REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

UNIVERSITE YAHIA FARES DE MEDEA



Faculté de **T**echnologie

Mémoire de fin d'études de Master

Filière : Génie des procédés
Spécialité : Génie chimique

Modélisation de la tension superficielle des mélanges binaires de liquides ioniques par les méthodes d'intelligence computationnelle

---

Présenté par :

BENMOULOUD Widad

TOUBAL Besma

Proposé et Dirigé par :

Dr. SI-MOUSSA Chérif

Promotion 2018/2019

Remerciement

*A*près avoir rendu grâce ALLAH le tous puissant de nous avoir donné la force et le courage de mener à bien ce modeste travail, nous tenons à remercier vivement tous ce qui, de prés ou de loin, il s'agit particulièrement :

*N*otre promoteur **DR CHERIF SI-MOUSSA** pour son encadrement, sa disponibilité, sa rigueur scientifique et son sens d'écoute et d'échange

*N*ous tenons à exprimer nos plus vifs remerciements au président du jury le professeur Salah HANINI et à messieurs les membres de jury pour avoir accepté de discuter ce mémoire : monsieur le Dr. M. LAIDI.

*N*ous adressons aussi notre plus vive reconnaissance à tous nos enseignants du département des génies des procédés pour la formation qu'ils nous ont donné pendant nos études universitaires.

*N*ous remercions tous ceux qui ont contribué de près ou de loin, directement ou indirectement à la réalisation de ce mémoire.

*N*ous remercions vivement les étudiants de master génie chimique et toute la promotion master 2019 pour leur aide morale durant toute la période de préparation.

DÉDICACE :

Avant tout, je tiens à remercier DIEU le tout puissant de m'avoir aidé à réaliser se modeste travail J'ai l'immense plaisir de dédier se travail à:

Ceux que j'aime le plus au monde, mes très chers et affectueux parents qui m'encouragent et me poussent toujours vers la réussite, que DIEU les garde et les protège

Spécial dédicace à mon cher mari

À mes sœurs, Mes chers frères

Et a toute la famille TOUBAL et CHIKHI

Mon binôme Widad

Mes copines Asma et Lamia

À mon promoteur DR CHERIF SI-MOUSSA

Je remercie également tous mes ami(e)s, les camarades les étudiant(e)s que je n'ai pas citée et tous ceux qui, de prés ou de loin ont contribue à l'élaboration de ce travaille

J'espère que notre mémoire servira de support pour les années à venir.

TOUBAL Besma

DÉDICACE :

Avant tout, je tiens à remercier ALLAH le tout puissant de m'avoir aidé à réaliser ce modeste travail. Je dédie ce travail à qui n'aura jamais pu voir le jour sans les soutiens indéfectibles et sans limite de mes chers parents qui ne cessent de me donner avec amour le nécessaire pour que je puisse arriver à ce que je suis aujourd'hui. Que dieux vous protège et que la réussite soit toujours à ma portée pour que je puisse vous combler de bonheur.

Une spéciale dédicace à mon cher frère Abdelhak

À toute ma famille BENMOULOUD

À ma copine Meriem et Soumia et à tous mes amies et mes collègues.

À mon binôme Besma

À mes enseignants et enseignantes et spécialement à mon promoteur

Dr. CHERIF SI-MOUSSA

À tous les étudiants et étudiantes du génie chimique la promotion 2018-2019.

BENMOULOUD Widad

الخلاصة - يلعب التوتر السطحي للسوائل الأيونية (LI) ومخاليطها مع المركبات الأخرى دورًا مهمًا في تصميم وتطوير العديد من العمليات الصناعية. لذلك فإن تطوير نماذج تنبئية مهم للغاية من الناحية الصناعية. بحثت هذه الدراسة قدرة وجدوى أربعة أساليب الذكاء الحسابي (RNA، ANFIS، SVM و LSSVM) للتنبؤ التوتر السطحي للأنظمة الثنائية التي تحتوي على السوائل الأيونية. من أجل بناء واختبار النماذج ، تم جمع 672 نقطة بيانات تقابل قيم التوتر السطحي التجريبية للمزائج الثنائية المحتوية على السوائل الأيونية من التراث العلمي المنشور . كان التوتر السطحي بين 21.8 و 71.9 mNm^{-1}. درجة الحرارة (T) ،تركيب السائل الأيوني (X_{IL}) ، واثنين من الواصفات على أساس ملامح سيغما، واحدة تتعلق بالصفة المانحة لرابطة H ، والأخرى المتعلقة بصفة المتلقي لرابطة H، لكل من الأيون الموجب والأيون السالب والمذيب تم استخدام كل المعاملات كمتغيرات إدخال للنموذج للتمييز بين المركبات المختلفة المشاركة في الأنظمة الثنائية. أظهرت مقارنة بين البيانات التجريبية والقيم المتوقعة باستخدام الطرق الأربعة (من حيث العديد من المقاييس الإحصائية) اتفاق جيد. ومع ذلك ، كان التنبؤ من قبل الشبكات العصبونية الإصطناعية أفضل من الطرق الثلاثة الأخرى. كشف التشخيص بطرق إحصائية عن وجودعدد من المعطيات التجريبية المشبوهة ، كما سمح هذا التشخيص بتحديد مجال قابلية تطبيق النماذج المحسنة.

الكلمات المفتاحية: السوائل الأيونية ، التوتر السطحي ، الذكاء الحسابي ، RNA، ANFIS، SVM، LSSVM.

Résumé - La tension superficielle des liquides ioniques (LI) et leurs mélanges avec d'autres composés jouent un rôle important dans la conception et le développement de nombreux processus industriels. Par conséquent, sa modélisation est extrêmement importante d'un point de vue industriel. Cette étude a examiné la capacité et la faisabilité de quatre méthodes d'intelligence computationnelle (RNA, ANFIS, SVM et LSSVM) pour prédire la tension superficielle de systèmes binaires contenant des LI. Pour construire et tester les modèles, 672 points de données correspondant aux valeurs expérimentales de tension superficielle de mélanges binaires contenant des LI ont été rassemblées de la littérature. La tension superficielle était comprise entre 21.8 et 71.9 $mN.m^{-1}$. La température (T), la composition du LI (X_{IL}), 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 utilisés comme variables d'entrée du modèle pour différencier les divers composés impliqués dans les systèmes binaires. Une comparaison de données expérimentales et de valeurs prédites en utilisant les quatre méthodes (en termes plusieurs métriques statistiques) a montré un bon accord. Cependant, la prédiction par l'RNA était meilleure que les trois autres approches. Un diagnostic par des méthodes statistiques a révélé un certain nombre de données aberrantes, permettant ainsi d'établir le domaine d'applicabilité des modèles optimisés.

Mots clés: Liquides ioniques, tension superficielle, intelligence computationnelle, RNA, ANFIS, SVM, LSSVM.

Abstract - The surface tension of ionic liquids (IL) and their mixtures with other compounds play an important role in the design and development of many industrial processes. Therefore, its modeling is extremely important from an industrial point of view. This study examined the capacity and feasibility of four computational intelligence methods (RNA, ANFIS, SVM and LSSVM) to predict the surface tension of binary systems containing IL. To construct and test the models, 672 data points corresponding to experimental surface tension values of IL-containing binary mixtures were collected from the literature. The surface tension was between 21.8 and 71.9 mNm^{-1}. The temperature (T), the composition of the IL (X_{IL}), two descriptors based on the sigma profiles, one relating to the donor character of H-bond, the other relating to the character of acceptor of H-bond, of the anion of the cation and the solvent were used as input variables of the model to differentiate the various compounds involved in the binary systems. A comparison (in terms of several statistical metrics) of experimental data and predicted values using the four methods showed a good agreement. However, prediction by ANN was better than the other three approaches. Diagnosis by statistical methods revealed a number of outliers, which enabled the establishing of the domain of applicability of the optimized models.

Key words: Ionic liquids, surface tension, computational intelligence, ANN, ANFIS, SVM, LSSVM.

Table des matières

Table des matières ….…… i
Liste des figures…….…….....…..iv
Liste des tableaux……...……..…..viii
Abréviations ……..……..…....ix
Nomenclature et symbole……..…….....x
Introduction générale…….……..……1

Chapitre I : Généralités sur les liquides ioniques

I.1 Introduction..... 5
I.2 Définition des liquides ioniques 5
I.3 Structure des liquides ioniques 6
I.4 Historique..... 10
I.5 Synthèses 11
I.5.1 Réaction de quaternisation du noyau imidazole..... 12
I.5.2 Réaction d'échange de l'anion..... 13
I.5.2.1 Résine échangeuse d'ions..... 13
I.5.2.2 Traitement acide..... 13
I.5.2.3. Métathèse anionique 14
I.5.3 Purification et pureté des LI..... 14
I.5.4 Nomenclature et acronymes des LI..... 15
I.6 Propriétés des liquides ioniques..... 16
I.6.1 La tension superficielle..... 16
I.6.2 Point de fusion 17
I.6.3 Température de décomposition..... 18
I.6.4 Viscosité..... 19
I.6.5 Densité 20
I.6.6 Volatilité 21
I.6.7 Toxicité et biodégradabilité 21
I.6.7.1 La toxicité 21
I.6.7.2 La biodégradabilité 22
I.6.8 Solubilité..... 23
I.6.8.1 Solubilité dans l'eau et dans les solvants organiques..... 23
I.6.8 Non inflammabilité 23
I.6.9 La masse volumique..... 23

I.7 Utilisation des liquides ioniques 24
I.7.1 Application dans l'industrie chimique 25
I.7.1.1 Application en électrochimie 25
I.7.1.2 Chimie analytique 26
I.7.1.3 Synthèse et catalyse 27
I.7.2 Applications dans le domaine des procèdes de séparation et de l'analyse 28
I.7.2.1 Extraction liquide –liquide 28
I.7.3 L'utilisation des LI comme solvants de réaction catalytique 29
I.7.3.1 Leur intérêt en catalyse 30
I.8 Conclusion 30

Chapitre II : Méthodes d'Intelligence Computationnelle

II.1 Introduction 32
II.2 Réseaux de neurones artificiels 33
II.2.1 Définition 33
II.2.2 Du neurone biologique au neurone artificiel 33
II.2.3 Concept du réseau de neurones de type feedforward 36
II.2.4 Fonctions d'activation 37
II.2.5 Sélection d'un réseau de neurones artificiel 39
II.2.5.1 Sélection du type de réseau 39
II.2.5.2 Sélection de la structure du réseau 40
II.2.5.3 Sélection de la fonction d'activation 41
II.2.6 Avantages des réseaux de neurones 42
II.3 Système d'Inférence Neuro-Flou Adaptatif (ANFIS) 42
II.3.1 Architecture d'ANFIS : 43
II.4 Machines à vecteurs de support (Support Vector Machines) (SVM) 45
II.4.1 Historique de SVM 46
II.4.2 C'est quoi un support à vecteur machine ? 46
II.4.3 SVM : Principe de fonctionnement général 47
II.4.3.1 Notions de base : hyperplan, marge et support vecteur 47
II.4.3.2 Linéarité et non linéarité 50
II.4.3.3 Cas non linéaire 50
II.4.3.4 Illustration de transformation de cas non linéaire : le cas XOR 51
II.4.4 Fondement mathématique 52

II.4.4.1 Problème d'apprentissage 52
II.4.4.2 Classification a valeurs réelles 53
II.5 Les machines à vecteurs avec support des moindres carrés (LSSVM) 56
II.5.1 Introduction 56
II.5.2 Définition 56
II.6 Critères statistiques : 59
II.7 Conclusion 62

Chapitre III : Résultats et interprétations

III.1 Introduction 64
III.2 Elaboration de la base de données 64
III.2.1 Données expérimentales de la tension superficielle 64
III.2.2 Données des descripteurs de profil sigma 67
III.3 La modélisation par les méthodes d'intelligence computationnelle 70
III.3.1 Choix des variables d'entrée et de sortie 70
III.3.2 Division des données en ensemble d'apprentissage, test et validation 71
III.3.3 Critères statistiques : 71
III.3.4 Modélisation par RNA 71
III.3.4.1 Développement du modèle 71
III.3.4.2 Analyse de sensibilité 80
III.3.5 Modélisation par ANFIS 83
III.3.5.1 Développement de modèle 83
III.3.6 Modélisation par SVM 88
III.3.7 Modélisation par LSSVM 91
III.4 Analyse de sensibilité des modèles ANFIS, SVM et LSSVM 95
III.5 Domaine d'applicabilité 96
III.6 Comparaison des résultats des modèles optimisés 98
III.7 Conclusion 101
Conclusion générale 102
Références bibliographiques 104

Liste des figures

Figure	Titre de figure	Page
Introduction générale		
Figure I.1	Superposition du nombre total de traités de coopération en matière de brevets (PCT) et d'article de revues publiés par année relatifs aux liquides ioniques	2
Chapitre I		
Figure I.1	définition des liquides ioniques et quelques exemples d'anions et de cations	6
Figure I.2	cation des LIS	7
Figure I.3	Échange d'un ion chlorure avec un carboxylate	13
Figure I.4	Échange d'un chlorure avec un ion métallique	14
Figure I.5	Échange d'anion par métathèse	14
Figure I.6	Cation immidazolium	15
Figure I.7	Données de tension superficielle pour les liquides ioniques purs, de 2001 à 2010	17
Figure I.8	les domaines d'application des liquides	25
Chapitre II		
Figure II.1	Neurone biologique.	34
Figure II.2	Schéma d'un neurone artificiel	35
Figure II.3	Schéma de comparaison entre un neurone biologique et un neurone artificiel	35
Figure II.4	structure de Réseau de neurones artificiels de type feedforward multicouches	36
Figure II.5	(a) Un modèle flou Takagi-Sugeno de premier ordre avec deux entrées et deux règles. (b) L'architecture ANFIS.	44
Figure II.6	principe des techniques SVM	47
Figure II.7	Exemple d'un hyperplan séparateur	47
Figure II.8	Exemple de vecteurs de support	48
Figure II.9	Exemple de marge maximal (hyperplan valide)	49
Figure II.10	L'Hyperplan avec faible marge et le Meilleur hyperplan séparateur	49
Figure II.11	(a) cas linéairement séparable et (b) cas non linéairement séparable	50
Figure II.12	Exemple de changement de l'espace de données	51
Figure II.13	Illustration de cas non linéairement séparable (le cas XOR)	51
Figure II.14	Illustration de passage d'un espace 2D à un espace 3D	52
Figure II.15	Organigramme des étapes pour la création du modèle SVM.	55
Figure II.16	Démonstration schématique de l'algorithme LSSVM formé	58
Chapitre III		
figure III.1	Criblage de la densité de charge (sigma) et structures des conformères minimaux de 1-butanol, [C2mim] et [NTf2].	68
Figure III.2	Profils sigma des conformères minimum de 1-butanol, [C2mim] et [NTf2].	68
Figure III.3	descripteur moléculaire basé sur les Profils σ	69

Figure III.4	Diagramme de dispersion de la tension superficielle calcule en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble du test par le modèle RNA « trainbr, logsig »	73
Figure III.5	Diagramme de dispersion de la tension superficielle calculer en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle RNA « trainlm, logsig »	74
Figure III.6	Diagramme de dispersion de la tension superficielle calculer en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle RNA « trainlm, tansig »	75
Figure III.7	L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainbr, logsig »	76
Figure III.8	L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainlm, logsig »	76
Figure III.9	L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainlm, tansig »	76
Figure III.10	Diagrammes de distribution d'ER[%] du modèle RNA RNA « trainbr, logsig »	77
Figure III.11	Diagrammes de distribution d'ER [%] du modèle RNA RNA « trainlm, logsig »	77
Figure III.12	Diagrammes de distribution d'ER [%] du modèle RNA RNA « trainlm, tansig »	77
Figure III.13	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm] [BTI]-Dimethyl sulfoxide du modèle de RNA « trainbr,logsig »	78
Figure III.14	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,logsig »	78
Figure III.15	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,tansig »	79
Figure III.16	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle de RNA « trainbr,logsig »	79
Figure III.17	Isothermes de la variation de tension superficielle en	79

	fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,logsig »	
Figure III.18	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,tansig »	80
Figure III.19	L'importance relative des variables par la méthode de Garson du modèle RNA « trainbr,logsig »	81
Figure III.20	L'importance relative des variables par la méthode de Garson du modèle RNA « trainlm,logsig »	81
Figure III.21	L'importance relative des variables par la méthode de Garson du modèle RNA « trainlm,tansig »	81
Figure III.22	L'importance relative des variables par la méthode de con-weight du modèle RNA « trainbr,logsig »	82
Figure III.23	L'importance relative des variables par la méthode de con-weight du modèle RNA « trainlm,logsig »	82
Figure III.24	L'importance relative des variables par la méthode de con-weight du modèle RNA « trainlm,tansig »	82
Figure III.25	Diagramme de dispersion de la tension superficielle calculé en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle ANFIS	85
Figure III.26	L'erreur relative (%) de la tension superficielle par le modèle ANFIS	86
Figure III.27	Diagrammes de distribution d'ER [%]du modèle ANFIS	86
Figure III.28	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle ANFIS	87
Figure III.29	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du ANFIS	87
Figure III.30	Diagramme de dispersion de la tension superficielle calcule en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble du test par le modèle SVM.	90
Figure III.31	L'erreur relative (%) de la tension superficielle par le modèle de SVM	90
Figure III.32	Diagrammes de distribution d'ER [%]du modèle SVM	91
Figure III.33	diagramme de dispersion de la tension superficielle calculé en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble	93

	du test par le modèle LSSVM.	
Figure III.34	L'erreur relative (%) de la tension superficielle par le modèle LSSVM	94
Figure III.35	Diagramme de distribution d'ER[%]du modèle LSSVM.	94
Figure III.36	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle LSSVM	95
Figure III.37	Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2Mim][BTI]-Dimethyl sulfoxide du modèle LSSVM	95
Figure III.38	Résultats de l'analyse de sensibilité réalisée sur les modèles ANFIS, SVM et LSSVM	96
Figure III.39	Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.	98
Figure III.40	comparaison de l'ERAM	99
Figure III.41	Figure de distribution de l'ER des ensembles d'apprentissage	100
Figure III.42	Figure de distribution de l'ER des ensembles de test	100

Liste des tableaux

Tableau	Titre de tableau	Page
Chapitre I		
Tableau I.1	Exemple de structure de cations	8
Tableau I.2	Exemple d'anion de LIs	9
Tableau I.3	Exemple de structure d'anions	10
Tableau I.4	Valeurs de température de fusion de certains liquides ioniques	18
Tableau I.5	Valeurs de température de composition de quelques liquides ioniques	19
Tableau I.6	Viscosité de certains liquides ioniques	20
Tableau I.7	Valeurs de densité de quelques liquides ioniques	20
Tableau I.8	Toxicité des LIs sur les microorganismes souvent utilisés en domaine de biotechnologie	22
Tableau I.9	La biodégradabilité de quelques liquides ioniques	22
Chapitre II		
Tableau II.1	Fonctions de transfert usuelles, a=f(n)	38
Chapitre III		
Tableau III.1	La liste des cations	64
Tableau III.2	La liste des anions	65
Tableau III.3	Source et domaines des données expérimentales des mélanges binaires étudiés	65
Tableau III.4	Descripteurs moléculaires ($S_{\sigma i}$) des cations des Lis étudiés	69
Tableau III.5	Descripteurs moléculaires ($S_{\sigma i}$) des anions des Lis étudiés	70
Tableau III.6	Descripteurs moléculaires ($S_{\sigma i}$) des solvants	70
Tableau III.7	Critères statistiques des modèles du RNA	72
Tableau III.8	Différentes erreurs des modèles de RNA	72
Tableau III.9	Détails des paramètres d'apprentissage du système d'inférence floue (FIS)	83
Tableau III.10	Critères statistiques du modèle ANFIS	84
Tableau III.11	Différentes erreurs par le modèles ANFIS	84
Tableau III.12	Les paramètres optimisés du modèle SVM	88
Tableau III.13	Critères statistiques de modèle SVM	88
Tableau III.14	différentes erreurs du modèles SVM	89
Tableau III.15	Les paramètres d'optimisation du modèles LSSVM	91
Tableau III.16	Critères statistiques du modèles LSSVM	92
Tableau III.17	Différentes erreurs de modèles LSSVM	92
Tableau III.18	Comparaison des écarts entre l'ERAM% de l'ensemble de test et de l'ensemble d'apprentissage pour les différents modèles développés	101

Abréviations

Abréviation	Signification
ANFIS	Système adaptif a inférence neuro-flou
CO_2	Dioxyde de carbone
COSMO-RS	Conductor Screenlike Model for Real Solvents
Covs	Composés organiques Volatiles
EAM	Erreur absolue moyenne
EQM	Erreur quadratique moyenne
ERAM	erreur relative absolue moyenne
IC	Intelligence computationnelle
Li	Liquide ionique
Liap	Liquide ionique aprotique
Lip	Liquide ionique protique
LSSVM	Machine à vecteur de support aux moindres carrés
MF	Membership functions
MLP	Multi-layer perception
RBF	Fonction de base radiale
REQM	Racine de l'erreur quadratique moyenne
RNA	Réseaux de neurones artificiels
SVM	Machine à vecteur de support
UV	Ultra violet

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
w	poids	[-]
X_{LI}	Fraction molaire	[-]
γ	Constante de régularisation	[-]
σ	Tension superficielle	mN/m
σ^2	Paramètre kernel	[-]

Introduction générale

Introduction générale

L'industrie chimique s'est considérablement développée au cours du vingtième siècle, et plus particulièrement après la deuxième guerre mondiale. La chimie fait aujourd'hui partie de notre quotidien : l'essentiel des produits que nous consommons ou utilisons, mais elle provoque beaucoup des risques toxiques de pollutions au cours des processus de fabrication, le rejet des déchets dans l'environnement contribuent à détériorer notre milieu. Ce changement a un impact sur la société, la santé humaine, l'économie, les espèces vivantes, la production alimentaire, le tourisme et l'écologie.

La chimie verte ou la chimie durable prévoit la mise en œuvre de principes pour réduire et éliminer l'usage ou la génération de substances néfastes pour l'environnement. Jusqu'à présent, encadrés par la législation, les industriels cherchaient à minimiser l'exposition aux dangers en contrôlant les substances toxiques à tous les stades des procédés, Cependant, la plupart de ces produits chimiques représentent un risque pour la santé humaine et sont peu écologiques en raison de leur volatilité. Une bonne illustration du caractère destructeur des produits chimiques modernes est donnée, par exemple la catastrophe de l'explosion à la centrale nucléaire de windscale (Royaume-Uni) en 1957, l'explosion de la mine de Courrrières (France) ; 1200 morts en 1906, accident de l'usine chimique de Seveso(Italie) ; pas de morts mais intoxication très importante des populations des environs en 1976 et explosion de l'usine d'AZF à Toulouse (France) en 2001.

Les liquides ioniques (LIs) ont suscité un vif intérêt ces dernières années, ils constituent une nouvelle classe des liquides considérés comme des solvants verts qui conservent leur état liquide dans de larges plages de température. Les LIs ont attiré beaucoup d'attention de la communauté scientifique en raison de leur grande variété d'applications dans l'industrie et la chimie appliquée. La nature ionique de ces composés résulte en une combinaison unique de propriétés. Parmi ceux-ci, une pression de vapeur négligeable peut être la plus attrayante du point de vue de l'environnement et implique également un potentiel de recyclage élevé. La plupart des LIs présentent d'autres propriétés intéressantes telles qu'une large plage de liquides, une stabilité chimique et physique, une non inflammabilité, une conductivité thermique élevée et la capacité de dissoudre une large gamme de composés. De plus, les propriétés physiques, chimiques et biologiques ajustables obtenues par la combinaison du cation et de l'anion d'un liquide ionique génèrent un très grand nombre d'applications spécifiques pour ces sels.

