003	1.0004	1.0003

(f): Modèle SVM-PSO			
Critères Statiques	**Train**	**Test**	**Global**
Q2F1	0.9721	0.9376	0.9614
Q2F2	0.9722	0.9374	0.9613
Q2F3	0.9721	0.9343	0.9607
Q2CCC	0.9855	0.9677	0.9800
R^2	0.9721	0.9374	0.9613
R	0.9863	0.9687	0.9808
K	1.0064	1.0081	1.0069
k’	0.9915	0.9870	0.9901
R_0^2	0.9995	0.9991	0.9994
$R_0'^2$	0.9991	0.9979	0.9988
M	-0.0282	-0.0658	-0.0396
N	-0.0278	-0.0645	-0.0389
R_m	0.8112	0.7046	0.7739
A_f	1.0084	1.0217	1.0124
B_f	0.9967	0.9928	0.9955

Tableau III.2 : Différentes erreurs des quatre topologies de modèle RNA, le modèle SVM-PSO et le modèle LSSVM.

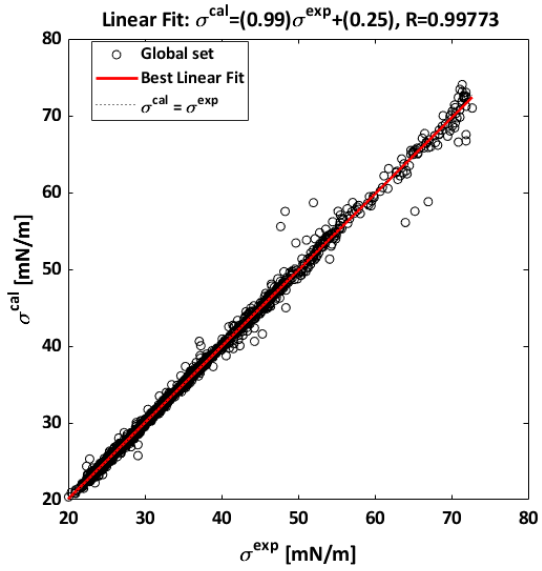
trainbr, logsig				**trainbr, tansig**			
Errors	**Train**	**Test**	**Global**	**Errors**	**Train**	**Test**	**Global**
ERAM%	0.7992	0.9812	**0.8466**	ERAM%	0.8256	1.1526	0.9108
EQRM	0.7235	0.6441	0.7037	EQRM	0.7416	0.6917	0.7289
EQM	0.5235	0.4149	**0.4952**	EQM	0.5499	0.4785	0.5313
trainlm, logsig				**trainlm, tansig**			
Errors	**Train**	**Test**	**Global**	**Errors**	**Train**	**Test**	**Global**
ERAM%	1.0111	1.4779	1.1511	ERAM%	0.9102	1.2568	1.0141
EQRM	0.7936	1.0421	0.8756	EQRM	0.7346	0.8414	0.7682
EQM	0.6298	1.0860	0.7667	EQM	0.5396	0.7079	0.5901
Modèle LSSVM				**Modèle SVM-PSO**			
Errors	**Train**	**Test**	**Global**	**Errors**	**Train**	**Test**	**Global**
ERAM%	0.8150	2.1653	1.2201	ERAM%	0.8050	2.0399	1.1755
EQRM	1.0430	2.6514	1.6942	EQRM	1.7923	2.7482	2.1247
EQM	1.0878	7.0301	2.8705	EQM	3.2122	7.5524	4.5143

A la lumière de ces résultats, tous les critères statistiques calculés vérifient les conditions d'acceptabilité. Cependant, les résultats du modèle optimisé en utilisant l'algorithme d'entraînement ***trainbr*** et la fonction d'activation ***logsig*** ont montré les meilleurs résultats significatifs : un coefficient de corrélation global R de 0,9979, le coefficient de détermination R^2 de 0,9958 et le facteur de précision Af de 1,0085. De plus, les paramètres d'erreur tels que : l'EQM (0,4952) et L'ERAM% (0,8466). Cela permet de prouver l'efficacité et la robustesse du modèle développé, qui présente des performances de corrélation et des capacités de généralisation et d'interpolation très significatives. D'après le tableau III.3, la plus petite valeur obtenue de l'AIC représente le modèle le mieux adapté aux six modèles étudiés.

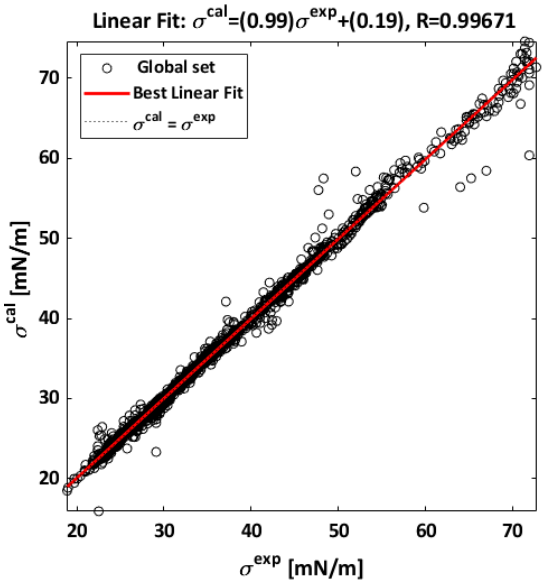
Tableau III.3 : Les valeurs de AIC des quatre topologies de modèle RNA, du modèle SVM-PSO et du modèle LSSVM.

Modèle	**RNA**				**LSSVM**	**SVM-PSO**
	trainbr, logsig	**trainbr, tansig**	**trainlm, logsig**	**trainlm, tansig**		
AIC	**-1178.675**	-1058.967	-435.523	-880.721	1808.694	2578.408

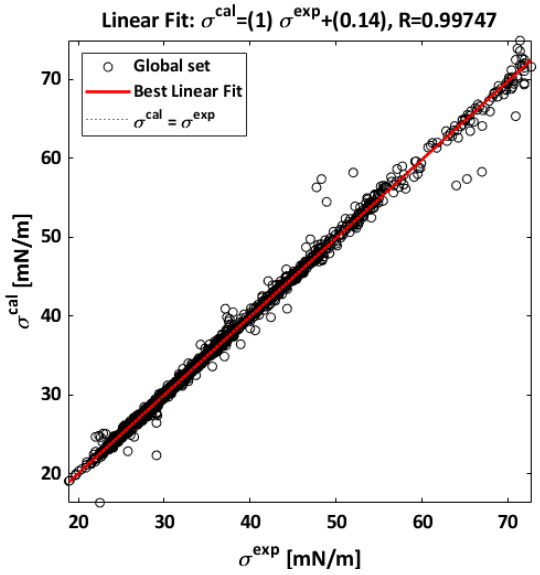
L'évaluation graphique a été réalisée à l'aide de diagramme de dispersion (pour les ensembles d'apprentissage, de test, de validation et global), le diagramme d'erreur relative (ER %) et le diagramme de distribution d'erreur relative (ER %). La figure III.6 montre le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale pour les quatre topologies du modèle RNA, du modèle LSSVM et du modèle SVM-PSO pour la base globale. La figure III.7 représente le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale pour le modèle optimisé **trainbr** et **logsig** qui est bien illustré sur la figure III.8. Ce diagramme montre clairement la concordance entre les valeurs calculées et les valeurs expérimentales par une forte dispersion des points sur la première bissectrice. Cependant, l'existence d'un certain nombre de valeurs aberrantes a été constatée.



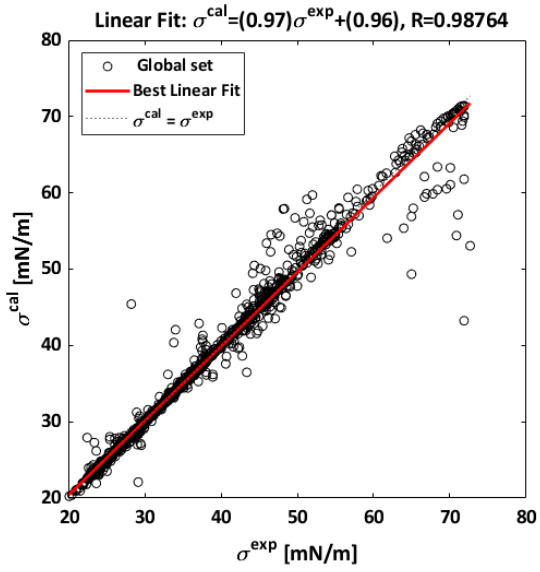
Trainbr, tansig



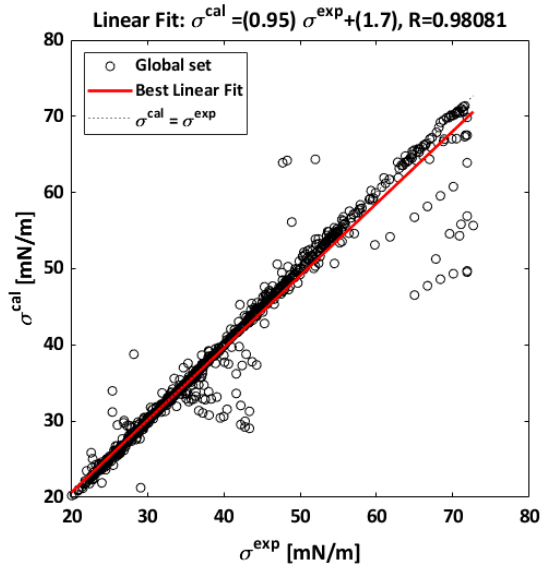
Trainlm, logsig



Trainlm, tansig

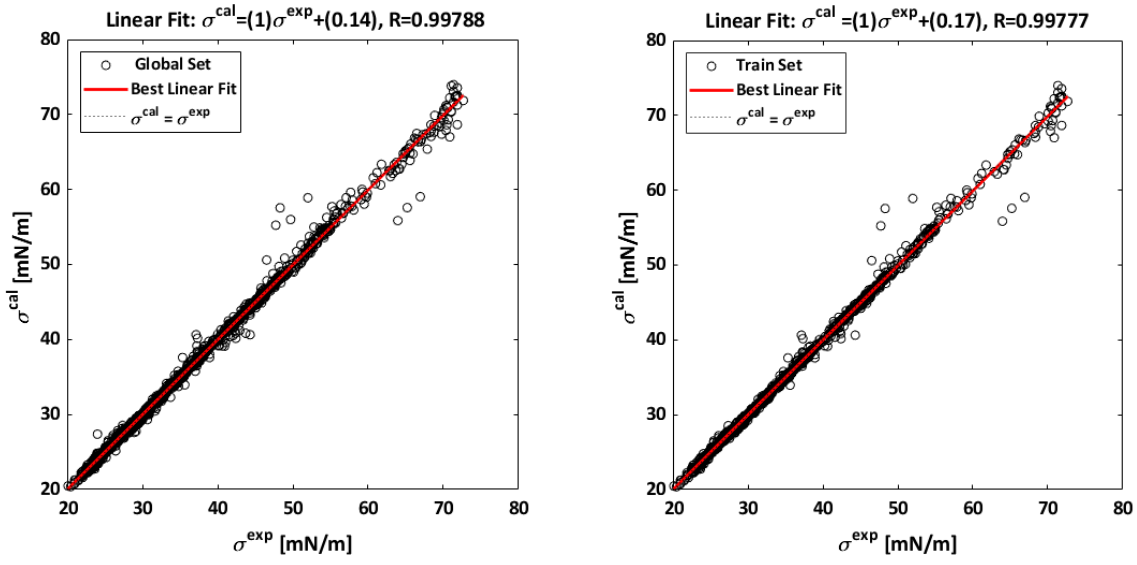


LSSVM

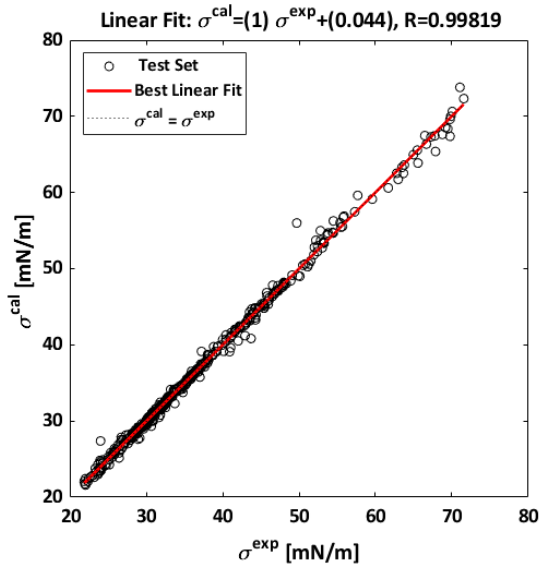


SVM-PSO

Figure III.6: Diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale pour la base globale par les trois topologies du modèle RNA, le modèle LSSVM et le modèle SVM-PSO.



(a) (b)



(c)

Figure III.7 : Diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test par le meilleur modèle de RNA « trainbr, logsig ».

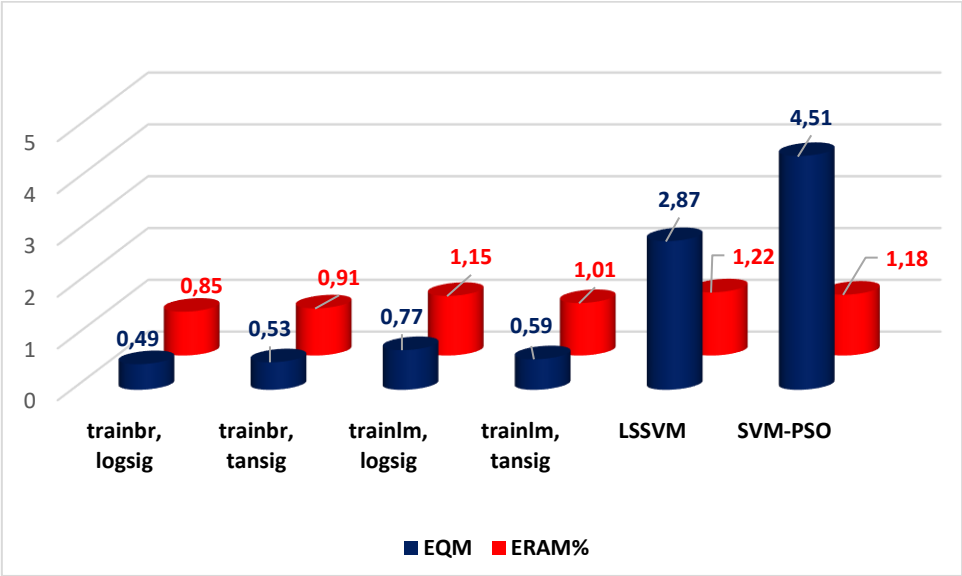


Figure III.8 : Histogramme représentant la différence en termes d'erreur (EQM et ERAM%) entre les modèles étudiés.

Afin de visualiser quantitativement la distribution des erreurs relatives, deux types de diagrammes ont été adoptés. Le premier diagramme montre l'erreur relative en fonction de la tension superficielle représentée sur la figure III.9 et le deuxième montre la distribution de cette erreur représentée sur la figure III.10 Cela montre que la majorité des points sont situés

autour de la ligne zéro. Quantitativement, l'erreur relative pour les quelques valeurs aberrantes ne dépassait pas 20 %.

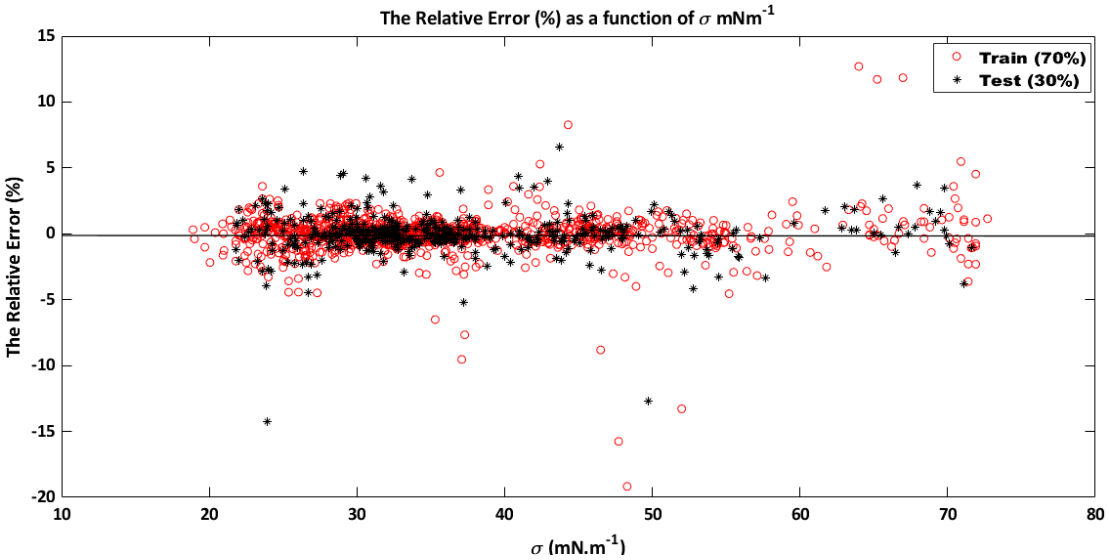


Figure III.9 : L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainbr, logsig ».

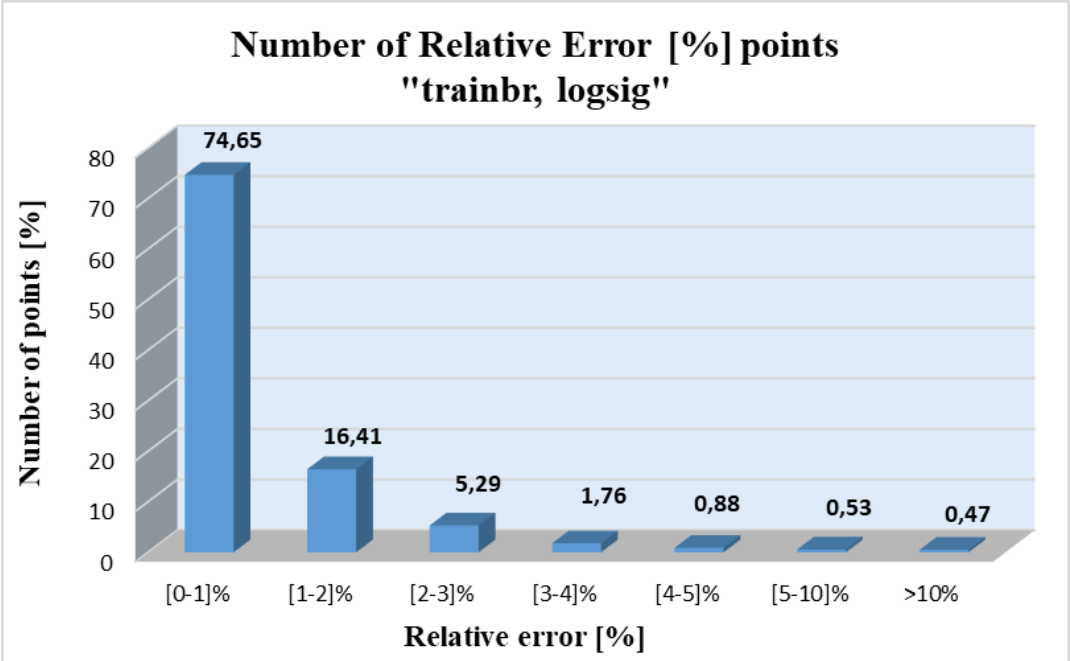


Figure III.10 : Diagramme de distribution d'ER[%] du modèle RNA « trainbr, logsig ».

Pour confirmer la capacité de généralisation du modèle développé, des calculs d'interpolation et d'extrapolation ont été réalisés pour la composition et la température. La figure III.11montre les isothermes de variation de la tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Diméthylsulfoxyde pour le modèle établi, les points interpolés suivent l'allure les points calculés. De la même façon, la figure III.12 représente le

calcul d'interpolation et d'extrapolation de la température. Les isothermes d'interpolation (305,15 K) et d'extrapolation (290,15 K et 315,15 K) ont les mêmes allures des isothermes expérimentales adjacentes.

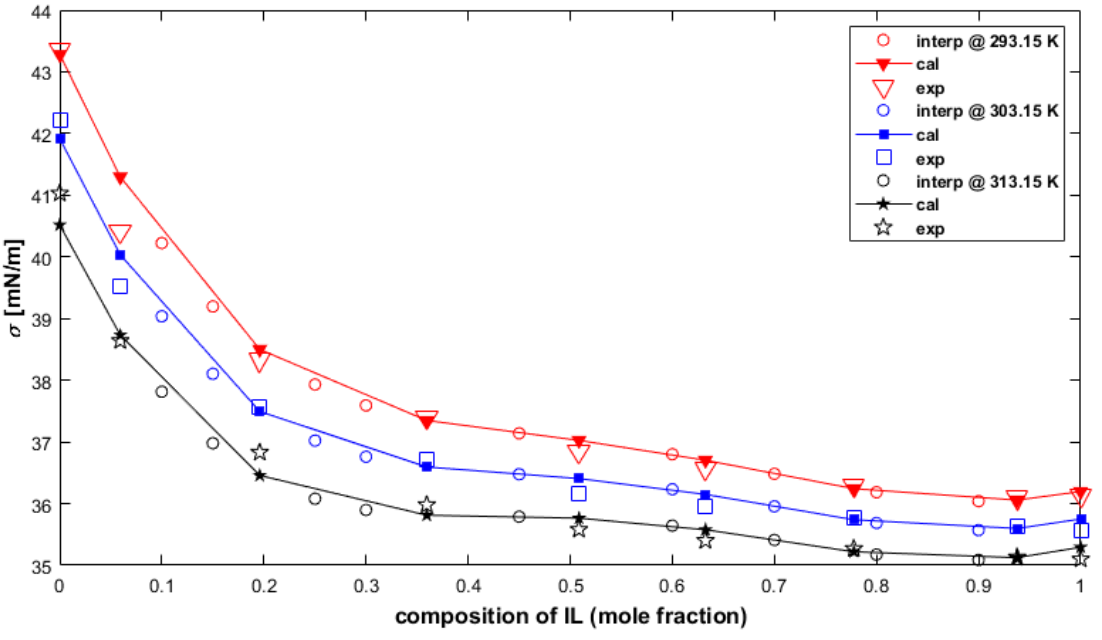


Figure III.11 : Calculs d'interpolation et d'extrapolation de la composition pour les isothermes de variation de tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Diméthylsulfoxyde pour le modèle RNA « trainbr, logsig ».

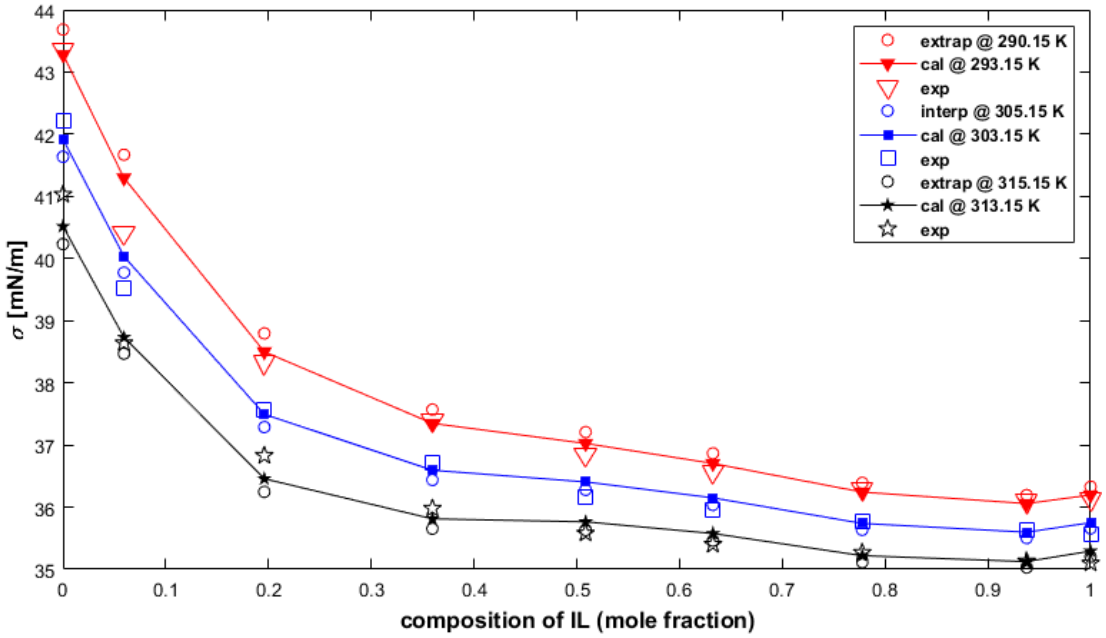


Figure III.12 : Calculs d'interpolation et d'extrapolation de la température pour les isothermes de variation de tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Diméthylsulfoxyde pour le modèle RNA « trainbr, logsig ».

III.2.3 Analyse de la sensibilité des paramètres d'entrée

Pour déterminer l'importance relative des variables d'entrée, différentes méthodes statistiques ont été appliquées.[210,211]Dans le présent travail, la méthode de Garson a été utilisée pour diviser les poids de connexion de sortie cachées en composants associés à chaque neurone d'entrée en utilisant des valeurs de poids de connexion absolues. Ainsi que le facteur de pertinence r, qui est compris entre −1 et +1 et est donné par l'équation suivante[212] :

$$r_k = \frac{\sum_{i=1}^{n}(X_{k,i}-\bar{X}_k)(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_{k,i}-\bar{X}_k)^2 \sum_{i}^{n}(Y_i-\bar{Y})^2}} \qquad \text{III.2}$$

Où $X_{k,i}$ est le i^{ème} paramètre d'importation, Y_i est la i^{ème} valeur d'exportation, $\bar{X}_k$est la valeur moyenne de la k^{ième} entrée, $\bar{Y}$ est la valeur moyenne du paramètre d'exportation et n est le nombre d'ensembles.

Comme le montre la figure III.13 pour le modèle établi, toutes les variables d'entrée ont presque la même importance selon la méthode de Garson, à l'exception de la température. Cela peut être expliqué par la faible variabilité des données de température, c'est-à-dire que pour certains mélanges binaires, les données expérimentales ne sont disponibles qu'à une ou deux températures. Cependant, le descripteur de profil sigma lié au caractère accepteur de liaison H du solvant, Sσ2S, a la plus grande importance relative (20,26 %). Sur la figure III.14, la tension superficielle a montré une dépendance directe aux descripteurs nommés : S_σ2A, S_σ1S, et S_σ2S et une dépendance opposée au reste des descripteurs. Les descripteurs les plus pertinents sont le S_σ2S et la température avec un facteur de pertinence de +0,021 et - 0,0235 respectivement. Par contre, S_σ1A et S_σ2A étaient les moins pertinents, avec un facteur de pertinence de - 0,0052 et +0,0024, respectivement.

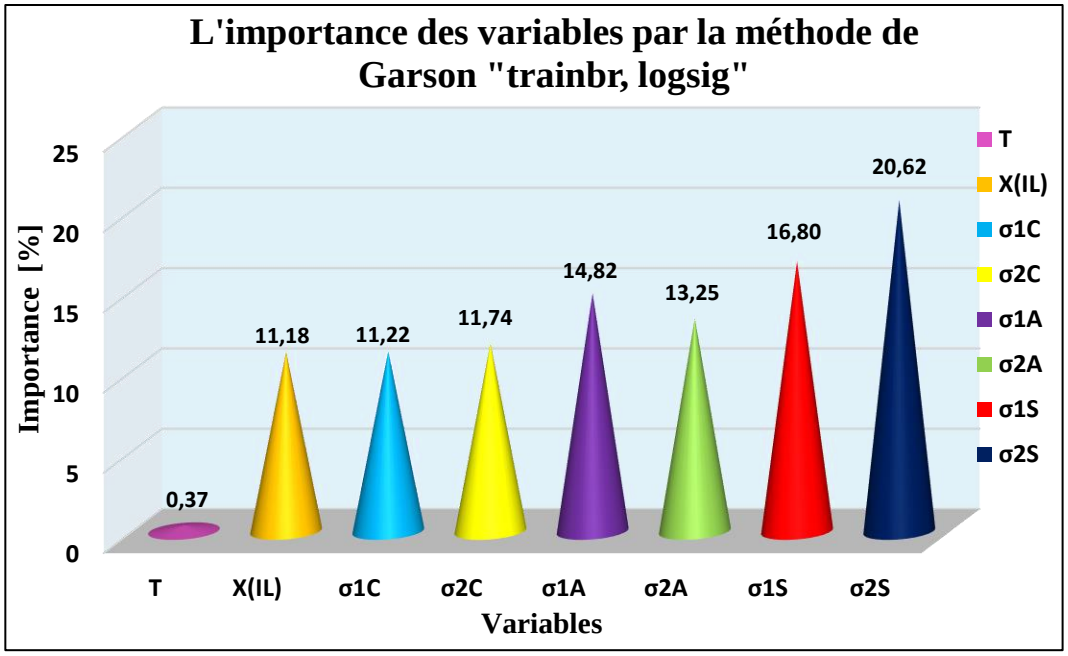


Figure III.13 : Importance relative des variables par la méthode Garson du modèle RNA«trainbr, logsig ».