Les liquides ioniques continuent de faire l'objet de nombreuses recherches dans un nombre croissant de plusieurs domaines comme par exemples, la synthèse (organiques et matériaux), les études et prévisions de propriétés physiques et les transformations enzymatiques …etc., ceci est traduit par la tendance de croissance exponentielle du nombre de demandes de brevets et d'articles de journaux sur les liquides ioniques durant les deux dernières décennies (Figure 1).

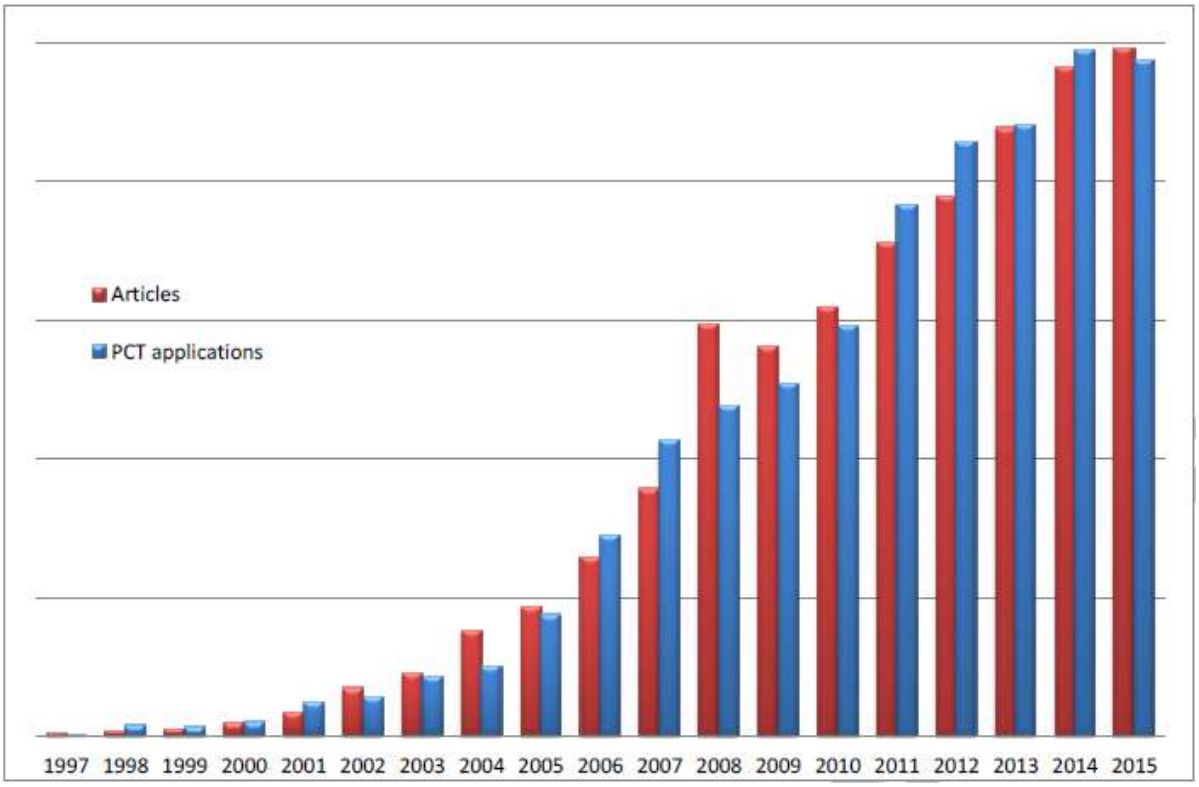


Figure1 : Superposition du nombre total de traités de coopération en matière de brevets (PCT) et d'articles de revues publiés par année relatifs aux liquides ioniques. *(Morton et Hamer ,2017)*

Pour concevoir tout processus impliquant des LIs à l'échelle industrielle, il est nécessaire de connaître une gamme de propriétés physiques, notamment la densité, la viscosité, la tension superficielle, etc. La tension superficielle d'un mélange liquide joue un rôle important dans le transfert la masse et de chaleur à l'interface. Il est nécessaire dans de nombreuses applications industrielles telles que l'extraction liquide-liquide, l'absorption de gaz, la distillation, la condensation, 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 thermophysiques prévues et des

équilibres de phase est un facteur clé dans l'évaluation des programmes de simulation et de conception. Malheureusement, ces données sont limitées et parfois contradictoires.

De plus, la tension superficielle révèle des informations 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. Dans la littérature, rares sont les travaux collectant des données de tension superficielle pour des mélanges de LIs et d'alcools. De plus, les interactions suivent des schémas très différents. L'agrégation apparaît dans plusieurs cas et de fortes interactions intermoléculaires conduisant à des variations continues se produisent dans plusieurs autres systèmes

Cependant, il est difficile de réaliser plusieurs des expériences à cause du nombre élevés de LIs (10^{18}) peuvent être obtenus par les différentes combinaisons cation-anion et comme les mesures expérimentales sont la plupart du temps difficiles, coûteuses et prennent beaucoup de temps, 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.

L'utilisation de méthodes alternatives et complémentaires à l'expérimentation, parmi les quelles 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. L'élaboration des modèles mathématiques QSPR/QSAR reliant les propriétés physicochimiques et les activités biologiques à la structure moléculaire permet, d'une part, d'expliquer l'origine de ces activités/propriétés et, d'autre part, de les prédire pour des molécules dont les données expérimentales ne sont pas disponibles.

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 telles les réseaux de neurones artificiels (RNA), les machines à vecteur de support (SVM), Machine à vecteurs de support pour les moindres carrés (LSSVM) et Système adaptatif d'inférence neuro- flou (ANFIS) 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.

L'objectif de ce travail est la collecte d'une base de données des mesures expérimentales de la tension superficielle des mélanges binaires (LI-Eau et divers solvants organiques) et le développent de modèles de prédiction et/ou de corrélation de cette propriété par l'utilisation des techniques d'intelligence computationnelle.

Ce mémoire est constitué de trois chapitres principaux :

- Un premier chapitre qui donne un aperçu sur les liquides ioniques.
- Un deuxième chapitre consacré aux outils et méthodes utilisées dans la modélisation de la tension superficielle des mélanges binaires solvant ionique par les méthodes d'apprentissage automatique
- Et un dernier chapitre présentant les résultats de la modélisation et interprétation de ces résultats
- Enfin, ce mémoire s'achève sur une conclusion générale et les perspectives envisageables de ce travail.

Chapitre I :

Généralités sur les liquides ioniques

Chapitre I : Généralités sur les liquides ioniques

I.1 Introduction

L'intérêt accordée aux liquides ioniques durant les dernières définies, due aux divers avantages que possèdent ces composés, est traduite par l'abondance des sources bibliographies (livres, sites web, bases de données, articles de recherches, brevets) relatives à leurs propriétés physico-chimiques, biologiques toxicologiques et leurs applications diverses. En plus les liquides ioniques sont de plus en plus étudiés dans les milieux académiques et industriels à travers les conférences et les colloques organisés à travers le monde pour exposer les résultats de recherche et de débattre les points relatifs à leurs propriétés et leurs application potentielles. Ce chapitre vise à présenter des généralités sur les liquides ioniques (définition, propriétés et applications).

I.2 Définition des liquides ioniques

Les liquides ioniques (LIs) sont des sels fondus qui sont formés exclusivement d'ions, qui possèdent un point de fusion sous les 100 C ? Ils ont une pression de vapeur pratiquement nulle, ils sont non inflammables, possédant une grande stabilité thermique, une large fenêtre électrochimique et une conductivité élevée. Toutefois, ils sont visqueux et hygroscopiques. De plus, ils sont miscibles avec la plupart des solvants organiques polaires et la dernière génération de LIs est stable avec l'air et l'humidité .Pour définir les LIs simplement, il s'agit de sels composés de larges ions présentant une asymétrie et un système de résonance dans leur structure ce qui a pour conséquence de diminuer l'énergie réticulaire (Bruno, 2017).

Il existe deux catégories de (LIs) protiques (LIP) et aprotiques (LIAP). Un LIP est produit par le transfert de proton entre une base et un acide de Brønsted ce qui n'est pas le cas pour un LIAP.

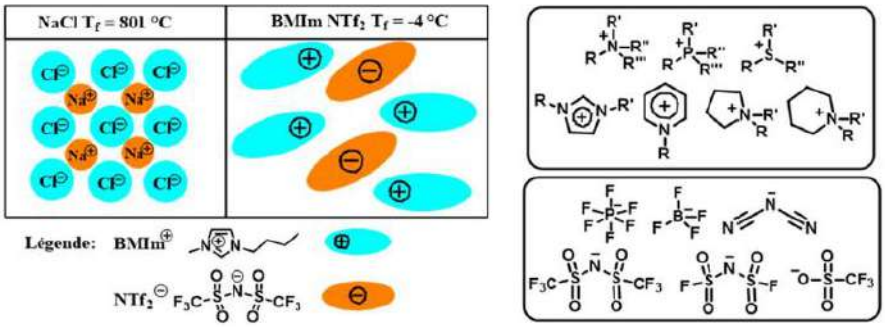


Figure I.1 : Définition des liquides ioniques et quelques exemples d'anions et de cations (Bruno, 2017).

Les liquides ioniques sont constitués d'un cation le plus souvent organique, associé à un anion organique ou inorganique et les combinaisons cations/anions possibles sont très nombreuses. Les liquides ioniques les plus utilisés ont une structure composée de cations centrés sur l'azote (tétraalkylammonium, alkylpyridinium, alkylimidazolium), le phosphore (phosphonium), ou le soufre (sulfonium), dialkylpiperidinium, et dialkylpyrrolidinium. Les anions ont la plus grande diversité structurale qui peut être trouvée et sont régulièrement employés comme des ions halogénures, acétate, trifluoroacétate, triflate, bitrifilimide, alkylsulfate, sulfonate tetrafluoroborate, tétraarylborate, hexafluorophosphate.

I.3 Structure des liquides ioniques

Les liquides ioniques (LIs) ou les liquides ioniques à température ambiante (RTIL) font toujours partie des classes de composés les plus populaires et sont étudiés de manière intensive, Cependant, les connaissances sur les liquides ioniques et leurs mélanges sont encore incomplètes. La complexité de la structure des cations et des anions donne plus que quelques formes possibles d'interactions intermoléculaires à l'état pur et en mélanges. (Huang, 2015)

Les cations rencontrés sont généralement volumineux et dissymétriques. Les plus classiques sont des ammoniums ou phosphoniums quaternaires, tels que les tétraalkylammoniums ou tétraalkylphosphoniums mais de nombreux LIs sont à base de systèmes hétéroaromatiques comme les alkylpyridiniums, les triazoliums ou encore des alkylpyrrolidiniums. Les plus étudiés sont les sels d'imidazoliums diversement substitués sur

les atomes d'azote et de carbone. De plus, toute une série de cations fonctionnalisés a été récemment développée notamment des cations portant des groupements amines[5], alcools ou éthers, acides carboxyliques ou esters, thiols, vinyl et allyl, alcynes ou encore nitriles. Des cations chiraux ont également été synthétisés (Figure I.2). *(Gutel, 2007)*

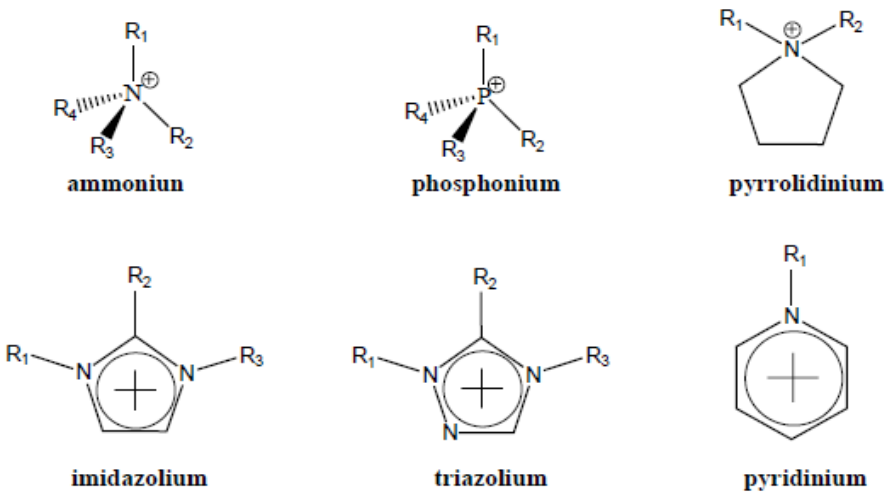
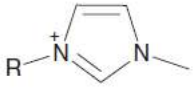
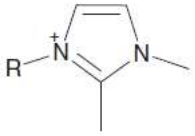
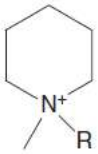
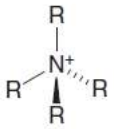
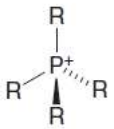
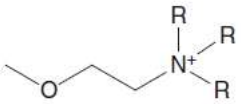
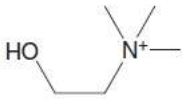
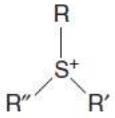
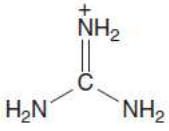


Figure I.2 : Cation des LIS

Tableau I.1 : Exemple de structure de cations (Kar et al, 2017).

Cations	Abréviation	Structure
Alkylmethylimidazolium	$[C_n mim]^+$	
Alkyldimethylimidazolium	$[C_n dmim]^+$	
Alkylmethylpyrrolidinium	$[C_n mpyr]^+$	
Alkylmethylpiperidinium	$[C_n mpip]^+$	
Ammonium	$[N_{n,n,n,n}]^+$	
Phosphonium	$[P_{n,n,n,n}]^+$	
Ether-functionalized	$[N_{R,R,R,2O1}]^+$, etc.	
Cholinium	$[Ch]^+$	
Sulfonium	$[S_{R,R',R''}]^+$	
Guanidinium	$[Gdm]^+$	

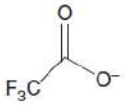
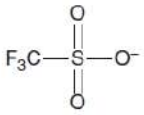
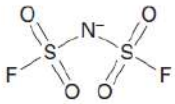
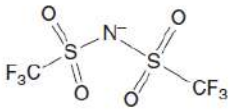
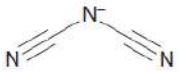
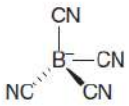
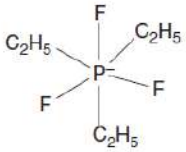
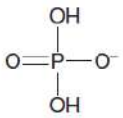
Les anions mis en œuvre sont des anions inorganiques ou organiques (tableau I.2). Les anions tétrafluoroborate ($BF4^-$) et hexafluorophosphate ($PF6^-$) très utilisés en chimie organique ou organométallique pour conférer une solubilité recherchée aux espèces ioniques

sont à la base de très nombreux sels liquides. Pour ce qui est des anions organiques, les anions fluorés ($CF3CO2^-$) sont très intéressants notamment en catalyse organométallique. Les dérivés sulfoniques : $CF3SO3^-$, $C4F9SO3^-$, $(CF3SO2)2N^-$ et $(CF3SO2)3C^-$ sont également très étudiés pour leur stabilité thermique et leur pouvoir faiblement coordonnant.

Tableau I.2 : Exemple d'anion de LIs (Gutel, 2007)

Anions inorganiques	Anions organiques
F^-, Cl^-, Br- , I	$CH_3CO_2^-$, $CH_3SO_4^-$,$C_6H_5SO_3^-$(=OTs)
BF_4^- , PF_6^-, SbF_6^- , AsF_6^-	$CF_3CO_2^-$, $C(CF_3SO_2)_3^-$
$NO3^-$, $ClO4^-$	$CFSO_3^-$(=OTF)
$AL_xCl_{(3x+1)}^-$, $AL_xEt_xCl_{(ex+1)}^-$	$N(SO_2CF_3)_2^-$ (=NTF_2)
$CuCl_2^-$, $ZnCl_3^-$, $SnCl_3$-	BR_4^-, R_3BOH^-

Tableau I.3 : Exemple de structure d'anions *(Kar et al, 2017).*

Anions	Abbreviation	Structure
Tetrafluoroborate	$[BF_4]^-$	
Hexafluorophosphate	$[PF_6]^-$	
Trifluoroacetate	[tfa]$^-$	
Triflate	$[OTf]^-$ or $[CF_3SO_3]^-$	
Bis(fluorosulfonyl)imide[a)]	[fsi]$^-$ or $[N(SO_2F)_2]^-$	
Bis(trifluoromethanesulfonyl)imide[a)]	$[NTf_2]^-$ or $[N(SO_2CF_3)_2]^-$	
Dicyanamide	$[N(CN)_2]^-$ or [dca]$^-$	
Tetracyanoborate	$[B(CN)_4]^-$	
Fluoroalkylphosphates	[fap]$^-$, [efap]$^-$, etc.	
Dihydrogenphosphate	$[H_2PO_4]^-$ or [dhp]$^-$	

I.4 Historique

Au début du 20ème siècle des sels liquides de nitrate d'alkylammoniums furent découverts en particulier le nitrate d'éthylammonium par (Walden, 1914) dont le point de fusion est de 12°C. Ces découvertes furent le début de l'ère des LI tels qu'ils sont connus à ce jour. A cette époque, les scientifiques ont porté peu d'intérêt à cette découverte et ce n'est que quarante ans plus tard que le premier brevet faisant apparaître les LIs fut déposé. En effet, Hurley & Wier en 1951 ont mis au point la première synthèse de LIs à base d'anions

chloroaluminate permettant l'électro-dépôt à basse température de l'aluminium. Les chloroaluminates constituent la première génération de LIs.

Durant les années 1970 et 1980, ces LIs ont été utilisés principalement dans le domaine de l'électrochimie. Ce fut une importante avancée, cependant les liquides basés sur des chloroaluminates, qu'ils soient constitués de cations pyridinium ou de cations imidazolium, sont très sensibles à l'eau et s'hydrolysent. Une étape supplémentaire fut de préparer des LIs stables à l'air, qui soient simples à préparer et sans précautions particulières. En 1992, Wilkes & Zaworotko ont publié la préparation du liquide ionique 1-éthyl-3-méthylimidazolium tétrafluoroborate correspondant à ces conditions. *(Revelli, 2010).*

C'est à partir de ces derniers travaux que l'intérêt porté aux LIs a pris son essor, tant du point de vue industriel que du point de vue fondamental. Les LIs ont connu une évolution exponentielle en termes de publications. Une large gamme de LIs a été synthétisée incorporant différents anions tels que l'hexafluorophosphate, l'éthanoate, le trifluoroéthanoate, les alkylsulfate, l'hydrogensulfate, le nitrate, le dicyanamide, le trifluorométhanesulfonate ou encore le bis (trifluoromethylsulfonyl) amide. Ces dernières années, un autre groupe de LI a été développé ; les LIs fonctionnalisés également appelés liquides ioniques à tâches spécifiques (TaskSpecificIonicLiquids, TSILs) comportant un groupement fonctionnel attaché par liaison covalente au cation ou l'anion. *(Ayad, 2017)*

Ce n'est qu'au début de ces années que les LIs furent utilisés comme solvants. Depuis ils sont utilisés dans de nombreux domaines tels que la synthèse organique, la catalyse, la spectroscopie, l'électrochimie, l'extraction, la séparation ou encore la préparation de nanomatériaux.

I.5 Synthèses

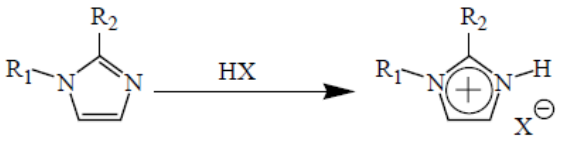
Ces dix dernières années connaissent une croissance importante du nombre de liquides ioniques synthétisés. Cependant, ils sont constitués pour la plupart d'un cation imidazolium, d'autres sont des sels de phosphonium ou de sulfonium mais leurs synthèses sont semblables. La synthèse de ces LIs s'effectue en deux étapes :

- La quaternisation du noyau imidazole
- L'échange d'anion

I.5.1 Réaction de quaternisation du noyau imidazole

La préparation du cation peut être effectuée soit par protonation en milieu acide soit par quaternisation d'une amine par un halogénure d'alcane.

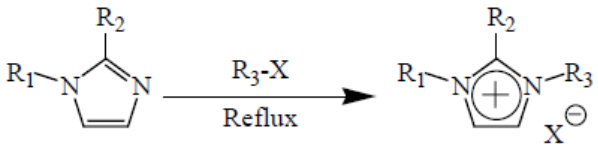
- La protonation des imidazoles par un acide conduit directement aux sels d'imidazoliums désirés. Cette technique ne permet pas la préparation de sels d'imidazoliums alkylés en position. *(Gutel, 2007)*



R_1 et R_2 = alkyl, H

X = Cl, NO_3, BF_4, PF_6

- La substitution nucléophile d'halogénures d'alcanes par les imidazoles conduit avec de bons rendements aux halogénures d'imidazoliums correspondants. Cette méthode a l'avantage d'utiliser des réactifs commerciaux et bon marché mais nécessite souvent la distillation des réactifs et des temps de réaction importants (plusieurs jours avec les chloroalcanes) même s'ils peuvent être réduits par utilisation des micro-ondes, des ultrasons ou en opérant sous pression. *(Gutel, 2007)*



R_1, R_2 et R_3 = alkyl

X= Cl, Br, I, OTf ou OTs

La réactivité des halogénures d'alcanes croit dans l'ordre : Cl >Br >I, les fluorures ne pouvant être préparés de cette manière. La réaction de quaternisation d'amine par des triflates ou des tosylates est également possible du fait de la présence d'un très bon groupe partant et peut être réalisée à température ambiante. Ces réactions peuvent être effectuées sans solvant mais dans tous les cas, cette étape est réalisée sous atmosphère inerte du fait du caractère extrêmement hygroscopique voire hydrolysable des réactifs et des produits.

La décantation en fin de réaction permet d'éliminer l'excès de solvant et de réactifs, les sels d'imidazolium étant généralement plus denses que les solvants organiques mais par

précaution, le produit est généralement traité sous vide avant usage pour éviter toutes traces d'eau ou de produits volatils. Le cation, une fois préparé, peut être également purifié par recristallisation ou lavé avec un solvant non-miscible.（Gutel, 2007)

I.5.2 Réaction d'échange de l'anion

La réaction d'échange de l'anion peut se diviser en trois catégories: traitement par une résine échangeuse d'ions, traitement direct du sel d'imidazolium par un acide de Lewis et la réaction d'échange par métathèse d'anions.

I.5.2.1 Résine échangeuse d'ions

Elle est généralement utilisée pour échanger un halogénure d'un liquide ionique avec un anion organique qui est souvent la base conjuguée de l'acide correspondant. Cette méthode consiste à échanger d'abord l'halogénure du LI pour un ion hydroxyde, puis de le faire réagir avec un acide pour obtenir un LI avec un contre ion organique et de l'eau suivant le schéma de la figure I.3. *(Diop, 2014)*

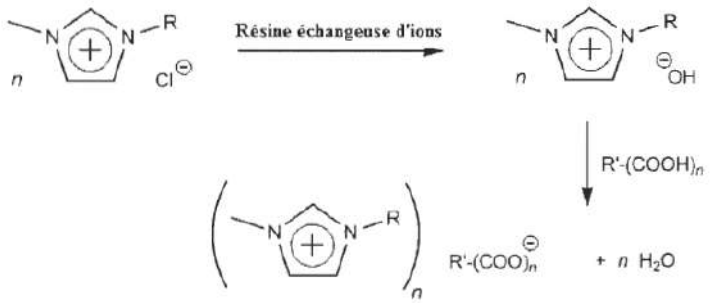
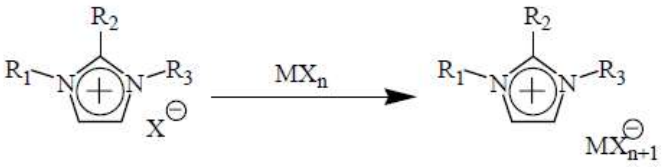


Figure I.3 : Échange d'un ion chlorure avec un carboxylate *(Diop, 2014)*

I.5.2.2 Traitement acide

Le contre ion halogénure d'un liquide ionique peut être remplacé par un contre ion métallique en faisant réagir 1 équation d'un halogénure de LI avec 1 équation d'un acide de Lewis selon la réaction présentée à la figure I.4. *(Diop, 2014)*



R_1, R_2 et R_3 = alkyl

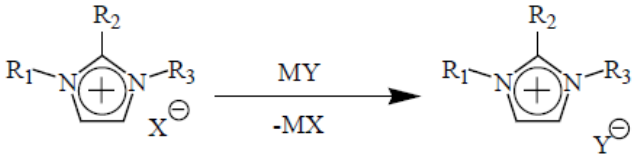
X= Cl, Br, I

M = Al, Cu, Sn, Fe, Zn

Figure I.4 : Échange d'un chlorure avec un ion métallique *(Diop, 2014)*

I.5.2.3. Métathèse anionique

L'anion inorganique d'un liquide Ionique peut être changé avec un autre anion inorganique en faisant réagir le sel de ce dernier avec le liquide ionique (figure I.5). (Diop, 2014)



R_1, R_2 et R_3 = alkyl

X= Cl, Br, I

MY=$LiNTf_2$, NaOTf, $NaPF_6$, $NaBF_4$

Figure I.5 : Échange d'anion par métathèse *(Diop, 2014)*

Les LI obtenus par ces voies de synthèse sont généralement des liquides incolores bien que les sels d'imidazoliums à base de PF6 ou BF4 puissent présenter une légère coloration jaune. Il est indispensable de caractériser la pureté de ces composés et il est parfois nécessaire de les purifier avant usage. *(Gutel, 2007)*

I.5.3 Purification et pureté des LI

La pureté des LIs est un paramètre important pour la plupart des applications car la présence d'impuretés peut gravement affecter les propriétés physico-chimiques des Lis et la réactivité des systèmes catalytiques. *(Gutel, 2007)*

Comme les LIs sont très peu volatils, la purification par distillation n'est pas envisageable excepté dans des conditions sévères. La purification par chromatographie est également délicate dans la mesure où les LIs ont tendance à s'absorber sur les phases stationnaires. Par conséquent, un grand soin doit être apporté lors de la synthèse des LIs afin

de limiter tant que possible les impuretés présentes dans le produit final notamment les traces d'eau, d'halogénures et d'imidazole de départ. *(Gutel, 2007)*

Il est possible de mesurer la quantité d'eau présente dans les LIs par dosage Karl-Fischer, par spectroscopie infra-rouge, par spectroscopie UV ou par spectrométrie de masse. La détection de traces d'halogénures est généralement réalisée par un test au nitrate d'argent mais ces impuretés peuvent être quantifiées par analyse élémentaire, par chromatographie ionique, ou par électrophorèse capillaire. La présence d'imidazole de départ peut être détectée par un test colorimétrique au chlorure de cuivre(II). *(Gutel, 2007)*

I.5.4 Nomenclature et acronymes des LI

À l'image de la définition des liquides ioniques, on trouve dans la littérature plusieurs nomenclatures. Nous prendrons comme nomenclature arbitraire celle en accord avec les pionniers des liquides ioniques de nouvelle génération. Il s'agit des liquides ioniques à base d'ion immidazolium plus particulièrement l'ion méthylimmidazolium (Figure I.6). Ils peuvent être associés à une large gamme d'anions. Nous avons adopté la nomenclature suivante:

- Le cation imidazolium est noté par le suffixe « im »
- Les alkyles Rl et R2 sont représentés respectivement par Cn et Cm ou n et m représente la longueur des chaines en position 1 et 2
- Les anions sont notés par leur formule chimique et les plus fréquents sont: cr, MeS04-, MeS03-, CF3S03-, BF4-, PF6- et N03-.

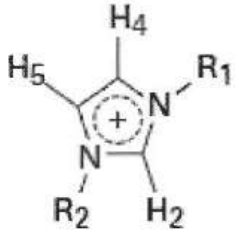


Figure I.6 : Cation immidazolium

Cette nomenclature sera appliquée à tous les types de liquides ioniques autres que les immidazoliums évoqués dans ce rapport en remplaçant les suffixes «im» par l'abréviation du cation du liquide en question.

I.6 Propriétés des liquides ioniques

Les liquides ioniques présentent un grand nombre des propriétés physico-chimiques importantes, qui font d'eux une classe de solvants très convoitée pour de nombreuses applications.

Les premiers travaux sur les propriétés des LIs purs, se sont d'abord concentrés sur la compréhension de la relation entre la structure du liquide ionique et les propriétés physico-chimique. La pureté du liquide ionique est apparue comme un facteur très important influençant les mesures de ces propriétés. (*Amichi et Benrouissi, 2017)*

La caractérisation des propriétés physicochimiques des LIs telles que la tension superficielle, le point de fusion, la viscosité, densité et la conductivité est fondamentale. Il est primordial que ces caractérisations soient effectuées avant même d'être utilisées comme composante d'un électrolyte.

I.6.1 La tension superficielle

L'énergie de surface d'un liquide est la quantité d'énergie nécessaire pour créer une unité de surface d'une nouvelle surface.

Parmi les nombreuses propriétés uniques des liquides ioniques, leur tension superficielle joue un rôle particulier pour différentes raisons. Même dans des conditions bien au-dessus de la température ambiante, la tension superficielle est une mesure des forces de cohésion exercées entre les molécules de liquide présentes à la surface. Elle représente la quantification de la force par unité de longueur d'énergie libre par unité de surface. Ainsi, la mesure de la tension superficielle de liquides ioniques est l'un des moyens les plus efficaces d'accéder (indirectement) à l'énergétique intrinsèque impliquée dans les interactions entre les ions. *(Tariq et al, 2012)*

D'une part, les mesures de tension superficielle des liquides ioniques, sont effectuées par un tensiomètre Système KSV Sigma 701 (Finlande) utilisant un anneau Du-Noüy prenant en compte la correction Zuidema Waters. (*Domanska et al, 2012*)

D'autre part, la tension de surface est liée à l'énergie de surface (et aux forces intermoléculaires) mais elles ne sont pas identiques. La première est liée à l'excès d'énergie de Helmholtz (ou Gibbs), tandis que la dernière est liée à l'excès d'énergie interne. La magnitude de la pente de la température dépend de la tension superficielle d'un liquide pur est l'entropie

de surface, qui est une mesure du désordre de surface par rapport à la masse liquide (Tariq et al, 2012)

Finalement, comme indiqué dans les trois paragraphes précédents, la tension superficielle des liquides purs est une propriété physique importante qui, contrairement à d'autres propriétés globales, n'est pas facilement corrélée aux caractéristiques des substances au niveau moléculaire.

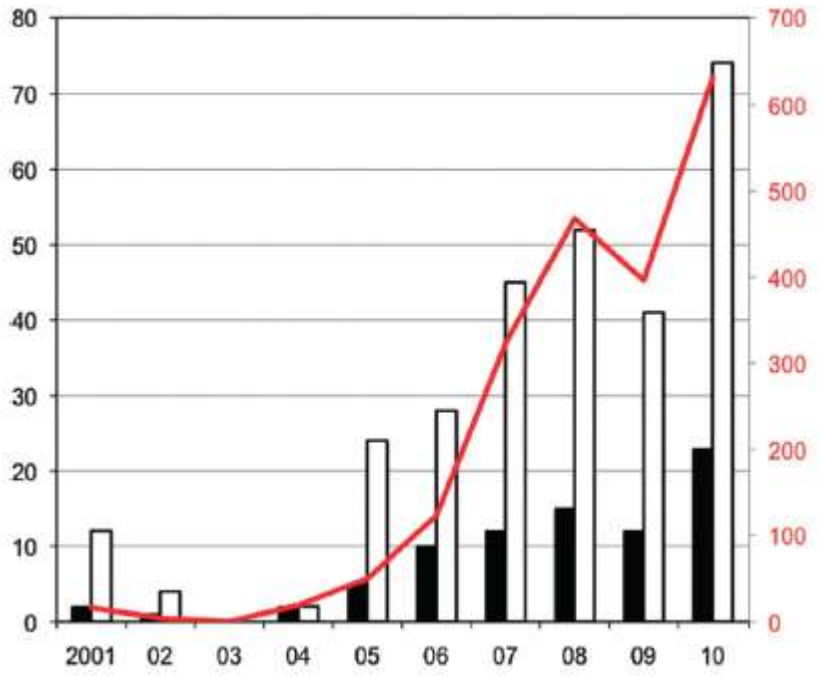


Figure I.7 : Données de tension superficielle pour les liquides ioniques purs, de 2001 à 2010 (*Tariq et al ,2012*)

I.6.2 Point de fusion

Un sel fondu est défini comme liquide ionique lorsque son point de fusion est inférieur à 100 °C. La température de fusion est influencée par la distribution de charge sur les ions, les capacités des liaisons hydrogène, la symétrie des ions et les interactions de VAN DER WAALS. On rencontre avec une grande incertitude sur la valeur de nombreux de point de fusion des liquides ioniques, c'est-à-dire, qu'ils possèdent une plage de température dans laquelle ils passent par une phase cristalline vitreuse. En effet, le point de fusion de certains liquides ioniques peut fortement varier selon qu'ils soient chauffés ou qu'ils soient refroidis. L'état surfondu est une caractéristique associée à des nombreux liquides ioniques à base de cations imidazolium. Pour exemple, la littérature indique que les cations imidazolium combinés avec les anions NTf_2 sont généralement des liquides au-dessus de -30 °C à -50 °C,

mais ils deviennent très visqueux jusqu'à l'état vitreux sans que l'on puisse observer leur point de fusion. *(Revelli, 2010).*

Le point de fusion des LIs diminue lorsqu'il y a une forte asymétrie des substituant de cation imidazolium. En plus, lorsque la longueur de la chaine d'alkyl augmente, en effet La longueur de la chaine alkyl greffée sur le cation a une grande influence sur leur point de fusion et lorsqu'il y a une forte asymétrie des substituant de cation imidazolium.

L'effet de l'anion sur le point de fusion (tableau) est plus difficile à expliquer. Dans le cas des LIs à base de cations imidazolium combinés à des anions TfO^- et NTf_2^-, les faibles valeurs des points de fusion sont attribuées à une importante délocalisation de la charge sur l'anion, et à une faible interaction par liaisons hydrogène. *(Revelli, 2010).*

Tableau I.4 : Valeurs de température de fusion de certains liquides ioniques *(Moutiers et Billard, 2011).*

Liquide ionique	**T fusion (°C)**
[C2MIm] [Cl]	87 89 (sec)
[C2MIm] [BF4]	15 11 (sec)
[C2MIm] [PF6]	61 (sec)
[C2MIm] [BTI]	-16 -15 (sec)
[C4MIm][Cl]	41 (sec) 65
[C4MIm] [BTI]	16
[N(n-C4)4] [BTI]	96
[N(n-C6)(C2)3] [BTI]	20
h[N(n-C6)(n-C4)4] [BTI]	26
L'indication (sec) signifie que le liquide a été séché	

I.6.3 Température de décomposition

Étant donné que les liquides ioniques présentent une faible tension de vapeur, leur température maximale d'utilisation est fixée par la température de décomposition.

Généralement, les cations imidazolium ont des températures de décomposition supérieures à celles des cations ammonium, permettant leur utilisation à des températures supérieures à 250 °C et dans certains cas à 400 °C. Mais la nature de l'anion a aussi une influence sur la température de dégradation des liquides ioniques. En effet, les températures de décomposition les plus élevées sont observées pour des anions engendrant les plus faibles interactions intermoléculaires selon l'ordre : $PF_6^- > NTf_2^- > BF_4^- > I^- > Br^- > Cl^-$.

Tableau I.5 : Valeurs de température de composition de quelques liquides ioniques (*Moutiers et Billard, 2011)*

Liquide ionique	**T décomposition (°C)**
[C2MIm] [Cl]	285 (sec)
[C2MIm] [BF4]	391
[C2MIm] [BTI]	417 455 (sec)
[C4MIm][Cl]	254 (sec)
[C4MIm][BF4]	403 (sec)
[C4MIm][PF6]	349
[C4MIm][BTI]	439 (sec) 400
[C6MIm][Cl]	253 (sec)
[C6MIm][PF6]	417
[C8MIm][Cl]	243 (sec)
L'indication (sec) signifie que le liquide a été séché	

I.6.4 Viscosité

La viscosité des liquides ioniques est généralement supérieure à celle des solvants organiques conventionnels. Les valeurs de viscosité des liquides ioniques de type 1,3-dialkylimidazolium sont comprises entre 40 à 1000 mPa.s à température ambiante. Elles sont plus élevées que les viscosités de l'eau, de l'éthylène glycol et du glycérol à 25 °C qui sont respectivement : 0,89 ; 16,1 et 934 mPa.s. (Bouchardi). La viscosité est évaluée essentiellement par leur capacité à former des liaisons hydrogène et par la force des interactions de Van der Waals.et la longue de la chaine alkyle provoque l'augmentation de la viscosité. La viscosité des liquides ioniques est donc modulable mais leurs valeurs restent toujours élevées, ce facteur doit être pris en compte dans les procédés car des lenteurs inhérentes à cette viscosité sont inévitables (dissolution, homogénéisation, établissement des équilibres, etc. (*Amichi et Benrouissi, 2017)*

Tableau I.6 : Viscosité de certains liquides ioniques. (*Amichi et Benrouissi, 2017)*

Liquide ionique	**Viscosité (cP)**
[C2MIm] [BTI]	25
[C2MIm] [BF4]	66.5 (20°C) 37.7 (22°C)
[C2MIm] [CH_3CO_2]	160 (20°C)
[C2MIm] [CF_3CO_2]	35 (20°C)
[C2MIm] [CF_3SO_2]	45 (20°C)
[C4MIm] [I]	1110 (sec) (25°C)
[C4MIm] [BF4]	219 (sec) (25°C)
[C4MIm] [PF6]	207 (sec) (25°C) 450 (sec) (25°C) ≈ 275 (25°C)
[C4MIm] [BTI]	52 (20°C) 69 (sec) (25°C) ≈ 42 (25°C)
[C4MIm] [CF_3CO_2]	73 (20°C)
[C4MIm] [CF_3SO_2]	90 (20°C)
[C6MIm] [BTI]	≈ 60 (25°C)
[C6MIm] [PF6]	≈ 410 (25°C) 585 (sec) (25°C)
La notation (sec) indique que la viscosité a été mesurée sur un liquide ionique séché	

I.6.5 Densité

La plupart des liquides ioniques sont plus denses que l'eau, avec des valeurs de densité comprises entre 1,0 et 1,6 g.cm^{-3}. Dans le cas des cations imidazolium, la densité diminue quand la longueur de la chaîne alkyle augmente *(Bouchardy ,2016)* et elle augmente si la masse molaire de l'anion augmente, leur densité est reliée à l'encombrement du cation et de l'anion mais aucune règle empirique n'existe à ce jour.

Tableau I.7 : Valeurs de densité de quelques liquides ioniques. (*Amiche et Benrouissi, 2017)*

Liquide ionique	**Densité à T = 25°C**
[C2MIm] [BF4]	1.28 1.24 (22°C) (sec)
[C2MIm] [BTI]	1.52
[C2MIm] [CF_3CO_2]	1.285
[C2MIm] [CF_3SO_2]	1.39
[C4MIm] [I]	1.08
[C4MIm] [PF6]	1.37 1.36 (sec)
[C4MIm] [BTI]	1.44
[C4MIm] [CF_3CO_2]	1.29
[C4MIm] [CF_3SO_2]	1.29 1.31 (sec)
[C6MIm] [BTI]	1.37
[C8MIm] [BF4]	1.08

[C8MIm] [PF6]	1.37 1.24 (sec)
[C8MIm] [BTI]	1.32
[N(n-C6)(C1)3] [BTI]	1.33 (20°C)
[N(n-C6)(C2)3] [BTI]	1.27 (20°C)
[N(n-C6)(n-C4)3] [BTI]	1.15 (20°C)
[N(n-C8)(C1)3] [BTI]	1.27 (20°C)
[N(n-C8)(C2)3] [BTI]	1.25 (20°C)
[N(n-C8)(n-C4)3] [BTI]	1.12 (20°C)
La notation (sec) indique que la densité a été mesurée sur un liquide ionique séché	

I.6.6 Volatilité

Les liquides ioniques sont des solvants dont la pression de vapeur saturante est négligeable à température ambiante. Cela permet de les recycler facilement en distillant les autres produits directement à partir du mélange. Cependant cette pression de vapeur saturante est négligeable et non nulle. Ainsi, dans certains cas précis, il est possible de séparer des liquides ioniques par distillation fractionnée à très faible pression et très haute température.

Ces cas sont extrêmement rares et la majorité des liquides ioniques se dégradent au chauffage avant de pouvoir être distillé *(Http://www.perso.univ rennes1.fr/mathieu.pucheault/_PDF/Publis/2010_Pucheault_Dunod.pdf (Consulté le 03/05/2017).*

I.6.7 Toxicité et biodégradabilité

I.6.7.1 La toxicité

Des études de toxicité entreprises sur les liquides ioniques ont été répertoriées par Thuy Pham *et Al.* Les recherches montrent que dans le cas des liquides ioniques constitués du cation alkylimidazolium, la toxicité augmente avec une élongation de la chaîne alkyle. *(Bouchardy ,2016)*

Le caractère toxique des LIs est relié à leur hydrophobicité *(Quijano et al, 2010)*

L'introduction d'une chaine polaire réduit leur toxicité et augmente leur biodégradabilité, les cations pyridinium semblent être plus respectueux de l'environnement que les cations imidazolium, l'anion a un effet moins important sur la toxicité, cependant, les LIs constitués de l'anion NTf_2^- ont un impact toxicologique sur l'environnement *(Revelli, 2010)*

Tableau I.8 : Toxicité des LIs sur les microorganismes souvent utilisés en domaine de biotechnologie (*Quijano et al, 2010)*

Liquide ionique	**IC_{50}(% v/v)**	**Microorganisme**	**Indicateur de toxicité**	**Temps exposé (h)**
[C4MIm] [PF6]	5	Lactobacillus rhamnosus	Inhibition de la production d'acide lactique	48
	0.5	Escherichia coli	Inhibition augmentée	8
	13	Klebsiella pneumoniae	Inhibition augmentée	24
[C4MIm] [BTI]	20	Saccharomyces cervisiae	Inhibition augmentée	80
[C4MIm] [BF4]	7	Photobacterium Phosphoreum	Inhibition augmentée de lumière	0.25
	15	Proteus vulgaris	Inhibition augmentée	24
[C6MIm] [PF6]	15	Candida albicans	Inhibition augmentée	48
	15	Escherichia coli	Inhibition augmentée	24
[C6MIm] [BF4]	16	Klebsiella pneumoniae	Inhibition augmentée	24
	16	Pseudomonas aeruginosa	Inhibition augmentée	24
[C8MIm] [PF6]	20	Saccharomyces cervisiae	Inhibition augmentée	80
[C1MIm] [PF6]	20	Saccharomyces cervisiae	Inhibition augmentée	80

IC_{50} : concentration inhibitoire à laquelle la croissance, la viabilité ou le métabolisme de cellule est inhibé à 50% relative au un control défavorisé de LI

I.6.7.2 La biodégradabilité

La non-biodégradabilité des LIs est une caractéristique nécessaire pendant les procédés biotechnologiques. C'est important spécialement quand le solvant est couteux et doit être réutilisé comme les LIs. Comme il est présenté dans le (tableau 9), dans la plus part des cas les LIs ne sont pas biodégradables même en utilisant des sédiments activés *(Moutiers et Billard, 2010)*

Tableau I.9 : La biodégradabilité de quelques liquides ioniques *(Quijano et al, 2010)*

Liquide ionique	Microorganisme utilisé	Test[a]	T (°C)	Temps d'expérience (J)	Biodégradable
[C4MIm][PF6]	Sédiment activé	DBO	20	10	Non
	Sédiment activé	Sturm test	20	28	Non
	Sédiment activé	DBO	20	28	Non
	Fusarium sp	Mesure directe de concentration des LIs	30	28	Oui
[C4MIm][BTI]	Sédiment activé	DBO	20	28	Non
[C4MIm][BF4]	Sédiment activé	DBO	20	28	Non
[C4Py][Br]	Sédiment activé	Mesure directe de concentration des LIs	25	28	Oui
[C4MIm][BF4]	Microorganisme de sol indigène	Sturm test [b]	25	180	Oui[c]
[C4MIm][PF6]	Sédiment activé	DBO			Non

A- tous les tests ont été réalisés aux conditions aerobiques.

B- non microorganismes additionnels ont été ajoutés au le système hors de sol indigène.

C- seulement deux atomes de carbone de la chaine alkyle ont été biodégradés.

Sédiment activé (activated sludge): flacon fermé inoculé avec les microorganismes des eaux polluées.

DBO: Demande Biochimique en Oxygène.

Sturm test : basé sur la production de CO_2 après l'inoculation avec le sédiment activé.

I.6.8 Solubilité

Les LI possèdent un grand pouvoir savatant, qui leur permet de solubiliser une large gamme de composés organiques, inorganiques et également organométalliques, et sont liquides dans une large gamme de température (*Amichi et Benrouissi, 2017)*

I.6.8.1 Solubilité dans l'eau et dans les solvants organiques

Le caractère hydrophile ou hydrophobe des liquides ioniques est un paramètre très important dans l'étude de leur solvatation. Il est gouverné par la nature de l'anion et la longueur de la chaîne alkyle greffée sur le cation. En règle générale, les liquides ioniques sont miscibles avec les solvants organiques polaires (alcools à courte chaîne, acétonitrile, THF, acétone, dichlorométhane et chloroforme). En revanche, la plupart des liquides ioniques sont non-miscibles avec les alcanes, le dioxane, le toluène et l'éther (*Bouchardy ,2016).*

De plus, les liquides ioniques ont la propriété d'être hygroscopiques, c'est-à-dire qu'ils peuvent absorber l'eau contenue dans l'atmosphère. Les molécules d'eau absorbées dans les liquides ioniques sont à l'état 'libre' en interaction par liaisons hydrogènes avec les anions pour des concentrations comprises entre 0.2 et 1.e-3 mol.litre^{-1}(*Amiche et Benrouissi, 2017).*

Du fait de leur riche variabilité, les liquides ioniques peuvent être utilisés dans divers domaines d'application.

I.6.8 Non inflammabilité

Les liquides ioniques sont en générales non inflammables, en exception les LIs inflammable appelée énergétique de composée d'ions de nitrate et de picrate.