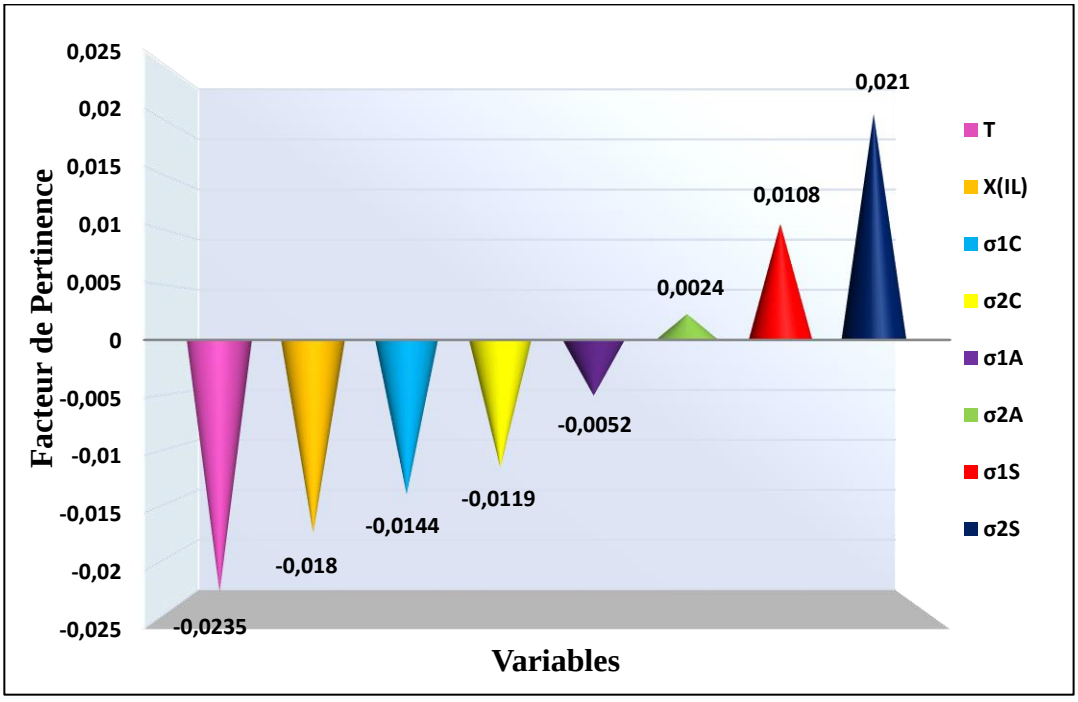


Figure III.14 : Résultats du facteur de pertinence effectué sur le modèle RNA «trainbr, logsig».

III.2.4 Domaine d'applicabilité

Les diagrammes de dispersions des modèles optimisés ont montré un certains nombres des points éloignés de la première bissectrice qui peuvent être attribués à l'existence probable

de points aberrants. Ceci a permis de réaliser un diagnostic des données potentiellement suspectes selon la méthode décrite dans les deux références suivantes[157,213]

Les valeurs aberrantes sont bien connues en tant que points de données distincts de la tendance générale des principaux points de données. La détection des valeurs aberrantes, qui joue un rôle important dans la proposition de modèles, est utilisée pour reconnaître des données individuelles (ou des groupes de données) pouvant différer de la masse des données constituant un ensemble de données. À cette fin, il est essentiel de trouver des méthodes rigoureuses pour détecter les valeurs potentiellement suspectes afin d'éliminer des données expérimentales inexactes et d'améliorer la précision des modèles[157,213]. Dans cette étude, la technique mathématique ‘‘Leverage’’ a été utilisée pour rechercher les valeurs aberrantes. Dans cette technique, les valeurs résiduelles sont d'abord calculées. Ensuite une matrice H de points de données d'entrée a été créée sur la base de la formule suivante :

$$H = X(X^tX)^{-1}X^t \qquad \text{III.3}$$

Où X fait référence à la matrice m × n (m et n représentent respectivement le nombre d'échantillons et les paramètres (variables d’entrée) du modèle). Les valeurs de Hat sont obtenues à partir de la diagonale principale de la matrice de H.

$$Hat = diagonale(H) \qquad \text{III.4}$$

Le diagramme de William représente les résidus normalisés en fonction du ‘’Leverage’’ levier (Hat) pour détecter graphiquement les valeurs aberrantes. La figure III.15 montre le graphique de William basé sur les résultats du modèle RNA. Dans ce graphique, le levier critique (H*) (seuil) est généralement fixé à la valeur donnée par l'équation suivante :

$$H^* = \frac{3(n+1)}{m} \qquad \text{III.5}$$

Ou m et n représentent respectivement le nombre d'échantillons et les paramètres (variables d’entrée) du modèle.

Les résidus normalisés ont été calculés à partir des données de la tension superficielle expérimentale et celle calculée par le modèle.

$$(R_Norm)_i = \frac{\left(\sigma_i^{exp} - \sigma_i^{cal}\right)}{\sqrt{Var\left(\sigma^{exp} - \sigma^{cal}\right)}} i = 1, \dots m \qquad \text{III.6}$$

Le résidu normalisé égal à trois est considéré comme une valeur ‘’seuil’’ pour accepter les points de données dans les écarts types compris sur une plage de ± 3 par rapport à la moyenne (pour couvrir 99% des données normalement distribuées). Si la majorité des points de données se situent dans les plages de $0 \leq \text{Hat} \leq H^*$ et de $-3 \leq R_Norm \leq 3$, cela indique que le développement du modèle et ses prédictions sont effectués dans le domaine d'applicabilité, ce qui conduit à un modèle statistiquement valide. Ainsi on peut affirmer qu'il existe des points ‘’Good High Leverage’’ dans le domaine de $0 \leq \text{Hat} \leq H^*$ et $-3 \leq R_Norm \leq 3$. Les points situés dans la plage de $R_Norm<-3$ ou $R_Norm>3$ (qu’ils soient plus grands ou plus petits que la valeur H*) sont désignés comme points aberrants du modèle ou comme points ‘’Bad High Leverage’’.

La figure III.15 montre que la majorité des points de données se situent dans les domaines d'applicabilité, à l'exception de l'existence de 2,82 % de données suspectes. Ceci confirme les observations faites lors des interprétations des nuages de points et des diagrammes d’erreur relative.

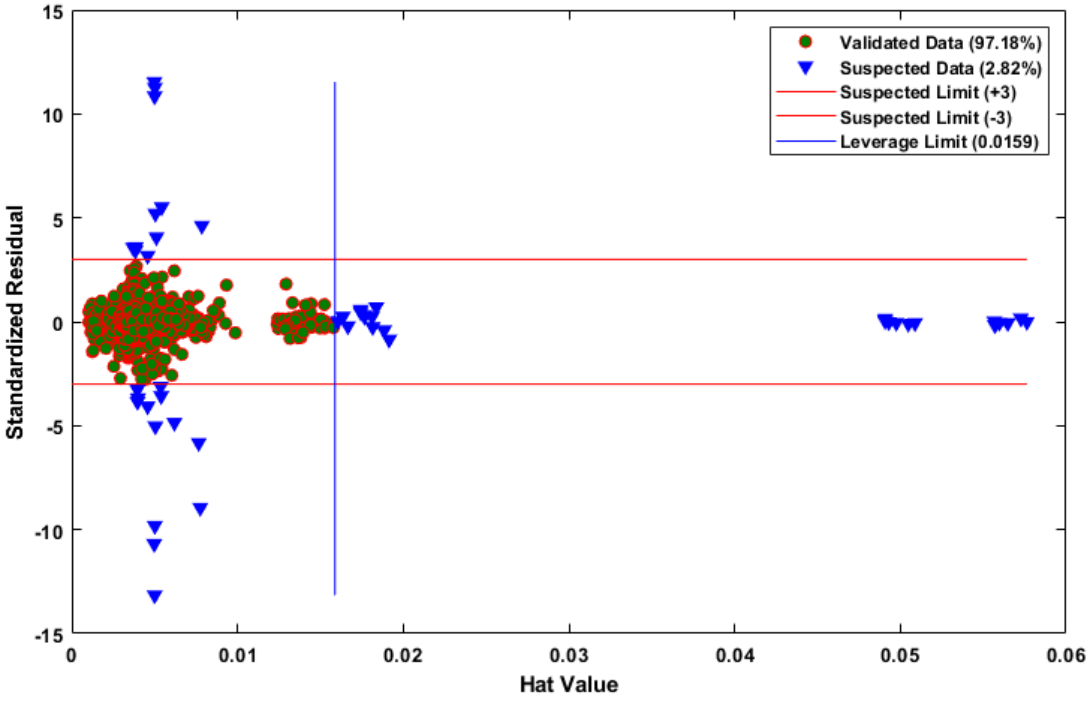


Figure III.15 : Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.

III.2.5 Comparaison avec d'autres modèles de la littérature

Dans cette section, une étude comparative a été réalisée pour évaluer les performances du modèle par rapport aux articles précédents en termes de systèmes communs parce qu’il

n'existe aucune étude contenant une base de données similaire à l'ensemble de données utilisées dans cette étude.

Une comparaison, en termes d'erreur (ERAM%), entre ce travail et les travaux de Huang et al.[41], Cardona et Valderrama.[114]est donnée dans les tableaux III.4 et III.5, respectivement. Les figures III.16 et III.17 représentent une autre comparaison entre la tension superficielle expérimentale et calculée de certains systèmes entre le modèle établi dans ce travail et les articles mentionnés précédemment.

Tableau III.4 : Comparaison entre le présent travail et celui de Huang et al. (2015) en termes d'ERAM%.

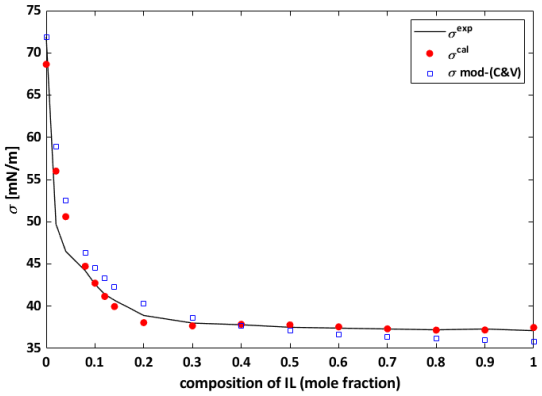
Systèmes {LI+ solvants}	Nombre de points	Notre Modèle RNA « trainbr, logsig »		Modèle de Huang et al			
		moyenne|%Δσ|	|%Δσ|$_{max}$	moyenne|%Δσ$_{SM}$|	|%Δσ|$_{max}$	Moyenne|%Δσ$_{RNA}$|	|%Δσ|$_{max}$
[C4MIm][BTI]+ 1-butanol	13	2.13690	14.2458	2.39918	4.83586	2.99069	5.56928
[C4MIm][BTI]+ 1-propanol	11	0.94144	2.64873	1.35199	3.04797	1.23155	2.18982
[C2MIm][BTI]+ acétonitrile	45	0.58221	2.03760	1.10309	3.61209	0.35043	1.63301
[C4MIm][BTI]+ acétonitrile	45	0.28893	1.22930	1.32803	4.15549	0.4058	0.99813
[C2MIm][BTI]+ Diméthylesulfoxide	45	0.48975	2.16922	1.59518	4.79112	0.13389	0.49756
[C4MIm][BTI]+ Diméthylesulfoxide	50	0.90014	19.16957	2.29169	8.29236	0.15671	0.53247
[C2MIm][C8SO4] + éthanol	18	1.01942	2.10372	1.84492	4.20381	0.44156	1.42752
[C4MIm][BF4]+ éthanol	12	0.86002	2.50953	0.70218	1.78356	1.0015	6.5699
[C6MIm][BF4]+ éthanol	10	3.20687	2.82941	1.14056	1.50424	0.74439	3.40772
[C8MIm][BF4]+ éthanol	10	0.76441	2.28718	0.82575	1.84567	1.17232	2.60078
[C1MIm][MeSO4]+ méthanol	9	0.89797	19.1695	1.04938	9.30710	0.56188	3.46873
[C2MIm][MeSO4]+ méthanol	7	0.69485	1.75784	0.61541	2.47781	0.67376	3.26845
[C2MIm][BTI]+ tetrahydrofuran	44	0.93363	3.58668	0.46162	1.46741	0.46749	1.4991
[C4MIm][BTI]+ tetrahydrofuran	40	1.22362	4.48224	0.58329	2.60047	0.63473	11.623
[C1MIm][MeSO4]+ Eau	10	0.86816	3.80506	0.95716	2.41303	0.33636	1.47093
[C2MIm][BF4]+ Eau	9	0.63134	1.20231	0.18729	1.07727	0.19068	0.50178

[C2MIm][EtSO4] + Eau	12	0.57436	2.21932	1.59173	3.93446	0.60338	1.77915
[C2MIm][MeSO 3]+ Eau	27	1.73691	3.70604	3.89171	6.62818	0.33508	1.02659
[C4MIm][BF4]+ Eau	15	0.73874	1.46092	0.55320	1.96164	0.4154	1.07962
[C4Py][NO3]+ Eau	15	1.04161	6.56224	3.71712	9.3071	1.6599	11.623
[C6MIm][BF4]+ Eau	8	0.33916	0.57460	0.19374	0.38724	0.20127	0.43478
[P666(14)][BTI]+ water	6	0.16640	0.34975	1.41799	2.47305	0.11722	0.22418
[P666(14)][Dca]+ Eau	6	0.16995	0.38829	1.24882	2.36076	0.33602	0.54937

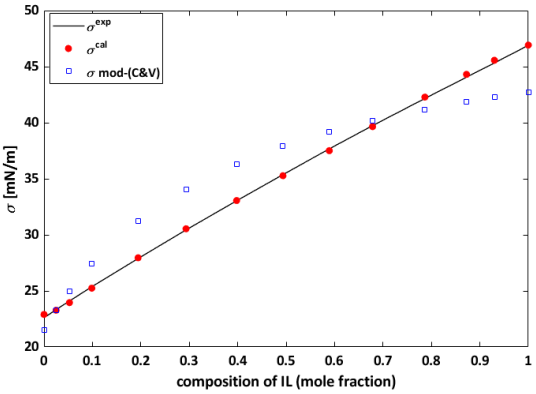
Tableau III.5 : Comparaison entre le présent travail et le travail de Cardona et Valderrama en (2020) termes d'ERAM%.

Systèmes {LI+ solvants}	N	Notre Modèle RNA « trainbr, logsig »		Modèlede Cardona etValderrama	
		moyenne $\lvert\%\Delta\sigma\rvert$	$\lvert\%\Delta\sigma\rvert_{max}$	moyenne $\lvert\%\Delta\sigma\rvert$	$\lvert\%\Delta\sigma\rvert_{max}$
[C2MIm] [AC]+1-propanol	77	0.727311	3.279583	10.73871	22.48719
[C8MIm] [BTI]+ 1-propanol	33	0.513259	2.078608	6.64114	14.11421
[PYR-4.1] [BTI]+acétonitrile	27	0.805194	2.216206	1.708758	4.497877
[C2MIm] [AC]+éthanol	77	0.419793	2.068807	7.379528	20.68375
[C2MIm][DEP]+ éthanol	11	1.369486	4.561168	11.7884	22.08567
[C2MIm][EtSO4] + éthanol	11	0.994978	2.051358	13.25353	22.58346
[C4MIm][BTI] + éthanol	36	0.38637	0.784236	6.433625	18.33271
[C4MIm][DMP] + éthanol	11	1.343429	2.666525	14.05437	26.31037
[C6MIm] [EtSO4] + éthanol	13	1.091658	2.951107	3.927588	8.77661
[C8MIm] [BTI] + éthanol	39	0.44928	1.518148	5.317242	15.36972
[C1MIm][DMP]+méthanol	11	0.562193	1.096063	6.937792	11.1255
[C2MIm] [AC] +méthanol	55	0.728288	3.611441	2.137411	9.412672
[C2MIm][EtSO4] +méthanol	11	0.955441	2.273106	5.599972	10.49703
[C2MIm][MeSO4] +méthanol	14	0.519645	2.936068	6.746018	15.85075
[C4MIm][BTI] +méthanol	24	0.194314	0.717207	2.579008	6.910141
[C4MIm][DMP] +méthanol	11	0.63623	1.737604	2.965154	5.431559
[C4Py] [BF4] +méthanol	56	0.438397	2.181546	7.621787	18.83593
[C8MIm] [BTI] +méthanol	20	1.102813	4.23808	2.704397	6.313109
[C1MIm][DMP] + Eau	16	1.289812	2.932375	1.380303	71.90
[C2MIm][DEP]+ Eau	16	2.236424	12.66962	4.54	18.54
[C2MIm][DMP]+ Eau	11	1.186484	3.990389	4.665865	9.781574
[C4MIm] [AC]+ Eau	49	0.385264	1.20122	3.213681	9.603819
[C4MIm] [CL]+ Eau	46	0.882057	2.076755	2.79283	7.822873
[C4MIm] [PF6]+Eau	9	0.416549	1.264713	3.662293	7.550244
[C4MIm][DBP]+ Eau	11	0.941997	4.423875	5.830528	71.90
[C4Py] [BF4]+Eau	72	0.694389	3.243486	8.413314	24.03831
[C6MIm] [AC]+ Eau	49	0.345018	1.013274	3.203462	10.76396
[C6MIm] [CL]+ Eau	10	2.258253	5.479549	17.48413	26.5232

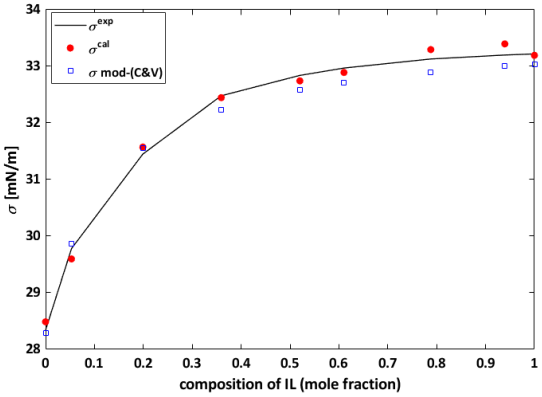
[C6MIm][BF4]+ Eau	8	0.339157	0.574605	6.808245	10.50152



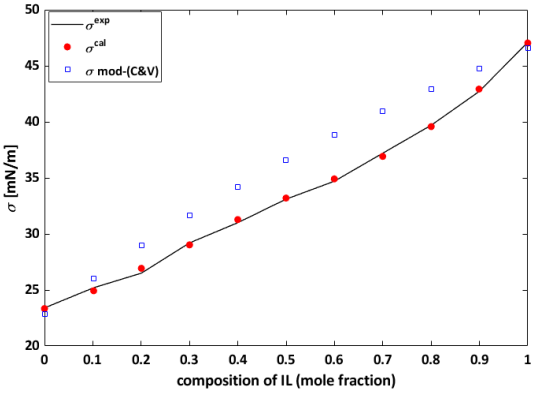
[C2MIm][DEP]-water at T=298.15K



[C4Py] [BF4]-methanol at T=293.15K

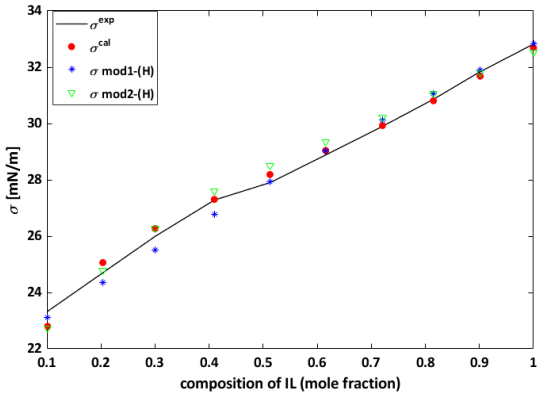


[PYR-4,1] [BTI]-acetonitrile at T=298.15K

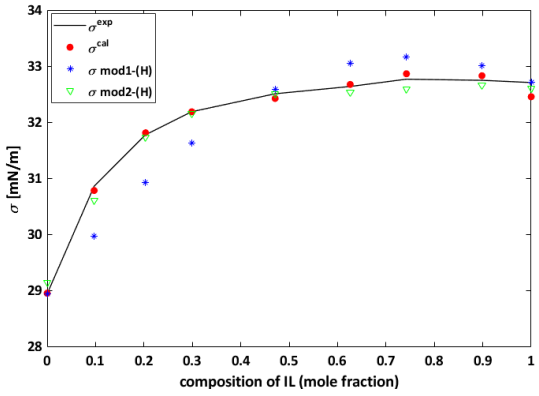


[C2MIm] [AC]-1-propanol at T=298.15K

Figure III.16 : Comparaison entre les valeurs expérimentales (lignes) et calculées de la tension superficielle par le modèle proposé (points rouges) et le modèle de Valderrama (2020) (carrés bleus) pour certains systèmes.



[C8MIm][BF4]-ethanol at T=298.15K

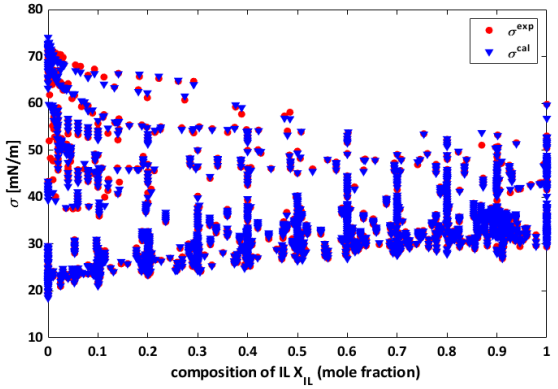


[C4MIm][BTI]-acetonitrileat T=293.15K

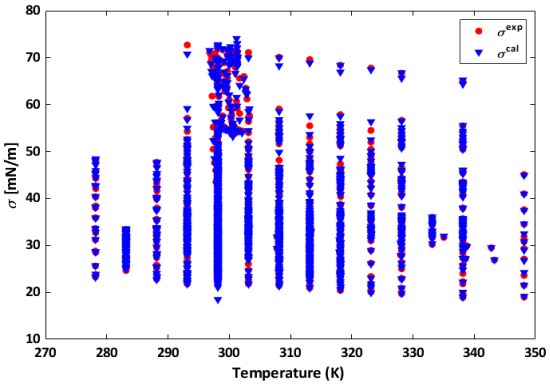
Figure III.17 : Comparaison entre les valeurs expérimentales (lignes) et calculées de la tension superficielle par le modèle proposé (points rouges) et les modèles de Huang

et al. (2015) (triangle vert) pour son modèle RNA et (étoiles bleues) pour son modèle SM pour certains systèmes.

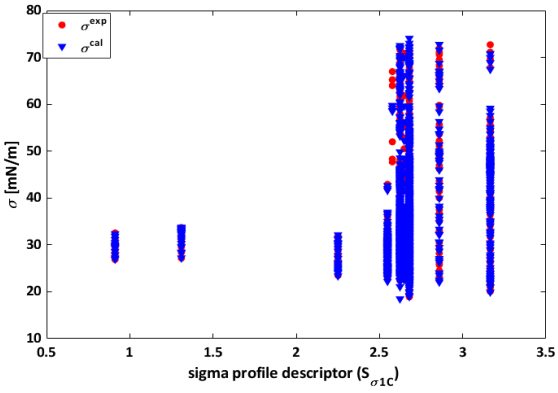
La figure III.18 illustre une comparaison entre les valeurs calculées et expérimentales du modèle RNA (trainbr, logsig) en fonction des entrées.



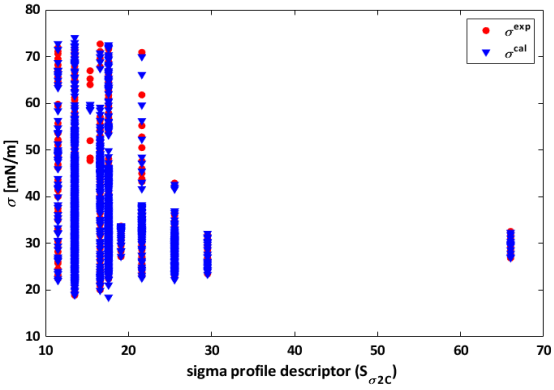
(a)



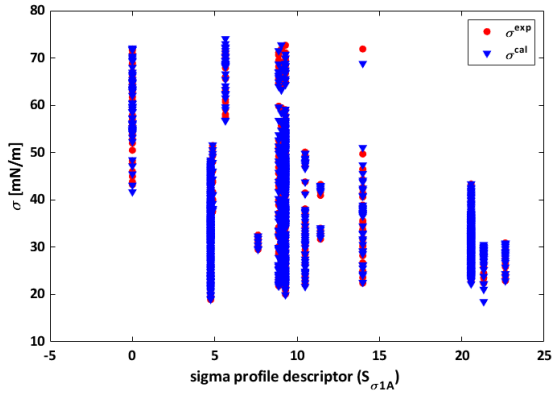
(b)



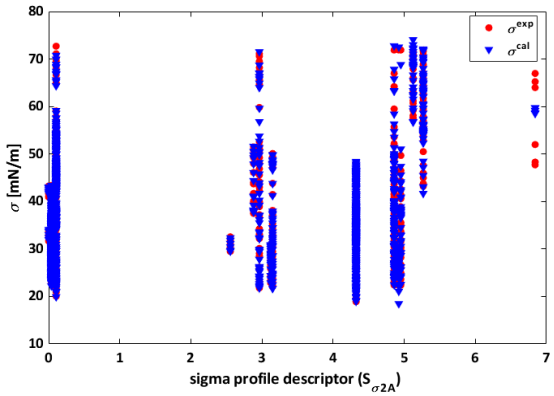
(c)



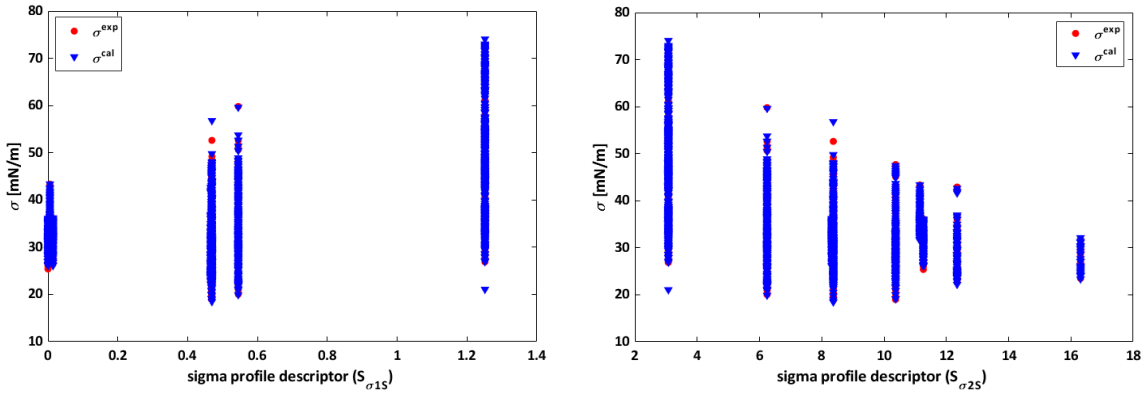
(d)



(e)



(f)



(g) (h)

Figure III.18 : Comparaison entre les données basées sur les prévisions et les données expérimentales par rapport aux données d'entrée : (a) composition de LI, (b) Température, (c) S_σ1C, (d) S_σ2C, (e) S_σ1A, (f) S_σ2A, (g) S_σ1S, (h) S_σ2S.

III.3 Contribution à la modélisation de la solubilité de CO_2 dans les liquides ioniques par les méthodes de RNA et DA-SVM.

III.3.1 Elaboration de la base de données

III.3.1.1 Données expérimentales de la solubilité

Un total de 10 116 points de données sur la solubilité du CO_2 (fraction molaire 0,0000648-0,9516) mesurés dans différents types de LIs (au total 124 LIs différents) de 243,2 K à 453,15 K et de 0,00798 bar à 499,9 bars ont été collectés. L'ensemble de données collectées d'un travail antérieur réalisé par Song et al. (2020)[164] des LIs contenant une combinaison de cations : imidazolium, pyrrolidinium, pyridinium, pipéridinium, ammonium, phosphonium et sulfonium, et d'anions : tétrafluoroborate [BF4], chlorure [Cl], dicyanamide [DCA], nitrate [NO3], hexafluorophosphate [PF6], thiocyanate [SCN], tricyanométhanide [C(CN)3], hydrogénosulfate [HSO4], bis(trifluorométhylsulfonyl)amide [Tf2N], méthylsulfate [MeSO4]. Les données expérimentales et leurs domaines de la solubilité de CO_2 dans les LIs étudiés sont indiqués dans le tableau 1 (Annexe B).

III.3.1.2 Les descripteurs QSPR

Un modèle de relation quantitative structure-propriété (QSPR) peut révéler une relation mathématique entre les attributs structurels et la propriété d'intérêt au niveau de la chimie quantique. Les grandes phases de développement d'un modèle QSPR peuvent être décrites comme suit :

- Choisir des descripteurs adaptés au problème structure propriété.
- Exploiter les valeurs des descripteurs comme variables, afin de définir une relation qui les corrèle à la propriété en question, à l'aide de machines d'apprentissage. C'est la fouille de données.
- Établir des critères de performance et de validation qui permettent de choisir les meilleurs modèles pour le problème posé et d'estimer les incertitudes de prédiction[214].

Aussi la forme générale de modèle est:

$$Propriété = f(Descripteurs) \quad \text{III.7}$$

III.3.1.2.1Obtention des descripteurs QSPR

Les descripteurs moléculaires ont été calculés à partir de la structure chimique des molécules au moyen du logiciel PaDEL pour obtenir un modèle basé sur les relations quantitatives structure propriété (QSPR). Le logiciel de dessin Chembioa a été utilisé pour déterminer les SMILES (The simplified molecular-input line-entry system) des 124 LIs et avoir une représentation numérique des caractéristiques structurelles des molécules. Leurs descripteurs moléculaires « 1975 » de 1 544 descripteurs 2D et 431 descripteurs 3D ont été calculés après optimisation de la structure des composés à l'aide de PM3.

Plusieurs méthodes ont été utilisées pour réduire les descripteurs de 1975 à 61, en commençant par supprimer les entités ayant des valeurs nulles. Ensuite, les descripteurs invariants, c'est-à-dire ceux avec des valeurs absentes ou ceux avec presque le même « descripteur avec toutes les valeurs constantes », ont été supprimés manuellement. Tous les descripteurs ayant des valeurs identiques pour 75 % des échantillons et tous les descripteurs ayant un écart type relatif < 0,05 ont également été supprimés. Enfin, la moitié des descripteurs affichant une valeur absolue du coefficient de corrélation de Pearson > 0,9 ont également été supprimés.