I.6.9 La masse volumique

Les LIs présentent une masse volumique plus grande que celle de l'eau ou des solvants organiques, avec des valeurs entre 1 à 1,6 g/cm3. Normalement, l'augmentation du numéro

carbone dans la chaîne du LI réduit la masse volumique. L'addition d'un troisième substituant pour le cation imidazolium entraîne une diminution de la masse volumique .Pour un cation donné, la densité augmente avec une augmentation de la masse molaire de l'anion $Cl^- < BF_4^- < CF_3CO_2^- < NTf_2^-$. La présence d'impuretés comme l'eau ou les ions halogénures entraîne une diminution de la masse volumique (*Amiche et Benrouissi, 2017).*

I.7 Utilisation des liquides ioniques

A cause de la variété des propriétés physico-chimique des liquides ioniques, ils sont affirmés au cours de cette dernière décennie comme de bon solvant pour la synthèse organique.les LIs sont utilisé dans plusieurs domaine tel que l'électrochimie, la synthèse organique, l'extraction, la biocatalyse, la spectroscopie comme il est présenté dans la (figure I.8).

Ils permettent en effet, de par la grande variété de leurs propriétés physico chimiques, de solubiliser à la fois les molécules ioniques et polaires ainsi que les composés organiques classiques. Il parait actuellement plus de 40 articles par jour utilisant les liquides ioniques comme solvants, et environ 150 revues sur le sujet par an. Il est donc impensable de confiner de manière exhaustive les différentes utilisations des liquides ioniques en tant que solvant (*Amiche et Benrouissi, 2017*)

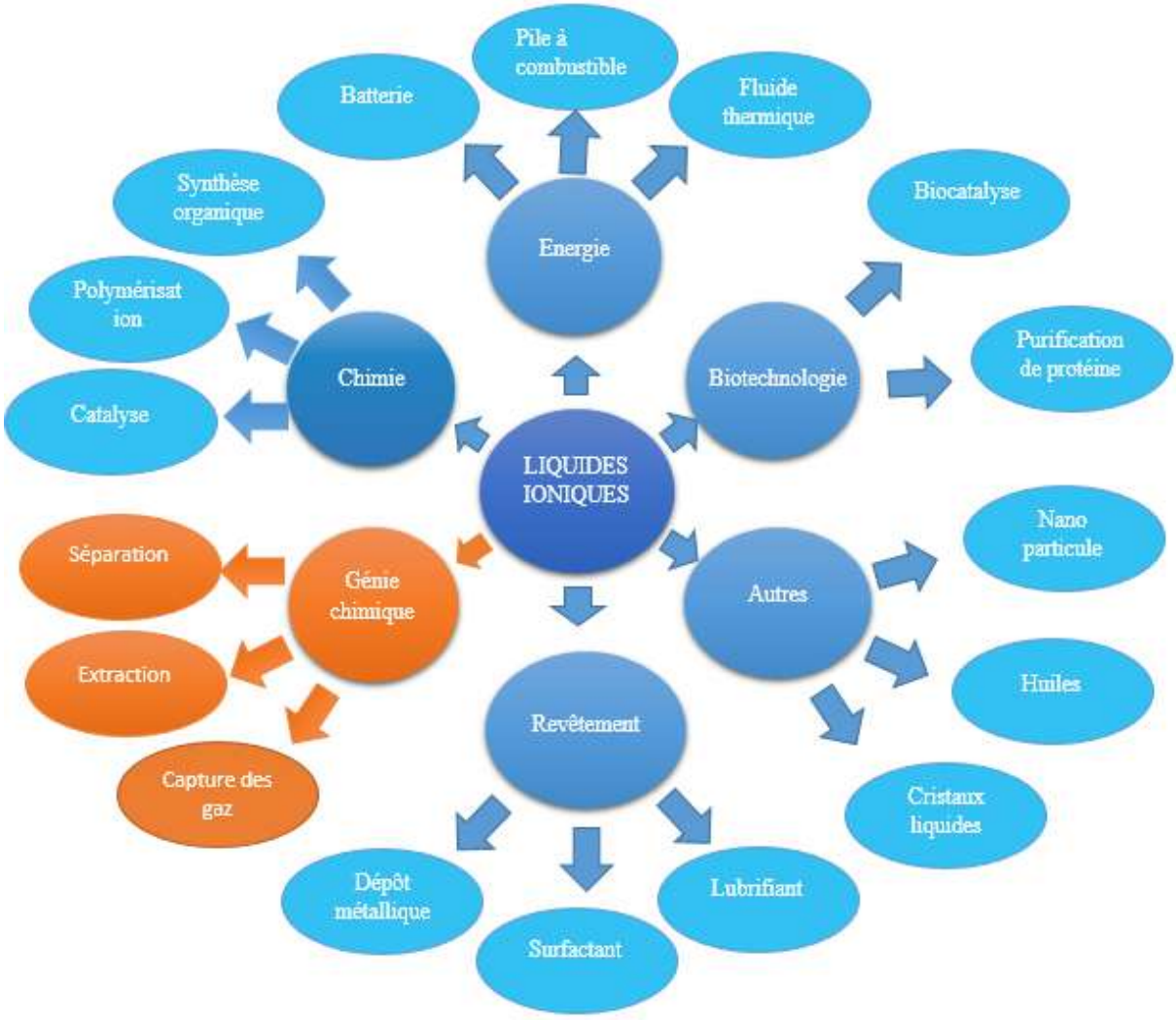


Figure I.8 : Les domaines d'application des liquides ioniques (*Chouih, 2018)*

I.7.1 Application dans l'industrie chimique

I.7.1.1 Application en électrochimie

La fenêtre électrochimique d'un solvant est un paramètre très important en électrochimie, les LIs ont des propriétés telles que leur très large domaine d'électro activité, leur forte conductivité, et leur grande stabilité thermique ont fait de ces nouveaux milieux des candidats de choix dans la recherche de nouveaux systèmes d'énergie (cellule photovoltaïque batterie...). Pour exemple, Caja et coll. ont montré, dans la recherche des batteries aux ions Li+ qui constitue un axe très important des batteries rechargeables, que les liquides ioniques, utilisés comme électrolyte, permettaient d'avoir une conductivité cinq fois supérieure aux électrolytes à base de solvant organique et de sel. De plus, leurs domaines d'électroactivité peuvent atteindre 4,5 V comparé à 1,2 V pour les électrolytes aqueux et ils offrent une très

bonne stabilité thermique et une meilleure solubilité que les composés usuels à base d'ammonium quaternaire... (*Zang et al, 2009).*

Parmi les applications des LIs dans le domaine électrochimique, on trouve :

- Les batteries Li-ion: les LIs sont utilisées comme électrolytes dans les batteries Li-ion à cause de leur nature non inflammable pour remplacer les électrolytes organiques très volatiles et inflammables.

- Piles à combustible : Certains LIs présentent une conductivité supérieure à 10^{-2}S/cm à température ambiante, valeur proche du niveau requis pour les applications de piles à combustible. La plupart des investigations sont axés sur le développement de LIs à base d'acide de Bronsted, synthétisés par un simple mélange d'amines ternaires et d'acides résultant en LIs avec proton libre. Ces LIs nommés, LIs protiques ont été étudiés largement pour le transfert de protons dans les électrolytes des piles à combustible (*Chouih, 2018)*

I.7.1.2 Chimie analytique

De nombreuses applications ont été proposées dans le domaine de la chimie analytique. A ce jour, les liquides ioniques sont principalement utilisés en chromatographie en phase gazeuse et en électrophorèse capillaire (EC). Il reste encore de nombreuses possibilités à explorer dans le domaine de séparations analytiques. *(Ayad, 2017)*

- **Chromatographie en phase gazeuse (CPG)**

Les LIs sont très étudiés en tant que phase stationnaire par chromatographie inverse en phase gazeuse. Cette technique intéressante est très utile pour étudier les propriétés physico chimiques des analyses et les interactions entre solutés et phase stationnaire.ont été les premiers à étudier les comportements des LIs ([BMIM] + [PF6]- et [BMIM] + [Cl]-), avec différents composés organiques. Les résultats ont montré que chaque LI agissait comme une phase stationnaire faiblement polaire avec les composés non polaires et comme une phase polaire avec les composés polaires. Les phases stationnaires de LI permettent d'obtenir une très bonne sélectivité et des séparations plus rapides. Plusieurs études ont été consacrées à l'utilisation de la CPG pour la détermination du coefficient d'activité à dilution infinie dans les LIs. Ce paramètre permet d'évaluer les interactions moléculaires entre le solvant et le soluté. *(Ayad, 2017)*

- **Électrophorèse capillaire**

Depuis les années 2000, une vingtaine de publications est sortie dans la littérature sur des applications en électrophorèse capillaire mettant en jeu des LIs. Ces applications qui ont été réalisées en phase aqueuse et non-aqueuse ont permis d'observer de nouveaux types d'interaction. En effet, les liquides ioniques ont la faculté de solubiliser des composés organiques et inorganiques et de générer des interactions particulières en solution, ils présentent donc un intérêt certain dans la recherche de nouveaux constituants d'électrolyte pour l'électrophorèse capillaire. De nombreux travaux ont été effectués d'une part, pour mettre en place une méthode de caractérisation des liquides ioniques, et d'autre part, pour évaluer l'intérêt de ces composés comme nouveaux constituants d'électrolytes. (*Ayad ,2017)*

I.7.1.3 Synthèse et catalyse

Il existe dans la littérature de nombreuses applications des LIs dans les domaines de synthèse organique et de la catalyse. Ces solvants peuvent avoir plusieurs fonctions dans la catalyse, comme solvant, solvant et co-solvant, solvant et support, solvant et ligands. Ainsi, il est important de mieux comprendre les propriétés physico-chimiques qui peuvent affecter les réactions catalytiques et de choisir les meilleurs LIs pour une réaction donnée.

Les LIs peuvent être utilisés comme milieux réactionnels pour les réactions organocatalysées. Ce potentiel a été démontré dans les réactions Diels-Alder où les Lis peuvent former une liaison hydrogène importante avec les réactifs et peuvent par la suite orienter la sélectivité de la réaction. Un exemple typique qui montre l'influence du LI sur le résultat de la réaction chimique est celui donnée par l'étude de la réaction entre le toluène et l'acide nitrique. En fonction du LI utilisé, les résultats observés montrent, la nitration du toluène dans [BMIM] + [OTf]-, l'oxydation du toluène en l'acide benzoïque dans [BMIM] + [OMs]-, ou halogénation dans [BMIM] + [Cl]-.

D'autre part, les LIs ont reçu une attention particulière comme solvants alternatifs et stabilisants pour la synthèse de nanomatériaux en raison de leur synthèse facile, leur stabilité et leur faible pression de vapeur. La synthèse de nanoparticule (NP) dans le LI a été décrite en utilisant différentes techniques. Les plus largement décrites sont la réduction chimique de sels métalliques et de complexes organométalliques. (*Ayad, 2017)*

I.7.2 Applications dans le domaine des procèdes de séparation et de l'analyse

Les procédés de séparations sont considérés comme le cœur du génie des procédés, à cause de son intérêt dans les domaines de la pétrochimie, l'hydrométallurgie, les industries pharmaceutique, agroalimentaire ou encore nucléaire.

D'autre part, Les différentes propriétés originales des liquides ioniques, présentent un très intérêt dans le domaine des procédés de séparation et de l'analyse chimique. Leurs capacités dissoudre des composés organiques apolaires aussi bien que des composés ioniques en ont fait des milieux de choix pour les sciences séparatives. L'extraction liquide - liquide est certainement le domaine dans lequel l'évaluation des liquides ioniques est la plus avancée.

L'étude des liquides ioniques dans le domaine des techniques chromatographiques et électrocinétiques est encore peu investiguée, toutefois un grand intérêt lui est porté ces dernières années, (*Temperley et al, 1980)*

I.7.2.1 Extraction liquide –liquide

Les propriétés physico - chimiques des liquides ioniques telles que la pression de vapeur négligeable, la non miscibilité avec d'autres solvants, et la bonne solubilité des composés organiques et inorganiques, en font un milieu de choix pour les techniques d'extraction - liquide et de micro extraction en phase liquide (LPME) (*Amichi et Benrouissi, 2017)*

- **Extraction des alcools pour la synthèse des biocarburants**

Les biocarburants sont envisagés comme une alternative aux carburants d'origine fossile. Les alcools : méthanol, éthanol, 1-propanol et 1-butanol peuvent être produits à partir de la biomasse par fermentation et sont donc considérés comme de bons candidats pour ce type de carburant. Lors de sa production, l'alcool synthétisées trouve en mélange avec de l'eau et d'autres composés organiques.

Des procédés de séparation sont donc nécessaires afin d'extraire l'alcool. La distillation est la technique généralement utilisée, mais elle est très coûteuse énergétiquement et requiert une installation lourde pour arriver à la pureté désirée. L'extraction liquide-liquide à l'aide de liquides ioniques est donc proposée comme procédé alternatif. Pour cela, des liquides ioniques hydrophobes constitués généralement de l'anion [Tf_2N] ou [PF_6] sont utilisés. L'étude du système {1-butanol +eau + [HMIM][eFAP]} montre que ce liquide ionique présente une très bonne sélectivité (de

l'ordre de 300) ainsi qu'un important coefficient de partage pour le1-butanol. On peut donc envisager ce solvant pour ce type d'applications *(Revelli, 2010)*

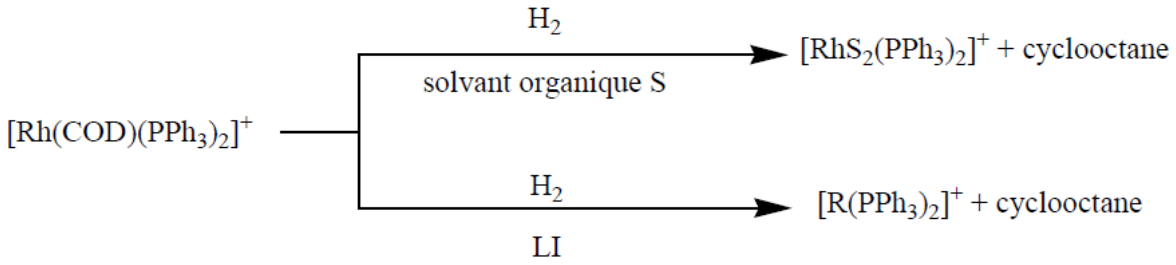
- **Extraction de composés aromatiques des vapeurs issues de procédés de raffinage**

La séparation des composés aromatiques (benzène, toluène, éthylbenzène ou xylène) des hydrocarbures aliphatiques (butane au décane) est difficile car ces composés ont des températures d'ébullitions voisines et certains mélanges forment des azéotropes. L'un des procédés utilisés pour la séparation de ces mélanges est l'extraction liquide –liquide avec des solvants organiques polaires comme le sulfolane. Cependant, la purification et la régénération du sulfolane nécessitent une importante quantité d'énergie. L'utilisation des liquides ioniques comme solvant de remplacement a donc été largement étudiée dans la littérature. La solubilité des n-alcanes est très faible dans les liquides ioniques, ce qui implique une sélectivité élevée. Ont démontré que le liquide ionique [EMIM][Tf2N] est plus performant que le sulfolane pour la récupération des composés aromatiques . *(Revelli, 2010)*

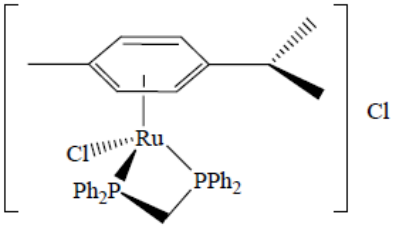
I.7.3 L'utilisation des LI comme solvants de réaction catalytique

L'utilisation des LI comme solvants de réaction catalytique est en plein essor ces dernières décennies. L'intérêt pour ces nouveaux milieux s'explique par la nécessité de trouver des procédés plus respectueux de l'environnement, les LI étant non-inflammables et très peu volatils. Cependant les effets de solvatation particuliers des LI induisent des différences de réactivités et de sélectivités des systèmes catalytiques mais restent mal compris et peu pris en compte.

Chauvin a envisagé que le caractère non-coordonnant des LI entraînait la formation d'intermédiaire « nu » lors de la réaction d'hydrogénation par des catalyseurs à base de rhodium expliquant la plus grande activité du système catalytique dans les LIs comparativement aux solvants organiques comme il est expliqué dans la figure .*(Gutel, 2007)*



Par contre, le caractère non-dissociant des LI explique la chute de réactivité des catalyseurs à base de ruthénium. En effet, l'absence de traces d'eau empêche la dissociation des ligands chlorure du catalyseur et inhibe la réaction *(Gutel, 2007).*



2

I.7.3.1 Leur intérêt en catalyse

Les LI offrent l'opportunité de :

- Travailler en phase liquide dans un large domaine de température
- Faciliter les échanges de chaleurs
- Utiliser des procédés électrochimiques
- Utiliser les micro-ondes pour accélérer les réactions103
- Travailler à pression réduite

De plus, la variation des anions et des cations du LIs permettent d'ajuster les propriétés physico-chimiques du système selon les besoins en particulier modifié :

- La miscibilité relative des substrats/produits vis-à vis des LIs
- La réactivité du LIs vis-à vis du catalyseur
- Le caractère acide et/ou protique des Lis

Mais les principaux avantages de l'utilisation des LI en tant que solvants de réaction catalytique sont : la possibilité d'immobilisation du catalyseur dans la phase LI et une séparation aisée des produits de réaction du système catalytique. *(Gutel, 2007).*

I.8 Conclusion

Cette étude bibliographique a souligné l'intérêt des liquides ioniques dans le génie des procédés. Les propriétés physico-chimiques, la stabilité thermique, le vaste domaine liquide et la très faible tension de vapeur en font des milieux propices pour leur utilisation dans les procédés industriels. Dans ce chapitre, on a essayé de présenter simplement le thème liquide

ionique on passant par leur définition, historique, propriétés et quelques applications dans le domaine industriel.

De plus le besoin d'une chimie plus propre est pleinement reconnu et les solvants verts peu à peu intégrés dans le processus normal de développement de composés et de processus chimiques, c'est pour ce là, on a basé sur les liquides ioniques qui ont des propriétés physico-chimique très importants.

Chapitre II : Méthodes d'Intelligence Computationnelle

Chapitre II : Méthodes d'Intelligence Computationnelle : Généralités

II.1 Introduction

L'informatique est la science du traitement automatique de l'information. Son développement est souvent confondu avec celui des machines de traitement : les ordinateurs. Depuis les débuts d'années 1946 jusqu'à aujourd'hui, les ordinateurs sont devenus de plus en plus puissants.

Cependant, cette augmentation de puissance ne permet pas de toujours résoudre les problèmes d'une application informatique dans un domaine particulier. L'idée s'est donc installée que ce n'était peut-être pas tant le matériel que le logiciel qui pêchait par manque de puissance. La construction de logiciels s'appuie sur plusieurs approches. Deux parmi les plus utilisées sont l'approche algorithmique et l'approche basée sur la connaissance.

Une approche algorithmique nécessite l'écriture du processus à suivre pour résoudre le problème. Lorsque le problème est complexe, ce peut être une étape coûteuse ou impossible. D'autre part, l'Intelligence Computationnelle (IC) est un successeur de l'intelligence artificielle. 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 (comme dans le cas des systèmes flous). Les réseaux de neurones artificiels sont une branche de l'intelligence computationnelle qui est très proche du domaine de l'apprentissage automatique.

Dans ce chapitre on va laisser tomber la lumière sur les méthodes d'intelligence computationnelle tel que :

- Réseaux de neurones artificiels (RNA)

- Système adaptif a inférence neuro-flou (ANFIS)
- Machine à vecteur de support (SVM)
- Machine à vecteur de support aux moindres carrés(LSSVM)

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), ils infèrent des propriétés émergentes permettant de solutionner des problèmes jadis qualifiés de complexes.

II.2.1 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. *(Rizkalla, 2005)*

II.2.2 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).

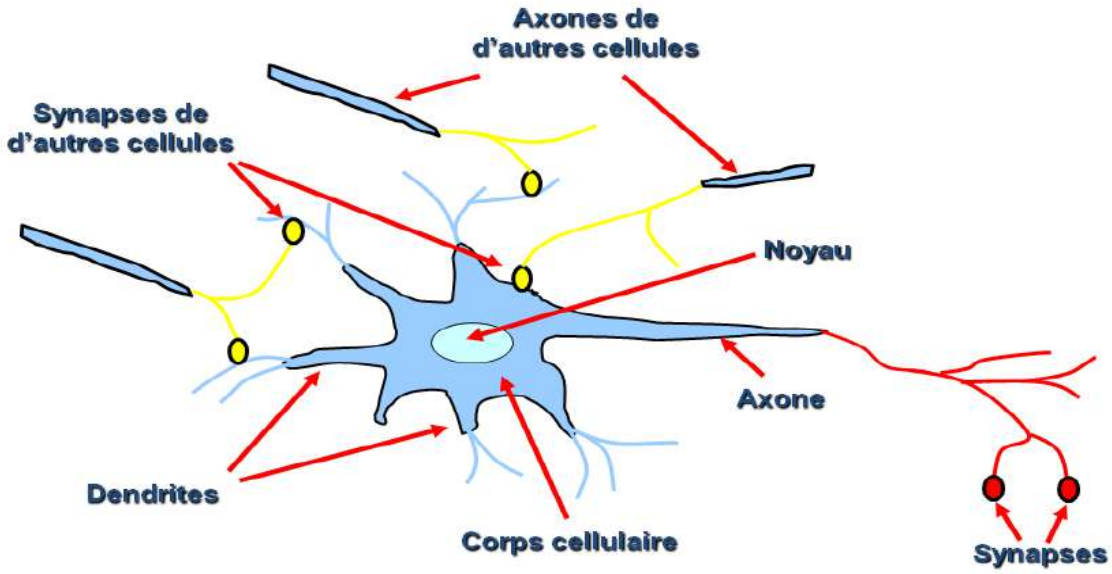


Figure II.1 : Neurone biologique. *(Kasri, 2010)*

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. *(Si-Moussa, 2010).*

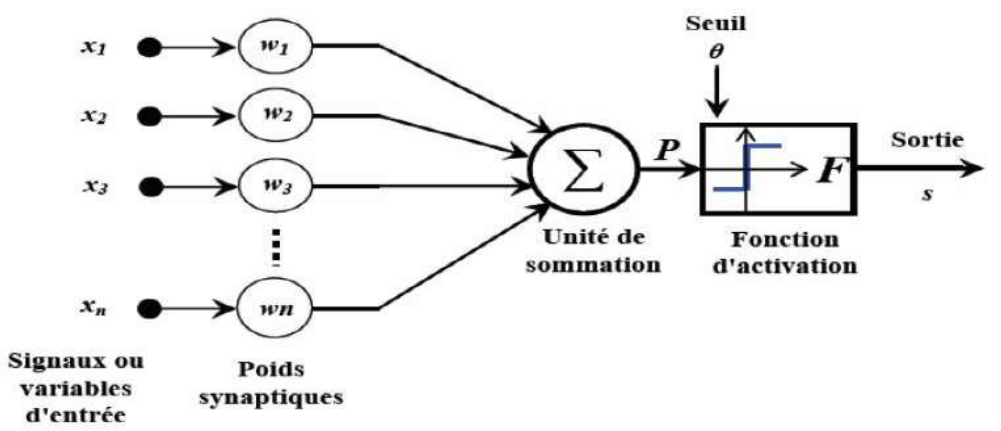


Figure II. 2: Schéma d'un neurone artificiel *(Si-Moussa, 2010).*

La figure II. 3 résume la correspondance entre le neurone artificiel et le neurone biologique. On remarque dans cette description schématique, que chaque composant ou constituant du neurone artificiel trouve son inspiration biologique.