III.3.1.2.2 Sélection de descripteurs pertinents

La méthode de l'algorithme génétique (GA-MLR) a été appliquée pour réduire davantage les entrées à la plus petite combinaison significative pouvant prédire la propriété cible avec une erreur acceptable. L'algorithme a pris en entrée le fichier d'entraînement global (80% de la base de donnée), avec une intercorrélation de 0,9 par défaut. La sélection du nombre approprié de descripteurs a été effectuée sur un intervalle de 5 à 20 caractéristiques, formant 15 combinaisons où chaque combinaison contenait la température, la pression et 3 à

18 descripteurs. Afin de sélectionner la meilleure combinaison, les 15 modèles obtenus ont été examinés statistiquement. Les résultats ont prouvé la surperformance du modèle, qui combine 15 caractéristiques. La multicolinéarité entre les descripteurs du modèle sélectionné a été détectée en calculant leurs facteurs d'inflation de variation (VIF). Un modèle est considéré comme acceptable si les valeurs VIF sont comprises entre 1 et 5 (tableau 1). Toutes les valeurs VIF des descripteurs étaient inférieures à 4.5100, ce qui indique que le modèle résultant a une signification statistique et que les descripteurs se sont révélés raisonnablement orthogonaux. Par conséquent, les descripteurs sélectionnés étaient : E2M, MATS8S, TDB6I, TDB1S, ATSC4V, MATS8M, ATSC7V, Gats2S, Gats5S, Gats5C, ATSC6V, DE, Lobmax avec l'ajout de T (K) et P (bar). Plus de détails sur la caractérisation des descripteurs sélectionnés peuvent être trouvés dans le tableau III.6.

Tableau III.6 : Caractéristiques des descripteurs sélectionnés dans le modèle optimal et leurs valeurs VIF.

Descripteur	**Type**	**Description**	**Classe de descripteur**	**VIF**
E2m	3D	Indice WHIM directionnel d'accessibilité 2e composante / pondéré par les descripteurs WHIM massiques	Descripteurs WHIM	1.6760
MATS8s	2D	Autocorrélation Moran du décalage 8 pondérée par l'état I	Autocorrélations 2D	4.0897
TDB6i	3D	Autocorrélation topologique 3D basée sur la distance - décalage 6 / pondéré par le premier potentiel d'ionisation	Autocorrélation 3D	1.4114
TDB1s	3D	Autocorrélation topologique 3D basée sur la distance - décalage 1 / pondérée par l'état I	Autocorrélation 3D	1.3069
ATSC4v	2D	Autocorrélation de Broto-Moreau centrée du décalage 4 pondérée par le volume de van der Waals	Autocorrélations 2D	3.8448
MATS8m	2D	Autocorrélation Moran du décalage 8 pondéré en masse	Autocorrélations 2D	3.8043
ATSC7v	2D	Autocorrélation de Broto-Moreau centrée du décalage 7 pondérée par le volume de van der Waals	Autocorrélations 2D	4.5100
GATS2s	2D	Autocorrélation Geary - décalage 2 / pondéré par l'état I	Autocorrélations 2D	1.5702
GATS5s	2D	Autocorrélation Geary - décalage 5 / pondéré par l'état I	Autocorrélations2D	1.7270

GATS5c	2D	Autocorrélation Geary - décalage 5 / pondéré par les charges	Autocorrélations 2D	2.4773
ATSC6v	2D	Autocorrélation de Broto-Moreau centrée - décalage 6 / pondérée par les volumes de van der Waals	Autocorrélations 2D	1.5188
De	3D	Indice d'accessibilité totale D / pondéré par l'électronégativité de Sanderson	Descripteurs WHIM	1.4079
LOBMAX	3D	Le rapport L/B maximum	Descripteur longueur sur largeur	1.4891
T (K)	-	Température	-	1.0948
P (bar)	-	Pression	-	1.0984

III.3.2 Modélisation de la solubilité du CO_2 dans Lis par les méthodes d'intelligence computationnelle

III.3.2.1 Choix des variables d'entrée et de sortie

Pour développer les modèles d'intelligence computationnelle nous avons exprimé la solubilité de CO_2 dans LIs (en fraction molaire) comme fonction de la Température (K) et la pression (bar) du LI et les 13 descripteurs PaDEL significatifs (E2M, MATS8S, TDB6I, TDB1S, ATSC4V, MATS8M, ATSC7V, Gats2S, Gats5S, Gats5C, ATSC6V, DE et Lobmax)

Solubilité_CO_2 = f (Température, Pression, E2M, MATS8S, TDB6I, TDB1S, ATSC4V, MATS8M, ATSC7V, Gats2S, Gats5S, Gats5C, ATSC6V, DE,Lobmax).III.8

III.3.2.2 Division des données en ensemble d'apprentissage, test et validation

L'ensemble de donnée d'apprentissage se compose de 80 % de données choisies au hasard à l'aide de la fonction ***dividerand*** de MATLAB 2018. Les 20 % restants ont été utilisés comme ensemble de test pour RNA et DA-SVM.

III.3.2.3 Critères statistiques :

Pour juger la performance des modèles nous avons élaboré une fonction MATALAB pour calculer 15 valeurs statistiques et 3 types d'erreur pour les ensembles d'apprentissage de test et global. Les formules et significations de ces critères de validité du modèle ont été mentionnées dans le chapitre II.

III.3.2.4 Modélisation par RNA-QSPR

Le réseau a été formé à l'aide de la fonction de formation trainbr « rétro-propagation de régularisation bayésienne ». La fonction tansig "Tan-sigmoid" a été choisie comme fonction de transfert pour la couche cachée et purelin pour la couche de sortie. Le nombre de neurones cachés a été choisi pour être de 27 avec 15 nœuds d'entrée et un neurone de sortie pour représenter la fraction de solubilité expérimentale (yexp). Afin d'arriver au modèle donnant la meilleure erreur moyenne absolue (MAE), plusieurs itérations de 10 à 30 ont été réalisées. Par conséquent, le RNA a l'architecture suivante {15-27-1}.

L'évaluation du modèle a été réalisée à l'aide d'outils statistiques et graphiques. Les tableaux III.7 et III.8 résument respectivement le calcul des différents critères statistiques et les différentes erreurs du modèle RNA. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité.

Tableau III.7: Critères statistiques du modèle RNA

Critères Statistiques	**Train**	**Test**	**Global**
Q2F1	0.9807	0.9828	0.9812
Q2F2	0.9807	0.9828	0.9812
Q2F3	0.9807	0.9824	0.9811
Q2CCC	0.9903	0.9913	0.9905
R^2	0.9807	0.9828	0.9812
R	0.9903	0.9914	0.9906
K	1.0001	1.0000	0.9996
k'	0.9935	0.9943	0.9942
R_0^2	1.0000	1.0000	1.0000
$R_0'^2$	0.9999	0.9999	0.9999
M	-0.0197	-0.0175	-0.0192
N	-0.0195	-0.0174	-0.0191
R_m	0.8445	0.8539	0.8467
A_f	1.1553	1.1530	1.1692
AIC	-69308.30		

Tableau III.8: Différentes erreurs par le modèles RNA.

Erreurs	**Train**	**Test**	**Global**
EQRM	0.0327	0.0313	0.0324
EQM	0.0011	9.7866e-04	0.0010
EAM	0.0199	0.0195	0.0199

Les résultats obtenus révèlent une précision globale du modèle développé dans la corrélation des données totales. Où le modèle RNA non linéaire a donné de bons résultats avec les coefficients de corrélation globaux R (0,9906), ainsi que les coefficients de détermination R2 0,9812. De plus, les paramètres d'erreur ont montré des résultats très acceptables, tels que le EQRM global (0,0324) et le EAM (0,0199). Ceux-ci indiquent la robustesse du modèle établi qui est en bon accord avec les données expérimentales sur la solubilité du CO_2.

L'évaluation graphique est réalisée à l'aide de diagramme de dispersion (pour les ensembles d'apprentissage, de test et global), le diagramme d'erreur relative (ER %) et l'histogramme de distribution d'erreur relative.La figureIII.19 montre le diagramme de dispersion de la solubilité du CO_2 prédite par rapport à la solubilité expérimentale du CO_2 (pour les ensembles d'apprentissage, de test et global). Ce diagramme montre clairement l'accord entre les valeurs prédites par le modèle RNA et les valeurs expérimentales par une dispersion serrée des points sur la première bissectrice, à l'exception de certains points aberrants, considérés comme acceptés hors gamme.

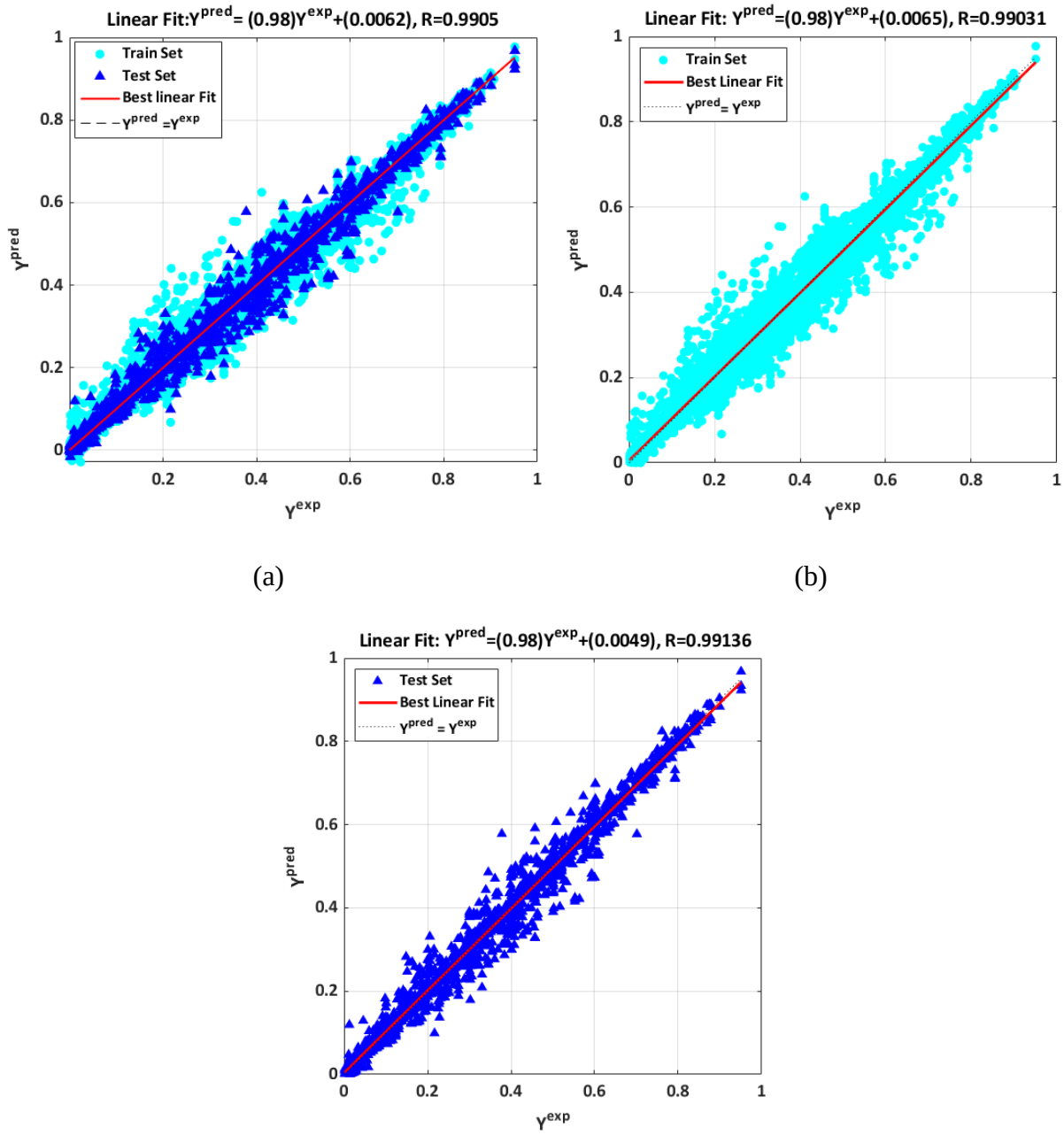


Figure III.19 : Diagramme de dispersion de la solubilité prédite du CO_2 en fonction de la solubilité expérimentale du CO_2 : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test du modèle RNA.

Pour visualiser quantitativement la distribution des erreurs relatives, deux types de diagrammes ont été adoptés. Le premier diagramme montre l'erreur relative en fonction de la solubilité du CO_2 représentée sur la figure III.20 et le deuxième montrel'histogramme de la distribution de cette erreur représentée sur la figure III.21. Chaque sous-ensemble de données, d'apprentissage et de test, est illustré avec une couleur différente pour donner une meilleure

idée de la distribution des erreurs à chaque étape du développement du modèle de processus. Comme on peut le constater, la majorité des points de données se situent autour de la ligne zéro et une très petite proportion présente une erreur absolue supérieure à 0,05. Quantitativement, l'erreur relative pour les quelques valeurs aberrantes n'a pas dépassé 0,2.

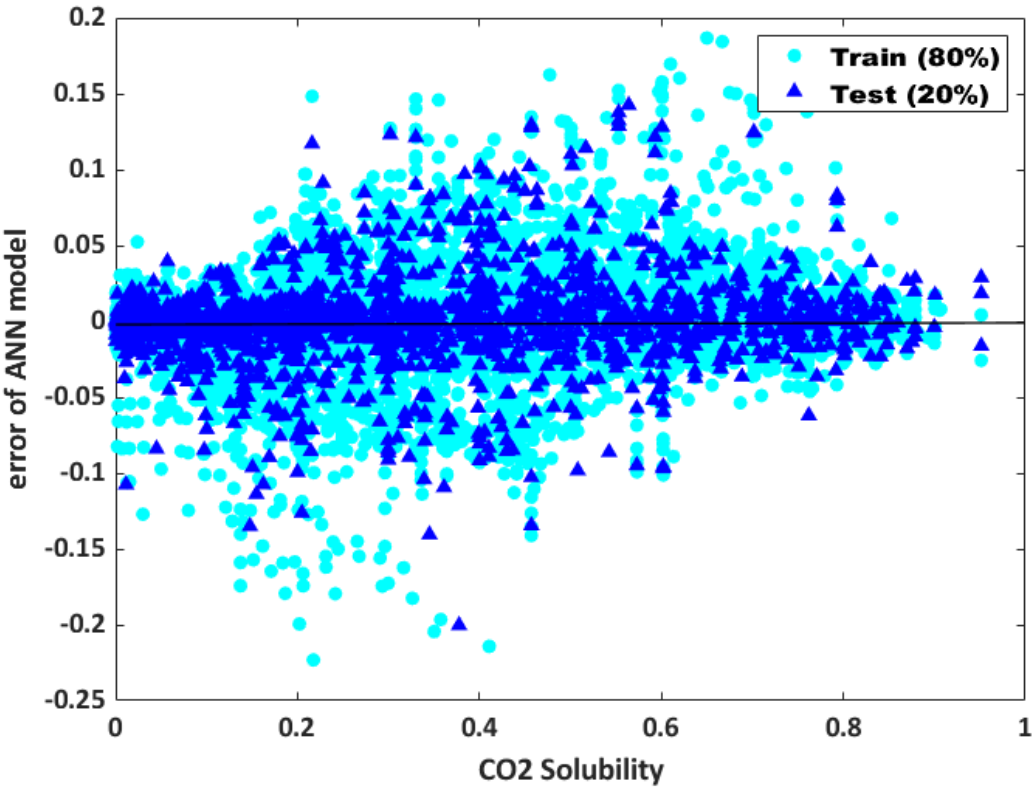


Figure III.20 : L'erreur relative de la solubilité du CO_2 prédite par le modèle de RNA.

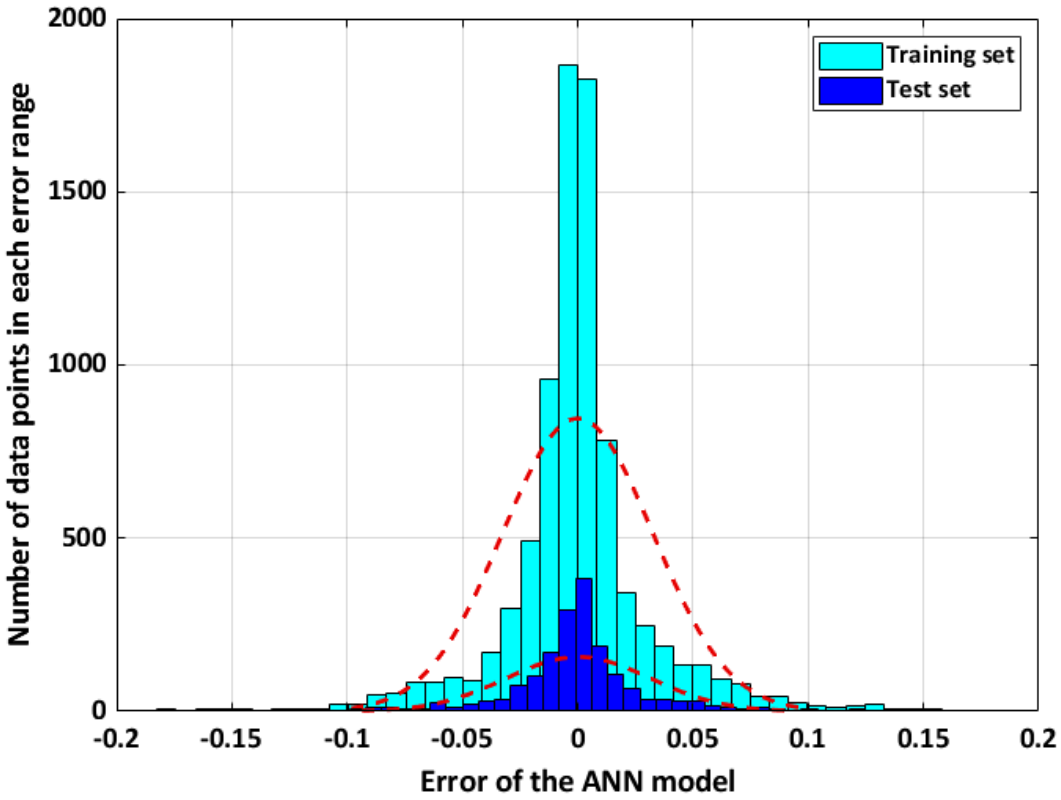


Figure III.21 : Histogramme de la distribution de l'erreur de prédiction du modèle RNA.

III.3.2.5 Modélisation par DA-SVM

Le développement du modèle DA-SVM a été réalisé par la fonction fitrsvm de l'environnement MATLAB. Il dispose de plusieurs options de calcul relatives au choix de la méthode de validation croisée, de la fonction noyau (Gaussien, rbf et polynôme), des valeurs optimales de la constante C (case Constraint), de l'epsilon et du paramètre de la fonction noyau (Kernel Scale) du modèle ont été optimisés avec l'algorithme DA (100 itérations). Les étapes conduisant au développement de l'algorithme hybride DA-SVM optimal sont illustrées dans l'organigramme présenté sur la figure III.22.

Dans cette étude, la méthode de validation croisée « holdout » fixée à 0,2 a été choisie. Cela signifie que 80 % des données ont été utilisées pour la formation et 20 % pour les tests. Les valeurs optimales sont 53,270, 1,372 et 0,008 pour c, k et ε respectivement avec la fonction Gaussienne comme fonction du noyau.

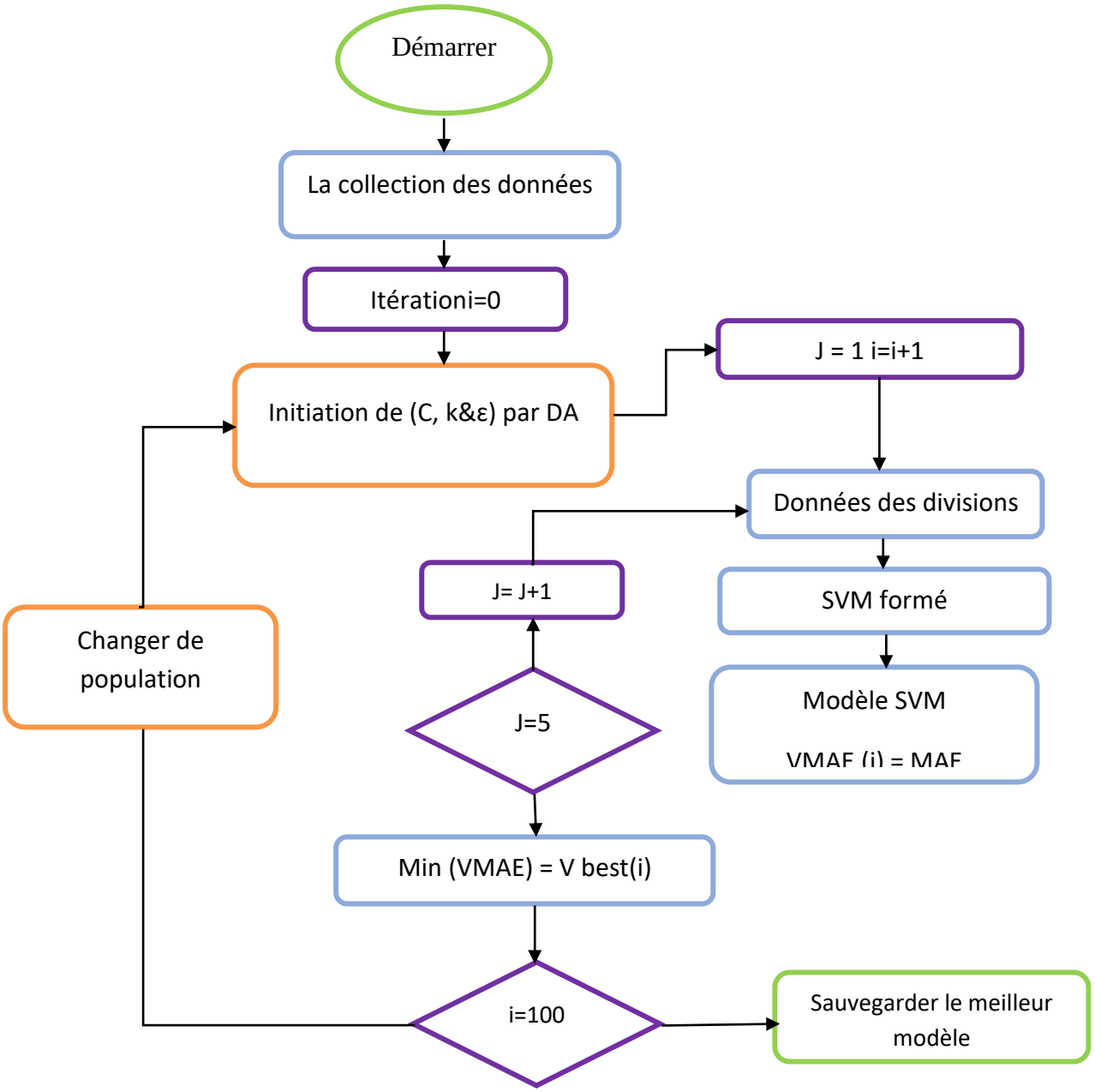


Figure III.22 : Organigramme du processus d'optimisation DA-SVM.

Les tableaux III.9 et III.10 résument respectivement le calcul des différents critères statistiques et les différentes erreurs du modèle DA-SVM. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité.

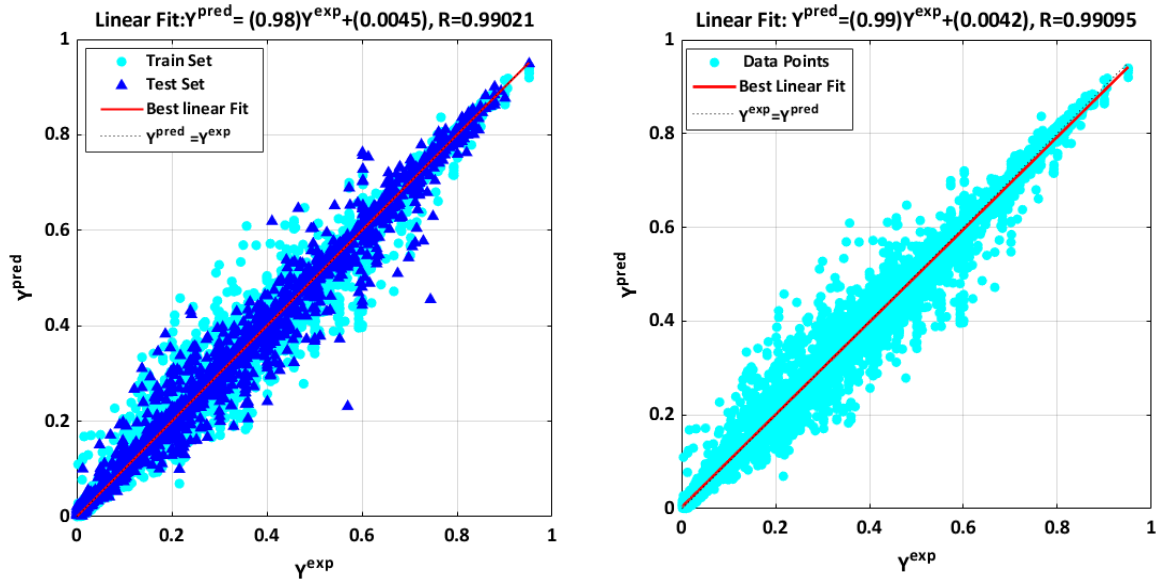
Tableau III.9: Critères statistiques du modèle DA-SVM.

Critères statistiques	**Train**	**Test**	**Global**
Q2F1	0.9820	0.9745	0.9805
Q2F2	0.9820	0.9745	0.9805
Q2F3	0.9820	0.9748	0.9805
Q2CCC	0.9909	0.9871	0.9902
R^2	0.9820	0.9745	0.9805
R	0.9910	0.9872	0.9902
K	0.9999	1.0004	1.0000
k’	0.9940	0.9913	0.9935
R_0^2	1.0000	1.0000	1.0000
$R_0{}'^2$	0.9999	0.9998	0.9999
M	-0.0184	-0.0262	-0.0199
N	-0.0183	-0.0260	-0.0198
R_m	0.8501	0.8188	0.8435
A_f	1.1429	1.1416	1.1426
AIC	-68992.86		

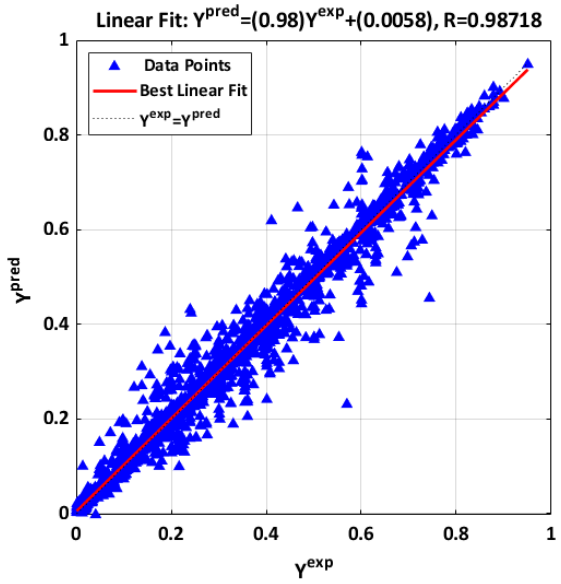
Tableau III.10: Différentes erreurs par le modèle DA-SVM.

Erreurs	Train	Test	Global
EQRM	0.0318	0.0375	0.0330
EQM	0.0010	0.0014	0.0011
EAM	0.0179	0.0219	0.0187

La figure III.23 montre le diagramme de dispersion de la solubilité du CO_2 prédite en fonction de la solubilité expérimentale du CO_2 (pour les ensembles d’apprentissage, de test et global) pour le modèle DA-SVM. Les résultats trouvés montrent que, tous les points sont accumulés autour de la bissectrice du premier quadrant pour tous les ensembles de données. Cependant, l’existence d’un certain nombre de valeurs aberrantes peut être constatée.



(a) (b)



(c)

Figure III.23 : Diagramme de dispersion de la solubilité du CO_2 prédite en fonction de la solubilité du CO_2 expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de test du modèle DA-SVM.

De même, les deux types de diagrammes ont été retenus afin de fournir une connaissance plus complète sur les erreurs de prédiction du modèle. Les figures III.24 et III.25 montrent l'erreur relative en fonction de la solubilité du CO_2 et de sa distribution. Comme indiqué, la plupart des écarts sont proches de zéro et se situent dans la plage de [-

0,03, 0,03]. Une très faible proportion présente une erreur absolue supérieure à 0,1. Avec le plus grand écart absolu d'environ 0,35 a été observé.

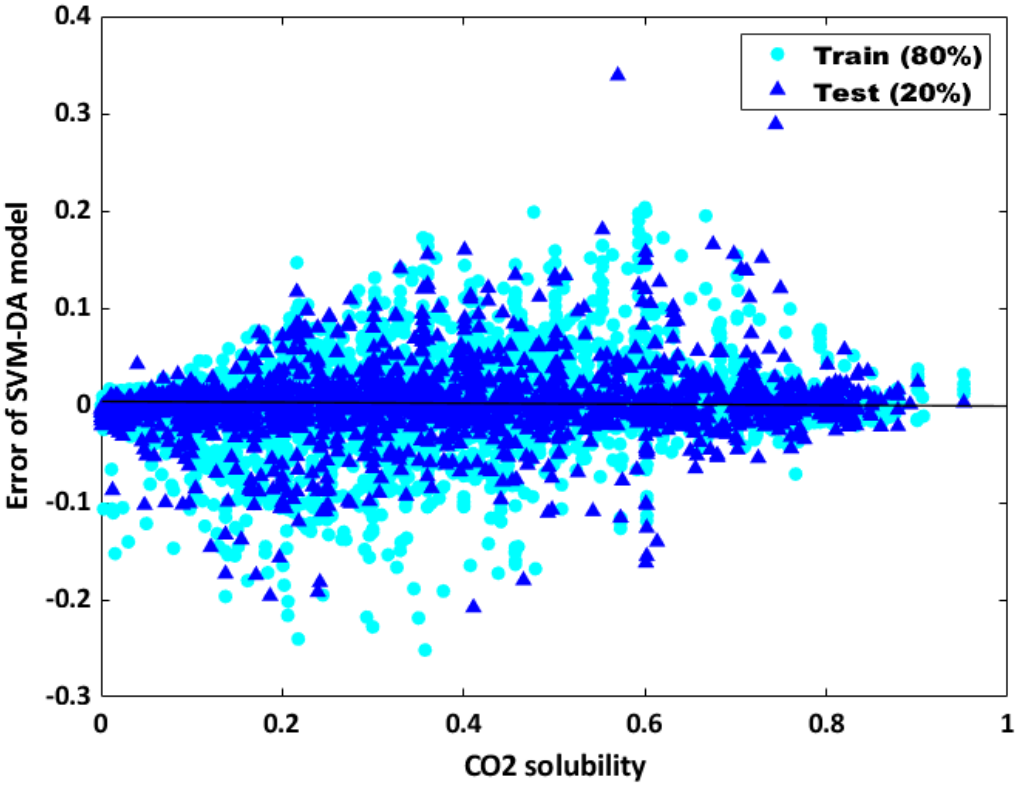


FIGURE III.24 : L'erreur relative de la solubilité du CO_2 prédite par le modèle de DA-SVM.

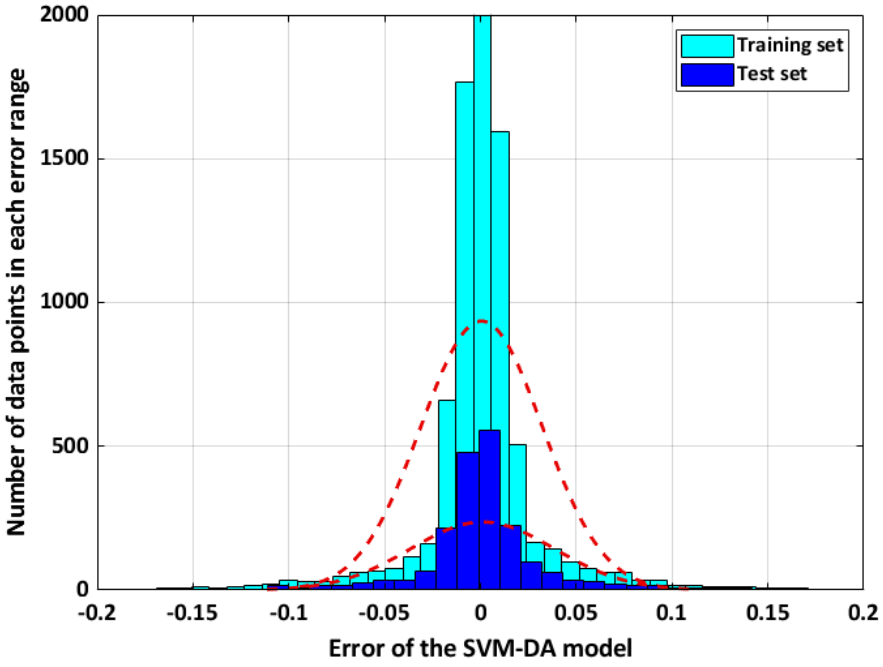


Figure III.25 : L'histogramme de la distribution de l'erreur de prédiction du modèle DA-SVM.

A la lumière de ces résultats, l'ensemble des critères statistiques calculées vérifient les conditions d'acceptabilité. Cependant, les résultats du modèle RNA optimisé à l'aide de l'algorithme d'entraînement trainbr et de la fonction d'activation tansig ont montré des résultats très significatifs avec un coefficient de corrélation global R de 0,9906, et un coefficient de détermination R2 de 0,9812. De plus, des paramètres d'erreur tels que : EQRM (0,0324) et EAM (0,0199). Ceux-ci permettent de prouver l'efficacité et la robustesse du modèle développé, qui démontre des performances de corrélation précises et de meilleures capacités de généralisation et d'interpolation. Selon les valeurs AIC obtenues, la plus petite valeur représente le modèle le mieux adapté aux deux modèles étudiés.

III.3.3 Analyse de la sensibilité des paramètres d'entrée

Comme illustré sur la figure III.26, pour le modèle RNA établi, toutes les variables d'entrée ont presque la même importance selon la méthode de Garson, à l'exception de la pression qui représente la plus importante variable avec Ir ('importance relative) = 14,1081%. Sur la figure III.27, la solubilité duCO_2 a montré une dépendance directe sur les descripteurs nommés : T, P, E2m, MATS8s, GATS2s, GATS5c, ATSC6v et De et une dépendance opposée au reste des descripteurs. Les descripteurs les plus pertinents sont la pression et le TDB1s avec un facteur de pertinence de +0,6650 et -0,1788 respectivement. Contrariety, MATS8s et ATSC4v étaient les moins pertinents, avec un facteur de pertinence de +0,0297 et -0,0046, respectivement.

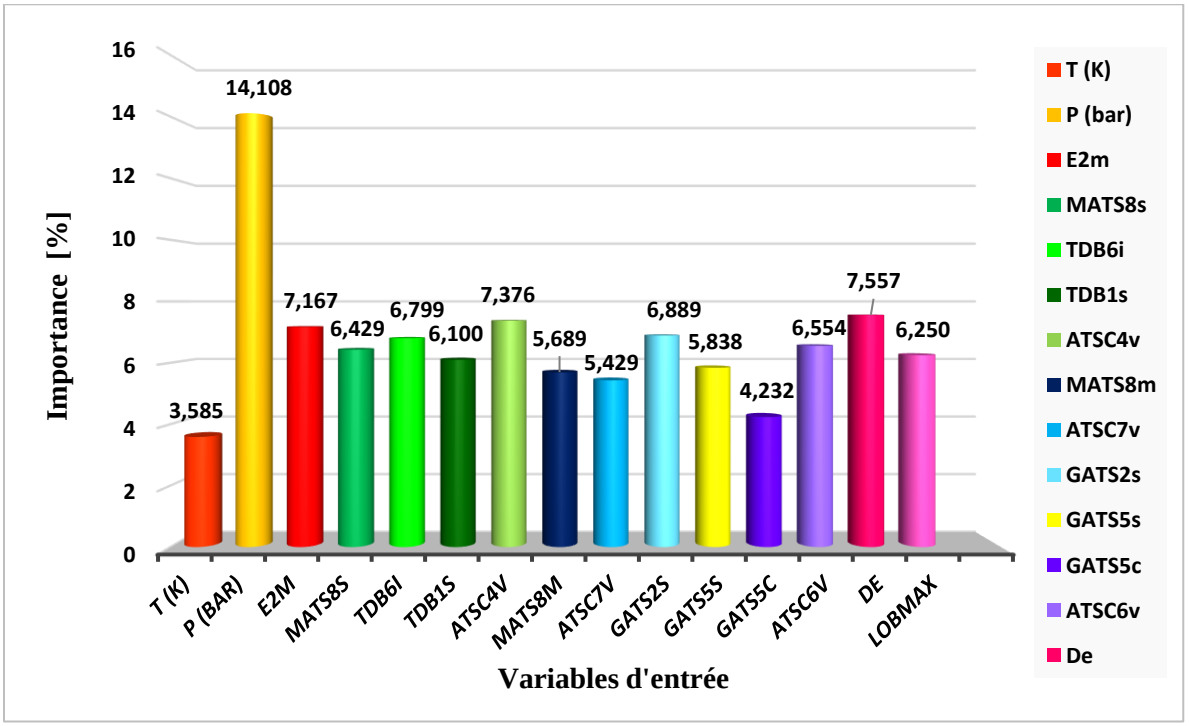


Figure III.26 : Importance relative des variables d'entrée par la méthode Garson du modèle RNA.

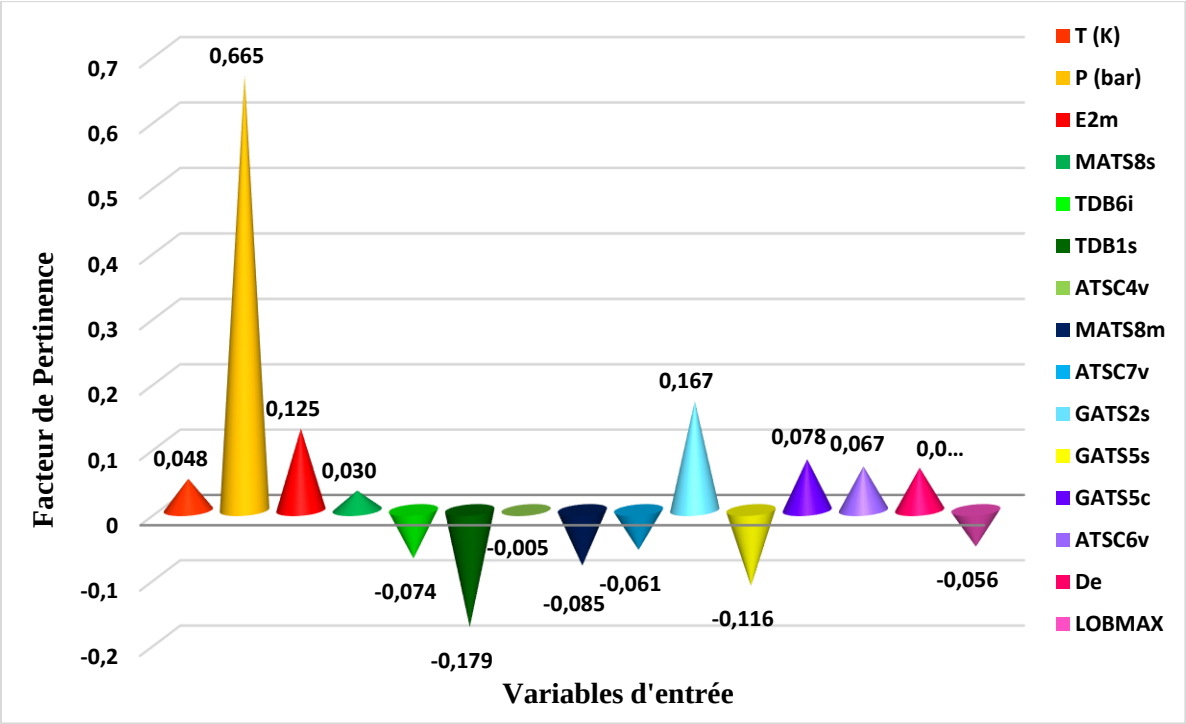


Figure III.27 : Les Résultats du facteur de pertinence effectué sur le modèle RNA.

III.3.4 Domaine d'applicabilité

La figure III.28 représente les zones/données valides et suspectes obtenues par la méthode de levier pour la base de données expérimentale collectée. La zone valide se situe entre -3<résidus normalisés<3 et 0<indice de Hat<H*=0,0059 a été constatée. Sur la base de la figure III.28, seules les données représentées par des cercles verts sont considérées valides et pour lesquelles la précision du modèle a été confirmée. Cet ensemble contient 9589 points de donnée sur les 10116 données globales initialement utilisées pour développer le modèle RNA et représente un ratio de 94,79% des données expérimentales globales. Les données douteuses sont signalées par des triangles rouges et contiennent 527 valeurs aberrantes correspondant à 5,2 % de l'ensemble des données expérimentales. Cela indique que l'incertitude de prévision pour chacun d'eux peut être considérée comme acceptable. Ainsi, le modèle développé peut être considéré comme statistiquement correct et valide pour la prédiction de la solubilité du CO_2.

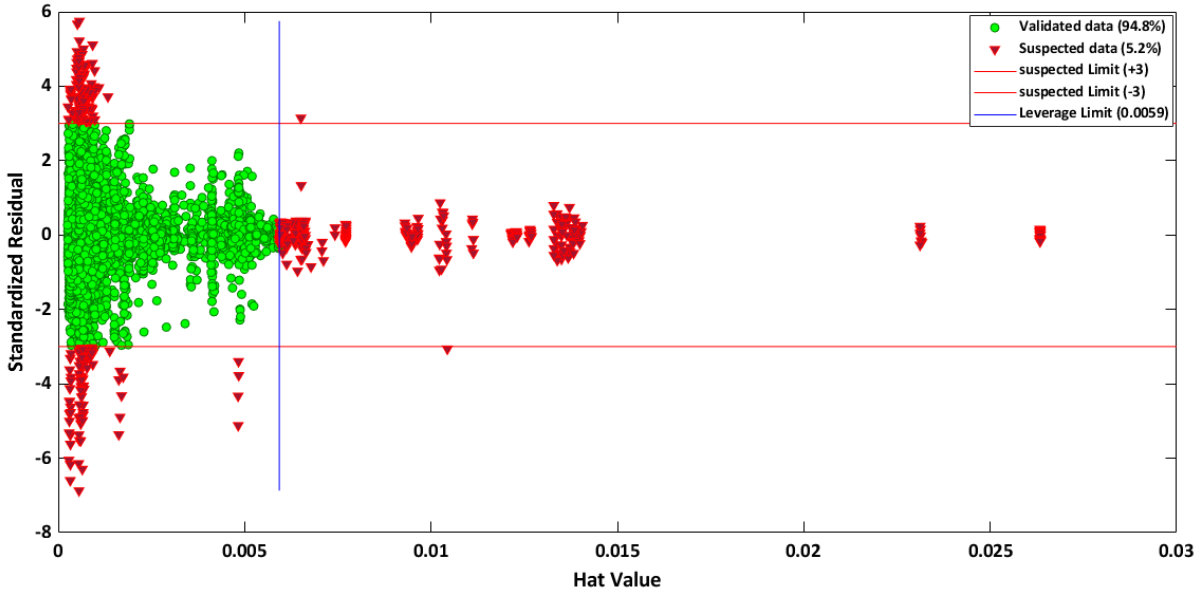


Figure III.28 : Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.

La méthode du graphique Insubria a été utilisée qui représente un graphique des valeurs hi diagonales par rapport aux valeurs prédites. Elle permet de fournir une visualisation des prédictions interpolées et extrapolées. Les LIs avec h>h* sont considérées en dehors du domaine structurel de l'ensemble d'apprentissage, par conséquent, leurs données prédites sont extrapolées par le modèle et pourraient donc être moins fiables. Sur le graphe Insubria, une zone de plus grande fiabilité pour les prédictions de structures (<h*) et de Y (entre Ymax et Ymin, la valeur maximale et la valeur minimale de la solubilité de CO_2 dans les LIs prédites de l'ensemble d'apprentissage) est toujours fournie.[215,216][215]

La figure III.29 représente le graphe Insubria utilisé pour illustrer la distribution des LIs dans l'espace des descripteurs employés et fournir les relations entre les valeurs de la solubilité de CO_2 prédites pour l'ensemble d'apprentissage et les actualités dans l'ensemble de prédiction. Tout comme le graphique Williams obtenu par la méthode de levier, le graphique Insubria est également un nuage de points en deux dimensions. L'axe des x est la valeur de levier (Hat) de l'ensemble de prédiction, qui est défini de la même manière que celui du graphique de Williams, et l'axe des y est la valeur la solubilité de CO_2prédite du modèle. Dans cette figure, Ymax et Ymin font référence aux valeurs maximale et minimale de la solubilité de CO_2 expérimentales de l'ensemble d'apprentissage, respectivement.

Comme on peut voir que la plupart des points de données de l'ensemble de prédiction étaient presque similaires à l'ensemble d'apprentissage (effets de levier inférieurs à h*) et que

les valeurs de solubilité du CO_2 prédites étaient comprises entre Ymax et Ymin (0,9516 et 0,000065) de l'ensemble d'apprentissage, ce qui indique que les valeurs de solubilité du CO_2 prédites sont le résultat d'une interpolation et ont été bien prédites, à l'exception de l'existence de certaines données suspectées étaient hors du domaine d'applicabilité (effet de levier supérieur à h*). Ainsi, les valeurs de solubilité du CO_2 prédites sont le résultat d'une extrapolation et ont été considérées moins fiables.

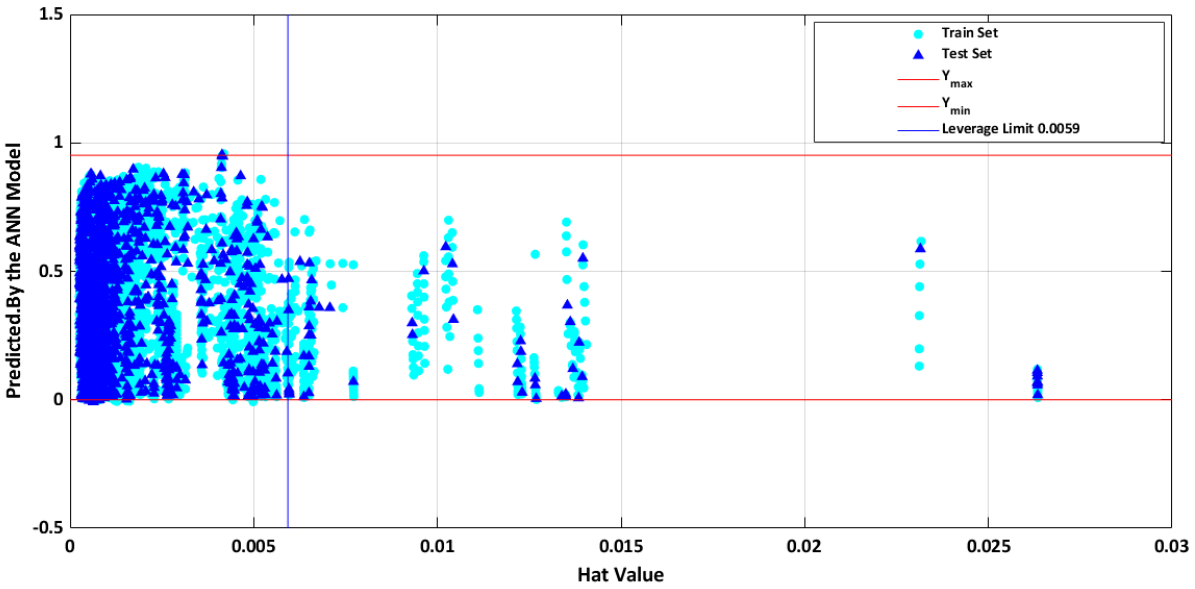


Figure III.29: Graphique Insubria de la solubilité du CO_2 par le modèle RNA.

III.3.5 Validation externe

Pour effectuer une validation externe, une autre base de donnée comprenant 77 points de donnée expérimentales a été collectée à partir des travaux de Han et al.[217]et de Qazi et al.[218].Le même nombre sélectionné de fonctionnalités a été utilisé pour le développement de modèles, c'est-à-dire pour la prédiction de la solubilité du CO_2 dans les LIs à partir d'un ensemble de données inutilisées, appelé ensemble de validation ou test externe.

La figure III.30 montre la solubilité expérimentale et prédites du CO_2 de l'ensemble de validation, prouvant que le modèle établi peut représenter la solubilité expérimentale et prédites du CO_2 de l'ensemble de validation avec une corrélation positive où R = 0,96568 ; tandis que le tableau III.11 montre une erreur acceptable. On peut affirmer que le modèle est valide pour prédire les données futures.

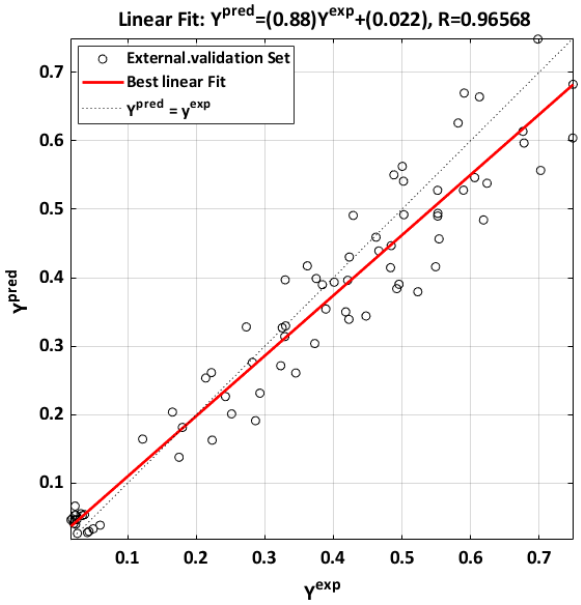


Figure III.30: Diagramme de dispersion des Valeurs prédites et expérimentales de la solubilité du CO_2 de validation externe du modèle RNA.

Tableau III.11 : Différentes erreurs pour l'ensemble de validation externe

Erreurs	**EQRM**	**EQM**	**EAM**
Ensemble de validation externe	0.0613	0.0038	0.0492

III.3.6 Comparaison avec d'autres modèles de la littérature

Une étude comparative a été réalisée entre le modèle proposé avec celui de Song et al. (2020) [164]. Les auteurs ont utilisé une méthode de contribution de groupe (GC) pour prédire la solubilité du CO_2 dans les Lis. La même base de données a été utilisée dans cette étude pour évaluer les performances du modèle obtenu.

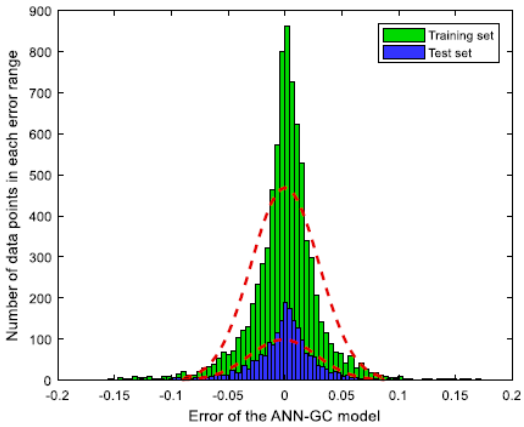
Une comparaison, en termes de paramètres statistiques, a été donnée dans le tableau III.12. La figure III.31 montre une autre comparaison entre la présente étude et celle mentionnée précédemment en termes de distribution d'erreurs.

L'analyse des critères statistiques des modèles développés et de la distribution a permis de conclure que le modèle établi est comparable au modèle précédent. En particulier, compte tenu des comparaisons des critères de prédictivité entre elles, et la dispersion de l'erreur centrée autour de zéro, qui pourraient montrer les performances de prédiction des modèles développés, il apparaît clairement que les performances du modèle présenté sont meilleures que celles du modèle précédent. Cependant, il convient de noter que le modèle de Song et al a été développé à partir de 53 variables d'entrées, alors que seulement 15 variables d'entrées

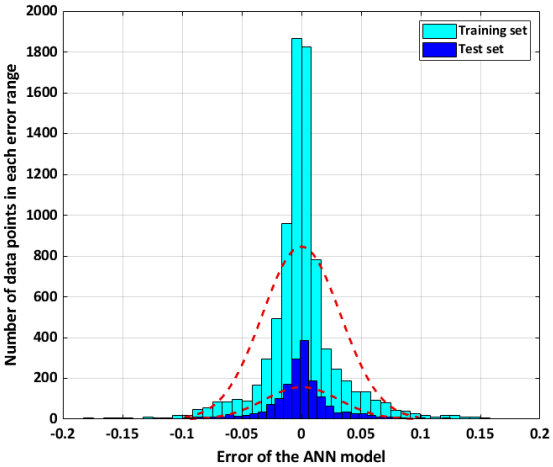
ont été utilisés dans cette étude. Ce qui présente une supériorité du modèle proposé en robustesse.

Tableau III.12: Comparaison des critères statistiques entre la présente étude et la précédente.

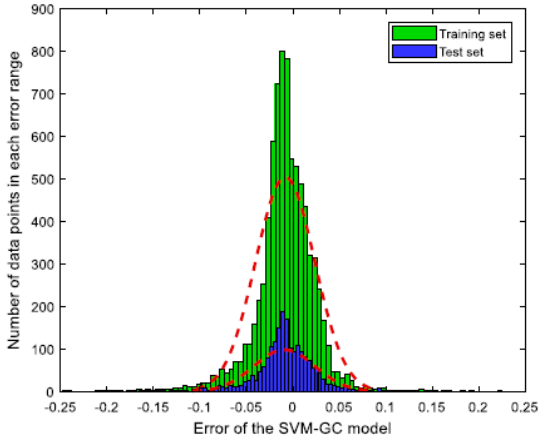
		Modèle RNA				Modèle SVM			
Etude	Nombre de variables d'entrées	Train		Test		Train		Test	
		EAM	R^2	EAM	R^2	EAM	R^2	EAM	R^2
Song et al.[164]	53	0.0200	0.9842	0.0202	0.9836	0.0231	0.9807	0.0240	0.9783
Cette étude	15	0.0199	0.9807	0.0195	0.9828	0.0179	0.9820	0.0219	0.9745



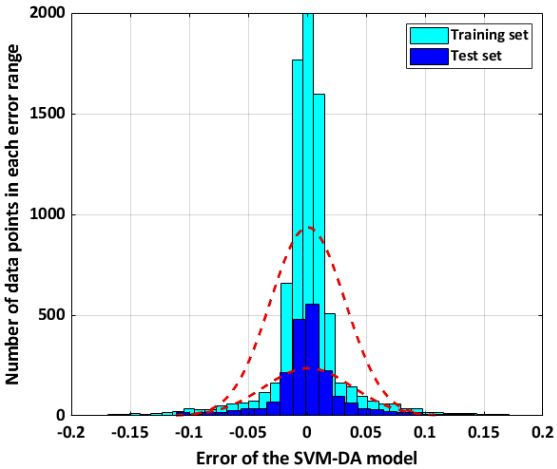
Modèle RNA-GC de Song et al.



Modèle RNA de cette étude



Modèle SVM-GC de Song et al.

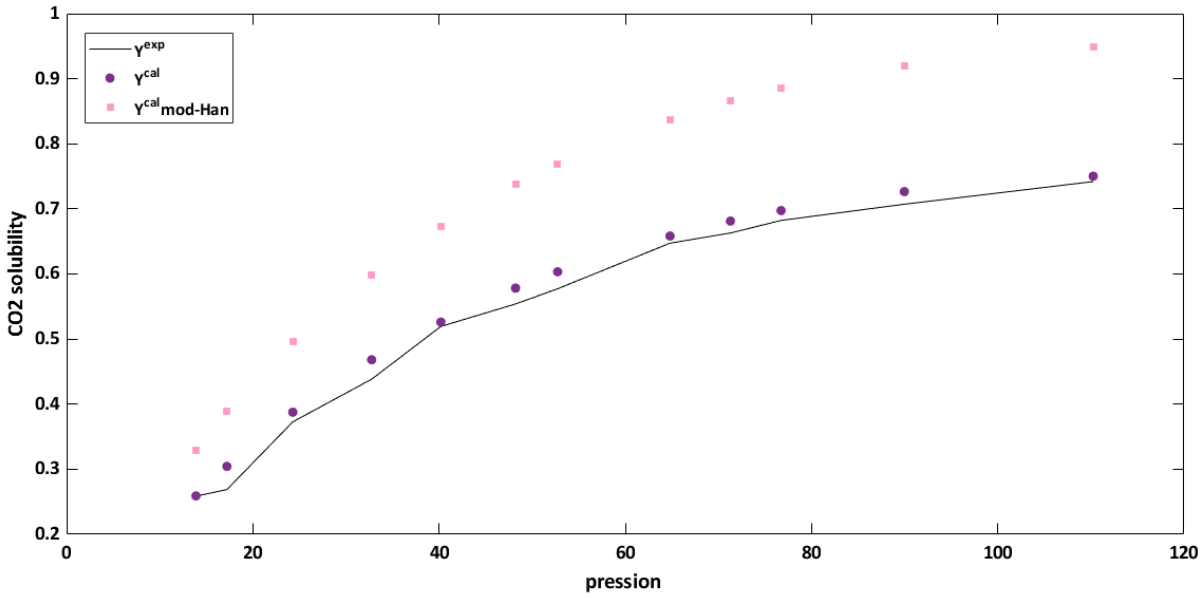


Modèle SVM-DA de cette étude

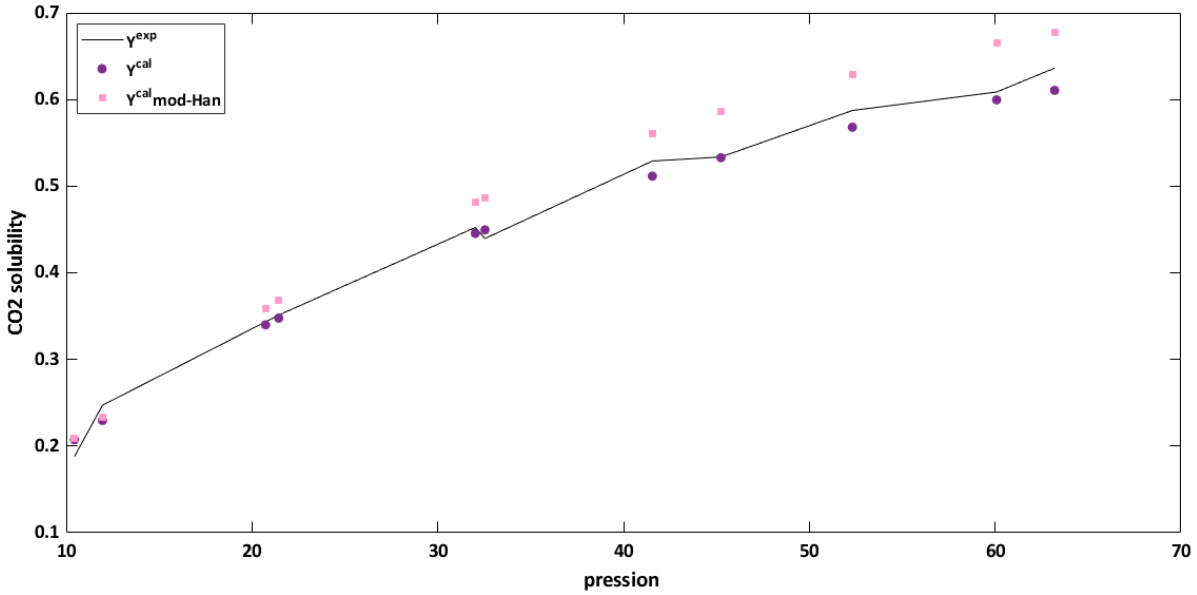
Figure III.31: Comparaison en termes de distribution d'erreurs.