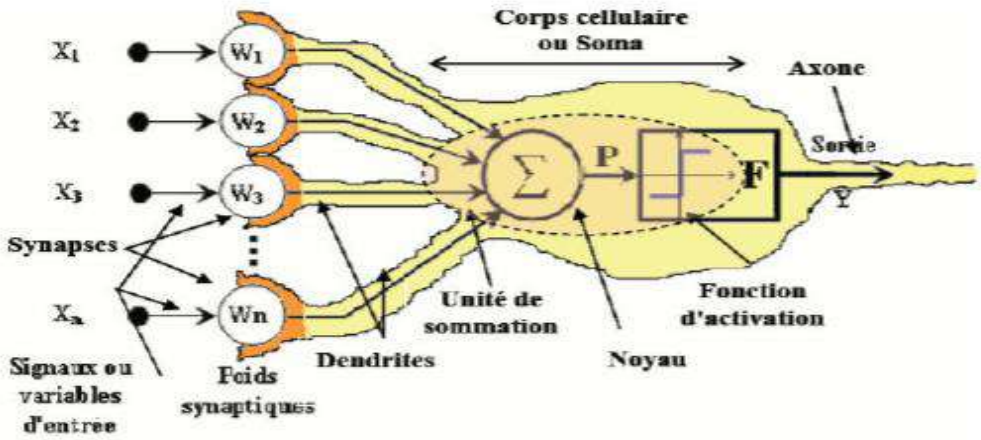


Figure II.3: Schéma de comparaison entre un neurone biologique et un neurone artificiel *(Si-Moussa, 2010)*

II.2.3 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 de type feedforward (cas de perceptrons multicouches), dans lequel chaque sortie d'une couche est alimentée en avant à la couche qui la succède où elle est traitée. Les sorties de la dernière couche constituent les sorties au monde réel *(Si-Moussa, 2010).*

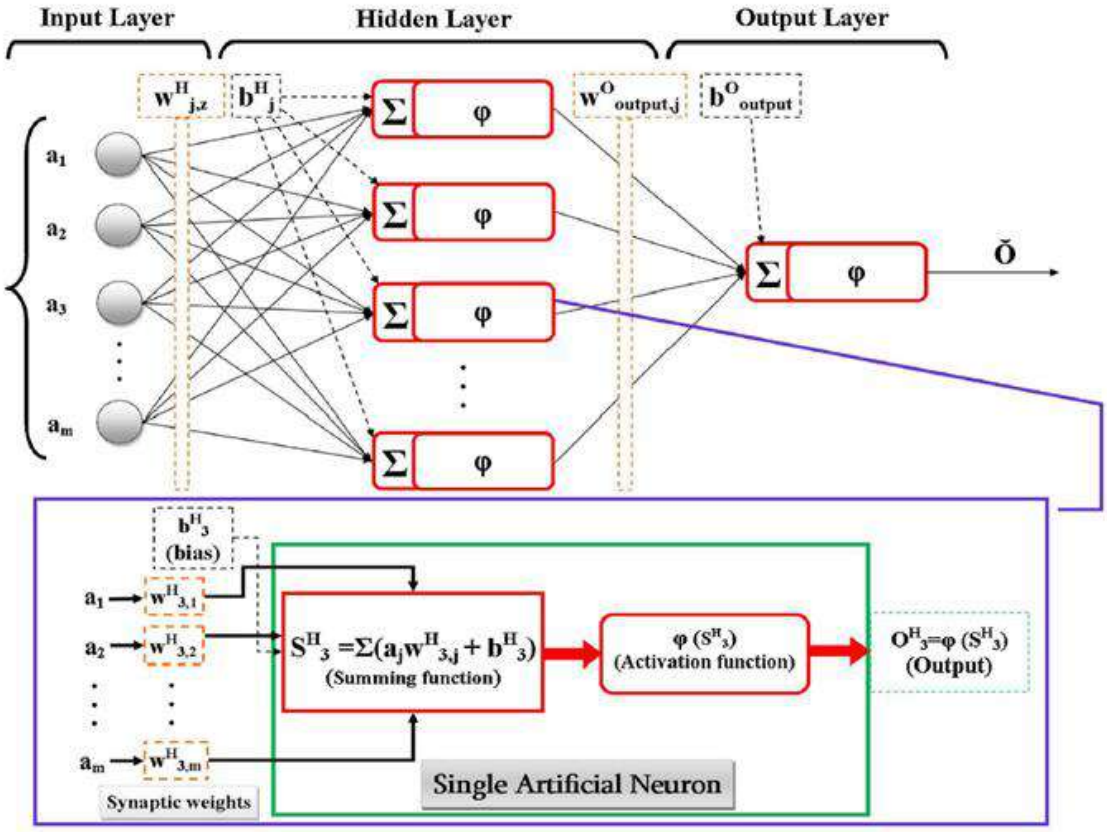


Figure II.4: Structure de Réseau de neurones artificiels de type feedforward multicouches *(Baghban, 2015)*

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 cette somme 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 de sortie comprises dans le domaine des données d'apprentissage. Par conséquent, l'objectif de la modélisation neuronale est de réduire au minimum les erreurs de prédiction de l'ensemble des données de validation présentées au réseau à la fin de l'étape d'apprentissage. *(Si-Moussa, 2010)*

II.2.4 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 tableau II.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. *(Si-Moussa, 2010)*

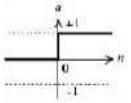
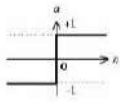
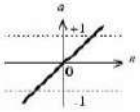
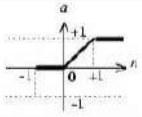
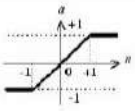
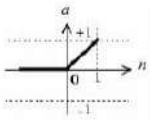
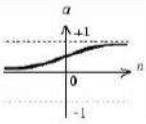
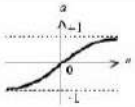
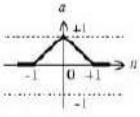
Nom de la fonction	Relation entrée / sortie	Graphe de la fonction	Fonction MATLAB® correspondante
Seuil	$a = 0$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlim
Seuil symétrique	$a = -1$ si $n < 0$ $a = 1$ si $n \geq 0$		hardlims
Linéaire	$a = n$		purelin
Linéaire saturée	$a = 0$ si $n < 0$ $a = n$ si $0 \leq n \leq 1$ $a = 1$ si $n > 1$		satlin
Linéaire saturée symétrique	$a = -1$ si $n < -1$ $a = n$ si $-1 \leq n \leq 1$ $a = 1$ si $n > 1$		satlins
Linéaire positive	$a = 0$ si $n < 0$ $a = n$ si $n \geq 0$		poslin
Logistique (Sigmoïde logarithmique)	$a = \frac{1}{1+e^{-n}}$		logsig
Tangente hyperbolique	$a = \frac{e^{+n} - e^{-n}}{e^{+n} + e^{-n}}$		tansig
Triangulaire de base	$a = 1 - \lvert n \rvert$, $-1 \leq n \leq +1$ $a = 0$, *autrement*		tribas

Tableau II.1 : Fonctions de transfert usuelles, a=f(n) *(Si-Moussa, 2010)*

II.2.5 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.2.5.1 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 préférence linéaire.

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. *(Si-Moussa, 2010).*

II.2.5.2 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 de généralisation. 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 *(Si-Moussa, 2010)*.

II.2.5.3 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 rétro propagation, 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 préférable 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 continues 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 continues sans bornes connues, la fonction d'activation linéaire ***(purelin)*** est préférable.

II.2.6 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 Système d'Inférence Neuro-Flou Adaptatif (ANFIS)

Le Système d'inférence neuro-flou adaptatif (Adaptif Neuro Fyzzy Inference System, ANFIS) est une combinaison de deux méthodes informatisées de RNA et de logique floue (Jang, 1993). La logique floue a la capacité de changer les aspects qualitatifs de la connaissance humaine et de mieux comprendre le processus d'analyse quantitative précise. Cependant, il n'a pas de méthode définie qui puisse être utilisée comme guide dans le processus de transformation et de pensée humaine en système d'inférence floue à base de règles (FIS), et il faut également beaucoup de temps pour ajuster les fonctions d'appartenance (MF) (Jang 1993). Contrairement à ANN, il a une plus grande capacité d'apprentissage à s'adapter à son environnement. Par conséquent, ANN peut être utilisé pour ajuster automatiquement les fonctions d'appartenance et réduire le taux d'erreurs dans la détermination des règles en logique floue. Cette section décrit en détail l'architecture d'ANFIS, les SIF, la flexibilité du réseau et l'algorithme d'apprentissage hybride. *(Suparta et Maulana Alhasa, 2016)*

II.3.1 Architecture d'ANFIS :

L'architecture ANFIS est un réseau adaptatif qui utilise un algorithme d'apprentissage supervisé sur apprentissage, qui a une fonction similaire au modèle du système d'inférence flou Takagi – Sugeno. La figure II.5 a, b montres le mécanisme de raisonnement flou du schéma pour le modèle Takagi-Sugeno et l'architecture ANFIS. Pour simplifier, supposons qu'il existe deux entrées x et y et une sortie f. Deux règles ont été utilisées dans la méthode «si-Alors» pour le modèle Takagi-Sugeno, comme suit:

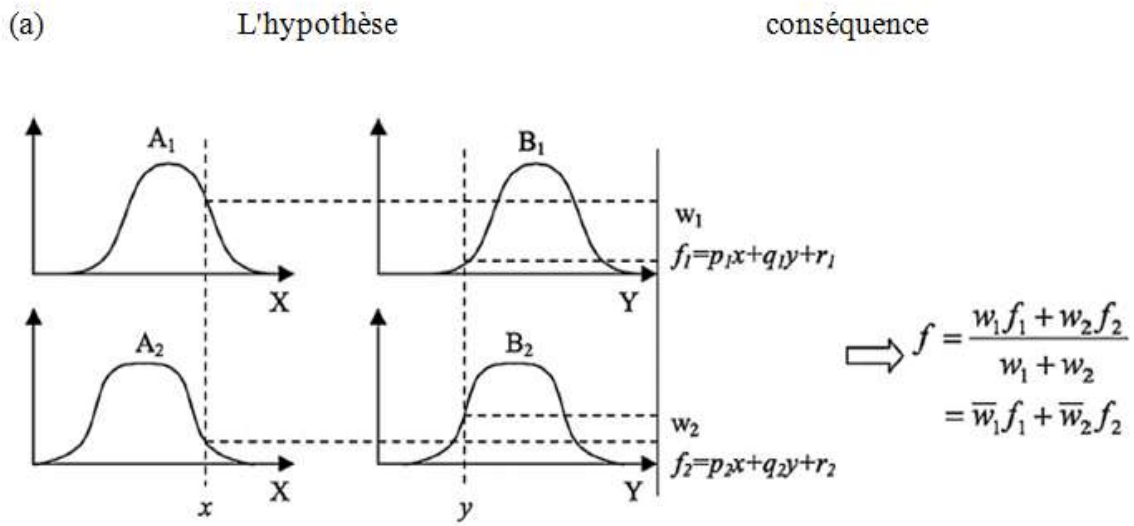
Règle1 : si $x = A_1$ et $y = B_1$ Alors $f_1 = p_1x + q_1x + r_1$ II.4

Règle2 : si $x = A_2$ et $y = B_2$ Alors $f_2 = p_2x + q_2x + r_2$ II.5

Où A_1, A_2 et B_1, B_2 sont les fonctions d'appartenance de chaque entrée x et y (partie des locaux), tandis que p_1, q_1, r_1 et p_2, q_2, r_2 sont des paramètres linéaires dans la partie -Alors (partie consécutive) de Takagi– Modèle d'inférence floue Sugeno. *(Suparta et Maulana Alhasa, 2016)*

En référence à la Figure II.5, l'architecture ANFIS a cinq couches. Les première et quatrième couches contiennent un nœud adaptatif, tandis que les autres couches contiennent un nœud fixe. Une brève description de chaque couche est la suivante:

Couche 1: chaque nœud de cette couche s'adapte à un paramètre de fonction. La sortie de chaque nœud est un degré de valeur d'appartenance donné par l'entrée de la Figure II.5.



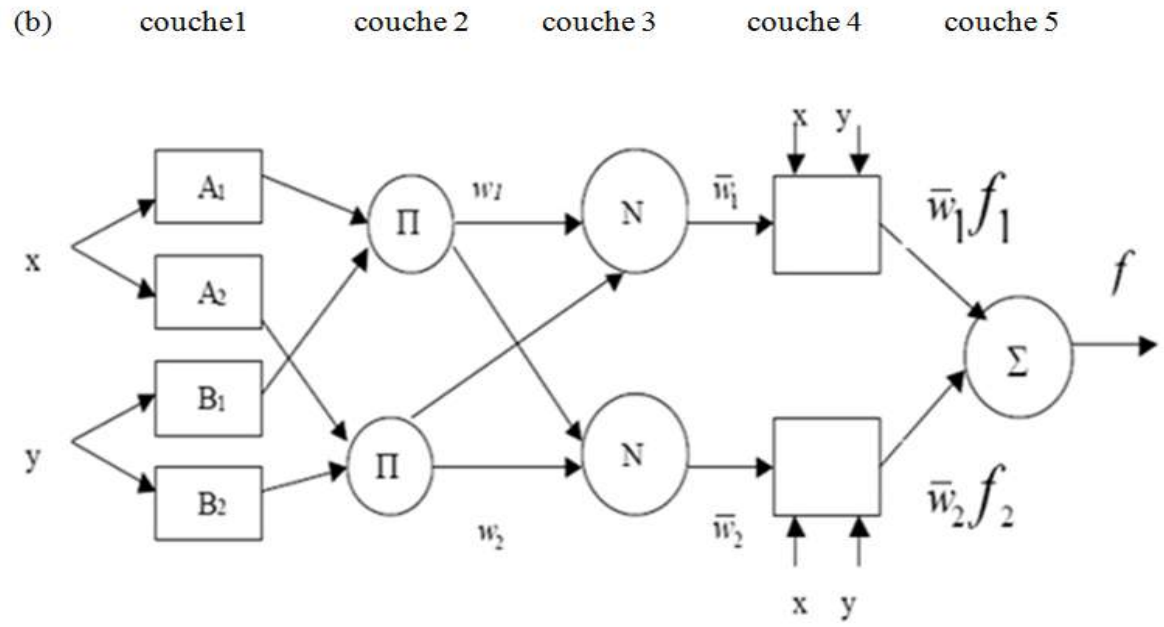


Figure II.5 : (a) Un modèle flou Takagi-Sugeno de premier ordre avec deux entrées et deux règles. (b) L'architecture ANFIS. *(Suparta et Maulana Alhasa, 2016)*

La fonction d'appartenance (Membership Function) peut être une fonction d'appartenance gaussienne (équation II.6), une fonction d'appartenance à une cloche généralisée (équation II.7) ou un autre type de fonction d'appartenance.

$$\mu_{Ai} = exp\left[-\left(\frac{x-c_i}{2a_i}\right)^2\right] \qquad \text{II.6}$$

$$\mu_{Ai}(x) = \frac{1}{1+\left|\frac{x-c_i}{a_i}\right|^{2b}} \qquad \text{II.7}$$

$$O_{1,i} = \mu_{Ai}(x), \quad i = 1,2 \qquad \text{II.8}$$

$$O_{1,i} = \mu_{Bi-2}(y), \quad i = 3,4 \qquad \text{II.9}$$

Où μ_{Ai} et μ_{Bi-2} sont le degré de fonctions d'appartenance pour les ensembles flous A_i et B_i, respectivement, et $\{a_i, b_i, c_i\}$ sont les paramètres d'une fonction d'appartenance pouvant modifier la forme de la fonction d'appartenance. Les paramètres de cette couche sont généralement appelés paramètres de site.

Couche 2: chaque nœud de cette couche est fixe ou non adaptable, et le nœud de cercle est étiqueté P. Le nœud de sortie est le résultat de la multiplication du signal entrant dans le nœud et transmis au nœud suivant. Chaque nœud de cette couche représente la puissance de

déclenchement de chaque règle. Dans la deuxième couche, l'opérateur de norme T à performances générales, tel qu'AND, est appliqué pour obtenir la sortie.

$$O_{2i} = w_i = \mu_{Ai}(x) * \mu_{Bi}(y), \quad i = 1,2 \qquad \text{II.10}$$

Où w_i est la sortie qui représente la force de frappe de chaque règle.

Couche 3: chaque nœud de cette couche est fixe ou non adaptable et le nœud de cercle est étiqueté N. Chaque nœud est un calcul du rapport entre la i-ème puissance de tir des règles et la somme de toutes les forces de tir de toutes les règles. Ce résultat est connu sous le nom de force de tir normalisée.

$$O_{3i} = \overline{w}_i = \frac{w_i}{\sum_i w_i} \qquad \text{II.11}$$

Couche 4: chaque nœud de cette couche est un nœud adaptatif d'une sortie, avec une fonction de nœud définie comme :

$$O_{4i} = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \qquad \text{II.12}$$

Où $\overline{w}_i$ est la puissance de tir normalisée de la couche précédente (troisième couche) et $(p_i x+q_i y+r_i)$ est un paramètre dans le nœud. Les paramètres de cette couche sont appelés paramètres conséquents.

Couche 5: le nœud unique de cette couche est un nœud fixe ou non adaptatif qui calcule la sortie globale en tant que somme de tous les signaux entrants provenant du nœud précédent. Dans cette couche, un nœud de cercle est étiqueté comme Σ. *(Suparta et Maulana Alhasa, 2016)*

$$O_{5i} = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \qquad \text{II.13}$$

II.4 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. 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. *(Mohamadally et Fomani, 2006)*

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.4.1 Historique de SVM

Les séparateurs à vastes marges reposent sur deux idées clés : la notion de marge maximale et la notion de fonction noyau. Ces deux notions existaient depuis plusieurs années avant qu'elles ne soient mises en commun pour construire les SVM.

L'idée des hyperplans à marge maximale a été explorée dès 1963 par Vladimir Vapnik et A. Lerner, et en 1973 par Richard Duda et Peter Hart dans leur livre *Pattern Classification*. Les fondations théoriques des SVM ont été explorées par Vapnik et ses collègues dans les années 70 avec le développement de la théorie de Vapnik-Chervonenkis, et par Valiant et la théorie de l'apprentissage PAC.

L'idée des fonctions noyaux n'est pas non plus nouvelle : le théorème de Mercer date de 1909, et l'utilité des fonctions noyaux dans le contexte de l'apprentissage artificiel a été montrée dès 1964 par Aizermann, Bravermann et Rozoener.

Ce n'est toutefois qu'en 1992 que ces idées seront bien comprises et rassemblées par Boser, Guyon et Vapnik dans un article, qui est l'article fondateur des séparateurs à vaste marge. L'idée des variables ressorts, qui permet de résoudre certaines limitations pratiques importantes, ne sera introduite qu'en 1995. À partir de cette date, qui correspond à la publication du livre de Vapnik, les SVM gagnent en popularité et sont utilisés dans de nombreuses applications.

Un brevet américain sur les SVM est déposé en 1997 par les inventeurs originels. *(https://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support)*

II.4.2 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. *(Kharroubi, 2002)*

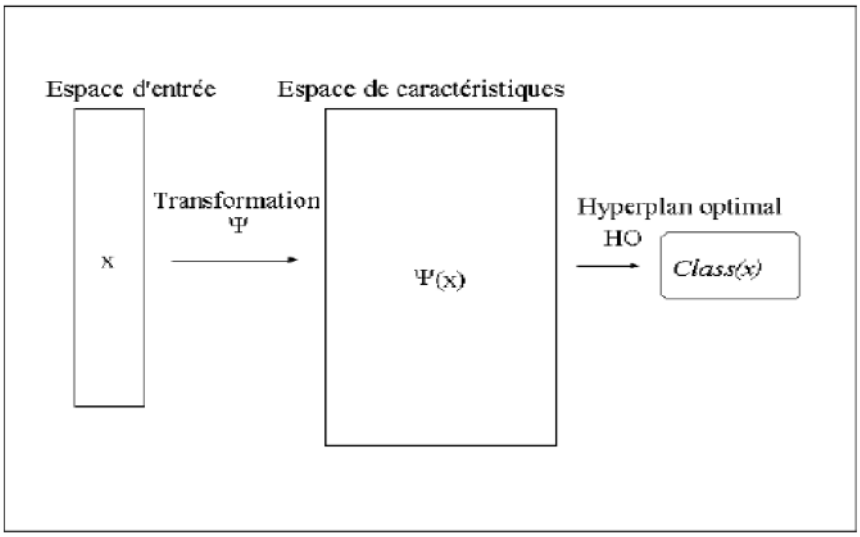


Figure II.6 : Principe des techniques SVM *(Chouih , 2018)*

II.4.3 SVM : Principe de fonctionnement général

II.4.3.1 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. *(Mohamadally et Fomani 2006)*

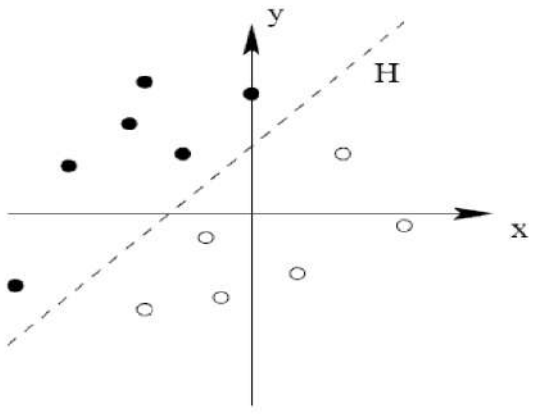


Figure II.7: Exemple d'un hyperplan séparateur *(Mohamadally et Fomani , 2006)*

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

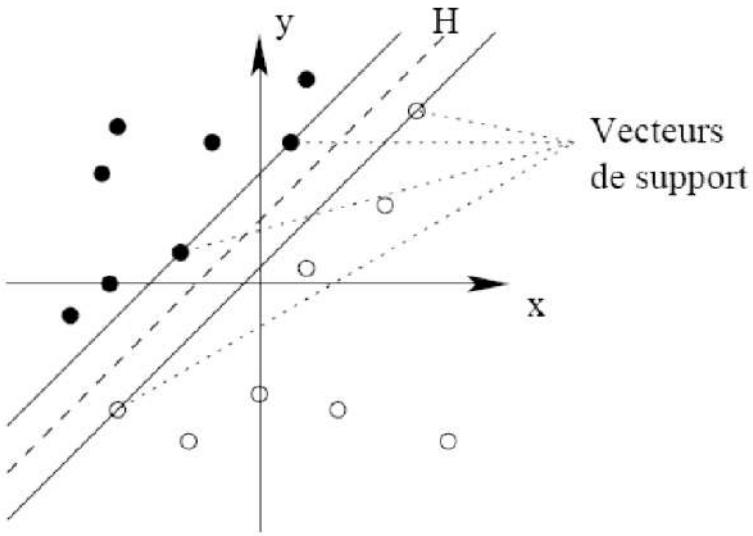


Figure II.8: Exemple de vecteurs de support *(Mohamadally et Fomani, 2006)*

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. *(Mohamadally et Fomani, 2006)*

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. *(Mohamadally et Fomani, 2006)*