De plus, la figure III.32 représente une autre comparaison entre la solubilité du CO_2 dans les LIs (en fraction molaire) expérimentale et celle prédite par le meilleure modèle RNA de cette étude et la solubilité du CO_2 prédite par l'étude de Han et al.(2018)[217] pour certains systèmes communs à des températures spécifiques. Les solubilités prédictives des modèles

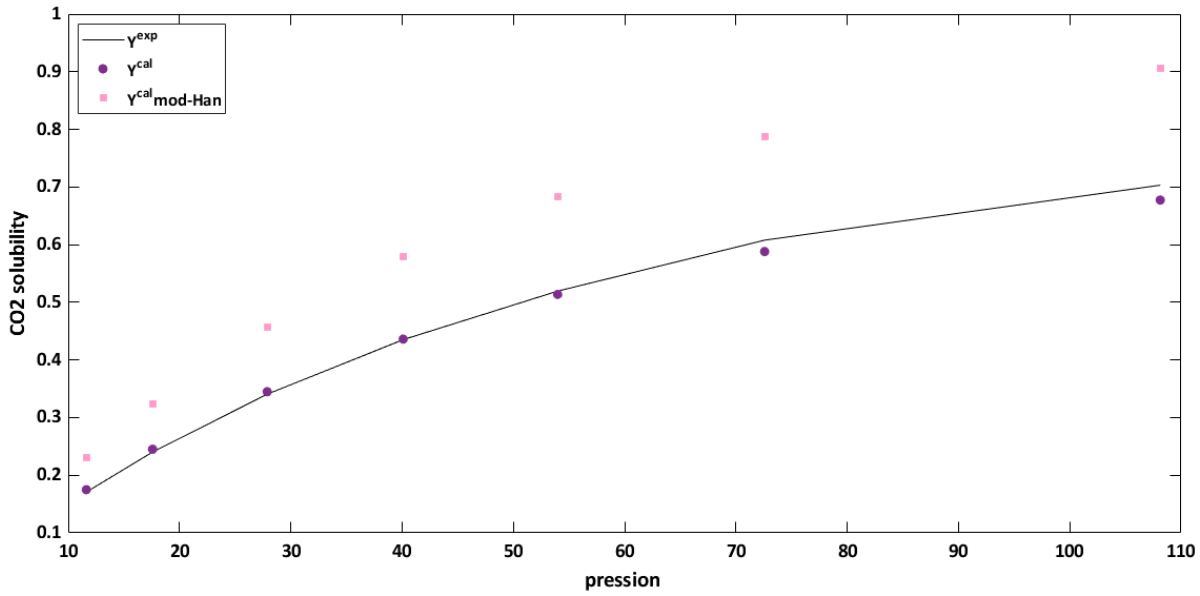
proposés étaient plus satisfaisantes. La figure III.32 confirme donc la grande capacité prédictive des modèles proposés.



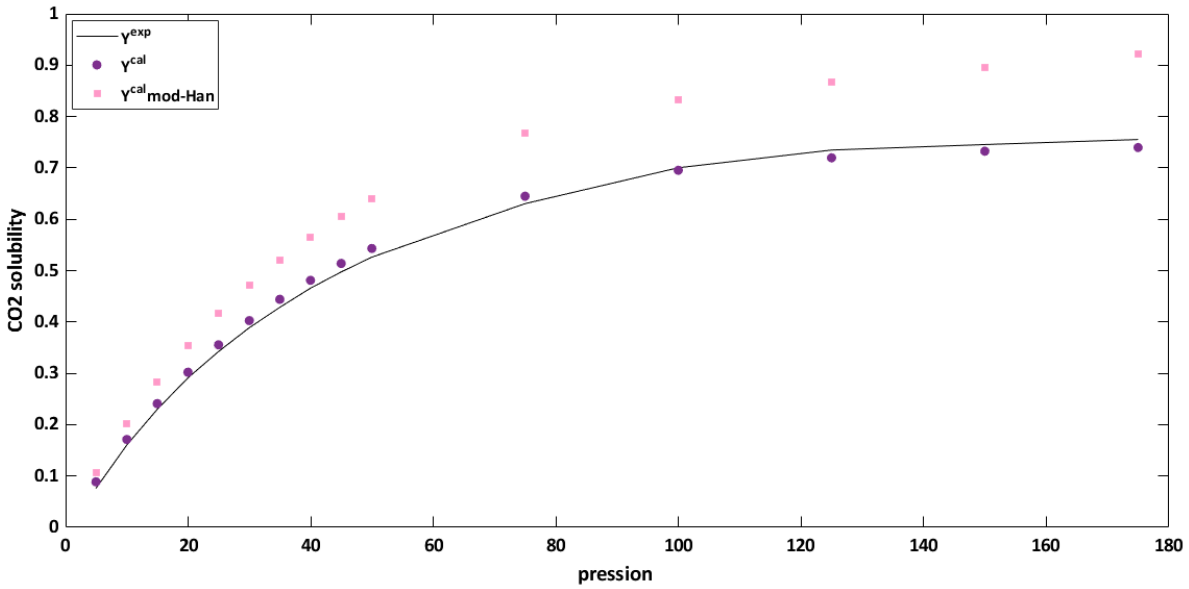
(a) : [BMIM][Tf2N] à 313,3



(b): [BMIM][TfO] à 298.2



(c): [BMPYR][Tf2N] à 333,1



(d): [EMIM][Tf2N] à 318

Figure III.32: Comparaison entre les valeurs expérimentales (lignes) et calculées de la solubilité de CO_2 dans LI par notre modèle (points mauve) et le modèle Han (carrés rose) pour certains systèmes.

III.4 Conclusion

Cette étude a permis de démontrer :

- Que l'utilisation des descripteurs basés sur les profils sigma du modèle de chimie quantique COSMO-RS pour la contribution à la modélisation de la tension superficielle de mélanges binaires LI+Eau et LI+différents solvants organiques en utilisant un ensemble de 1623 points à partir de 62 mélanges binaires d'IL par les méthodes d'intelligence computationnelle donne des résultats très satisfaisants. Les modèles développés par les méthodes de RNA, LSSVM et SVM-PSO sont tous statistiquement valides avec une ERAM globale comprise entre 0.8466% et 1.2201%. Cependant, les résultats du modèle optimisé en utilisant l'algorithme d'entraînement trainbr et la fonction d'activation logsig a montré les meilleurs résultats significatifs.
- Que l'utilisation des descripteurs moléculaires ont été calculés par le logiciel PaDEL à partir de la structure chimique des molécules dans le but de développer deux modèles QSPR en utilisant RNA et DA-SVM comme méthodes d'intelligence computationnelle pour la contribution à la modélisation la solubilité du CO_2 dans différents types de LIs en utilisant un total de 10 116 points de données sur la solubilité du CO_2 (en fraction molaire) donne des bons résultats. Cependant, les résultats du modèle RNA optimisé à l'aide de l'algorithme d'entraînement trainbr et de la fonction d'activation tansig ont montré des résultats très significatifs avec un coefficient de corrélation global R de 0,9906, et un erreur absolue moyenne EAM de 0,0199. Ceux-ci permettent de prouver l'efficacité et la robustesse du modèle développé, qui démontre des performances de corrélation précises et de meilleures capacités de généralisation et d'interpolation.

Conclusion Générale

Conclusion Générale et Perspectives

La présente étude consiste en la détermination de l'importance accordée actuellement aux : (i) liquides ioniques comme de nouveaux classes de matériaux dont les applications potentielles ne cessent de croitre; (ii) modèles COSMO-RS de chimie quantique comme un outil largement utilisé et efficace de prédiction et de filtrage des propriétés physico-chimiques des liquides ioniques; (iii) modèles QSPR comme un outil qui révéle une relation mathématique entre la structure moléculaire d'un ensemble de composés et les propriétés physico-chimiques au niveau de la chimie quantique. ;(vi) méthodes d'intelligence computationnelle comme outils de modélisation ayant la capacité d'apprendre une tâche spécifique à partir des données ou d'observations expérimentales. L'intérêt croissant accordé aux liquides ioniques peut être attribué aux caractéristiques exceptionnelles les plus polyvalentes, telles que des pressions de vapeur négligeables, une excellente capacité de solvatation des composés organiques et inorganiques, une stabilité thermique et chimique élevée, une capacité thermique élevée, un manque d'inflammabilité, une résistance thermique élevée.

Ce projet a été entrepris pour mettre en évidence l'étude de deux de propriétés de liquide ionique, la première est la tension superficielle qui est une propriété fondamentale importante qui révèle des informations sur les interactions superficielles, intermoléculaires et les forces dans le liquide et exerce une influence considérable sur le transfert de masse et d'énergie à travers l'interface. Ainsi, la tension superficielle a un effet déterminant dans divers procédés de génie chimique, tels que la séparation, l'adsorption, la distillation et les performances des membranes biologiques, …etc. La seconde est la solubilité du dioxyde de carbone qui est une caractéristique importante des LIs dans le contexte de la séparation et du captage de CO_2. Pour surmonter le problème des émissions du CO_2 et de réduire leurs taux dans l'atmosphère, qui affecte le changement climatique et le réchauffement climatique. Par conséquent le développement de modèles pour prédire ou corréler ces propriétés est fondamental pour la conception et l'exploitation de ces procédés.

Dans ce travail la capacité des modèles mathématiques QSPR liant les propriétés physico-chimiques à la structure moléculaire en utilisant les méthodes d'intelligence computationnelle. Premièrement pour corréler la tension superficielle de systèmes binaires contenant divers LI avec l'eau et LI avec divers solvants organiques ont été évalué à l'aide des données expérimentales tirées de la littérature dans un ensemble de données contenants 1623

points de données expérimentales. 70% des données, choisies aléatoirement, de l'ensemble de données ont été utilisé comme un ensemble d'apprentissage et les données restantes ont été utilisé comme un ensemble de test où la température (T), la composition du LI deux descripteurs basés sur les profils sigma ; l'un relatif au caractère donneur de liaison H, l'autre relatif au caractère accepteur de liaison H de l'anion, du cation et du solvant ont été utilisé comme variables d'entrées du modèle. En second lieu pour corréler la solubilité du CO_2 dans différents types de LIs en utilisant un total de 10 116 points de données divisé en deux ensembles aléatoires comme suit : un ensemble d'apprentissage (8 093 points, 80 % de données) pour construire le modèle et un ensemble de test (2 023 points, 20 % de données restantes) pour évaluer le modèle en considérant les 13 descripteurs PaDEL significatifs (E2M, MATS8S, TDB6I, TDB1S, ATSC4V, MATS8M, ATSC7V, Gats2S, Gats5S, Gats5C, ATSC6V, DE et Lobmax), la température et la pression comme données d'entrée du modèle. Plusieurs critères et métriques statistiques ont été calculés afin de justifier la validité et le domaine d'applicabilité des modèles optimisés. Les conclusions suivantes peuvent être tirées des résultats de ce travail :

1. Cette étude a démontré que l'utilisation des descripteurs basés sur les profils sigma du modèle de chimie quantique COSMO-RS pour corréler la tension superficielle de 62 mélanges binaires LI+Eau et LI+différents solvants organiques par les méthodes d'intelligence computationnelle donne des résultats très satisfaisants. Les modèles développés par les méthodes de RNA, LSSVM et SVM-PSO sont tous statistiquement valides avec une ERAM globale comprise entre 0.8466% et 1.2201%. Cependant, les résultats du modèle optimisé en utilisant l'algorithme d'entraînement trainbr et la fonction d'activation logsig ont montré des résultats très significatifs. Compte tenu de leur précision et de leur simplicité, ces modèles peuvent être utilisés efficacement dans la pratique.
2. Cette étude a démontré aussi que l'utilisation des descripteurs moléculaires calculés par le logiciel PaDEL à partir de la structure chimique des molécules dans le but de développer deux modèles QSPR en utilisant RNA et DA-SVM comme méthodes d'intelligence computationnelle pour corréler la solubilité du CO_2 dans 124 types de LIs a donné de bons résultats. Cependant, les résultats du modèle RNA optimisé à l'aide de l'algorithme d'entraînement trainbr et de la fonction d'activation tansig ont montré des résultats très significatifs avec un coefficient de corrélation global R de 0,9906, et une erreur absolue moyenne EAM de 0,0199. Ceux-ci, ont permis de

prouver l'efficacité et la robustesse du modèle développé, qui a démontré des performances de corrélation précises et de meilleures capacités de généralisation et d'interpolation.

3. L'analyse des données a révélé l'existence d'un certain nombre de données suspectes. Cette analyse a permis aussi de définir le domaine d'applicabilité des modèles développés.
4. Les modèles QSPR obtenus peuvent fournir des estimations rapides des propriétés physico-chimiques et thermodynamiques des liquides ioniques dans le cas où les mesures expérimentales sont difficiles, coûteuses ou même irréalisables. D'autre part, ces modèles peuvent être intégrés dans un cadre CAMD (conception moléculaire assistée par ordinateur) pour identifier les meilleurs LIs pour les processus à appliquer.

En perspectives nous proposons de :

- Promouvoir ces méthodes pour d'autres propriétés physicochimiques des LIs visant à surmonter les limitations du développement des industries et des technologies basées sur les LI,
- Utiliser des améliorations qui ne cessent d'être apportés aux méthodes d'intelligence computationnelle, telles que le '' deeplearning'' pour les RNA dans le but de prédiction/corrélation de ces propriétés ou d'autres propriétés physicochimiques des LIs.

Références bibliographiques

Références bibliographiques

[1] S.K. Singh, A.W. Savoy, Ionic liquids synthesis and applications: An overview, J. Mol. Liq. 297 (2020) 112038. https://doi.org/10.1016/j.molliq.2019.112038.

[2] G. Kaur, H. Kumar, M. Singla, Diverse applications of ionic liquids: A comprehensive review, J. Mol. Liq. 351 (2022) 118556. https://doi.org/10.1016/j.molliq.2022.118556.

[3] A.J. Greer, J. Jacquemin, C. Hardacre, Industrial Applications of Ionic Liquids, 2020. https://doi.org/10.3390/molecules25215207.

[4] A. Szyma, Ionic Liquids, Green Process Eng. (2015) 291–315. https://doi.org/10.1201/b18533-17.

[5] E. Ghasemian, R. Zobeydi, Ionic liquids surface tension prediction based on enthalpy of vaporization, Fluid Phase Equilib. 358 (2013) 40–43. https://doi.org/10.1016/j.fluid.2013.08.008.

[6] A.E. Andreatta, E. Rodil, A. Arce, A. Soto, Surface tension of binary mixtures of 1-alkyl-3-methyl-imidazolium bis(trifluoromethylsulfonyl)imide ionic liquids with alcohols, J. Solution Chem. 43 (2014) 404–420. https://doi.org/10.1007/s10953-014-0128-9.

[7] M.B. Oliveira, M. Domínguez-Pérez, M.G. Freire, F. Llovell, O. Cabeza, J.A. Lopes-Da-Silva, L.F. Vega, J.A.P. Coutinho, Surface tension of binary mixtures of 1-alkyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide ionic liquids: Experimental measurements and soft-SAFT modeling, J. Phys. Chem. B. 116 (2012) 12133–12141. https://doi.org/10.1021/jp3059905.

[8] G.V.S.M. Carrera, J. Inês, C.E.S. Bernardes, K. Klimenko, K. Shimizu, J.N. Canongia Lopes, The Solubility of Gases in Ionic Liquids: A Chemoinformatic Predictive and Interpretable Approach, ChemPhysChem. 22 (2021) 2190–2200. https://doi.org/10.1002/cphc.202100632.

[9] N. Noorani, A. Mehrdad, Fluid Phase Equilibria CO 2 solubility in some amino acid-based ionic liquids : Measurement , correlation and DFT studies, Fluid Phase Equilib. 517 (2020) 112591. https://doi.org/10.1016/j.fluid.2020.112591.

[10] N. Thangarajoo, S.A.A. Taqvi, P. Matheswaran, K. Johari, M.H. Noh, Prediction of infinite dilution activity coefficient of alcohol in ionic liquids using group contribution method, J. Mol. Liq. 324 (2021) 114723. https://doi.org/10.1016/j.molliq.2020.114723.

[11] M. Hashemkhani, R. Soleimani, H. Fazeli, M. Lee, A. Bahadori, M. Tavalaeian, Prediction of the binary surface tension of mixtures containing ionic liquids using Support Vector Machine algorithms, J. Mol. Liq. 211 (2015) 534–552. https://doi.org/10.1016/j.molliq.2015.07.038.

[12] Y. Benguerba, I.M. Alnashef, A. Erto, M. Balsamo, B. Ernst, A quantitative prediction of the viscosity of amine based DESs using Sσ-profile molecular descriptors, J. Mol. Struct. 1184 (2019) 357–363. https://doi.org/10.1016/j.molstruc.2019.02.052.

[13] A. Manuscript, www.rsc.org/pccp, (n.d.).

[14] A. Barati-Harooni, A. Najafi-Marghmaleki, A.H. Mohammadi, ANFIS modeling of

ionic liquids densities, J. Mol. Liq. 224 (2016) 965–975. https://doi.org/10.1016/j.molliq.2016.10.050.

[15] G.W. Meindersma, E. Quijada-Maldonado, M.T.G. Jongmans, J.P.G. Hernandez, B. Schuur, A.B. de Haan, Extractive Distillation with Ionic Liquids: Pilot Plant Experiments and Conceptual Process Design, 2016. https://doi.org/10.1007/978-3-662-48520-0_2.

[16] N. V. Plechkova, K.R. Seddon, Applications of ionic liquids in the chemical industry, Chem. Soc. Rev. 37 (2008) 123–150. https://doi.org/10.1039/b006677j.

[17] R. Setiawan, R. Daneshfar, O. Rezvanjou, S. Ashoori, M. Naseri, Surface tension of binary mixtures containing environmentally friendly ionic liquids: Insights from artificial intelligence, Environ. Dev. Sustain. 23 (2021) 17606–17627. https://doi.org/10.1007/s10668-021-01402-3.

[18] G. Gonfa, M.A. Bustam, A.M. Shariff, N. Muhammad, S. Ullah, Quantitative structure–activity relationships (QSARs) for estimation of activity coefficient at infinite dilution of water in ionic liquids for natural gas dehydration, J. Taiwan Inst. Chem. Eng. 66 (2016) 222–229. https://doi.org/10.1016/j.jtice.2016.06.027.

[19] L. Xi, H. Sun, J. Li, H. Liu, X. Yao, P. Gramatica, Prediction of infinite-dilution activity coefficients of organic solutes in ionic liquids using temperature-dependent quantitative structure-property relationship method, Chem. Eng. J. 163 (2010) 195–201. https://doi.org/10.1016/j.cej.2010.07.023.

[20] H. Malik, H.W. Khan, M.U. Hassan Shah, M.I. Ahmad, I. Khan, A.A. Al-Kahtani, M. Sillanpää, Screening of ionic liquids as green entrainers for ethanol water separation by extractive distillation: COSMO-RS prediction and aspen plus simulation, Chemosphere. 311 (2023). https://doi.org/10.1016/j.chemosphere.2022.136901.

[21] M. Diedenhofen, F. Eckert, A. Klamt, C. Gmbh, C. Kg, B. Strasse, D.- Leverkusen, Compounds in Ionic Liquids Using COSMO-RS, Engineering. 48 (2003) 475–479.

[22] T. Banerjee, A. Khanna, Infinite dilution activity coefficients for trihexyltetradecyl phosphonium ionic liquids: Measurements and COSMO-RS prediction, J. Chem. Eng. Data. 51 (2006) 2170–2177. https://doi.org/10.1021/je0602925.

[23] Z. Lei, W. Arlt, P. Wasserscheid, Separation of 1-hexene and n-hexane with ionic liquids, Fluid Phase Equilib. 241 (2006) 290–299. https://doi.org/10.1016/j.fluid.2005.12.024.

[24] F. Nami, F. Deyhimi, Prediction of activity coefficients at infinite dilution for organic solutes in ionic liquids by artificial neural network, J. Chem. Thermodyn. 43 (2011) 22–27. https://doi.org/10.1016/j.jct.2010.07.011.

[25] R. Anantharaj, T. Banerjee, R. Anantharaj and Tamal B-COSMO-RS-Based Screening of Ionic Liquids as Green Solvents in Denitrification.pdf, (2010) 8705–8725.

[26] H. Passos, I. Khan, F. Mutelet, M.B. Oliveira, P.J. Carvalho, L.M.N.B.F. Santos, C. Held, G. Sadowski, M.G. Freire, J.A.P. Coutinho, Vapor-liquid equilibria of water + alkylimidazolium-based ionic liquids: Measurements and perturbed-chain statistical associating fluid theory modeling, Ind. Eng. Chem. Res. 53 (2014) 3737–3748. https://doi.org/10.1021/ie4041093.

[27] H. Benimam, C. Si-Moussa, M. Laidi, S. Hanini, Modeling the activity coefficient at infinite dilution of water in ionic liquids using artificial neural networks and support vector machines, Neural Comput. Appl. 32 (2020) 8635–8653. https://doi.org/10.1007/s00521-019-04356-w.

[28] H. Benimam, C.S. Moussa, M. Hentabli, S. Hanini, M. Laidi, Dragonfly-Support Vector Machine for Regression Modeling of the Activity Coefficient at Infinite Dilution of Solutes in Imidazolium Ionic Liquids Using σ-Profile Descriptors, J. Chem. Eng. Data. 65 (2020) 3161–3172. https://doi.org/10.1021/acs.jced.0c00168.

[29] S. ullah Rather, A.M. Shariff, A.A. Sulaimon, H.S. Bamufleh, A. Qasim, M. Saad Khan, H. Alhumade, U. Saeed, W. M. Alalayah, Prediction of carbon-dioxide activity coefficient for solubility in ionic liquids using multi-non-linear regression analysis., Chemosphere. 311 (2023) 137102. https://doi.org/10.1016/j.chemosphere.2022.137102.

[30] J.G. Rittig, K. Ben Hicham, A.M. Schweidtmann, M. Dahmen, A. Mitsos, Graph neural networks for temperature-dependent activity coefficient prediction of solutes in ionic liquids, Comput. Chem. Eng. 171 (2023) 108153. https://doi.org/10.1016/j.compchemeng.2023.108153.

[31] K. Li, F. Chang, S. Shi, C. Jiang, Y. Bai, H. Dong, X. Meng, J.C.S. Wu, X. Zhang, A new method of Ionic Fragment Contribution-Gradient Boosting Regressor for predicting the infinite dilution activity coefficient of dichloromethane in ionic liquids, Fluid Phase Equilib. 564 (2023) 113622. https://doi.org/10.1016/j.fluid.2022.113622.

[32] B. Kabane, R. de P. Soares, N. Deenadayalu, I. Bahadur, Pre-screening of 1-ethyl-3-methylimidazolium tetrachloroaluminate and influence of diethylene glycol on the ionic liquid in separation of molecular solutes: Activity coefficients at infinite dilution and COSMO-SAC modelling, Fluid Phase Equilib. 571 (2023) 113808. https://doi.org/10.1016/j.fluid.2023.113808.

[33] H. Xu, H. Li, J. Wang, Density prediction of pyridinium, imidazolium, pyrrolidinium, ammonium, and phosphonium-based ionic liquids using an equation of state, Fluid Phase Equilib. 427 (2016) 406–413. https://doi.org/10.1016/j.fluid.2016.07.027.

[34] A. Najafi-Marghmaleki, A. Tatar, A. Barati-Harooni, A.H. Mohammadi, A GEP based model for prediction of densities of ionic liquids, J. Mol. Liq. 227 (2017) 373–385. https://doi.org/10.1016/j.molliq.2016.11.072.

[35] R.L. Gardas, J.A.P. Coutinho, Extension of the Ye and Shreeve group contribution method for density estimation of ionic liquids in a wide range of temperatures and pressures, Fluid Phase Equilib. 263 (2008) 26–32. https://doi.org/10.1016/j.fluid.2007.09.016.

[36] J.O. Valderrama, K. Zarricueta, A simple and generalized model for predicting the density of ionic liquids, Fluid Phase Equilib. 275 (2009) 145–151. https://doi.org/10.1016/j.fluid.2008.10.002.

[37] J.O. Valderrama, A. Reátegui, R.E. Rojas, Density of ionic liquids using group contribution and artificial neural networks, Ind. Eng. Chem. Res. 48 (2009) 3254–3259. https://doi.org/10.1021/ie801113x.

[38] J.A. Lazzús, A group contribution method to predict ρ-T-P of ionic liquids, Chem. Eng. Commun. 197 (2010) 974–1015. https://doi.org/10.1080/00986440903359376.

[39] K. Paduszy, U. Doma, I&EC density.pdf, (2012) 591–604.

[40] N.K. Patel, M.H. Joshipura, Generalized PSRK model for prediction of liquid density of ionic liquids, Procedia Eng. 51 (2013) 386–394. https://doi.org/10.1016/j.proeng.2013.01.054.

[41] Y. Huang, X. Zhang, Y. Zhao, S. Zeng, H. Dong, S. Zhang, New models for predicting thermophysical properties of ionic liquid mixtures, Phys. Chem. Chem. Phys. 17 (2015) 26918–26929. https://doi.org/10.1039/c5cp03446a.

[42] X. Ji, C. Held, Modeling the density of ionic liquids with ePC-SAFT, Fluid Phase Equilib. 410 (2016) 9–22. https://doi.org/10.1016/j.fluid.2015.11.014.

[43] M.H. Keshavarz, H.R. Pouretedal, E. Saberi, A simple method for prediction of density of ionic liquids through their molecular structure, J. Mol. Liq. 216 (2016) 732–737. https://doi.org/10.1016/j.molliq.2016.01.095.

[44] A. Najafi-Marghmaleki, M.R. Khosravi-Nikou, A. Barati-Harooni, A new model for prediction of binary mixture of ionic liquids + water density using artificial neural network, J. Mol. Liq. 220 (2016) 232–237. https://doi.org/10.1016/j.molliq.2016.04.085.

[45] A. Barati-Harooni, A. Najafi-Marghmaleki, A.H. Mohammadi, ANFIS modeling of ionic liquids densities, J. Mol. Liq. 224 (2016) 965–975. https://doi.org/10.1016/j.molliq.2016.10.050.

[46] R. Farzi, F. Esmaeilzadeh, Prediction of densities of pure ionic liquids using Esmaeilzadeh-Roshanfekr equation of state and critical properties from group contribution method, Fluid Phase Equilib. 423 (2016) 101–108. https://doi.org/10.1016/j.fluid.2016.04.020.

[47] A. Barati-Harooni, A. Najafi-Marghmaleki, A.H. Mohammadi, A reliable radial basis function neural network model (RBF-NN) for the prediction of densities of ionic liquids, J. Mol. Liq. 231 (2017) 462–473. https://doi.org/10.1016/j.molliq.2017.01.052.

[48] H. Shekaari, M.T. Zafarani-Moattar, S. Faraji, M. Mokhtarpour, Prediction of vapor pressure and density for nonaqueous solutions of the ionic liquid 1-ethyl-3-methylimidazolium ethyl sulfate using PC-SAFT equation of state, Fluid Phase Equilib. 506 (2020) 112320. https://doi.org/10.1016/j.fluid.2019.112320.

[49] A.S. Zimmermann, S. Mattedi, Density and speed of sound prediction for binary mixtures of water and ammonium-based ionic liquids using feedforward and cascade forward neural networks, J. Mol. Liq. 311 (2020) 113212. https://doi.org/10.1016/j.molliq.2020.113212.

[50] M. Nait Amar, M.A. Ghriga, A. Hemmati-Sarapardeh, Application of gene expression programming for predicting density of binary and ternary mixtures of ionic liquids and molecular solvents, J. Taiwan Inst. Chem. Eng. 117 (2020) 63–74. https://doi.org/10.1016/j.jtice.2020.11.029.

[51] H. Bagheri, M. Sadegh Hosseini, H. Ghayoumi Zadeh, B. Notej, A. Fayazi, A novel

modification of ionic liquid mixture density based on semi-empirical equations using laplacian whale optimization algorithm, Arab. J. Chem. 14 (2021) 103368. https://doi.org/10.1016/j.arabjc.2021.103368.

[52] S.S. Ashrafmansouri, Modeling the density and the second-order thermodynamic derivative properties of imidazolium-, cyano-based ionic liquids using the SAFT-γ EoS, Fluid Phase Equilib. 548 (2021) 113190. https://doi.org/10.1016/j.fluid.2021.113190.

[53] M. Hosseinzadeh, A. Hemmati-Sarapardeh, Toward a predictive model for estimating viscosity of ternary mixtures containing ionic liquids, J. Mol. Liq. 200 (2014) 340–348. https://doi.org/10.1016/j.molliq.2014.10.033.

[54] A.D. Boualem, K. Argoub, A.M. Benkouider, A. Yahiaoui, K. Toubal, Viscosity prediction of ionic liquids using NLR and SVM approaches, J. Mol. Liq. 368 (2022) 120610. https://doi.org/10.1016/j.molliq.2022.120610.

[55] F. Gharagheizi, P. Ilani-Kashkouli, A.H. Mohammadi, D. Ramjugernath, D. Richon, Development of a group contribution method for determination of viscosity of ionic liquids at atmospheric pressure, Chem. Eng. Sci. 80 (2012) 326–333. https://doi.org/10.1016/j.ces.2012.06.045.

[56] B.K. Chen, M.J. Liang, T.Y. Wu, H.P. Wang, A high correlate and simplified QSPR for viscosity of imidazolium-based ionic liquids, Fluid Phase Equilib. 350 (2013) 37–42. https://doi.org/10.1016/j.fluid.2013.04.009.

[57] D.M. Bajić, S.P. Šerbanović, E.M. Živković, J. Jovanović, M.L. Kijevčanin, Prediction and correlation of viscosity of binary mixtures of ionic liquids with organic solvents, J. Mol. Liq. 197 (2014) 1–6. https://doi.org/10.1016/j.molliq.2014.04.005.

[58] M.R. Fatehi, S. Raeissi, D. Mowla, Estimation of viscosity of binary mixtures of ionic liquids and solvents using an artificial neural network based on the structure groups of the ionic liquid, Fluid Phase Equilib. 364 (2014) 88–94. https://doi.org/10.1016/j.fluid.2013.11.041.

[59] K. Paduszyński, U. Domańska, Viscosity of ionic liquids: An extensive database and a new group contribution model based on a feed-forward artificial neural network, J. Chem. Inf. Model. 54 (2014) 1311–1324. https://doi.org/10.1021/ci500206u.

[60] R. Haghbakhsh, S. Raeissi, Two simple correlations to predict viscosities of pure and aqueous solutions of ionic liquids, J. Mol. Liq. 211 (2015) 948–956. https://doi.org/10.1016/j.molliq.2015.08.036.

[61] J.A. Lazzús, G. Pulgar-Villarroel, A group contribution method to estimate the viscosity of ionic liquids at different temperatures, J. Mol. Liq. 209 (2015) 161–168. https://doi.org/10.1016/j.molliq.2015.05.030.

[62] M. Sattari, A. Kamari, H. Hashemi, A.H. Mohammadi, D. Ramjugernath, A group contribution model for prediction of the viscosity with temperature dependency for fluorine-containing ionic liquids, J. Fluor. Chem. 186 (2016) 19–27. https://doi.org/10.1016/j.jfluchem.2016.04.001.

[63] N. Zhao, J. Jacquemin, R. Oozeerally, V. Degirmenci, New method for the estimation of viscosity of pure and mixtures of ionic liquids based on the UNIFAC-VISCO

Model, J. Chem. Eng. Data. 61 (2016) 2160–2169. https://doi.org/10.1021/acs.jced.6b00161.

[64] Y. Zhao, X. Zhang, L. Deng, S. Zhang, Prediction of viscosity of imidazolium-based ionic liquids using MLR and SVM algorithms, Comput. Chem. Eng. 92 (2016) 37–42. https://doi.org/10.1016/j.compchemeng.2016.04.035.

[65] M.R. Fatehi, S. Raeissi, D. Mowla, Estimation of viscosities of pure ionic liquids using an artificial neural network based on only structural characteristics, J. Mol. Liq. 227 (2017) 309–317. https://doi.org/10.1016/j.molliq.2016.11.133.

[66] A. Baghban, M.N. Kardani, S. Habibzadeh, Prediction viscosity of ionic liquids using a hybrid LSSVM and group contribution method, J. Mol. Liq. 236 (2017) 452–464. https://doi.org/10.1016/j.molliq.2017.04.019.

[67] X. Kang, Z. Zhao, J. Qian, R.M. Afzal, Predicting the viscosity of ionic liquids by the ELM intelligence algorithm, Ind. Eng. Chem. Res. 56 (2017) 11344–11351. https://doi.org/10.1021/acs.iecr.7b02722.

[68] J.O. Valderrama, L.F. Cardona, R.E. Rojas, Correlation and prediction of ionic liquid viscosity using Valderrama-Patel-Teja cubic equation of state and the geometric similitude concept. Part I: Pure ionic liquids, Fluid Phase Equilib. 497 (2019) 164–177. https://doi.org/10.1016/j.fluid.2019.04.031.

[69] J.O. Valderrama, L.F. Cardona, R.E. Rojas, Correlation of ionic liquid viscosity using Valderrama-Patel-Teja cubic equation of state and the geometric similitude concept. Part II: Binary mixtures of ionic liquids, Fluid Phase Equilib. 497 (2019) 178–194. https://doi.org/10.1016/j.fluid.2019.04.034.

[70] K. Paduszyński, Extensive Databases and Group Contribution QSPRs of Ionic Liquids Properties. 2. Viscosity, Ind. Eng. Chem. Res. 58 (2019) 17049–17066. https://doi.org/10.1021/acs.iecr.9b03150.

[71] N. Nabipour, J. Sasanipour, A. Baghban, A.H. Mohammadi, Towards ANFIS-PSO strategy for estimating viscosity of ternary mixtures containing ionic liquids, J. Mol. Liq. 298 (2020) 111802. https://doi.org/10.1016/j.molliq.2019.111802.

[72] L. Darabi, M. Zare, High correlate simple equations for temperature and pressure dependence of the viscosity of ionic liquids, Chem. Phys. 539 (2020) 110933. https://doi.org/10.1016/j.chemphys.2020.110933.

[73] I.Y. López-Cortés, G.A. Iglesias-Silva, M. Ramos-Estrada, J.L. Rivera-Rojas, A correlation for the viscosity of binary mixtures of ionic liquids with organic solvents and water, Fluid Phase Equilib. 514 (2020). https://doi.org/10.1016/j.fluid.2020.112543.

[74] T. Boldoo, M. Lee, Y.T. Kang, H. Cho, Development of artificial neural network model for predicting dynamic viscosity and specific heat of MWCNT nanoparticle-enhanced ionic liquids with different [HMIM]-cation base agents, J. Mol. Liq. 341 (2021) 117356. https://doi.org/10.1016/j.molliq.2021.117356.

[75] Y. Chen, B. Peng, G.M. Kontogeorgis, X. Liang, Machine learning for the prediction of viscosity of ionic liquid–water mixtures, J. Mol. Liq. 350 (2022) 118546. https://doi.org/10.1016/j.molliq.2022.118546.

[76] A.S. Zimmermann, S. Mattedi, Feedforward and cascade forward networks for viscosity prediction for binary mixtures of ammonium-based ionic liquids and water, Fluid Phase Equilib. 556 (2022) 113416. https://doi.org/10.1016/j.fluid.2022.113416.

[77] M. Huang, J. Deng, G. Jia, Predicting viscosity of ionic liquids - water mixtures by bridging UNIFAC modeling with interpretable machine learning, J. Mol. Liq. 383 (2023) 122095. https://doi.org/10.1016/j.molliq.2023.122095.

[78] H. Moslehi, S.M. Hosseini, M. Pierantozzi, M.M. Alavianmehr, B. Haghighi, Modeling the viscosity of ionic liquids using combined friction theory with perturbed hard-chain equation of state and neural network approaches, J. Mol. Liq. 383 (2023) 122004. https://doi.org/10.1016/j.molliq.2023.122004.

[79] W. He, F. Yan, Q. Jia, S. Xia, Q. Wang, Description of the Thermal Conductivity λ(T, P) of Ionic Liquids Using the Structure-Property Relationship Method, J. Chem. Eng. Data. 62 (2017) 2466–2472. https://doi.org/10.1021/acs.jced.7b00422.

[80] J.A. Lazzús, G. Pulgar-Villarroel, Estimation of thermal conductivity of ionic liquids using quantitative structure-property relationship calculations, J. Mol. Liq. 211 (2015) 981–985. https://doi.org/10.1016/j.molliq.2015.08.037.

[81] A.Z. Hezave, S. Raeissi, M. Lashkarbolooki, Estimation of thermal conductivity of ionic liquids using a perceptron neural network, Ind. Eng. Chem. Res. 51 (2012) 9886–9893. https://doi.org/10.1021/ie202681b.

[82] K.J. Wu, C.X. Zhao, C.H. He, Development of a group contribution method for determination of thermal conductivity of ionic liquids, Fluid Phase Equilib. 339 (2013) 10–14. https://doi.org/10.1016/j.fluid.2012.11.024.

[83] S.A. Shojaee, S. Farzam, A.Z. Hezave, M. Lashkarbolooki, S. Ayatollahi, A new correlation for estimating thermal conductivity of pure ionic liquids, Fluid Phase Equilib. 354 (2013) 199–206. https://doi.org/10.1016/j.fluid.2013.06.004.

[84] J. Albert, K. Müller, Thermal conductivity of Ionic Liquids: An estimation approach, Chem. Eng. Sci. 119 (2014) 109–113. https://doi.org/10.1016/j.ces.2014.08.023.

[85] J. Albert, K. Müller, A group contribution method for the thermal properties of ionic liquids, Ind. Eng. Chem. Res. 53 (2014) 17522–17526. https://doi.org/10.1021/ie503366p.

[86] Q.L. Chen, K.J. Wu, C.H. He, Thermal conductivity of ionic liquids at atmospheric pressure: Database, analysis, and prediction using a topological index method, Ind. Eng. Chem. Res. 53 (2014) 7224–7232. https://doi.org/10.1021/ie403500w.

[87] This is a repository copy of Speed of sound of ionic liquids : Database , estimation , and its application for thermal conductivity prediction . White Rose Research Online URL for this paper : Version : Accepted Version Article : Wu , K-J orcid . org / 00, (2014).

[88] J.A. Lazzús, Estimation of the thermal conductivity λ(T, P) of ionic liquids using a neural network optimized with genetic algorithms, Comptes Rendus Chim. 19 (2016) 333–341. https://doi.org/10.1016/j.crci.2015.09.010.

[89] J.A. Lazzús, A group contribution method to predict the thermal conductivity λ(T,P) of

ionic liquids, Fluid Phase Equilib. 405 (2015) 141–149. https://doi.org/10.1016/j.fluid.2015.07.015.

[90] S. Atashrouz, M. Mozaffarian, G. Pazuki, Modeling the Thermal Conductivity of Ionic Liquids and Ionanofluids Based on a Group Method of Data Handling and Modified Maxwell Model, Ind. Eng. Chem. Res. 54 (2015) 8600–8610. https://doi.org/10.1021/acs.iecr.5b00932.

[91] S. Atashrouz, A. Hemmati-Sarapardeh, H. Mirshekar, B. Nasernejad, M. Keshavarz Moraveji, On the evaluation of thermal conductivity of ionic liquids: Modeling and data assessment, J. Mol. Liq. 224 (2016) 648–656. https://doi.org/10.1016/j.molliq.2016.09.106.

[92] D.P. Soman, P. Kalaichelvi, T.K. Radhakrishnan, Thermal conductivity enhancement of aqueous ionic liquid and nanoparticle suspension, Brazilian J. Chem. Eng. 36 (2019) 855–868. https://doi.org/10.1590/0104-6632.20190362s20180436.

[93] L.F. Cardona, J.O. Valderrama, Physical and transport properties of ionic liquids using the geometric similitude concept and a cubic equation of state. Part 1: Thermal conductivity and speed of sound of pure substances, J. Mol. Liq. 315 (2020) 113681. https://doi.org/10.1016/j.molliq.2020.113681.

[94] L.F. Cardona, J.O. Valderrama, Physical and transport properties of ionic liquids using geometric similitude and a cubic equation of state. Part 2: Thermal conductivity, and speed of sound of water + ionic liquid mixtures, J. Mol. Liq. 317 (2020) 113926. https://doi.org/10.1016/j.molliq.2020.113926.

[95] S.P. Mousavi, S. Atashrouz, F. Rezaei, M.E. Peyvastegan, A. Hemmati-Sarapardeh, A. Mohaddespour, Modeling thermal conductivity of ionic liquids: A comparison between chemical structure and thermodynamic properties-based models, J. Mol. Liq. 322 (2021) 114911. https://doi.org/10.1016/j.molliq.2020.114911.

[96] S. Zhang, Q. Jia, F. Yan, S. Xia, Q. Wang, Evaluating the properties of ionic liquid at variable temperatures and pressures by quantitative structure–property relationship (QSPR), Chem. Eng. Sci. 231 (2021) 116326. https://doi.org/10.1016/j.ces.2020.116326.

[97] R. Wan, M. Li, F. Song, Y. Xiao, F. Zeng, C. Peng, H. Liu, Predicting the Thermal Conductivity of Ionic Liquids Using a Quantitative Structure-Property Relationship, Ind. Eng. Chem. Res. 61 (2022) 12032–12039. https://doi.org/10.1021/acs.iecr.2c01964.

[98] D. Mital, S. Liaqat, P. Nancarrow, T. Ibrahim, N.A. Jabbar, M. Khamis, Thermal Conductivities of NTf2-Based Ionic Liquids: Experimental Measurement and Prediction via Group Contribution Methods, J. Chem. Eng. Data. 67 (2022) 2136–2148. https://doi.org/10.1021/acs.jced.2c00123.

[99] S.P. Mousavi, S. Atashrouz, M. Nait Amar, F. Hadavimoghaddam, M.R. Mohammadi, A. Hemmati-Sarapardeh, A. Mohaddespour, Modeling surface tension of ionic liquids by chemical structure-intelligence based models, J. Mol. Liq. 342 (2021) 116961. https://doi.org/10.1016/j.molliq.2021.116961.

[100] R.L. Gardas, J.A.P. Coutinho, Applying a QSPR correlation to the prediction of surface

tensions of ionic liquids, Fluid Phase Equilib. 265 (2008) 57–65. https://doi.org/10.1016/j.fluid.2008.01.002.

[101] F. Gharagheizi, P. Ilani-Kashkouli, A.H. Mohammadi, Group contribution model for estimation of surface tension of ionic liquids, Chem. Eng. Sci. 78 (2012) 204–208. https://doi.org/10.1016/j.ces.2012.05.008.

[102] K.J. Wu, C.X. Zhao, C.H. He, A simple corresponding-states group-contribution method for estimating surface tension of ionic liquids, Fluid Phase Equilib. 328 (2012) 42–48. https://doi.org/10.1016/j.fluid.2012.05.021.

[103] Q. Shang, F. Yan, S. Xia, Q. Wang, P. Ma, Predicting the surface tensions of ionic liquids by the quantitative structure property relationship method using a topological index, Chem. Eng. Sci. 101 (2013) 266–270. https://doi.org/10.1016/j.ces.2013.05.053.

[104] S.A. Mirkhani, F. Gharagheizi, N. Farahani, K. Tumba, Prediction of surface tension of ionic liquids by molecular approach, J. Mol. Liq. 179 (2013) 78–87. https://doi.org/10.1016/j.molliq.2012.11.018.

[105] E. Ghasemian Lemraski, R. Zobeydi, Applying parachor method to the prediction of ionic liquids surface tension based on modified group contribution, J. Mol. Liq. 193 (2014) 204–209. https://doi.org/10.1016/j.molliq.2013.12.043.

[106] E.G. Lemraski, Z. Pouyanfar, Prediction of surface tension, surface mole fraction and thickness of the surface layer in the ionic liquid binary mixtures, J. Mol. Liq. 203 (2015) 52–58. https://doi.org/10.1016/j.molliq.2014.12.030.

[107] S. Atashrouz, H. Mirshekar, A. Hemmati-Sarapardeh, M.K. Moraveji, B. Nasernejad, Implementation of soft computing approaches for prediction of physicochemical properties of ionic liquid mixtures, Korean J. Chem. Eng. 34 (2017) 425–439. https://doi.org/10.1007/s11814-016-0271-7.

[108] R. Soleimani, A.H. Saeedi Dehaghani, N.A. Shoushtari, P. Yaghoubi, A. Bahadori, Toward an intelligent approach for predicting surface tension of binary mixtures containing ionic liquids, Korean J. Chem. Eng. 35 (2018) 1556–1569. https://doi.org/10.1007/s11814-017-0326-4.

[109] S. Atashrouz, H. Mirshekar, A. Mohaddespour, A robust modeling approach to predict the surface tension of ionic liquids, J. Mol. Liq. 236 (2017) 344–357. https://doi.org/10.1016/j.molliq.2017.04.039.

[110] J.A. Lazzús, G. Pulgar-Villarroel, F. Cuturrufo, P. Vega, Development of a group contribution method for estimating surface tension of ionic liquids over a wide range of temperatures, J. Mol. Liq. 240 (2017) 522–531. https://doi.org/10.1016/j.molliq.2017.05.095.

[111] S. Shahsavari, M. Mesbah, E. Soroush, H. Farhangian, S. Alizadeh, S. Soltanali, A simple group contribution correlation for modeling the surface tension of pure ionic liquids, J. Mol. Liq. 265 (2018) 292–298. https://doi.org/10.1016/j.molliq.2018.06.004.

[112] H. Esmaeili, H. Hashemipour, A simple correlation for determining ionic liquids surface tension, J. Mol. Liq. 272 (2018) 692–696. https://doi.org/10.1016/j.molliq.2018.10.011.

[113] A. Shojaeian, M. Asadizadeh, Prediction of surface tension of the binary mixtures containing ionic liquid using heuristic approaches; an input parameters investigation, J. Mol. Liq. 298 (2020) 111976. https://doi.org/10.1016/j.molliq.2019.111976.

[114] L.F. Cardona, J.O. Valderrama, Surface tension of mixtures containing ionic liquids based on an equation of state and on the geometric similitude concept, Ionics (Kiel). 26 (2020) 6095–6118. https://doi.org/10.1007/s11581-020-03697-0.

[115] H. Esmaeili, H. Hashemipour, A simple correlation to predict surface tension of binary mixtures containing ionic liquids, J. Mol. Liq. 324 (2021) 114660. https://doi.org/10.1016/j.molliq.2020.114660.

[116] J. Wang, X. Zheng, M. Hong, Y. Kong, J. Tong, Prediction of viscosity and surface tension properties of aqueous solution of ionic liquids EMIES, J. Mol. Liq. 320 (2020) 114460. https://doi.org/10.1016/j.molliq.2020.114460.

[117] K. Paduszyński, Extensive Databases and Group Contribution QSPRs of Ionic Liquid Properties. 3: Surface Tension, Ind. Eng. Chem. Res. 60 (2021) 5705–5720. https://doi.org/10.1021/acs.iecr.1c00783.

[118] Y. Ge, Q. Zhu, Y. Li, H. Dong, J. Ma, An electrostatic-variable coarse-grained model for predicting enthalpy of vaporization, surface tension, diffusivity, conductivity, and dielectric constant of aqueous ionic liquid, J. Mol. Liq. 346 (2022) 118230. https://doi.org/10.1016/j.molliq.2021.118230.

[119] I. Baskin, A. Epshtein, Y. Ein-Eli, Benchmarking machine learning methods for modeling physical properties of ionic liquids, J. Mol. Liq. 351 (2022) 118616. https://doi.org/10.1016/j.molliq.2022.118616.

[120] M. A. S. Abourehab, A.M. Shawky, K. Venkatesan, S. Yasmin, A.M. Alsubaiyel, K.M. AboRas, Efficiency development of surface tension for different ionic liquids through novel model of Machine learning Technique: Application of in-thermal engineering, J. Mol. Liq. 367 (2022) 120391. https://doi.org/10.1016/j.molliq.2022.120391.

[121] S.M. Alshahrani, J. Uddin, S. Yasmin, S. Alshehri, A.M. Alsubaiyel, Prediction of novel ionic liquids' surface tension via Bagging KNN predictive model: Modeling and simulation, J. Mol. Liq. 368 (2022) 120748. https://doi.org/10.1016/j.molliq.2022.120748.

[122] Y. Fu, Y. Chen, C. Zhang, Y. Lei, X. Liu, Prediction surface tension of ionic liquid–water mixtures using a hybrid group contribution and artificial neural network method, Fluid Phase Equilib. 563 (2022) 113571. https://doi.org/10.1016/j.fluid.2022.113571.

[123] M. Mohan, M.D. Smith, O.N. Demerdash, B.A. Simmons, S. Singh, M.K. Kidder, J.C. Smith, Quantum Chemistry-Driven Machine Learning Approach for the Prediction of the Surface Tension and Speed of Sound in Ionic Liquids, ACS Sustain. Chem. Eng. 11 (2023) 7809–7821. https://doi.org/10.1021/acssuschemeng.3c00624.

[124] M. Mohan, M.D. Smith, O. Demerdash, M.K. Kidder, J.C. Smith, Predictive understanding of the surface tension and velocity of sound in ionic liquids using machine learning, J. Chem. Phys. 158 (2023). https://doi.org/10.1063/5.0147052.

[125] D. Type, A.A. Faculty, M. Engineering, Prediction of Density , Dynamic Viscosity , and Surface Tension of seven, 222 (2023) 7–8.

[126] R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Bostani, A. Hemmati-Sarapardeh, A. Mohaddespour, Solubility of gaseous hydrocarbons in ionic liquids using equations of state and machine learning approaches, Sci. Rep. 12 (2022) 1–26. https://doi.org/10.1038/s41598-022-17983-6.

[127] Y.S. Kim, W.Y. Choi, J.H. Jang, K.P. Yoo, C.S. Lee, Solubility measurement and prediction of carbon dioxide in ionic liquids, Fluid Phase Equilib. 228–229 (2005) 439–445. https://doi.org/10.1016/j.fluid.2004.09.006.

[128] A.Y. Mark B. Shiflett, Solubilities and Diffusivities of Carbon Dioxide in Ionic Liquids: [bmim][PF_6] and [bmim][$BF4_4$], AIChE Annu. Meet. Conf. Proc. (2005) 4453–4464.

[129] Y.S. Kim, J.H. Jang, B.D. Lim, J.W. Kang, C.S. Lee, Solubility of mixed gases containing carbon dioxide in ionic liquids: Measurements and predictions, Fluid Phase Equilib. 256 (2007) 70–74. https://doi.org/10.1016/j.fluid.2006.11.019.

[130] E.K. Karakatsani, I.G. Economou, M.C. Kroon, C.J. Peters, G.J. Witkamp, tPC-PSAFT modeling of gas solubility in imidazolium-based ionic liquids, J. Phys. Chem. C. 111 (2007) 15487–15492. https://doi.org/10.1021/jp070556+.

[131] J.S. Andreu, L.F. Vega, Modeling the solubility behavior of CO2, H2, and Xe in [Cn-mim] [Tf2N] ionic liquids, J. Phys. Chem. B. 112 (2008) 15398–15406. https://doi.org/10.1021/jp807484g.

[132] T.K. Carlisle, J.E. Bara, C.J. Gabriel, R.D. Noble, D.L. Gin, Interpretation of CO2 solubility and selectivity in nitrile-functionalized room-temperature ionic liquids using a group contribution approach, Ind. Eng. Chem. Res. 47 (2008) 7005–7012. https://doi.org/10.1021/ie8001217.

[133] Z. Qiao, Z. Wang, C. Zhang, S. Yuan, Y. Zhu, J. Wang, PVAm–PIP/PS composite membrane with high performance for CO_2/N_2 separation, AIChE J. 59 (2012) 215–228. https://doi.org/10.1002/aic.

[134] Y. Shimoyama, A. Ito, Predictions of cation and anion effects on solubilities, selectivities and permeabilities for CO2 in ionic liquid using COSMO based activity coefficient model, Fluid Phase Equilib. 297 (2010) 178–182. https://doi.org/10.1016/j.fluid.2010.03.026.

[135] A. Yokozeki, M.B. Shiflett, Gas solubilities in ionic liquids using a generic van der Waals equation of state, J. Supercrit. Fluids. 55 (2010) 846–851. https://doi.org/10.1016/j.supflu.2010.09.015.

[136] A. Eslamimanesh, F. Gharagheizi, A.H. Mohammadi, D. Richon, Artificial Neural Network modeling of solubility of supercritical carbon dioxide in 24 commonly used ionic liquids, Chem. Eng. Sci. 66 (2011) 3039–3044. https://doi.org/10.1016/j.ces.2011.03.016.

[137] J.E. Kim, J.S. Lim, J.W. Kang, Measurement and correlation of solubility of carbon dioxide in 1-alkyl-3-methylimidazolium hexafluorophosphate ionic liquids, Fluid Phase Equilib. 306 (2011) 251–255. https://doi.org/10.1016/j.fluid.2011.04.017.

[138] M.D. Bermejo, T.M. Fieback, Á. Martín, Solubility of gases in 1-alkyl-3methylimidazolium alkyl sulfate ionic liquids: Experimental determination and

modeling, J. Chem. Thermodyn. 58 (2013) 237–244. https://doi.org/10.1016/j.jct.2012.11.018.

[139] Y. Chen, F. Mutelet, J.N. Jaubert, Modeling the solubility of carbon dioxide in imidazolium-based ionic liquids with the PC-SAFT equation of state, J. Phys. Chem. B. 116 (2012) 14375–14388. https://doi.org/10.1021/jp309944t.

[140] M. Gonzalez-Miquel, M. Talreja, A.L. Ethier, K. Flack, J.R. Switzer, E.J. Biddinger, P. Pollet, J. Palomar, F. Rodriguez, C.A. Eckert, C.L. Liotta, COSMO-RS Studies: Structure-property relationships for CO2 capture by reversible ionic liquids, Ind. Eng. Chem. Res. 51 (2012) 16066–16073. https://doi.org/10.1021/ie302449c.

[141] R. MacÍas-Salinas, J.A. Chávez-Velasco, M.A. Aquino-Olivos, J.L. Mendoza De La Cruz, J.C. Sánchez-Ochoa, Accurate modeling of CO2 solubility in ionic liquids using a cubic EoS⊥, Ind. Eng. Chem. Res. 52 (2013) 7593–7601. https://doi.org/10.1021/ie400106g.

[142] R. Haghbakhsh, H. Soleymani, S. Raeissi, A simple correlation to predict high pressure solubility of carbon dioxide in 27 commonly used ionic liquids, J. Supercrit. Fluids. 77 (2013) 158–166. https://doi.org/10.1016/j.supflu.2013.02.022.

[143] M. Mirarab, M. Sharifi, M.A. Ghayyem, F. Mirarab, Prediction of solubility of CO2 in ethanol-[EMIM][Tf2N] ionic liquid mixtures using artificial neural networks based on genetic algorithm, Fluid Phase Equilib. 371 (2014) 6–14. https://doi.org/10.1016/j.fluid.2014.02.030.

[144] M.A. Sedghamiz, A. Rasoolzadeh, M.R. Rahimpour, The ability of artificial neural network in prediction of the acid gases solubility in different ionic liquids, J. CO2 Util. 9 (2015) 39–47. https://doi.org/10.1016/j.jcou.2014.12.003.

[145] A. Baghban, M.A. Ahmadi, B.H. Shahraki, Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches, J. Supercrit. Fluids. 98 (2015) 50–64. https://doi.org/10.1016/j.supflu.2015.01.002.

[146] H.F. Hizaddin, M.K. Hadj-Kali, I.M. Alnashef, F.S. Mjalli, M.A. Hashim, Prediction of CO2 solubility in ionic liquids using the PSRK model, J. Supercrit. Fluids. 100 (2015) 184–193. https://doi.org/10.1016/j.supflu.2015.02.009.

[147] M.E. Hamzehie, M. Fattahi, H. Najibi, B. Van Der Bruggen, S. Mazinani, Journal of Natural Gas Science and Engineering Application of arti fi cial neural networks for estimation of solubility of acid gases (H 2 S and CO 2) in 32 commonly ionic liquid and amine solutions, 24 (2015) 106–114. https://doi.org/10.1016/j.jngse.2015.03.014.

[148] Z. Bazargani, F. Sabzi, Prediction of CO2 solubility in ionic liquids with [HMIM] and [OMIM] cations by equation of state, J. Mol. Liq. 216 (2016) 87–93. https://doi.org/10.1016/j.molliq.2015.12.092.

[149] I. Mehraein, S. Riahi, The QSPR models to predict the solubility of CO2 in ionic liquids based on least-squares support vector machines and genetic algorithm-multi linear regression, J. Mol. Liq. 225 (2017) 521–530. https://doi.org/10.1016/j.molliq.2016.10.133.

[150] A. Kamgar, M.R. Rahimpour, Prediction of CO2 solubility in ionic liquids with QM and UNIQUAC models, J. Mol. Liq. 222 (2016) 195–200.

https://doi.org/10.1016/j.molliq.2016.06.107.

[151] A. Tatar, S. Naseri, M. Bahadori, A.Z. Hezave, T. Kashiwao, A. Bahadori, H. Darvish, Prediction of carbon dioxide solubility in ionic liquids using MLP and radial basis function (RBF) neural networks, J. Taiwan Inst. Chem. Eng. 60 (2016) 151–164. https://doi.org/10.1016/j.jtice.2015.11.002.

[152] E. Ali, M.K. Hadj-Kali, I. Alnashef, Modeling of CO2 Solubility in Selected Imidazolium-Based Ionic Liquids, Chem. Eng. Commun. 204 (2017) 205–215. https://doi.org/10.1080/00986445.2016.1254086.

[153] H. Al-fnaish, L. Lue, Modelling the solubility of H2S and CO2 in ionic liquids using PC-SAFT equation of state, Fluid Phase Equilib. 450 (2017) 30–41. https://doi.org/10.1016/j.fluid.2017.07.008.

[154] I. Mehraein, S. Riahi, The QSPR models to predict the solubility of CO 2 in ionic liquids based on least-squares support vector machines and genetic algorithm-multi linear regression, J. Mol. Liq. 225 (2017) 521–530. https://doi.org/10.1016/j.molliq.2016.10.133.

[155] C. Li, H. Liu, M. Xiao, X. Luo, H. Gao, Z. Liang, International Journal of Greenhouse Gas Control Thermodynamics and ANN models for predication of the equilibrium CO 2 solubility in aqueous 3-dimethylamino-1-propanol solution, Int. J. Greenh. Gas Control. 63 (2017) 77–85. https://doi.org/10.1016/j.ijggc.2017.05.007.

[156] V. Venkatraman, B.K. Alsberg, Predicting CO 2 capture of ionic liquids using machine learning, J. CO2 Util. 21 (2017) 162–168. https://doi.org/10.1016/j.jcou.2017.06.012.

[157] A. Baghban, A.H. Mohammadi, M.S. Taleghani, Rigorous modeling of CO2 equilibrium absorption in ionic liquids, Int. J. Greenh. Gas Control. 58 (2017) 19–41. https://doi.org/10.1016/j.ijggc.2016.12.009.

[158] H. Soltani Panah, Modeling H2S and CO2 solubility in ionic liquids using the CPA equation of state through a new approach, Fluid Phase Equilib. 437 (2017) 155–165. https://doi.org/10.1016/j.fluid.2017.01.023.

[159] A. Akbari, M.R. Rahimpour, Prediction of the solubility of carbon dioxide in imidazolium based ionic liquids using the modified scaled particle theory, J. Mol. Liq. 255 (2018) 135–147. https://doi.org/10.1016/j.molliq.2018.01.085.

[160] A. Dashti, H. Riasat Harami, M. Rezakazemi, S. Shirazian, Estimating CH4 and CO2 solubilities in ionic liquids using computational intelligence approaches, J. Mol. Liq. 271 (2018) 661–669. https://doi.org/10.1016/j.molliq.2018.08.150.

[161] M. Mesbah, S. Shahsavari, E. Soroush, N. Rahaei, Accurate prediction of miscibility of CO 2 and supercritical CO 2 in ionic liquids using machine learning, J. CO2 Util. 25 (2018) 99–107. https://doi.org/10.1016/j.jcou.2018.03.004.