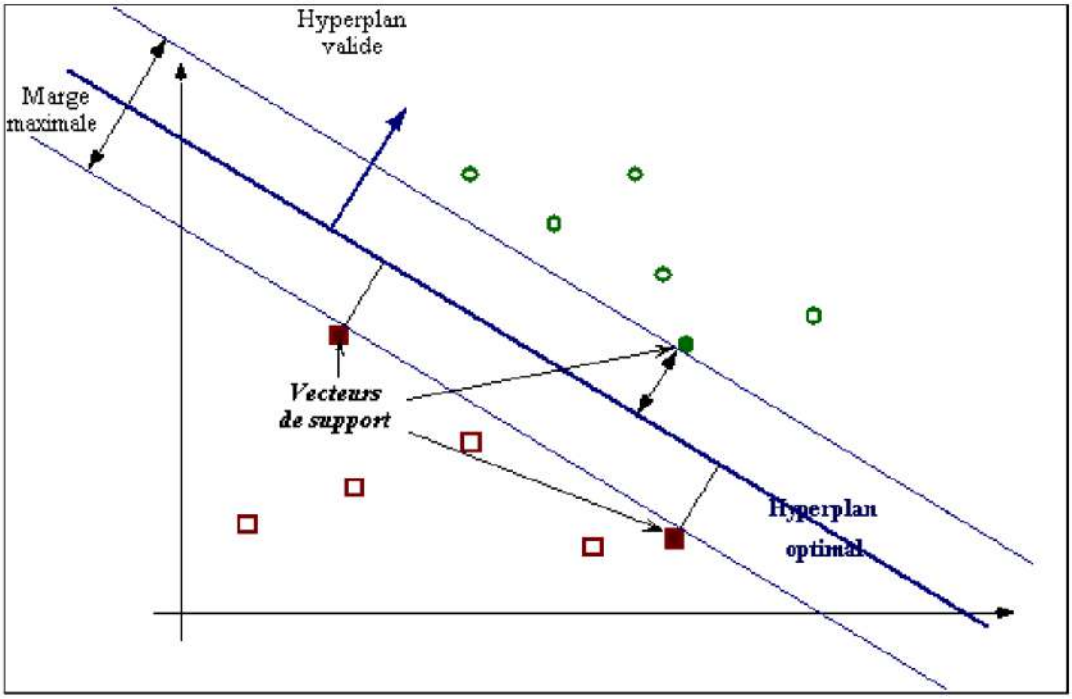


Figure II.9 : Exemple de marge maximal (hyperplan valide) *(Mohamadally et Fomani, 2006)*

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. *(Mohamadally et Fomani, 2006)*

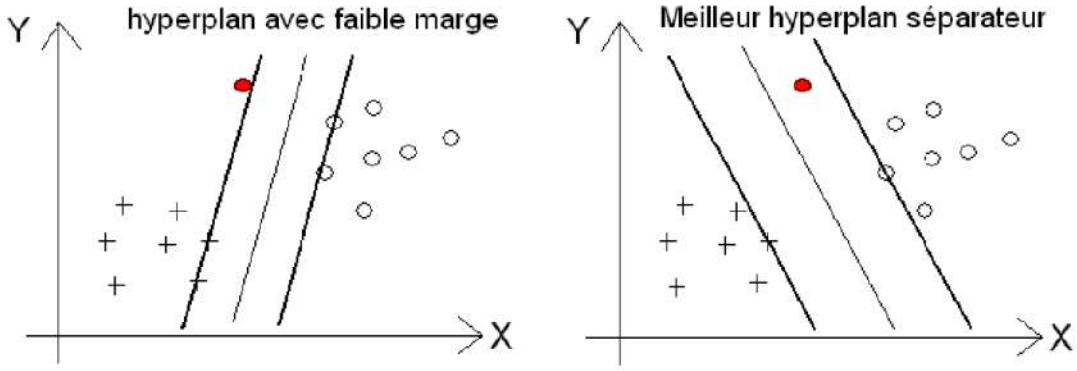


Figure II.10 : L'Hyperplan avec faible marge et le Meilleur hyperplan séparateur *(Mohamadally et Fomani, 2006)*

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 ≪ + ≫. *(Mohamadally et Fomani 2006)*

II.4.3.2 Linéarité et non linéarité

Parmi les modèles des SVM, on constate les cas linéairement séparable 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. *(Mohamadally et Fomani, 2006)*

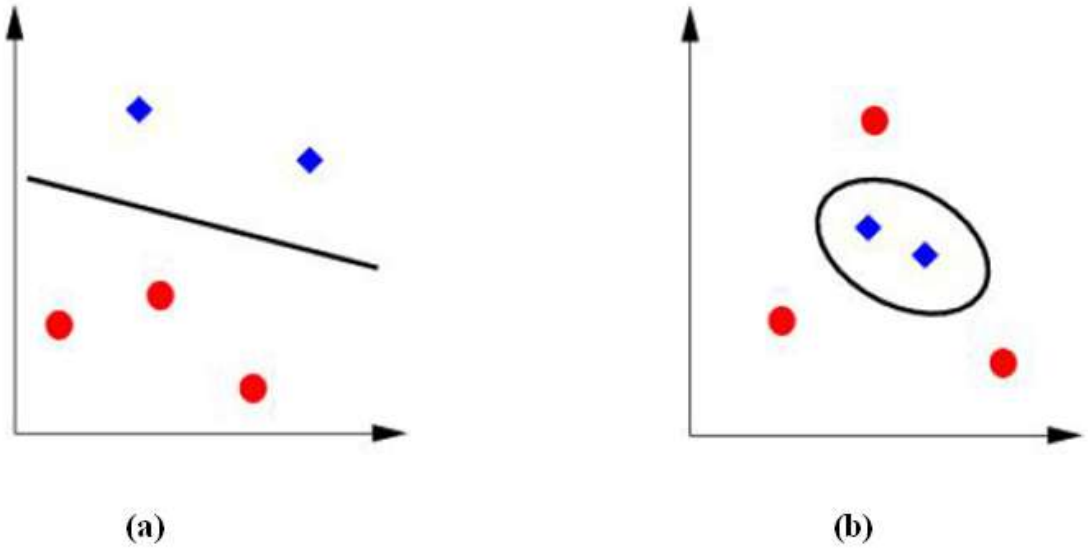


Figure II.11 : (a) cas linéairement séparable et (**b)** cas non linéairement séparable *(Mohamadally et Fomani, 2006)*

II.4.3.3 Cas non linéaire

Pour surmonter les inconvénients des cas non linéairement séparable, l'idée des SVM est de changer l'espace des données. La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouvel espace. On va donc avoir un changement de dimension. Cette nouvelle dimension est appelé ≪ espace de re-description ≫.

En effet, intuitivement, plus la dimension de l'espace de re-description est grande, plus la probabilité de pouvoir trouver un hyperplan separateur entre les exemples est élevée. Ceci est illustre par le schéma suivant : *(Mohamadally et Fomani, 2006).*

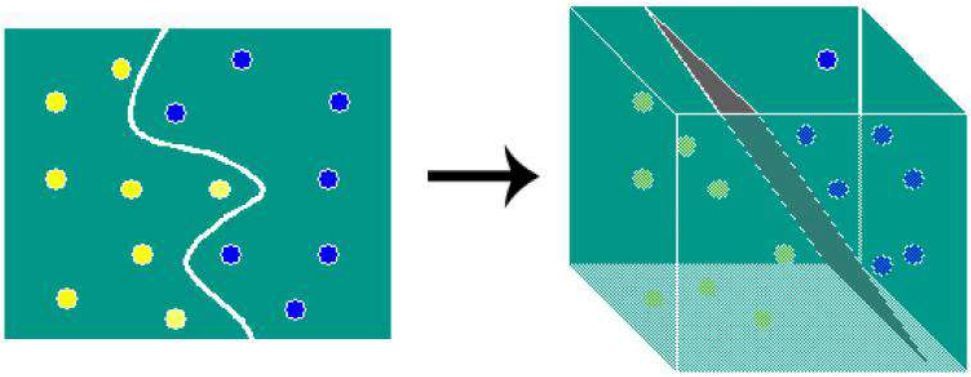


Figure II.12: Exemple de changement de l'espace de données. *(Mohamadally et Fomani, 2006)*

On a donc une transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de plus grande dimension. Cette transformation non linaire est réalisée *via* une fonction noyau. En pratique, quelques familles de fonctions noyau paramétrables sont connues et il revient à l'utilisateur de SVM d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. On peut citer les exemples de noyaux suivants : polynomiale, gaussien, sigmoïde et laplacien. *(Mohamadally et Fomani, 2006)*

II.4.3.4 Illustration de transformation de cas non linéaire : le cas XOR

Le cas de XOR n'est pas linéairement séparable, si on place les points dans un plan à deux dimensions, on obtient la figure suivante : Coordonnées des points : (0,0) ; (0,1) ; (1,0) ; (1,1)

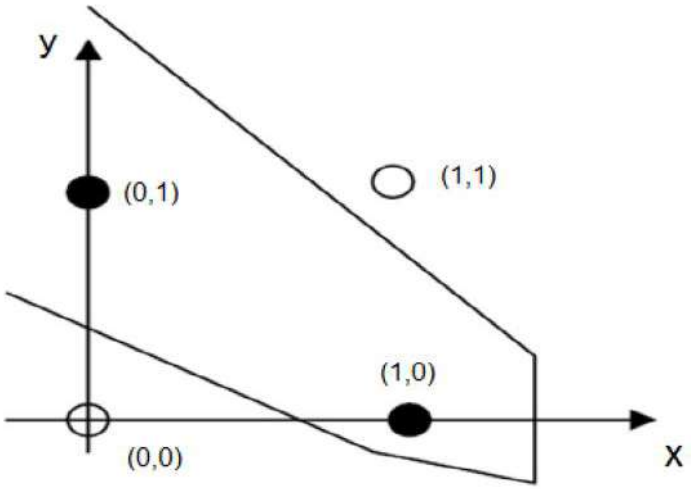


Figure II.13 : Illustration de cas non linéairement séparable (le cas XOR) *(Chouih, 2018)*

Si on prend une fonction polynomiale (x, y) → (x, y, z) qui fait passer d'un espace d dimensions 2 à un espace de dimension 3, on obtient un problème en trois dimensions linéairement séparable :

(0,0) → (0, 0,0) (1,0) → (1, 0,0)
(0,1) → (0, 1,0) (1,1) → (1, 1,1)

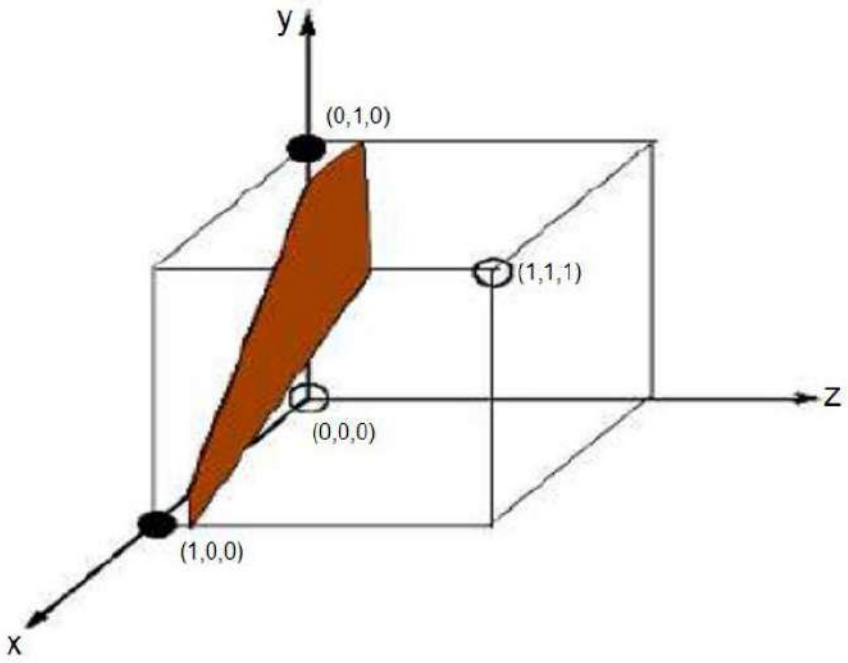


Figure II.14 : Illustration de passage d'un espace 2D à un espace 3D *(Chouih, 2018)*

II.4.4 Fondement mathématique

Nous allons détailler dans les paragraphes ci-dessous les principes mathématiques sur lesquels reposent les SVM.

II.4.4.1 Problème d'apprentissage

La notion d'apprentissage étant importante, nous allons commencer par effectuer un rappel. L'apprentissage par induction permet d'arriver à des conclusions par l'examen d'exemples particuliers. Il se divise en apprentissage supervisé et non supervisé. Le cas qui concerne les SVM est l'apprentissage supervisé. Les exemples particuliers sont représentés par un ensemble de couples d'entrée/sortie. Le but est d'apprendre une fonction qui correspond aux exemples vus et qui prédit les sorties pour les entrées qui n'ont pas encore été vues. Les entrées peuvent être des descriptions d'objets et les sorties la classe des objets donnés en entrée. *(Chouih, 2018)*

On s'intéresse à un phénomène f (éventuellement non déterministe) qui, a partir d'un certain jeu d'entrées x, produit une sortie y = f(x).

Le but est de retrouver cette fonction f à partir de la seule observation d'un certain nombre de couples entrée-sortie {(xi; yi) : i = 1, .., n} afin de " prédire " d'autres événements.

On considère un couple (X, Y) de variables aléatoires a valeurs dans ***X*** **x** ***Y***.

Seul le cas ***Y*** = {-1, 1} (classification) nous intéresse ici (on peut facilement étendre au cas Card (***Y***) = m > 2 et au cas ***Y*** =R). La distribution jointe de (X, Y) est inconnue.

Sachant qu'on observe un échantillon S = {(X1, Y1),..., (Xn, Yn)} de n copies indépendantes de (X, Y), on veut: construire une fonction h : $X \rightarrow Y$ telle que P (h(X) != Y) soit minimale. *(Mohamadally et Fomani, 2006)*

II.4.4.2 Classification a valeurs réelles

Plutôt que de construire directement h : $\boldsymbol{X} \rightarrow$ {-1, 1}, on construit :

f : $\boldsymbol{X} \rightarrow \boldsymbol{R}$ (ensemble des réels).

La classe est donnée par le signe de f ;

h = signe(f).

L'erreur se calcule avec P (h(X) != Y) = P(Yf(X) ≤0). Ceci donne une certaine idée de la confiance dans la classification. Idéalement, |Yf(X)| est proportionnel a P(Y|X).

Yf(X) représente la marge de f en (X, Y).

Le but à atteindre est la construction de f et donc h. Nous allons voir comment y parvenir. (*Mohamadally et Fomani 2006)*

- **Transformation des entrées**

Il est peut être nécessaire de transformer les entrées dans le but de les traiter plus facilement.

X est un espace quelconque d'objets.

On transforme les entrées en vecteurs dans un espace de caractéristiques ***F*** (feature space) par une fonction :

Φ: $\boldsymbol{X} \rightarrow \boldsymbol{F}$

F n'est pas nécessairement de dimension finie mais dispose d'un produit scalaire (espace de Hilbert). L'espace de Hilbert est une généralisation de l'espace euclidien qui peut avoir un nombre infini de dimensions.

La non-linéarité est traitée dans cette transformation, on peut donc choisir une séparation linéaire (on verra plus loin comment on arrive à ramener un problème non linéaire en un problème linéaire classique).

Des lors, il s'agit de choisir l'hyperplan optimal qui classifie correctement les données (lorsque c'est possible) et qui se trouve le plus loin possible de tous les points à classer. *(Mohamadally et Fomani, 2006)*

- **Problème primal**

Un point (x; y) est bien classe si et seulement si yf(x) > 0

Comme le couple (w, b) est défini à un coefficient multiplicatif prés, on s'impose yf(x) ≥ 1

On en déduit (en s'appuyant également sur le paragraphe précédent), le problème de minimisation sous contraintes suivantes :

$$\begin{cases} \min \frac{1}{2}||W||^2 \\ \forall i, y_i \ \ (w.x_i + b) > 1 \end{cases}$$

Il peut être en effet plus aisé de minimiser ||w||2 plutôt que directement ||w||. (*Mohamadally et Fomani , 2006)*

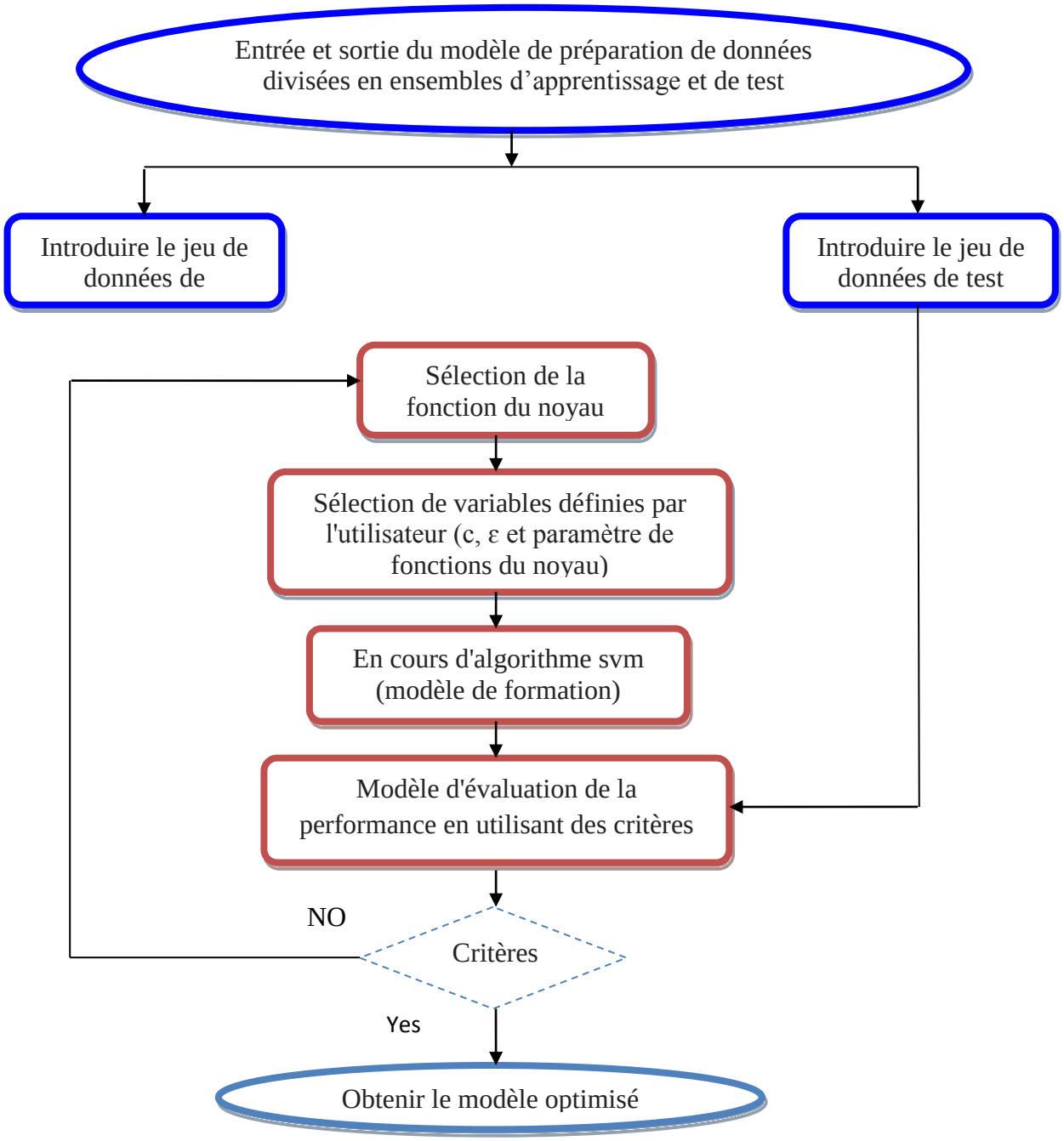


Figure II.15 : Organigramme des étapes pour la création du modèle SVM. *(Hashemkhani et al, 2015)*

II.5 Les machines à vecteurs avec support des moindres carrés (LSSVM)

II.5.1 Introduction

Suykens et Vandewalle (1999) 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. *(Shabri et Suhartono, 2012)*

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 *(Shabri et Suhartono, 2012)*

II.5.2 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. *(Shabri et Suhartono, 2012)*

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 :

$$f(x) = b + (\text{ø}(x), \omega) \qquad \text{II.14}$$

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 :

$$\text{Min}_{\text{ w, b, e}}\, j\,(w, e)= \frac{1}{2}\|w\|^2 +\frac{1}{2}\ \gamma\ \sum_{k=1}^{n} e_k^2 \qquad \text{II.15}$$

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

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:

$$L_{lssvm} = \frac{1}{2}\|w\|^2 +\frac{1}{2}\gamma \sum_{k=1}^{n} e_k^2-\sum_{k=1}^{n} a_k\{(\omega.\ \text{ø}\ (x_k)) +b+e_k -y_k\} \qquad \text{II.16}$$

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

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

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

$\frac{\partial Llssvm}{\partial w} =0 \rightarrow a_k = \gamma e \quad k=1,\ldots\ldots, N$

$\frac{\partial Llssvm}{\partial w} =0 \rightarrow (\omega.\ \text{ø}\ (x_k)) +b+ e_k-y_k =0 \quad k=1,\ldots..N$

En considérant $\alpha = [\alpha 1,\ldots, \alpha N]$, $lv = [1,\ldots l]$, $Y = [y1,\ldots yN]$ 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 *(Razavi et al, 2019)*

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.16

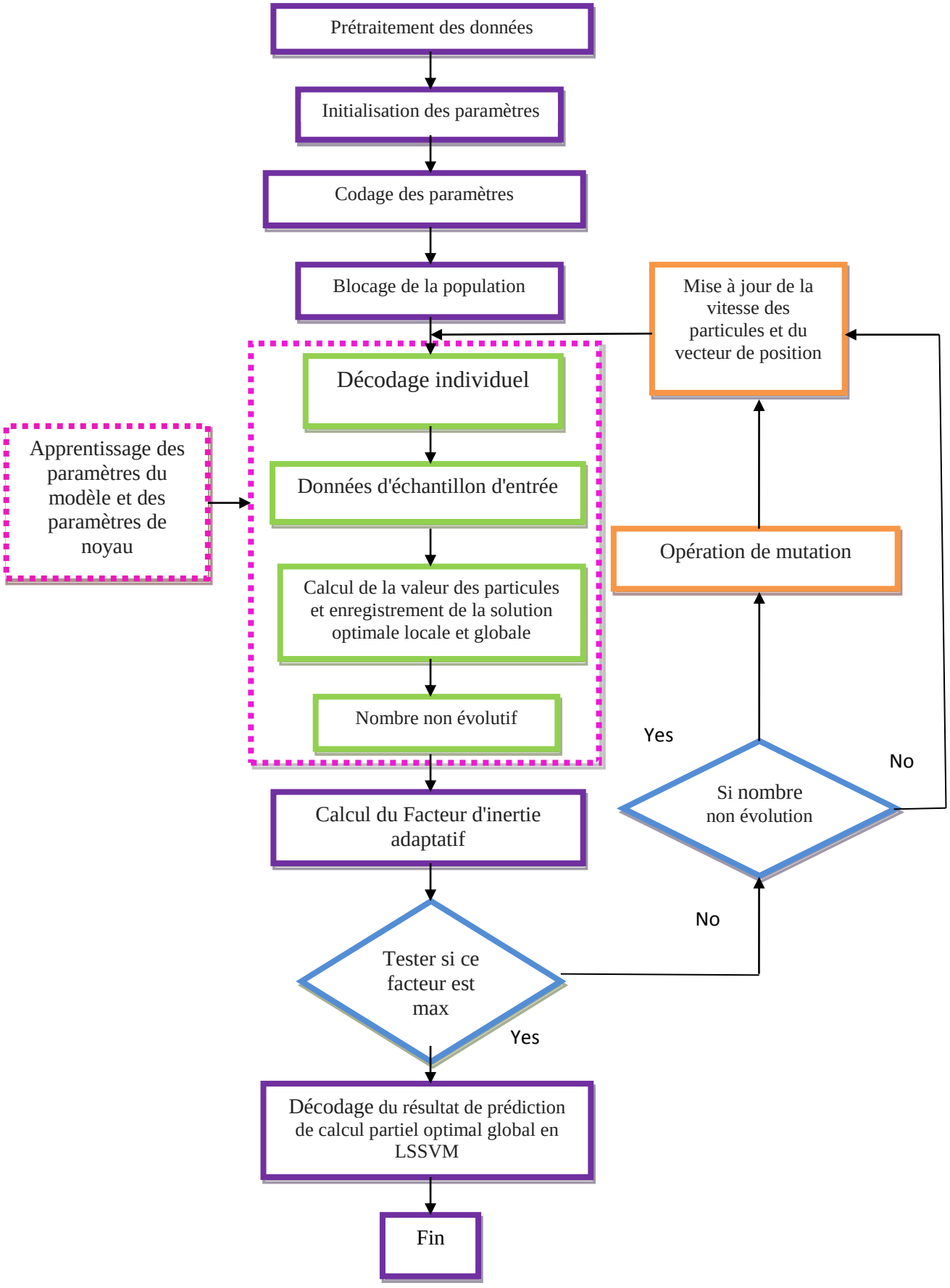


Figure II.16 : Démonstration schématique de l'algorithme LSSVM formé *(Razavi et al, 2019)*