[162] M.K. Hadj-Kali, M. Althuluth, S. Mokraoui, I. Wazeer, E. Ali, D. Richon, Screening of ionic liquids for gas separation using COSMO-RS and comparison between performances of ionic liquids and aqueous alkanolamine solutions, Chem. Eng. Commun. 207 (2020) 1264–1277. https://doi.org/10.1080/00986445.2019.1680363.

[163] T. Deng, F. Liu, G. Jia, Prediction carbon dioxide solubility in ionic liquids based on

deep learning, Mol. Phys. 0 (2019) 1–8. https://doi.org/10.1080/00268976.2019.1652367.

[164] Z. Song, H. Shi, X. Zhang, T. Zhou, Prediction of CO 2 solubility in ionic liquids using machine learning methods, Chem. Eng. Sci. 223 (2020) 115752. https://doi.org/10.1016/j.ces.2020.115752.

[165] M. Aghaie, S. Zendehboudi, Estimation of CO2 solubility in ionic liquids using connectionist tools based on thermodynamic and structural characteristics, Fuel. 279 (2020) 117984. https://doi.org/10.1016/j.fuel.2020.117984.

[166] M. Mesbah, S. Soltanali, M. Momeni, S. Pouresmaeil, N. Rahaei, Z. Amiri, Effective modeling methods to accurately predict the miscibility of CO2 in ionic liquids, Chem. Eng. Res. Des. 154 (2020) 262–272. https://doi.org/10.1016/j.cherd.2019.12.017.

[167] H. Moosanezhad-Kermani, F. Rezaei, A. Hemmati-Sarapardeh, S.S. Band, A. Mosavi, Modeling of carbon dioxide solubility in ionic liquids based on group method of data handling, Eng. Appl. Comput. Fluid Mech. 15 (2021) 23–42. https://doi.org/10.1080/19942060.2020.1842250.

[168] Y. Zhang, X. Xu, Solubility predictions through LSBoost for supercritical carbon dioxide in ionic liquids, New J. Chem. 44 (2020) 20544–20567. https://doi.org/10.1039/d0nj03868g.

[169] P. Valeh-e-Sheyda, M. Faridi Masouleh, P. Zarei-Kia, Prediction of CO2 solubility in pyridinium-based ionic liquids implementing new descriptor-based chemoinformatics models, Fluid Phase Equilib. 546 (2021) 113136. https://doi.org/10.1016/j.fluid.2021.113136.

[170] B. Daryayehsalameh, M. Nabavi, B. Vaferi, Modeling of CO2 capture ability of [Bmim][BF4] ionic liquid using connectionist smart paradigms, Environ. Technol. Innov. 22 (2021) 101484. https://doi.org/10.1016/j.eti.2021.101484.

[171] H. Feng, P. Zhang, W. Qin, W. Wang, H. Wang, Estimation of solubility of acid gases in ionic liquids using different machine learning methods, J. Mol. Liq. 349 (2022) 118413. https://doi.org/10.1016/j.molliq.2021.118413.

[172] P. Valeh-e-Sheyda, P. Heidarian, A. Rezvani, A novel molecular structure-based model for prediction of CO2 equilibrium absorption in blended imidazolium-based ionic liquids, J. Mol. Liq. 360 (2022) 119420. https://doi.org/10.1016/j.molliq.2022.119420.

[173] R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Abedi, K. Jabbour, A. Hemmati-Sarapardeh, A. Mohaddespour, Chemical structure and thermodynamic properties based models for estimating nitrous oxide solubility in ionic Liquids: Equations of state and Machine learning approaches, J. Mol. Liq. 367 (2022) 120445. https://doi.org/10.1016/j.molliq.2022.120445.

[174] S.C. Balchandani, A. Dey, Prediction of CO2 solubility in potential blends of ionic liquids with Alkanolamines using statistical non-rigorous and ANN based modeling: A comprehensive simulation study for post combustion CO2 capture, Int. Commun. Heat Mass Transf. 132 (2022) 105866. https://doi.org/10.1016/j.icheatmasstransfer.2021.105866.

[175] E.N. Fierro, C.A. Faúndez, A.S. Muñoz, P.I. Cerda, Application of a Single Multilayer

Perceptron Model to Predict the Solubility of CO2 in Different Ionic Liquids for Gas Removal Processes, Processes. 10 (2022). https://doi.org/10.3390/pr10091686.

[176] M. Yazdani, E. Salehi, S. Zilabi, G. Nikravesh, An insight into the analogy between solute-solvent binding energy and solubility of acid gases in ionic liquids: Thermodynamic modeling versus molecular dynamic simulation, J. Chem. Thermodyn. 184 (2023) 107092. https://doi.org/10.1016/j.jct.2023.107092.

[177] Y. Wang, S. Huang, X. Liu, M. He, Thermodynamic model for CO2 absorption in imidazolium-based ionic liquids using cubic plus association equation of state, J. Mol. Liq. 378 (2023) 121587. https://doi.org/10.1016/j.molliq.2023.121587.

[178] Y. Tian, X. Wang, Y. Liu, W. Hu, Prediction of CO2 and N2 solubility in ionic liquids using a combination of ionic fragments contribution and machine learning methods, J. Mol. Liq. 383 (2023) 122066. https://doi.org/10.1016/j.molliq.2023.122066.

[179] S. Heidari, A. Aleosfoor, M. Mohammadi, F. Esmaeilzadeh, Solubility prediction of CO2 in ionic liquids under subcritical and supercritical carbon dioxide conditions by the adaptive neuro-fuzzy inference system, Case Stud. Chem. Environ. Eng. 7 (2023) 100317. https://doi.org/10.1016/j.cscee.2023.100317.

[180] B. Bazooyar, F. Shaahmadi, A. Jomekian, S.S. Mirfasihi, Carbon capture via aqueous ionic liquids intelligent modelling, Case Stud. Chem. Environ. Eng. 8 (2023) 100444. https://doi.org/10.1016/j.cscee.2023.100444.

[181] M. Loreno, R.A. Reis, S. Mattedi, M.L.L. Paredes, Statistical analysis of parameterization strategies for GC-sPC-SAFT equation of state to predict carbon dioxide solubility in imidazolium-based ionic liquids, Fluid Phase Equilib. 574 (2023). https://doi.org/10.1016/j.fluid.2023.113899.

[182] C. Touzet, HAL Id : hal-01338010 INTRODUCTION AU Claude TOUZET Juillet 1992, Les Reseaux Neurones Artif. Introd. Au Connexionnisme. (2016) 130.

[183] A. IBRIR, Contribution À L'Étude En Trois Dimensions (3D) De La Dispersion D'Un Polluant Gazeux Dans Un Milieu Urbain, (2022).

[184] Amalia Yunia Rahmawati, 済無No Title No Title No Title, (2020) 1–23.

[185] C. SI-MOUSSA, Modélisation des équilibres de phases a pressions élevées par les réseaux de neurones artificiels, These Dr. (2003).

[186] R. Zemouri, C. Devalland, S. Valmary-Degano, N. Zerhouni, Neural network: A future in pathology?, Ann. Pathol. 39 (2019) 119–129. https://doi.org/10.1016/j.annpat.2019.01.004.

[187] N.H. Farhat, Photonit neural networks and learning mathines the role of electron-trapping materials, IEEE Expert. Syst. Their Appl. 7 (1992) 63–72. https://doi.org/10.1109/64.163674.

[188] J. Wang, H. Du, H. Liu, X. Yao, Z. Hu, B. Fan, Prediction of surface tension for common compounds based on novel methods using heuristic method and support vector machine, Talanta. 73 (2007) 147–156. https://doi.org/10.1016/j.talanta.2007.03.037.

[189] J. Kharroubi, L. Likforman-sulem, a vecteurs supports " pour la v ´ erification automatique du locuteur Thèse Présentée pour obtenir le grade de docteur de l ' Ecole nationale supérieure des télécommunica- tions Jamal Kharroubi Etude de Techniques de Classement Machines à Vecteurs Support, (2005).

[190] M. Hasan, F. Boris, SVM : Machines à Vecteurs de Support ou Séparateurs à Vastes Marges, Report, January. (2006) 2006.

[191] A. Cornuéjols, Une nouvelle méthode d'apprentissage : les SVM. Séparateurs à vaste marge., Bull. l'AFIA. 51 (2002) 14–23.

[192] Q.-S.X. Yizeng Liang, Support Vector, 2007.

[193] H. Moussa, F. Dahmoune, M. Hentabli, H. Remini, L. Mouni, Optimization of ultrasound-assisted extraction of phenolic-saponin content from Carthamus caeruleus L. rhizome and predictive model based on support vector regression optimized by dragonfly algorithm, Chemom. Intell. Lab. Syst. 222 (2022). https://doi.org/10.1016/j.chemolab.2022.104493.

[194] R. Eberhart, J. Kennedy, New optimizer using particle swarm theory, Proc. Int. Symp. Micro Mach. Hum. Sci. (1995) 39–43. https://doi.org/10.1109/mhs.1995.494215.

[195] H. Garg, A hybrid PSO-GA algorithm for constrained optimization problems, Appl. Math. Comput. 274 (2016) 292–305. https://doi.org/10.1016/j.amc.2015.11.001.

[196] M.A. Ahmadi, Z. Chen, Comparison of machine learning methods for estimating permeability and porosity of oil reservoirs via petro-physical logs, Petroleum. 5 (2019) 271–284. https://doi.org/10.1016/j.petlm.2018.06.002.

[197] S. Mirjalili, Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems, Neural Comput. Appl. 27 (2016) 1053–1073. https://doi.org/10.1007/s00521-015-1920-1.

[198] L.L. Li, X. Zhao, M.L. Tseng, R.R. Tan, Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm, J. Clean. Prod. 242 (2020) 118447. https://doi.org/10.1016/j.jclepro.2019.118447.

[199] J.A.K. SUYKENS and J. VANDEWALLE, Least squares support vector machine classifiers using PCNNs, 2008 IEEE Int. Conf. Cybern. Intell. Syst. CIS 2008. (2008) 290–295. https://doi.org/10.1109/ICCIS.2008.4670890.

[200] J. Suykens, LS-SVMlab: a MATLAB/C toolbox for Least Squares Support Vector Machines, (n.d.).

[201] A. Shabri, Suhartono, Prévision de débit à l'aide de machines à vecteurs de support en moindres carrés, Hydrol. Sci. J. 57 (2012) 1275–1293. https://doi.org/10.1080/02626667.2012.714468.

[202] R. Razavi, A. Bemani, A. Baghban, A.H. Mohammadi, S. Habibzadeh, An insight into the estimation of fatty acid methyl ester based biodiesel properties using a LSSVM model, Fuel. 243 (2019) 133–141. https://doi.org/10.1016/j.fuel.2019.01.077.

[203] O. Falyouna, O. Eljamal, I. Maamoun, A. Tahara, Y. Sugihara, Magnetic zeolite synthesis for efficient removal of cesium in a lab-scale continuous treatment system, J.

Colloid Interface Sci. 571 (2020) 66–79. https://doi.org/10.1016/j.jcis.2020.03.028.

[204] R. Todeschini, D. Ballabio, F. Grisoni, Beware of Unreliable Q2! A Comparative Study of Regression Metrics for Predictivity Assessment of QSAR Models, J. Chem. Inf. Model. 56 (2016) 1905–1913. https://doi.org/10.1021/acs.jcim.6b00277.

[205] K. Paduszyński, In Silico Calculation of Infinite Dilution Activity Coefficients of Molecular Solutes in Ionic Liquids: Critical Review of Current Methods and New Models Based on Three Machine Learning Algorithms, J. Chem. Inf. Model. 56 (2016) 1420–1437. https://doi.org/10.1021/acs.jcim.6b00166.

[206] M. Diedenhofen, A. Klamt, COSMO-RS as a tool for property prediction of IL mixtures-A review, Fluid Phase Equilib. 294 (2010) 31–38. https://doi.org/10.1016/j.fluid.2010.02.002.

[207] K. Paduszy, Kamil Paduszyński 1, (2017) 1–309.

[208] E. Mullins, R. Oldland, Y.A. Liu, S. Wang, S.I. Sandler, C.C. Chen, M. Zwolak, K.C. Seavey, Sigma-profile database for using COSMO-based thermodynamic methods, Ind. Eng. Chem. Res. 45 (2006) 4389–4415. https://doi.org/10.1021/ie060370h.

[209] E. Mullins, Y.A. Liu, A. Ghaderi, S.D. Fast, Sigma profile database for predicting solid solubility in pure and mixed solvent mixtures for organic pharmacological compounds with COSMO-based thermodynamic methods, Ind. Eng. Chem. Res. 47 (2008) 1707–1725. https://doi.org/10.1021/ie0711022.

[210] M. Gevrey, I. Dimopoulos, S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecol. Modell. 160 (2003) 249–264. https://doi.org/10.1016/S0304-3800(02)00257-0.

[211] J.D. Olden, M.K. Joy, R.G. Death, An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, Ecol. Modell. 178 (2004) 389–397. https://doi.org/10.1016/j.ecolmodel.2004.03.013.

[212] A. Baghban, A. Jalali, M. Shafiee, M.H. Ahmadi, K. wing Chau, Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids, Eng. Appl. Comput. Fluid Mech. 13 (2019) 26–39. https://doi.org/10.1080/19942060.2018.1542345.

[213] M. Hosseinzadeh, A. Hemmati-Sarapardeh, Toward a predictive model for estimating viscosity of ternary mixtures containing ionic liquids, J. Mol. Liq. 200 (2014) 340–348. https://doi.org/10.1016/j.molliq.2014.10.033.

[214] C. Hansch, A Quantitative Approach to Biochemical Structure-Activity Relationships, Acc. Chem. Res. 2 (1969) 232–239. https://doi.org/10.1021/ar50020a002.

[215] X. Zhao, Y. Pan, J. Jiang, S. Xu, J. Jiang, L. Ding, Thermal Hazard of Ionic Liquids: Modeling Thermal Decomposition Temperatures of Imidazolium Ionic Liquids via QSPR Method, Ind. Eng. Chem. Res. 56 (2017) 4185–4195. https://doi.org/10.1021/acs.iecr.6b04762.

[216] N. Bouarra, S. Kherouf, A. Bouakkadia, D. Messadi, Research Journal of Pharmaceutical, Biological and Chemical Sciences QSPR Application on Modeling of Boiling Point of Polycyclic Aromatic Hydrocarbons, Res. J. Pharm. Biol. Chem. Sci. 8

(2017) 19–28.

[217] J. Han, C. Dai, G. Yu, Z. Lei, Parameterization of COSMO-RS model for ionic liquids, Green Energy Environ. 3 (2018) 247–265. https://doi.org/10.1016/j.gee.2018.01.001.

[218] Q. Sohaib, M.A. Kazemi, C. Charmette, J. Cartier, M. Younas, A. Azarafza, M. Rezakazemi, J. Sanchez-Marcano, CO2 solubility and diffusivity in 1-ethyl-3-methylimidazolium cation-based ionic liquids; isochoric pressure drop approach, Fluid Phase Equilib. 563 (2022). https://doi.org/10.1016/j.fluid.2022.113581.

[219] N.N. Ren, Y.H. Gong, Y.Z. Lu, H. Meng, C.X. Li, Surface tension measurements for seven imidazolium-based dialkylphosphate ionic liquids and their binary mixtures with water (methanol or Ethanol) at 298.15 K and 1 atm, J. Chem. Eng. Data. 59 (2014) 189–196. https://doi.org/10.1021/je400004j.

[220] J.W. Russo, M.M. Hoffmann, Measurements of surface tension and chemical shift on several binary mixtures of water and ionic liquids and their comparison for assessing aggregation, J. Chem. Eng. Data. 56 (2011) 3703–3710. https://doi.org/10.1021/je200659c.

[221] E. Rilo, J. Pico, S. García-Garabal, L.M. Varela, O. Cabeza, Density and surface tension in binary mixtures of CnMIM-BF4 ionic liquids with water and ethanol, Fluid Phase Equilib. 285 (2009) 83–89. https://doi.org/10.1016/j.fluid.2009.07.010.

[222] A. Shojaeian, Surface tension measurements of aqueous 1-alkyle-3-methylimidazolume tetrafluoroborate {[C n mim][BF 4] (n = 2, 4, 6)} solutions and modeling surface tension of ionic liquid binary mixtures using six various models, Thermochim. Acta. 673 (2019) 119–128. https://doi.org/10.1016/j.tca.2019.01.018.

[223] J.S. Torrecilla, T. Rafione, J. García, F. Rodrígue, Effect of relative humidity of air on density, apparent molar volume, viscosity, surface tension, and water content of 1-ethyl-3-methylimidazolium ethylsulfate ionic liquid, J. Chem. Eng. Data. 53 (2008) 923–928. https://doi.org/10.1021/je700523b.

[224] J.Y. Wang, X.J. Zhang, Y.Q. Hu, G. Di Qi, L.Y. Liang, Properties of n-butylpyridinium nitrate ionic liquid and its binary mixtures with water, J. Chem. Thermodyn. 45 (2012) 43–47. https://doi.org/10.1016/j.jct.2011.09.003.

[225] U. Domańska, A. Pobudkowska, M. Rogalski, Surface tension of binary mixtures of imidazolium and ammonium based ionic liquids with alcohols, or water: Cation, anion effect, J. Colloid Interface Sci. 322 (2008) 342–350. https://doi.org/10.1016/j.jcis.2008.02.039.

[226] H.F.D. Almeida, J.A. Lopes-Da-Silva, M.G. Freire, J.A.P. Coutinho, Surface tension and refractive index of pure and water-saturated tetradecyltrihexylphosphonium-based ionic liquids, J. Chem. Thermodyn. 57 (2013) 372–379. https://doi.org/10.1016/j.jct.2012.09.004.

[227] J.Y. Wang, F.Y. Zhao, Y.M. Liu, X.L. Wang, Y.Q. Hu, Thermophysical properties of pure 1-ethyl-3-methylimidazolium methylsulphate and its binary mixtures with alcohols, Fluid Phase Equilib. 305 (2011) 114–120. https://doi.org/10.1016/j.fluid.2011.03.008.

[228] E. Rilo, M. Domínguez-Pérez, J. Vila, L.M. Varela, O. Cabeza, Surface tension of four

binary systems containing (1-ethyl-3-methyl imidazolium alkyl sulphate ionic liquid + water or + ethanol), J. Chem. Thermodyn. 49 (2012) 165–171. https://doi.org/10.1016/j.jct.2012.01.023.

[229] M. Geppert-Rybczyńska, J.K. Lehmann, A. Heintz, Surface tensions and the gibbs excess surface concentration of binary mixtures of the ionic liquid 1-ethyl-3-methylimidazolium bis[(trifluoromethyl) sulfonyl]imide with tetrahydrofuran and acetonitrile, J. Chem. Eng. Data. 56 (2011) 1443–1448. https://doi.org/10.1021/je101146f.

[230] M. Geppert-Rybczyńska, J.K. Lehmann, J. Safarov, A. Heintz, Thermodynamic surface properties of [BMIm][NTf2] or [EMIm][NTf2] binary mixtures with tetrahydrofuran, acetonitrile or dimethylsulfoxide, J. Chem. Thermodyn. 62 (2013) 104–110. https://doi.org/10.1016/j.jct.2013.02.021.

[231] A. Wandschneider, J.K. Lehmann, A. Heintz, Surface tension and density of pure ionic liquids and some binary mixtures with 1-propanol and 1-butanol, J. Chem. Eng. Data. 53 (2008) 596–599. https://doi.org/10.1021/je700621d.

[232] U. Domańska, M. Zawadzki, A. Lewandrowska, Effect of temperature and composition on the density, viscosity, surface tension, and thermodynamic properties of binary mixtures of N- octylisoquinolinium bis{(trifluoromethyl)sulfonyl}imide with alcohols, J. Chem. Thermodyn. 48 (2012) 101–111. https://doi.org/10.1016/j.jct.2011.12.003.

Annexe A

TABLEAU 1 : Source et domaines de données expérimentales de la Tension Superficielle des mélanges binaires étudiés ainsi que Descripteurs moléculaires ($S_{\sigma iC}$) des cations, ($S_{\sigma iA}$) des anions et ($S_{\sigma iS}$) des solvants.

Solvent	LI	T (K)	X_{LI}	σ_m^{exp} (mN/m)	*Sσ1C*	*Sσ2C*	*Sσ1A*	*Sσ2A*	*Sσ1S*	*Sσ2S*	N°	Ref
Eau	[C1MIm][DMP]	298,15	[0,0000 – 1,0000]	[48,4-71,9]	2,8591	11,5085	9,0433	4,8582	1,2513	3,0754	16	[219]
	[C1MIm][MeSO4]	[296,8 –298,1]	[0,0000 – 0,2920]	[64,7-71,1]	2,8591	11,5085	8,8817	2,9627	1,2513	3,0754	10	[220]
	[C2MIm][BF4]	[298,15-338,15]	[0,0000 – 1,0000]	[50,45-70,1]	2,6790	13,5263	9,2925	0,1052	1,2513	3,0754	53	[221,222]
	[C2MIm][DEP]	298,15	[0,0000 – 1,0000]	[37,1-71,9]	2,6790	13,5263	13,9877	4,9520	1,2513	3,0754	16	[219]
	[C2MIm][DMP]	298,15	[0,0000 –1,0000]	[44,1-71,9]	2,6790	13,5263	9,0433	4,8582	1,2513	3,0754	11	[219]
	[C2MIm][EtSO4]	298,15	[0,0062 – 0,5791]	[47,00-50,11]	2,6790	13,5263	10,4886	3,1491	1,2513	3,0754	12	[223]
	[C2MIm][MeSO3]	[300,2 –303,2]	[0,0000 – 0,4850]	[57,30-71,60]	2,6790	13,5263	5,6481	5,12535	1,2513	3,0754	27	[220]
	[C4MIm] [AC]	[298,15-338,15]	[0,8533-1,0000]	[32,20-38,70]	2,6230	17,6003	4,7504	4,3228	1,2513	3,0754	49	[114]
	[C4MIm] [CL]	[298,1-302,1]	[0,0000 –0,4800]	[53,70-71,90]	2,6230	17,6003	-0,0001	5,2676	1,2513	3,0754	46	[114]
	[C4MIm] [PF6]	303,15	[0,9268-0,9969]	[40,95-43,27]	2,62295	17,6003	11,4203	1,37E-08	1,2513	3,0754	9	[114]
	[C4MIm][BF4]	[298,15-338,15]	[0,0000 – 1,0000]	[41,55-70,10]	2,6230	17,60025	9,2925	0,1052	1,2513	3,0754	59	[221,222]
	[C4MIm][DBP]	298,15	[0,0000 –1,0000]	[28,20-71,90]	2,6230	17,60025	21,3317	4,9239	1,2513	3,0754	11	[219]

	[C4MIm][DMP]	298,15	[0,0000 –1,0000]	[38,00-71,90]	2,6230	17,6003	9,0433	4,8582	1,2513	3,0754	11	[219]
	[C4Py] [BF4]	[293,15-323,15]	[0,0000 –1,0000]	[45,10-72,70]	3,1653	16,5872	9,2925	0,1052	1,2513	3,0754	72	[114]
	[C4Py][NO3]	298,15	[0,0000 – 0,9903]	[37,50-51,50]	3,1653	16,5872	4,8726	2,8813	1,2513	3,0754	15	[224]
	[C6MIm] [AC]	[298,15-338,15]	[0,8597-1,0000]	[29,80-35,70]	2,6477	21,611	4,75044	4,3229	1,2513	3,0754	49	[114]
	[C6MIm] [CL]	[297,2-298,2]	[0,0000 –0,143]	[43,20-70,90]	2,6477	21,611	-0,0001	5,2675525	1,2513	3,0754	10	[220]
	[C6MIm][BF4]	[298,15-338,15]	[0, 3 –1,0000]	[34,52-37,93]	2,6477	21,6110	9,2925	0,1052	1,2513	3,0754	40	[221,222]
	[C8MIm] [PF6]	[298,05-335,05]	0,7827	[31,70-34,01]	2,5481	25,5567	11,4203	1,372E-08	1,2513	3,0754	5	[114]
	[N112(hoe)][Br]	298,15	[0,9915 – 0,9981]	[63,98-66,97]	2,5773	15,3787	0,0009	6,8408	1,2513	3,0754	3	[225]
	[N311(hoe)][Br]	298,15	[0,9900 –0,9979]	[47,72-51,99]	2,5773	15,3787	0,0009	6,8408	1,2513	3,0754	3	[225]
	[P666(14)][BTI]	[298,1 -343,3]	0,0891	[26,88-30,05]	0,9115	66,0228	20,5680	0,0376	1,2513	3,0754	6	[226]
	[P666(14)][Dca]	[298,2 – 342,8]	0,4932	[29,50-32,56]	0,9115	66,0228	7,6290	2,5549	1,2513	3,0754	6	[226]
Methanol	[C1MIm][DMP]	298,15	[0,0000 – 1,0000]	[22,40-48,40]	2,8591	11,50845	9,0433	4,8582	0,5449	6,2453	11	[219]
	[C1MIm][MeSO4]	298,15	[0,0000 – 1,0000]	[22,51-59,80]	2,8591	11,50845	8,8817	2,9631	0,5449	6,2453	9	[225]
	[C2MIm] [AC]	[278,2-318,15]	[0,0000 –1,0000]	[21,00-48,10]	2,6790	13,5263	4,7504	4,3228	0,5449	6,2453	55	[114]
	[C2MIm][DEP]	298,15	[0,0000 –1,0000]	[22,40-44,30]	2,6790	13,5263	13,9877	4,9520	0,5449	6,2453	22	[219]
	[C2MIm][EtSO4]	298,15	[0,0000 –1,0000]	[22,40-37,30]	2,679	13,5263	10,4886	3,1491	0,5449	6,2453	11	[114]
	[C2MIm][MeSO4]	[293,15-298,15]	[0,0000 –1,0000]	[21,80-52,62]	2,6790	13,5263	8,8817	2,9631	0,5449	6,2453	21	[114]

												[227]
	[C4MIm][BTI]	[283,15-298,15]	[0,1049-0,9577]	[25,52-33,44]	2,6230	17,6003	20,568	0,0376	0,5449	6,2453	24	[114]
	[C4MIm][DBP]	298,15	[0,0000 –1,0000]	[22,40-29,50]	2,6230	17,6003	21,3317	4,9239	0,5449	6,2453	11	[219]
	[C4MIm][DMP]	298 ,15	[0,0000 –1,0000]	[22,40-38,10]	2,6230	17,6003	9,0433	4,8582	0,5449	6,2453	11	[219]
	[C4Py] [BF4]	[293,15-323,15]	[0,0000 –1,0000]	[20,05-46,90]	3,1653	16,5872	9,2925	0,1052	0,5449	6,2453	56	[114]
	[C8MIm] [BTI]	[283,15-298,15]	[0,0978-0,9368]	[25,05-31,62]	2,5481	25,5567	20,568	0,0376	0,5449	6,2453	20	[114]
Ethanol	[C1MIm][DMP]	298,15	[0,0000 –1,0000]	[23,20-48,40]	2,8591	11,50845	9,0433	4,8582	0,4694	8,3713	11	[219]
	[C2MIm] [AC]	[278,15-338,15]	[0,0000 –1,0000]	[18,90-48,10]	2,679	13,5263	4,75044	4,3228	0,4694	8,3713	77	[114]
	[C2MIm][C8SO4]	298,15	[0,0399 –1,0000]	[22,89-30,91]	2,6790	13,5263	22,6460	3,1221	0,4694	8,3713	18	[228]
	[C2MIm][DEP]	298,15	[0,0000 – 1,0000]	[22,60-37,10]	2,6790	13,5263	13,9877	4,9520	0,4694	8,3713	11	[219]
	[C2MIm][DMP]	298,15	[0,0000 – 1,0000]	[22,90-44,30]	2,6790	13,5263	9,0433	4,8582	0,4694	8,3713	11	[219]
	[C2MIm][EtSO4]	298,15	[0,0000 –1,0000]	[22,01-47,16]	2,6790	13,5263	10,4886	3,1491	0,4694	8,3713	11	[114]
	[C2MIm][MeSO4]	293,15	[0,0000 –1,0000]	[22,00-52,62]	2,6790	13,5263	8,8817	2,9631	0,4694	8,3713	14	[114]
	[C4MIm][BF4]	298,15	[0,0986 –1,0000]	[22,83-45,33]	2,6230	17,6003	9,2925	0,1052	0,4694	8,3713	12	[222]
	[C4MIm][BTI]	[283,15-313,15]	[0,0999-0,9601]	[22,66-33,42]	2,62295	17,6003	20,568	0,0376	0,4694	8,3713	36	[114]
	[C4MIm][DBP]	298,15	[0,0000 –1,0000]	[22,50-29,40]	2,6230	17,60025	21,3317	4,9239	0,4694	8,3713	11	[219]
	[C4MIm][DMP]	298,15	[0,0000 –1,0000]	[22,80-38,00]	2,6230	17,60025	9,0433	4,8582	0,4694	8,3713	11	[219]

	[C6MIm] [EtSO4]	298,15	[0,03 –1,0000]	[22,98-34,77]	2,6477	21,611	10,4886	3,1491	0,4694	8,3713	13	[114]
	[C6MIm][BF4]	298,15	[0,0980 –1,0000]	[23,11-37,33]	2,6477	21,6110	9,2925	0,1052	0,4694	8,3713	10	[222]
	[C8MIm] [BTI]	[283,15-313,15]	[0,1042-0,9784]	[22,93-31,96]	2,5481	25,5567	20,568	0,0376	0,4694	8,3713	39	[114]
	[C8MIm][BF4]	298,15	[0,1012 –1,0000]	[23,33-32,82]	2,5481	25,5567	9,2925	0,1052	0,4694	8,3713	10	[222]
Tetrahydrofuran	[C2MIm][BTI]	[293,15 – 308,15]	[0,0000– 1,0000]	[25,38-36,14]	2,6790	13,5263	20,5680	0,0376	0,0000	11,2621	44	[229]
	[C4MIm][BTI]	[293,15 – 308,15	[0,0000 – 1 ,0000]	[25,38-32,72]	2,6230	17,6003	20,5680	0,0376	0,0000	11,2621	40	[230]
Acetonitrile	[C2MIm][BTI]	[293,15 –313,15]	[0,0000 –1,0000]	[26,43-36,14]	2,6790	13,5263	20,5680	0,0376	0,0143	8,3215	45	[229]
	[C4MIm][BTI]	[293,15 –313,15]	[0,0000 – 1,0000]	[26,43-32,78]	2,6230	17,60025	20,5680	0,0376	0,0143	8,3215	45	[230]
	[PYR-4,1] [BTI]	[288,15-308,15]	[0,0000 –1,0000]	[27,15-33,71]	1,30988	19,1008	20,568	0,0376	0,0143	8,3215	27	[114]
Dimethyl sulfoxide	[C2MIm][BTI]	[293,15 – 313,15]	[0,0000 – 1,0000]	[35,10-43,36]	2,6790	13,5263	20,5680	0,0376	0,0055	11,1500	45	[230]
	[C4MIm][BTI]	[293,15 – 313,15]	[0,0000 –1,0000]	[31,88-43,36]	2,6230	17,60025	20,5680	0,0376	0,0055	11,1500	50	[230]
1-propanol	[C2MIm] [AC]	[288,15-348,15]	[0,0000 –1,0000]	[19,00-47,70]	2,6790	13,5263	4,75044	4,3228	0,4678	10,371	77	[114]
	[C4MIm][BTI]	[283,15-313,15]	[0,0000 – 1,0000]	[23,11-33,23]	2,6230	17,60025	20,5680	0,0376	0,4678	10,3710	31	[114] [231]
	[C8MIm] [BTI]	[283,15-313,15]	[0,0974-0,9461	[23,28-31,72]	2,5481	25,5567	20,568	0,0376	0,4678	10,371	33	[114]

1-butanol	[C4MIm][BTI]	[283,15-313,15]	[0,0000 – 1,0000]	[23,52-33,36]	2,6230	17,60025	20,5680	0,0376	0,4658	12,3460	29	[114] [231]
	[C8MIm] [BTI]	[298,15 – 318,15]	[0,0000 –1,0000]	[22,62-42,92]	2,5481	25,5567	20,568	0,0376	0,4658	12,346	48	[114]
1-hexanol	[C8iQuin]BTI]	[298,15 – 318,15]	[0,0000 –1,0000]	[23,48-31,87]	2,2507	29,5226	20,5680	0,0376	0,4668	16,3093	24	[232]

Annexe B

Tableau 1 : domaines de données expérimentales pour la solubilité du CO2 dans les IL (en fraction molaire) et les descriptions moléculaires calculés.

N°	Systèmes IL-CO2	T (K)	P (bar)	E2m	MATS8s	TDB6i	TDB1s	ATSC4v	MATS8m	ATSC7v	GATS2s	GATS5s	GATS5c	ATSC6v	De	LOBMAX	Y^{exp}
1	[(ETO)2IM][Tf2N]-CO2	302,85-363,15	21,5-328,6	0,4024	0	947,4359	4,0920	-144,3067	0	0	1,0927	0,7791	0,96932	79,9395	1,36041	1,71354	0,3960-0,8120
2	[aMIM][Tf2N] -CO2	313,15-333,15	0,1-19,998	0,1292	0	946,59706	3,7938	-144,3067	0	0	1,0927	0,7791	0,96932	79,9395	1,2786	1,6933	0,0021-0,3150
3	[BBIM][MeSO3] -CO2	296,95-343,35	15-342	0,4802	0,0332	903,6565	2,922	47,2920	-0,0417	-113,6206	1,013	1,1468	1,43992	-642,3482	1,5688	1,3993	0,1389-0,7738
4	[BMIM][BETA] -CO2	298,25-313,65	0,4624-0,495	0,2326	0,6170	891,0686	3,1642	39,7780	0,0290	20,1596	1,1891	1,3840	1,3312	-25,9742	1,4015	1,6844	0,0118-0,0143
5	[BMIM][BF4] -CO2	278,47-383,15	0,00991-467,2	0,1351	-0,1049	995,3503	3,4050	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,2577	2,6472	0,0001-0,6100
6	[BMIM][C(CN)3] -CO2	303,15-373,15	4-347	0,2268	-0,1049	895,0499	3,3886	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,5445	1,6844	0,157-0,633
7	[BMIM][C3F7CO2] -CO2	303,15-333,15	5,5-42,7	0,2326	0,0000	891,0686	3,1642	0,8391	0,0000	0,0000	1,3715	0,0517	0,0309	0,0000	1,4015	1,6844	0,0734-0,5458
8	[BMIM][Cl] -CO2	353,15-373,15	24,54-369,47	0,2362	-0,1049	886,1836	3,1711	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,3828	1,6844	0,1306-0,406
9	[BMIM][DBPO4] -CO2	313,15-333,15	0,2234-1,754	0,1367	0,0832	810,0497	3,1500	92,0530	-0,0352	-383,1557	0,3441	2,0662	1,8254	-522,1309	1,2790	1,6668	0,0032-0,0353
10	[BMIM][DCA] -CO2	293,36-373,15	8,2-467,49	0,2326	-0,1049	891,0686	3,1642	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,4015	1,6844	0,1543-0,601
11	[BMIM][MDEGSO4] -CO2	313,15-363,15	14,3-264,8	0,2326	-0,1487	891,0686	3,1642	-320,6866	0,0662	-164,7411	0,3694	0,9305	0,8255	-83,8271	1,4015	1,6844	0,0932-0,6835

12	[BMIM][MeSO3] -CO2	302,35-343,15	5,5-357	0,1661	-0,1049	932,3439	3,6233	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,2093	1,8351	0,0517-0,5219
13	[BMIM][MeSO4] -CO2	293,2413,1	9,08-98,5	0,2025	-0,1049	977,9900	3,5489	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,3873	1,3508	0,0319-0,4527
14	[BMIM][methide] -CO2	298,2-333,3	12,58-114,27	0,1466	0,0000	883,5428	3,1972	-457,7498	0,0000	0,0000	1,1063	0,7602	0,9683	280,3523	1,2326	1,8134	0,2770-0,7740
15	[BMIM][NfO] -CO2	293,15-343,15	12,21-82,87	0,1290	0,0000	978,8366	4,2971	4,6424	0,0000	0,0000	1,3447	0,7799	0,7649	28,4430	1,4085	1,7785	0,1373-0,9000
16	[BMIM][NO3] -CO2	293,13-368,24	3,68-128,32	0,1466	-0,1049	883,5428	3,1972	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,2326	1,8134	0,0382-0,5300
17	[BMIM][PF6] -CO2	282,05-393,15	0,03-439	0,1351	-0,1049	995,3503	3,4050	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,2577	2,6472	0,0004-0,7290
18	[BMIM][SCN] -CO2	292,35-384,15	10,5-315	0,2326	-0,1049	891,0686	3,1642	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,4015	1,6844	0,1260-0,4300
19	[BMIM][TDfO] -CO2	298,2-323,2	19,1-141,1	0,1466	0,2831	883,5428	3,1972	3,5967	-0,1197	-24,9306	1,3894	0,8748	0,8774	3,2453	1,2326	1,8134	0,1043-0,9516
20	[BMIM][Tf2N] -CO2	273,15-453,15	0,0102-499,9	0,1466	0,0000	883,5428	3,1972	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2326	1,8134	0,0007-0,8032
21	[BMIM][TFA] -CO2	293,25-363,18	0,105-484,58	0,1466	-0,1049	883,5428	3,1972	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,2326	1,8134	0,0050-06790
22	[BMIM][TfO] -CO2	294,85-344,55	0,4835-375	0,1237	-0,1049	886,6767	3,1488	5,8032	0,0232	-20,4592	0,9191	0,9655	0,9325	-297,2848	1,4933	1,6287	0,0089-0,6720
23	[BMMIM][Tf2N] -CO2	298,15-343,15	0,099-18,997	0,2158	0,0000	974,2289	4,1341	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3533	2,2000	0,0020-0,3820
24	[BMPYR][MeSO4] -CO2	303,15-373,15	30,7-458,5	0,2158	0,4286	974,2289	4,1341	-128,2351	0,4979	186,3433	0,9494	0,6004	0,7239	-458,8964	1,3533	2,2000	0,2871-0,5319
25	[BMPYR][Tf2N] -CO2	283,15-413,20	0,0124-484	0,1464	0,0000	948,8429	3,3212	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3808	2,5173	0,0001-0,8530
26	[BMPYR][TfO] -CO2	303,15-373,35	18,8-486,5	0,2074	0,4286	851,2216	2,4903	-128,2351	0,4979	186,3433	0,9494	0,6004	0,7239	-458,8964	1,3279	1,9002	0,2583-0,7058

27	[C12MIM][BF4] -CO2	333,15-373,15	9,4-57,5	0,1077	-0,0251	985,5130	2,8418	-7,4500	0,0030	13,6197	0,9735	0,8238	0,7116	14,4682	1,3637	4,3583	0,0895-0,4586
28	[C12MIM][PF6] -CO2	333,15-373,15	9,,9-57,9	0,1077	-0,0251	985,5130	2,8418	-7,4500	0,0030	13,6197	0,9735	0,8238	0,7116	14,4682	1,3637	4,3583	0,0809-0,4472
29	[C12MIM][Tf2N] -CO2	333,15-373,15	7,6-56,7	0,1571	-0,0412	947,9642	2,7923	-5,2862	0,0033	14,7557	0,9553	0,8411	0,7226	15,8574	1,2892	2,3269	0,0870-0,5118
30	[C3MPYR][FSA] -CO2	298,15-343,15	0,099-18,999	0,2337	0,0000	0,0000	2,7371	-56,6232	0,0000	278,4343	0,9155	1,1535	0,1950	182,1667	1,6956	2,0091	0,0020-0,3460
31	[C3MPYR][Tf2N] -CO2	298,15-373,15	5,2-471	0,1540	0,0000	893,3774	2,5164	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3870	1,3885	0,0930-0,7870
32	[C3OMIM][C(CN)3] -CO2	298,15-353,25	0,93-19,93	0,0932	-0,2464	1044,1070	4,4091	-166,8969	-0,3233	147,5831	1,4815	0,9562	0,9936	-289,6721	1,1259	2,4777	0,0082-0,3106
33	[C3OMIM][PF6] -CO2	298,18-343,15	10,62-51,85	0,1160	-0,2464	1006,0599	3,9549	-166,8969	-0,3233	147,5831	1,4815	0,9562	0,9936	-289,6721	1,1670	1,4234	0,0998-0,4499
34	[C3OMIM][Tf2N] -CO2	298,15-343,15	0,93-46,12	0,3177	0,0000	938,3105	4,3892	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3118	2,0873	0,0181-0,5994
35	[C4MPY][BF4] -CO2	296,9-313,4	0,5104-0,5437	0,2131	-0,2465	870,1748	3,2876	38,9384	-0,1026	-80,0062	0,8630	1,0880	0,8965	-202,9480	1,3814	1,0802	0,0098-0,0118
36	[C4PY][BF4] -CO2	33,15-333,15	15,47-95,8	0,2117	0,0374	838,6019	3,3841	-114,7844	0,1523	-77,0461	1,0254	0,9332	0,9333	-314,4860	1,4105	1,2503	0,142-0,581
37	[C5MIM][Tf2N] -CO2	293,3-363,29	6,18-479,93	0,1462	0,0000	923,6605	3,0997	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2219	2,0244	0,2120-0,8020
38	[C5MPYR][Tf2N] -CO2	298,15-373,15	0,097-484,2	0,1605	0,0000	900,3970	2,4689	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3589	1,6228	0,0020-0,7850
39	[C5O2MIM][C(CN)3] -CO2	298,15-353,25	0,94-19,93	0,1600	0,0785	929,2018	3,6430	-316,5959	0,0102	-8,8574	1,6111	1,2208	1,1928	-134,6934	1,2348	2,0106	0,0083-0,3367
40	[C6F9MIM][Tf2N] -CO2	298,13-333,30	0,0157-92,3	0,3193	-0,5483	1137,6889	6,2353	23,4620	-0,5083	249,5215	0,6366	0,7658	0,7433	160,8203	1,5983	2,0453	0,0002-0,8530
41	[C7MIM][C(CN)3] -CO2	298,15-353,25	0,94-19,93	0,1787	-0,0106	939,7540	2,9304	-0,8846	-0,0148	16,5714	0,9305	0,8995	0,7634	18,3337	1,3354	2,0703	0,0104-0,3497

42	[C7MPYR][Tf2N] -CO2	303,15-373,15	2,6-476	0,1744	0,0000	926,5586	3,5103	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4034	3,2330	0,3020-0,8530
43	[C8F13MIM][Tf2N] -CO2	298,16-333,05	0,00798-36,2	0,3567	-0,5659	1183,6541	6,8639	150,3211	-0,6487	107,3767	0,7309	0,7638	0,9255	163,7333	1,5031	2,3517	0,0001-0,6550
44	[C9MPYR][Tf2N] -CO2	303,15-373,15	2,6-365,8	0,1472	0,0024	925,8016	2,4119	-118,8545	0,0009	-7,2681	1,0732	0,7678	0,6700	-7,1578	1,3380	2,2836	0,3230-0,8200
45	[COCMPYR][FSA] -CO2	298,15-333,15	5,924-49,171	0,3573	0,0000	1089,0440	1,5442	353,9234	0,0000	296,1147	1,3551	1,4095	0,6937	19,6245	1,5530	2,2109	0,0480-0,5690
46	[COCMPYR][Tf2N] -CO2	298,15-333,15	5,286-52,432	0,5131	0,0000	1199,8454	5,5099	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4233	2,1937	0,1030-0,6310
47	[deme][Tf2N] -CO2	313,15-333,15	0,0985-19,0044	0,1228	0,0000	897,7562	3,1955	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2601	2,1238	0,0023-0,3202
48	[DMIM][Tf2N] -CO2	298,15-343,15	1-201,5	0,1524	-0,0412	976,5675	3,5702	-5,2862	0,0033	14,7557	0,9553	0,8411	0,7226	15,8574	1,3405	2,7121	0,0260-0,8780
49	[EMIM][BETA] -CO2	313,15	5,01-60	0,5778	0,6170	1394,0073	8,0739	39,7780	0,0290	20,1596	1,1891	1,3840	1,3312	-25,9742	2,0057	2,6677	0,1120-0,6330
50	[EMIM][BF4] -CO2	298,15-343,2	2,51-150	0,1346	0,0000	1059,5164	3,7680	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,1727	2,3281	0,0156-0,6127
51	[EMIM][C(CN)3] -CO2	298,15-373,15	0,94-227,4	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024	0,0072-0,5530
52	[EMIM][DCA] -CO2	303,15-373,15	1,48-191,4	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024	0,0086-0,4180
53	[EMIM][DEPO4] -CO2	298,4-333,3	0,2379-71,2	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024	0,0031-0,6500
54	[EMIM][EtSO4] -CO2	298,04-353,15	1,22-94,61	0,1454	0,0000	916,0859	3,4798	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2243	1,7433	0,0080-0,4570
55	[EMIM][HSO4] -CO2	313,3-333,6	14,4-97	0,1216	0,0000	0,0000	4,8327	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,1670	2,4284	0,2000-0,6900
56	[EMIM][MDEGSO4] -CO2	303,2-343,2	8,42-67,1	0,1583	-0,1487	981,0762	3,5376	-320,6866	0,0662	-164,7411	0,3694	0,9305	0,8255	-83,8271	1,5981	3,3318	0,0608-0,4493

57	[EMIM][MeSO3] -CO2	298,4-342,85	4-461,5	0,1388	0,0000	894,4389	3,7602	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,3353	1,7433	0,0500-0,4518
58	[EMIM][MeSO4] -CO2	299,5-333	11,3-79,3	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024	0,0300-0,5100
59	[EMIM][NfO] -CO2	313,15	5-59,94	0,1922	0,0000	910,9727	3,7150	4,6424	0,0000	0,0000	1,3447	0,7799	0,7649	28,4430	1,2486	1,8601	0,0092-0,5920
60	[EMIM][PF6] -CO2	308,14-366,03	14,9-497,6	0,1346	0,0000	1059,5164	3,7680	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,1727	2,3281	0,1040-0,4590
61	[EMIM][SCN] -CO2	298,6-373,15	10,9-228,6	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024	0,0800-0,3640
62	[EMIM][Tf2N] -CO2	283,43-453,15	0,1-478,5	0,1958	0,0000	939,2526	3,4368	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2584	1,9024	0,0040-0,7820
63	[EMIM][TFA] -CO2	298,1-348,1	0,1-19,996	9,999	0,1958	0,0000	939,2526	3,4368	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,2584	1,9024
64	[EMIM][TfO] -CO2	303,15-353,15	1,362-378	0,1333	0,0000	991,5896	4,0714	-311,5624	0,0000	302,0570	0,9060	0,8524	0,8085	-18,1129	1,3916	2,3583	0,0139-0,6268
65	[he][FOR] -CO2	303-323	4,4-100,1	0,0721	0,0000	0,0000	4,0109	48,0957	0,0000	0,0000	0,9882	0,0000	0,1708	0,0000	1,5223	1,6727	0,0271-0,3083
66	[hea][FOR] -CO2	303-323	6,6-85,2	0,0956	0,2376	985,0725	3,8303	0,1246	0,5655	85,2738	1,1187	1,1523	0,8730	-0,8647	1,6089	2,6348	0,0202-0,1907
67	[HMIM][BF4] -CO2	293,18-373,15	3,12-489,8	0,1289	-0,1741	1004,9864	3,1824	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3165	2,9779	0,0710-0,6530
68	[HMIM][C(CN)3] -CO2	298,15-373,15	0,94-245,3	0,1371	-0,1741	930,1789	3,0206	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,2596	2,0328	0,0094-0,6290
69	[HMIM][DCA] -CO2	303,15-373,15	6,8-192,6	0,1959	-0,1741	935,1928	2,9937	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3647	1,8574	0,2090-0,5490
70	[HMIM][MeSO4] -CO2	303,15-373,15	8,7-461,9	0,1959	-0,1741	935,1928	2,9937	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3647	1,8574	0,1580-0,6020
71	[HMIM][PF6] -CO2	243,2-373,15	2,2-474,7	0,1289	-0,1741	1004,9864	3,1824	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3165	2,9779	0,0580-0,6910

72	[HMIM][SCN] -CO2	298,15-373,15	4,5-226,9	0,1959	-0,1741	935,1928	2,9937	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3647	1,8574	0,0500-0,4570
73	[HMIM][TCB] -CO2	283,56-364,04	2,71-123,36	0,1289	-0,1741	1004,9864	3,1824	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3165	2,9779	0,1009-0,7016
74	[HMIM][Tf2N] -CO2	273,15-413,2	0,00826-452,8	0,1959	0,0000	935,1928	2,9937	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3647	1,8574	0,0001-0,8333
75	[HMIM][TfO] -CO2	298,10-373,15	10,4-454,7	0,1959	-0,1741	935,1928	2,9937	1,0229	-0,1623	91,8513	0,9242	0,9366	0,7882	19,2351	1,3647	1,8574	0,0500-0,7600
76	[HMMIM][Tf2N] -CO2	298,2-333,3	14,79-118,04	0,1446	0,0000	920,2294	2,9898	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2311	1,8745	0,1991-0,7447
77	[HMPY][Tf2N] -CO2	283,18-323,15	0,0105-13	0,2127	0,0000	910,8883	2,9598	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4405	1,2184	0,0001-0,3760
78	[HMPYR][Tf2N] -CO2	303,15-373,15	10,6-475,5	0,1657	0,0000	908,8981	2,4511	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3612	1,7440	0,2778-0,8105
79	[HOC2MIM][BF4] -CO2	303,15-353,15	1,14-11,94	0,1152	0,2963	1035,2680	4,2691	-90,9495	0,7990	138,0102	0,9572	1,3721	0,9008	17,4014	1,2823	1,8643	0,0040-0,1020
80	[HOC2MIM][PF6] -CO2	303,15-353,15	1,333-11,27	0,1152	0,2963	1035,2680	4,2691	-90,9495	0,7990	138,0102	0,9572	1,3721	0,9008	17,4014	1,2823	1,8643	0,0060-0,1040
81	[HOC2MIM][Tf2N] -CO2	303,15-353,15	0,973-11,143	0,1211	0,0000	931,7533	3,9707	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3580	1,4841	0,0160-0,1910
82	[HOC2MIM][TFA] -CO2	299,2-333,6	11,4-82,3	0,1184	0,2963	954,8212	3,9298	-90,9495	0,7990	138,0102	0,9572	1,3721	0,9008	17,4014	1,5551	2,2111	0,0900-0,6300
83	[HOC2MIM][TfO] -CO2	303,15-353,15	1,012-12,8	0,1211	0,2963	931,7533	3,9707	-90,9495	0,7990	138,0102	0,9572	1,3721	0,9008	17,4014	1,3580	1,4841	0,0130-0,1880
84	[HOC3MIM][NO3] -CO2	293,47-363,55	8,51-116,21	-0,4563	976,9806	3,7205	42,7823	0,0062	6,0451	0,9284	0,6044	0,7939	-85,6714	1,6222	2,0464		0,0528-0,3081
85	[HOC3MIM][Tf2N] -CO2	298,15-323,15	0,93-19,93	0,1927	0,0000	1063,6531	4,6615	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,6171	1,9016	0,0145-0,3577
86	[HOEPY][Tf2N] -CO2	298,15-333,15	6,801-50,444	0,5374	0,0000	939,8822	4,3759	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4096	1,1107	0,0940-0,5890

87	[m-2-HEA][FOR] -CO2	293,21-363,42	4,94-485,11	0,0854	0,0000	917,0458	3,5390	-23,8532	0,0000	0,0000	0,9581	1,9082	1,0018	54,9055	1,6220	2,0335	0,0570-0,5340
88	[MDEA][Cl] -CO2	3113,15-333,15	12,2-46,3	0,1055	0,1932	945,3579	3,6981	-144,5059	0,5545	46,5232	0,9314	1,6722	1,1078	148,2540	1,3403	1,9920	0,4852-0,8924
89	[MMIM][DMPO4] -CO2	3113,15-333,15	0,4881-1,7464	0,1533	0,0000	954,3315	3,7132	0,1038	0,0000	0,0000	0,8466	0,9524	0,0513	317,7378	1,3805	1,9266	0,0035-0,0156
90	[MMIM][MeSO3] -CO2	332,85-363,65	16,5-431	0,1395	0,0000	1040,5044	3,6628	0,1038	0,0000	0,0000	0,8466	0,9524	0,0513	317,7378	1,3890	2,4758	0,0629-0,3690
91	[N1,1,1,2-OH][Tf2N] -CO2	333,24	12-83,6	0,2600	0,0000	971,6034	4,7892	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3142	1,8566	0,1560-0,6060
92	[N1,1,3,2-OH][Tf2N] -CO2	304,16-344,84	0,467-0,5439	0,1830	0,0000	1107,8713	4,3150	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4949	2,1708	0,0122-0,0238
93	[N1,4,4,4][Tf2N] -CO2	298,15	0,0473-5,5	0,1347	0,0000	922,2023	2,7282	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4131	1,7570	0,0004-0,1140
94	[N1,8,8,8][Tf2N] -CO2	298,2-323,2	2,98-205,6	0,1629	-0,0040	907,3817	2,6698	0,1282	0,0003	0,1282	1,1359	0,9115	1,2169	0,1282	1,0460	2,3229	0,1191-0,9070
95	[N1,8,8,8][TfO] -CO2	329,2-369,2	37-150	0,2365	-0,0040	911,7291	2,3650	0,1282	0,0003	0,1282	1,1359	0,9115	1,2169	0,1282	1,1114	1,7558	0,5700-0,8000
96	[N1,8,8,8][TOS] -CO2	329,2-369,2	55-150	0,1437	-0,0040	884,7908	2,4685	0,1282	0,0003	0,1282	1,1359	0,9115	1,2169	0,1282	0,9773	1,9283	0,6600-0,7500
97	[N2,1,1,3][FSA] -CO2	298,15-343,15	0,099-19,007	0,1284	0,0000	940,2170	2,6013	52,4675	0,0000	0,0000	0,4485	0,0000	0,0000	0,0000	1,3285	1,8205	0,0020-0,3540
98	[N2,1,1,3][Tf2N] -CO2	313,15-333,15	0,098-94,66	0,1284	0,0000	940,2170	2,6013	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3285	1,8205	0,0018-0,7390
99	[N4,1,1,1][Tf2N] -CO2	282,93-343,07	0,06-203,7	0,3440	0,0000	1066,2137	3,9608	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,6168	3,1955	0,0012-0,8790
100	[N4,4,4,4][BF4] -CO2	369,2	150	0,2818	0,1051	823,7854	2,4223	46,9527	-0,0087	-135,4444	1,0847	0,6706	1,5793	-1734,2909	1,3567	1,5373	0,5700
101	[N6,1,1,1][Tf2N] -CO2	296,45-313,65	0,48-0,5142	0,1303	0,0000	939,7543	2,5338	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4473	2,2810	0,0131-0,0159

102	[NMIM][PF6] -CO2	293,15-298,15	8,6-35,4	0,1650	-0,0544	963,4444	2,9629	-4,0008	0,0033	15,3590	0,9464	0,8548	0,7359	16,6322	1,3463	2,3569	0,1860-0,5540
103	[oleyl-MIM][Tf2N] -CO2	298,2-323,2	0,249-20,005	0,2351	-0,0251	862,0446	3,7936	-16,0109	-0,0010	3,0995	1,0712	0,9039	0,7482	4,4360	1,2171	1,8944	0,0105-0,4700
104	[OMIM][BF4] -CO2	243,2-363,29	2,37-496,4	0,1214	-0,0426	995,1086	3,0321	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3508	3,4445	0,1005-0,7498
105	[OMIM][C(CN)3] -CO2	298,15-353,15	0,94-19,93	0,1726	-0,0426	943,1749	2,8771	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3201	1,9883	0,0096-0,3567
106	[OMIM][NO3] -CO2	298,15-333,15	2,96-37,17	0,1726	-0,0426	943,1749	2,8771	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3201	1,9883	0,0475-0,4108
107	[OMIM][PF6] -CO2	303,15-353,15	1,287-92,88	0,1214	-0,0426	995,1086	3,0321	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3508	3,4445	0,0231-0,7550
108	[OMIM][SCN] -CO2	298,15-373,15	7-207,2	0,1269	-0,0426	948,0595	3,0210	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,2880	2,3441	0,0686-0,3810
109	[OMIM][Tf2N] -CO2	297,55-353,15	1-348	0,1726	-0,0426	943,1749	2,8771	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3201	1,9883	0,0250-0,8456
110	[OMIM][TfO] -CO2	303,85-344,55	6,8-340	0,1427	-0,0426	941,5759	2,8892	-2,5450	0,0031	15,9718	0,9381	0,8734	0,7448	17,4599	1,3250	2,3413	0,2166-0,7414
111	[OMPYR][Tf2N] -CO2	303,15-373,15	7,4-359,2	0,1564	0,0025	921,3745	2,4231	-120,1954	0,0011	-7,8350	1,0531	0,7582	0,6658	-7,7084	1,3476	2,0988	0,2409-0,8176
112	[P1,4,4,4][Tf2N] -CO2	298,15-343,15	0,099-18,999	0,3635	0,0000	939,2768	4,0396	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3630	1,5470	0,0030-0,3930
113	[P1,4,4,4][TOS] -CO2	323,15	0,0111-13	0,2358	0,2598	909,6274	2,5532	1654,6826	0,3797	-2699,4167	0,7183	1,7585	1,3400	1916,5338	1,2431	1,6870	0,0001-0,1210
114	[P2,4,4,4][DEPO4] -CO2	298,8-333,7	11,3-84,1	0,3750	0,1267	865,2192	3,3417	-419,6781	0,1087	-67,9411	1,3292	0,7097	1,0710	-1347,9877	1,3910	1,3754	0,1000-0,7100
115	[P4,4,4,4][FOR] -CO2	298,0-298,2	0,1-19,996	0,1951	0,0848	862,6545	3,1252	28,4508	0,0152	-59,8005	1,2854	0,6603	1,1644	-1777,4464	1,2541	1,5950	0,0080-0,3480

116	[P6,6,6,14][C12PhSO3] -CO2	303,15-363,15	3,9-113,7	0,4001	-0,0494	831,9263	2,6175	14,9196	-0,0597	-158,1565	1,3489	0,9332	1,0462	15,0532	1,4008	2,1172	0,0971-0,7435
117	[P6,6,6,14][Cl] -CO2	302,55-363,68	1,68-245,7	0,3461	-0,0494	823,7456	2,4917	14,9196	-0,0597	-158,1565	1,3489	0,9332	1,0462	15,0532	1,3619	1,3600	0,1190-0,8240
118	[P6,6,6,14][Tf2N] -CO2	292,88-375,35	0,4846-493,5	0,2101	-0,0494	861,7187	3,1376	14,9196	-0,0597	-158,1565	1,3489	0,9332	1,0462	15,0532	1,3508	1,4629	0,0115-0,8790
119	[PMIM][Tf2N] -CO2	313,15-333,15	0,099-19,003	0,2146	0,0000	920,2061	3,2827	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,2738	1,9495	0,0023-0,3160
120	[PMPip][FSA] -CO2	298,15-343,15	0,099-18,999	0,2084	0,4286	874,9057	2,4615	-196,9453	0,4979	128,2430	0,9494	1,1785	0,2243	326,8769	1,3127	1,4270	0,0020-0,3440
121	[PMPY][DCA] -CO2	298	0,06-9,94	0,2124	0,6500	878,1693	3,2134	112,9287	0,7108	-139,7043	0,8641	1,2708	0,2523	42,3391	1,3441	1,1401	0,0009-0,1218
122	[PMPY][Tf2N] -CO2	313,15-333,15	0,096-19,008	0,1673	0,0000	915,1239	3,2000	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,3165	1,4147	0,0022-0,3170
123	[S222][Tf2N] -CO2	313,15-333,15	0,095-19,006	0,2153	0,0000	846,8013	1,5130	-144,3067	0,0000	0,0000	1,0927	0,7791	0,9693	79,9395	1,4004	1,6096	0,0020-0,3033
124	[thma][MeSO4] -CO2	313,2-363,28	6,79-76,29	0,0848	0,2029	963,3987	4,1177	-450,9535	0,5916	165,1155	0,9237	1,5739	1,1496	303,4528	1,6677	2,7184	0,0300-0,1140