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

II.6 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 (EQMR); Erreur relative absolue moyenne (ERAM); Coefficient de détermination (R^2); Coefficient de corrélation (R); Facteur de précision (A_f) et facteur de biais (B_f) qui sont données par les formules suivantes *(Soleimani et al, 2017)* :

$$EQM = \frac{1}{N}\sum_{i=1}^{N}({y_i}^{exp} - {y_i}^{cal})^2 \quad \text{II.17}$$

$$EAM = \frac{1}{N}\sum_{i=1}^{N}|{y_i}^{exp} - {y_i}^{cal}| \quad \text{II.18}$$

$$EQMR = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{{y_i}^{exp} - {y_i}^{cal}}{{y_i}^{exp}}\right)^2 \quad \text{II.19}$$

$$ERAM\% = 100 * \frac{1}{N}\sum_{i=1}^{N}\left|\frac{{y_i}^{exp} - {y_i}^{cal}}{{y_i}^{exp}}\right| \quad \text{II.20}$$

$$ER\% = 100 * \left|\frac{{y_i}^{exp} - {y_i}^{cal}}{{y_i}^{exp}}\right| \quad \text{II.21}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}({y_i}^{exp} - {y_i}^{cal})^2}{\sum_{i=1}^{N}({y_i}^{exp} - \overline{y^{exp}})^2} \quad \text{II.22}$$

$$R = \frac{\sum_{i=1}^{N}({y_i}^{exp} - \overline{y^{exp}}) \times \left({y_i}^{cal} - \overline{y^{cal}}\right)}{\sqrt{\sum_{i=1}^{N}({y_i}^{exp} - \overline{y^{exp}})^2} \times \sqrt{\sum_{i=1}^{N}\left({y_i}^{cal} - \overline{y^{cal}}\right)^2}} \quad \text{II.23}$$

- **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

$$B_f = 10^{\left(\frac{\sum_{i=1}^{N} log\frac{y_i^{cal}}{y_i^{exp}}}{N}\right)} \qquad \text{II.24}$$

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:

$$A_f = 10^{\left(\frac{\sum_{i=1}^{N} \left|log\frac{y_i^{cal}}{y_i^{exp}}\right|}{N}\right)} \qquad \text{II.25}$$

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:

$$\overline{y^{exp}} = \frac{1}{N}\sum_{i=1}^{N} y_i{}^{exp} \qquad \text{II.26}$$

$$\overline{y^{cal}} = \frac{1}{N}\sum_{i=1}^{N} y_i{}^{cal} \qquad \text{II.27}$$

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:

$$k = \frac{\sum y_i^{exp} y_i^{cal}}{\sum (y_i^{cal})^2} \quad \text{II.28}$$

$$k' = \frac{\sum y_i^{exp} y_i^{cal}}{\sum (y_i^{exp})^2} \quad \text{II.29}$$

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

$$R_o{}^2 = 1 - \frac{\sum (y_i^{cal} - k y_i^{cal})^2}{\sum (y_i^{cak} - \overline{y_i^{cal}})^2} \quad \text{II.30}$$

$$R_o{}'^2 = 1 - \frac{\sum (y_i^{exp} - k' y_i^{exp})^2}{\sum (y_i^{exp} - \overline{y^{exp}})^2} \quad \text{II.31}$$

$$M = \frac{R^2 - R_0{}^2}{R^2} \leq 0.1 \quad \text{II.32}$$

$$N = \frac{R^2 - R_o{}'^2}{R^2} \leq 0.1 \quad \text{II.33}$$

$$R_m{}^2 = R^2 \times \left(1 - \sqrt{\left|R^2 - R_o{}^2\right|}\right) \geq 0.5 \quad \text{II.34}$$

- **Définitions des métriques Q^2 :** *(Todeschini et al, 2016)*

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:

$$Q_{F1}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{TR})^2} \quad \text{II.35}$$

$$Q_{F2}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2}{\sum_{i=1}^{n_{OUT}} (y_i - \bar{y}_{OUT})^2} \quad \text{II.36}$$

$$Q_{F3}{}^2 = 1 - \frac{\sum_{i=1}^{n_{OUT}} (y_i - \hat{y}_{i/i})^2 / n_{OUT}}{\sum_{i=1}^{n_{TR}} (y_i - \bar{y}_{TR})^2 / n_{TR}} \quad \text{II.37}$$

- **Le coefficient 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:

$$Q^2{}_{ccc} = \frac{2.\sum_{i=1}^{n_{OUT}}(y_i - \bar{y}_{EXP}).(\hat{y}_{i/i} - \bar{y}_{CAL})}{\sum_{i=1}^{n_{OUT}}(y_i - \bar{y}_{EXP})^2 + \sum_{i=1}^{n_{OUT}}(\hat{y}_{i/i} - \bar{y}_{CAL})^2 + n_{OUT}.(\bar{y}_{EXP} - \bar{y}_{CAL})^2} \qquad \text{II.38}$$

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:

$$Q^2{}_{ccc} = \frac{2.S_{XY}}{S_{XX} + S_{YY} + (\bar{X} - \bar{Y})^2} \qquad \text{II.39}$$

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:

$$Q^2{}_{Rm} = \frac{1}{2}.\left[R^2.\left(1 - \sqrt{R^2 - R^2{}_{0.XY}}\right) + R^2.\left(1 - \sqrt{R^2 - R^2{}_{0.YX}}\right)\right] = R^2.\left[1 - \frac{\sqrt{R^2 - R^2{}_{0.XY}} + \sqrt{R^2 - R^2{}_{0.YX}}}{2}\right] \qquad \text{II.40}$$

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.7 Conclusion

Dans ce chapitre, nous avons tenté de présenter de manière simple et résumé le concept de système d'apprentissage. Nous avons donné une vision générale et une vision purement mathématique sur les méthodes d'intelligence computationnelle qui sont des méthodes alternatives et complémentaires à l'expérimental et qui nous permet de facilité la résolution d'un problème mathématique.

Dans cette étude, quatre catégories des méthodes d'apprentissage de l'intelligence, à savoir RNA, ANFIS, SVM et LSSVM ont été utilisées pour estimer la tension superficielle de différents liquides ioniques.

Nous également présenté les critères et métriques statistiques que nous allons utiliser pour test la validité des modèles développés.

Chapitre III : Résultats et interprétations

Chapitre III : Résultats et interprétations

III.1 Introduction

Les techniques de modélisation décrites dans le chapitre 2 ont été appliquées aux données expérimentales de la tension superficielles de 40 mélanges binaires (LI+ Eau/Solvants organiques). Dans ce chapitre on présente la base de données élaborée ainsi que les résultats de modélisation par les méthodes de RNA, ANFIS, SVM et LSSVM.

III.2 Elaboration de la base de données

III.2.1 Données expérimentales de la tension superficielle

Les données expérimentales de la tension superficielle des mélanges binaires sont publiées à différentes températures et compositions du mélange. Les cations et les anions constituant les Lis considérés dans cette études sont représentés dans les tableaux III.1 et III.2 respectivement. La source et les domaines de température et de compostions des données des 40 mélanges binaires considérés dans cette étude sont résumés dans le tableau III.3.

Tableau III.1 : La liste des cations.

N^0	Nom du cation	Abréviation	Famille
1	1,3-dimethylimidazolium	[C1MIm]	Imidazolium
2	1-Ethyl-3-methylimidazolium	[C2MIm]	
3	1-Butyl-3-methylimidazolium	[C4MIm]	
4	1-Hexyl-3-methylimidazolium	[C6MIm]	
5	1-octyl-3-methylimidazolium	[C8MIm]	
6	1-butylpyridinium	[C4Py]	Pyridinium
7	N-octylisoquinolinium	[C8iQuin]	Quinolinium
8	Ethyl(2-hydroxyethyl)dimethylammonium	[N112(hoe)]	Ammonium
9	Propyl(2-hydroxyethyl)dimethylammonium	[N311(hoe)]	
10	trihexyltetradecylphosphonium	[P666(14)]	Phosphonium

Tableau III.2 : La liste des anions.

N⁰	Nom de l'anion	Abréviation
1	Dimethylphosphate	[DMP]
2	Methylsulfate	[MeSO4]
3	Tetrafluoroborate	[BF4]
4	Bis(trifluoromethylsulfonyl)imide	[BTI]
5	Octyl sulfate	[C8SO4]
6	Diethylphosphate	[DEP]
7	Ethylsulfate	[EtSO4]
8	Methane sulfonate	[MeSO3]
9	Methylsulfate	[MeSO4]
10	Dibutylphosphate	[DBP]
11	Nitrate	[NO3]
12	Bromide	[Br]
13	Dicyanamide	[Dca]

Tableau III.3 : Source et domaines des données expérimentales des mélanges binaires étudiés.

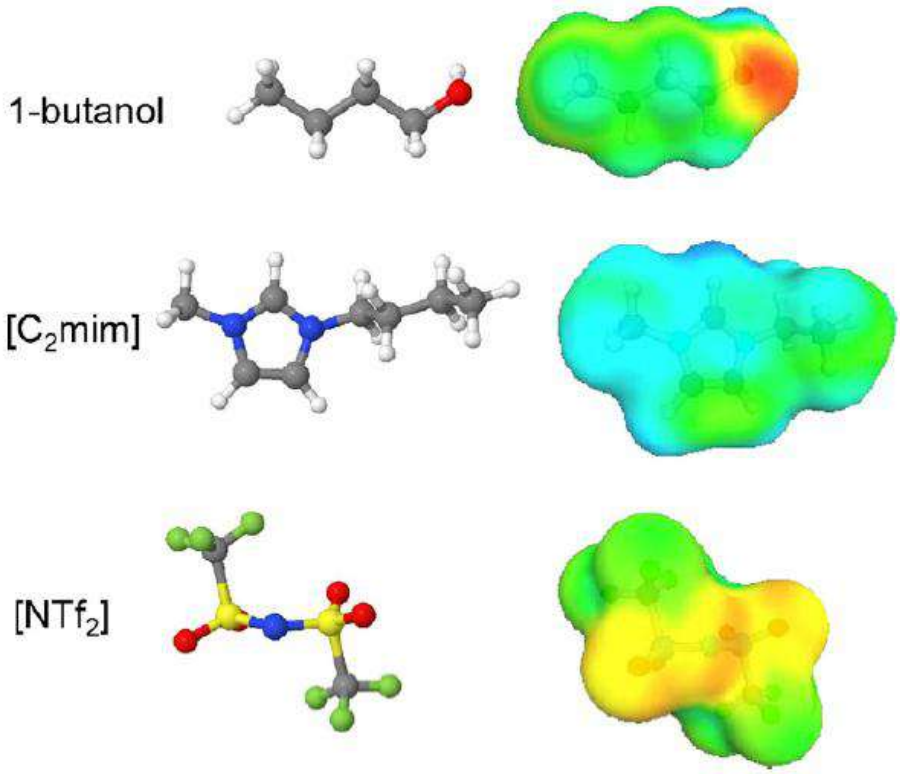
N	LI	Solvant	T (K)	X_{LI}	Np	Référence
1	[C1MIm][DMP]	Eau	298,15	[0,0000 – 1,0000]	16	*(Wan Normazlan et al.,2014)*
2	[C1MIm][DMP]	Méthanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al.,2014)*
3	[C1MIm][DMP]	Ethanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
4	[C1MIm][MeSO4]	Eau	[296,8 – 298,1]	[0,7080 – 1.0000]	10	*(Russo et Hoffmann, 2011)*
5	[C1MIm][MeSO4]	Méthanol	298,15	[0,0000 – 1,0000]	9	*(Domaoska et al, 2008)*
6	[C2MIm][BF4]	Eau	298,15	[0,0000 – 0,8582]	9	*(Rilo et al, 2009)*
7	[C2MIm][BTI]	Tetrahydrofuran	[293,15 – 308,15]	[0,0000– 1,0000]	44	*(Geppert-Rybczyoska et al, 2011)*
8	[C2MIm][BTI]	Acetonitrile	[293.15 – 313,15]	[0,0000 – 1,0000]	45	*(Geppert-Rybczyoska et al, 2011)*
9	[C2MIm][BTI]	Diméthyl sulfoxide	[293.15 – 313,15]	[0,0000 – 1,0000]	45	*(Geppert-Rybczyoska et al, 2013)*
10	[C2MIm][C8SO4]	Ethanol	298,15	[0,0000 – 0.9601]	18	*(Rilo et al, 2012)*
11	[C2MIm][DEP]	Eau	298,15	[0,0000 – 1,0000]	16	*(Wan Normazlan et al, 2014)*
12	[C2MIm][DEP]	Méthanol	298,15	[0,0000 – 1,0000]	22	*(Wan Normazlan et al, 2014)*
13	[C2MIm][DEP]	Ethanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
14	[C2MIm][DMP]	Eau	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*

15	[C2MIm][DMP]	Ethanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
16	[C2MIm][EtSO4]	Eau	298,15	[0,0062 – 0,5791]	12	*(Torrecilla et al, 2008)*
17	[C2MIm][MeSO3]	Eau	[300,2 – 303,2]	[0,5150 – 1,0000]	27	*(Russo et Hoffmann, 2011)*
18	[C2MIm][MeSO4]	Méthanol	298,15	[0.5988 – 1,0000]	7	*(Wang et al, 2011)*
19	[C4MIm][BF4]	Eau	298,15	[0,0000 – 0,8987]	15	*(Rilo et al, 2009)*
20	[C4MIm][BF4]	Ethanol	298,15	[0,0000 – 0,9014]	12	*(Rilo et al, 2009)*
21	[C4MIm][BTI]	1-propanol	298,15	[0,0000 – 1,0000]	11	*(Wandschneider et al, 2008)*
22	[C4MIm][BTI]	1-butanol	298,15	[0,0000 – 1,0000]	13	*(Wandschneider et al, 2008)*
23	[C4MIm][BTI]	Tetrahydrofuran	[293.15 – 308,15	[0,0000 – 1 ,0000]	40	*(Geppert-Rybczyoska et al, 2013)*
24	[C4MIm][BTI]	Acetonitrile	[293,15 – 313,15]	[0,0000 – 1,0000]	45	*(Geppert-Rybczyoska et al, 2013)*
25	[C4MIm][BTI]	Dimethyl sulfoxide	[293,15 – 313,15]	[0,0000 – 1,0000]	50	*(Geppert-Rybczyoska et al, 2013)*
26	[C4MIm][DBP]	Eau	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
27	[C4MIm][DBP]	Méthanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
28	[C4MIm][DBP]	Ethanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
29	[C4MIm][DMP]	Eau	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
30	[C4MIm][DMP]	Méthanol	298 ,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
31	[C4MIm][DMP]	Ethanol	298,15	[0,0000 – 1,0000]	11	*(Wan Normazlan et al, 2014)*
32	[C4Py][NO3]	Eau	298,15	[0,0000 – 0.9903]	15	*(Wang et al, 2012)*
33	[C6MIm][BF4]	Eau	298,15	[0,0000 – 0,6968]	8	*(Rilo et al, 2009)*
34	[C6MIm][BF4]	Ethanol	298,15	[0,0000 – 0,9020]	10	*(Rilo et al, 2009)*

35	[C8iQuin]BTI]	1-hexanol	[298,15 – 318,15]	[0,0000 – 1,0000]	24	*(Domanska et al, 2012)*
36	[C8MIm][BF4]	Ethanol	298,15	[0,0000 – 0.8988]	10	*(Rilo et al, 2009)*
37	[N112(hoe)][Br]	Eau	298,15	[0.9915 – 0.9981]	3	*(Domaoska et al, 2008)*
38	[N311(hoe)][Br]	Eau	298,15	[0.9900 – 0.9979]	3	*(Domaoska et al ,2008)*
39	[P666(14)][BTI]	Eau	[298,1 – 343,3]	0,0891	6	*(Almeida et al, 2013)*
40	[P666(14)][Dca]	Eau	[298,2 – 342,8]	0,4932	6	*(Almeida et al, 2013)*

III.2.2 Données des descripteurs de profil sigma

Suite aux résultats satisfaisant obtenu dans les travaux antérieurs *(Paduszynski, 2016),* nous avons adoptés les descripteurs moléculaires basés sur les profiles σ. Ses descripteurs dérivent de la théorie COSMO-RS qui est 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 sont calculées par des méthodes de chimie quantique. Les calculs de chimie quantique fournissent 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, ce 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, et il est considéré comme les propriétés caractéristiques de la molécule. Les figures III.1et III.2 montent 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 *(Diedenhofen et Klamt, 2010).*



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

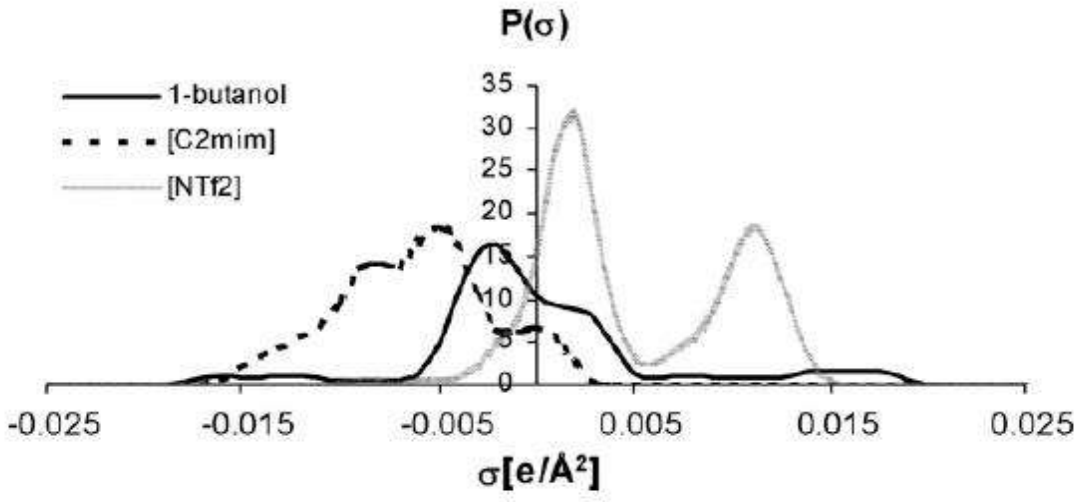
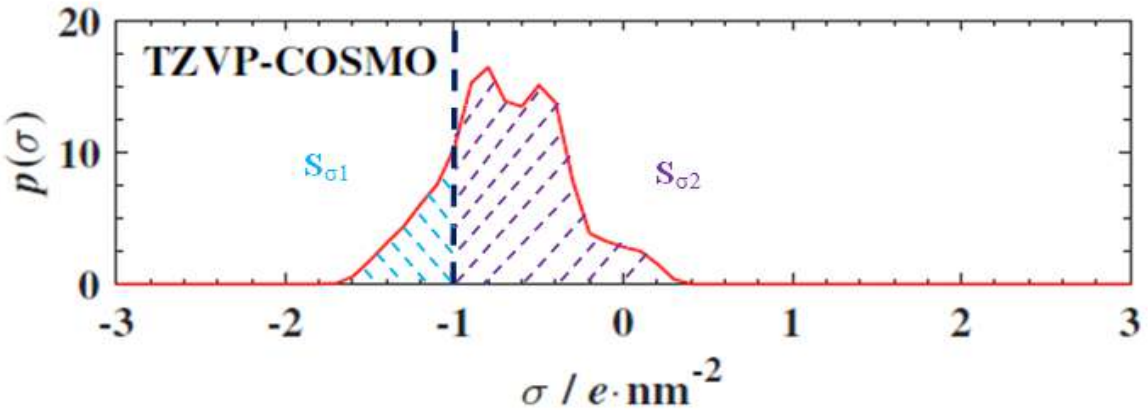


Figure III.2 : Profils sigma des conformères minimum de 1-butanol, [C2mim] et [NTf2].

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 possibles 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 actuel du modèle, est 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



La figure III.3 : descripteur moléculaire basé 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 nous avons élaboré un programme MATLAB 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 et les profiles σ numérisés pour l'eau et les solvants organiques obtenus à partir de bases VT-2005 et VT-2006 des références (Mullins et al. (2006), Mullins et al. (2008) respectivement) et le site électronique : https://www.design.che.vt.edu/VT-Databases.html.

Les descripteurs fondés sur le profil σ ($S_{\sigma i}$) des cations, des anions et des solvants étudiés sont regroupés dans les tableaux III.4, III.5 et III.6 respectivement.

Tableau III.4 : Descripteurs moléculaires ($S_{\sigma i}$) des cations des Lis étudiés.

Cations	$S_{\sigma 1C}$	$S_{\sigma 2C}$
[C1MIm]	2,85905	11,50845
[C2MIm]	2,67900	13,52630
[C4MIm]	2,62295	17,60025
[C6MIm]	2,64770	21,61100
[C8MIm]	2,54810	25,55670
[C4Py]	3,1653	16,58720
[C8iQuin]	2,25065	29,52255

[N112(hoe)]	2,5773	15,3787
[N311(hoe)]	2,5773	15,3787
[P666(14)]	0,9115	66,0228

Tableau III.5 : Descripteurs moléculaires ($S_{\sigma i}$) des anions des Lis étudiés.

Anions	**$S_{\sigma 1A}$**	**$S_{\sigma 2A}$**
[DMP]	9,0433	4,8582
[MeSO4]	8,8817	2,9631
[BF4]	9,2925	0,1052
[BTI]	20,568	0,0376
[C8SO4]	22,646	3,1221
[DEP]	13,9877	4,952
[EtSO4]	10,4886	3,1491
[MeSO3]	5,64805	5,12535
[MeSO4]	8,8817	2,9631
[DBP]	21,3317	4,9239
[NO3]	4,8726	2,8813
[Br]	0,0009	6,8408
[Dca]	7,629	2,5549

Tableau III.6 : Descripteurs moléculaires ($S_{\sigma i}$) des solvants étudiés.

solvants	**$S_{\sigma 1S}$**	**$S_{\sigma 2S}$**
Eau	1,2513	3,0754
Méthanol	0,5449	6,2453
Ethanol	0,4694	8,3713
1-propanol	0,4678	10,3710
1-butanol	0,4658	12,3460
1-hexanol	0,4668	16,3093
acetonitrile	0,0143	8,3215
Dimethyl sulfoxide	0,0055	11,150
tetrahydrofuran	0,0000	11,2621

III.3 La modélisation par les méthodes d'intelligence computationnelle

III.3.1 Choix des variables d'entrée et de sortie

Pour développer les modèles d'intelligence computationnelle nous avons exprimé la tension superficielle du mélange 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.3.2 Division des données en ensemble d'apprentissage, test et validation

Pour les 4 techniques l'ensemble des données d'apprentissage est constitué de 70% des données choisies aléatoirement par l'utilisation de la fonction ***dividerand*** de MATLAB 2018.

Les 30% restantes sont utilisées comme ensemble de test pour SVM et LSSVM. Ces mêmes 30% sont divisés aléatoirement en deux ensembles de 15% pour le test et la validation dans le cas de la modélisation par les RNA et ANFIS. Cette division a pour but d'éviter le sur-apprentissage.

III.3.3 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.3.4 Modélisation par RNA

Nous avons expérimenté 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). 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 nous avons expérimentés les fonctions logsig et tansig pour les couches cachées et la fonction purelin pour la couche de sortie.

III.3.4.1 Développement du modèle

Nous avons élaboré un programme sous MATLAB 2018, ou on fait varier de 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 donne la meilleure erreur quadratique moyenne, on exécute le programme 20 fois pour la même topologie. L'évaluation du modèle est 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 pour les trois topologies de modèle RNA. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité. Cependant le modèle optimisé en utilisant l'algorithme d'apprentissage ***trainlm*** et la fonction d'activation ***tansig*** donne l'égerment les meilleures valeurs des critères statistiques et des erreurs.

Tableau III.7 : Critères statistiques des modèles du RNA.

	trainbr, logsig NCC=10			**trainlm, logsig NCC=19**			**trainlm, tansig NCC=19**		
Critères statiques	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**
Q2F1	0.9897	0.9928	0.9905	0.9887	0.9901	0.9892	0.9926	0.9865	0.9906
Q2F2	0.9897	0.9928	0.9905	0.9887	0.9901	0.9892	0.9926	0.9864	0.9906
Q2F3	0.9897	0.9936	0.9907	0.9887	0.9881	0.9885	0.9926	0.9848	0.9902
Q2CCC	0.9948	0.9964	0.9952	0.9943	0.9950	0.9946	0.9963	0.9933	0.9953
R2	0.9897	0.9928	0.9905	0.9887	0.9901	0.9892	0.9926	0.9864	0.9906
R	0.9948	0.9964	0.9952	0.9943	0.9951	0.9946	0.9963	0.9934	0.9953
k	1.0002	1.0005	1.0003	1.0003	1.0031	1.0012	0.9999	0.9960	0.9987
k’	0.9989	0.9989	0.9989	0.9987	0.9960	0.9979	0.9995	1.0029	1.0005
R_0^2	1.0000	1.0000	1.0000	1.0000	0.9999	1.0000	1.0000	0.9998	1.0000
$R_0'^2$	1.0000	1.0000	1.0000	1.0000	0.9998	0.9999	1.0000	0.9999	1.0000
M	-0.0104	-0.0072	-0.0096	-0.0114	-0.0099	-0.0109	-0.0075	-0.0136	-0.0095
N	-0.0104	-0.0072	-0.0096	-0.0114	-0.0098	-0.0109	-0.0075	-0.0136	-0.0095
R_m	0.8893	0.9089	0.8937	0.8836	0.8921	0.8864	0.9070	0.8724	0.8945
A_f	1.0115	1.0158	1.0126	1.0125	1.0173	1.0140	1.0085	1.0137	1.0100
B_f	1.0004	0.9993	1.0001	0.9999	0.9956	0.9986	1.0003	1.0021	1.0008

Tableau III.8 : Différentes erreurs des modèles de RNA.

	trainbr, logsig			**trainlm, logsig**			**trainlm, tansig**		
Erreurs	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**
ERAM%	1.1516	1.5602	1.2580	1.2495	1.6928	1.3827	0.8457	1.3742	**1.0045**
REQM	1.1366	0.8956	1.0790	1.1403	1.1685	1.1489	0.9349	1.3378	**1.0721**
EQM	1.2919	0.8021	1.1643	1.3003	1.3653	1.3199	0.8741	1.7898	**1.1494**

L’évaluation graphique est réalisée à l’aide de diagramme de dispersion (pour les ensembles d’apprentissage, de test, de validation et global), le diagramme de ER% et le diagramme de distribution de ER%. Les figures III.4 à III.6 montrent respectivement le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale pour les trois topologies considérées. Ces diagrammes montrent clairement la concordance entre les valeurs calculées et les valeurs expérimentales à travers une dispersion serrée des points sur la première bissectrice. Cependant on peut remarquer

l'existence d'un certains nombre de points légèrement éloignés qui seront élucidés par la suite à travers une analyse de diagnostic des points aberrants.

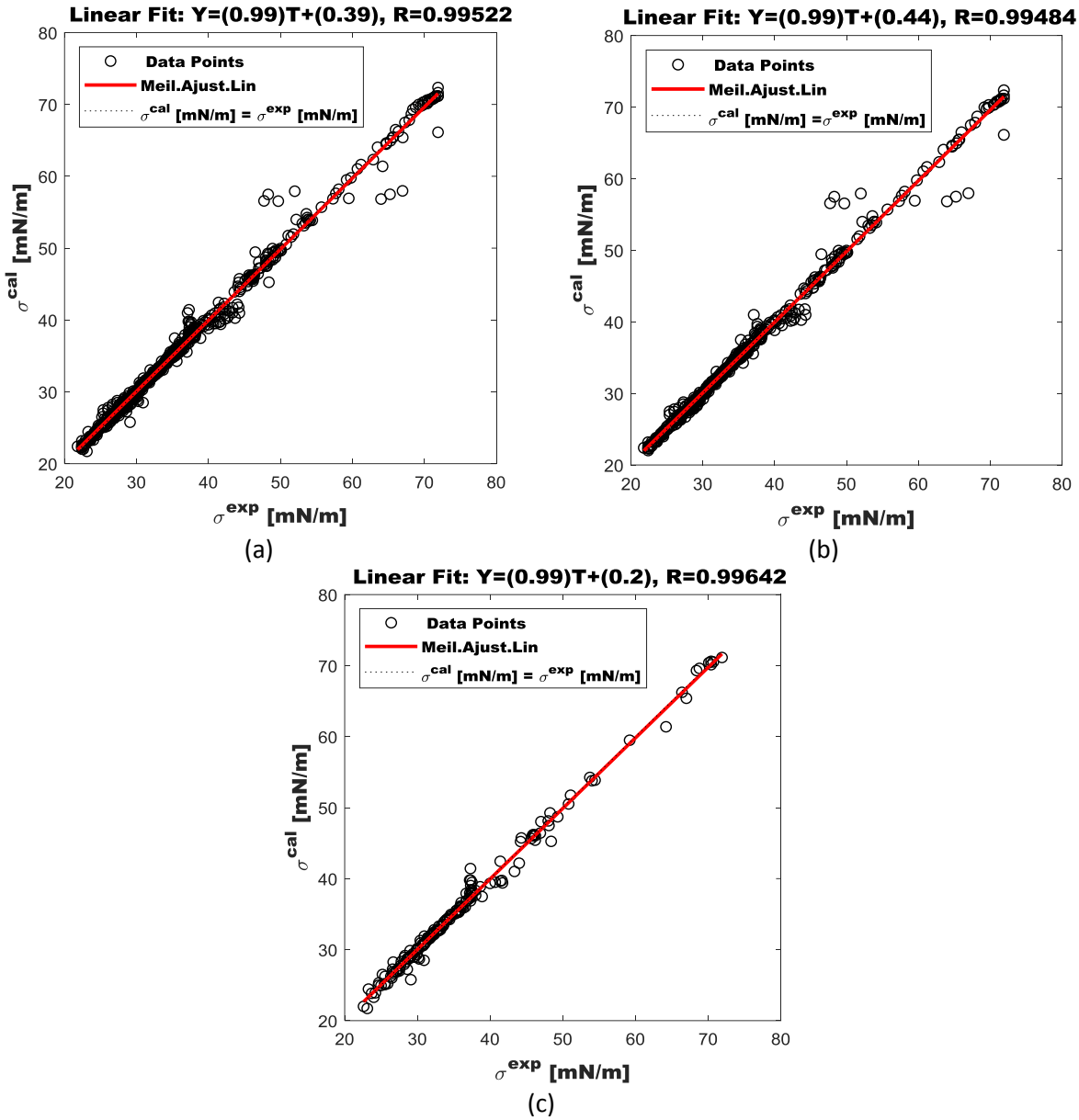


Figure III.4 : Diagramme de dispersion de la tension superficielle calcule en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble du test par le modèle RNA « trainbr, logsig ».

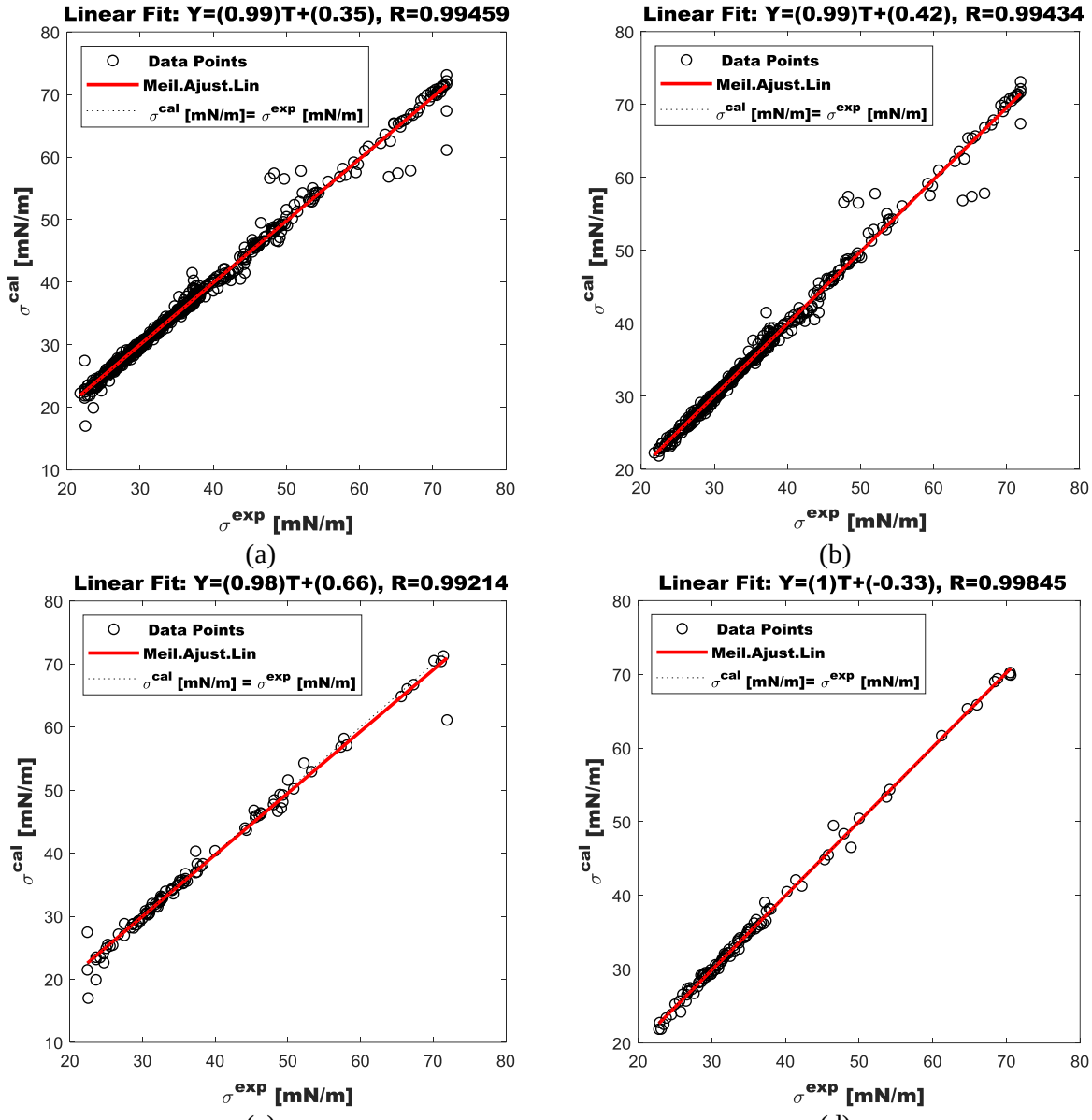


Figure III.5: Diagramme de dispersion de la tension superficielle calculer en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle RNA « trainlm, logsig ».

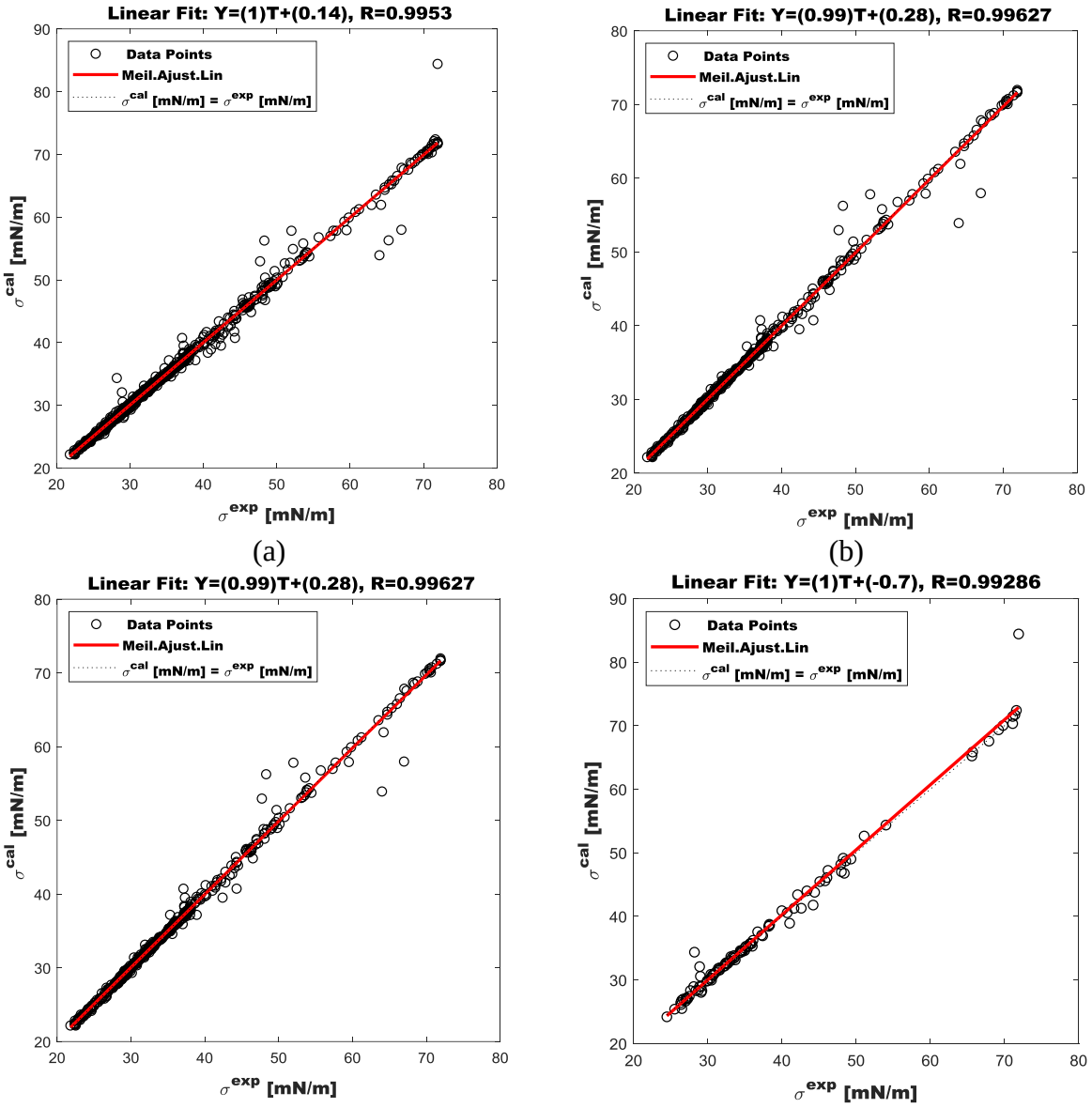


Figure III.6: Diagramme de dispersion de la tension superficielle calculer en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle RNA « trainlm, tansig ».

Afin de visualiser quantitativement la distribution des erreurs relatives nous avons adoptés deux types de diagrammes. Le premier diagramme montre l'erreur relative en fonction de la tension superficielle et le deuxième montre la distribution de cette erreur. Les figures III.7 à III.9 représentent erreur relative en fonction de la tension superficielle pour les trois topologies respectivement. Pour les trois modèles la majorité des points sont situés autour de la ligne zéro. Quantitativement l'ER pour les quelques points aberrant ne dépasse pas 25%.

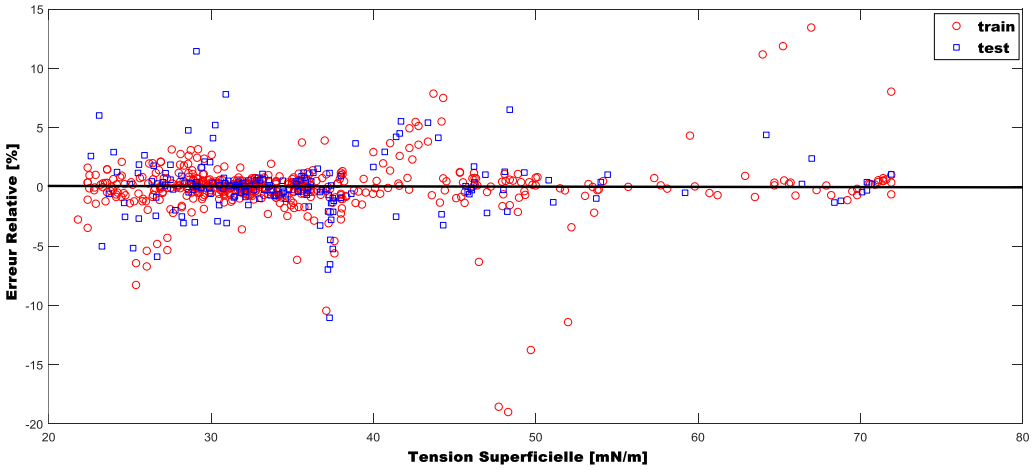


Figure III.7 : L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainbr, logsig ».

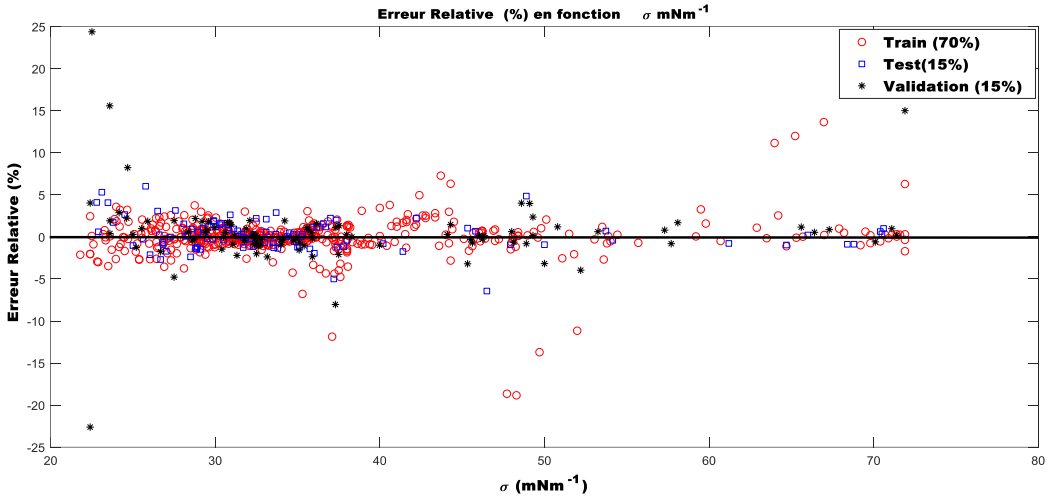


Figure III.8 : L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainlm, logsig ».

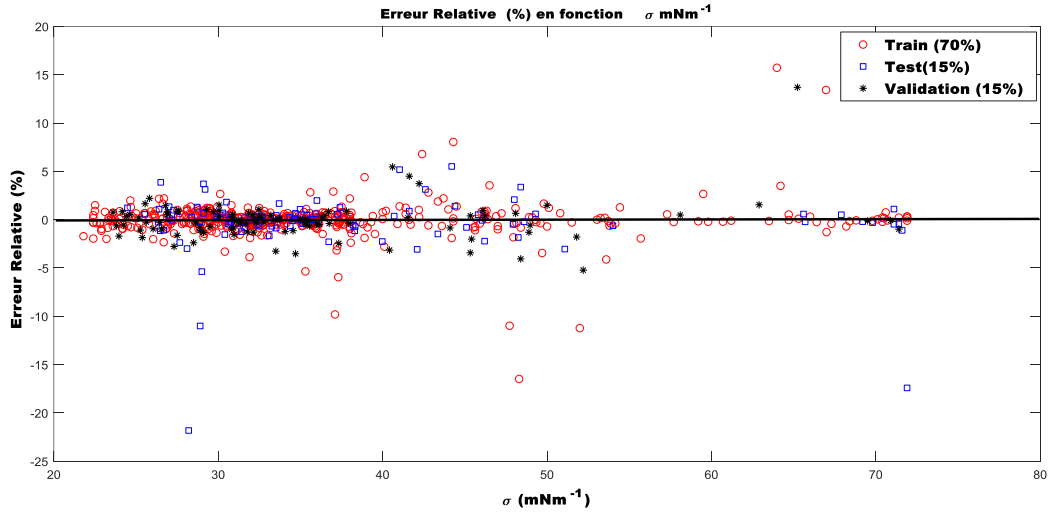


Figure III.9 : L'erreur relative (%) de la tension superficielle par le modèle de RNA « trainlm, tansig ».

Ces remarques relatives à l'ER sont mieux illustrées sur les diagrammes de distribution des figures III.10 à III.12.

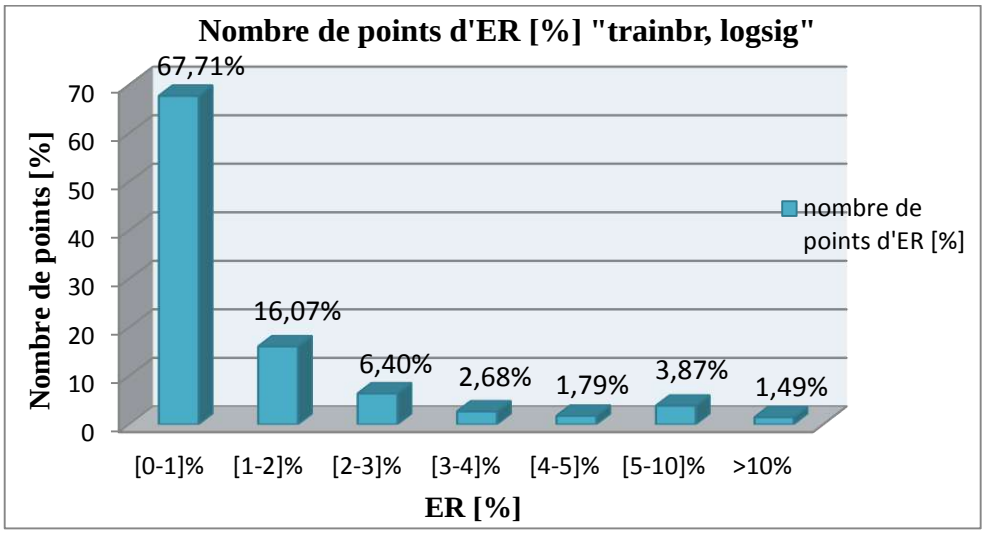


Figure III.10 : Diagrammes de distribution d'ER[%] du modèle RNA « trainbr, logsig ».

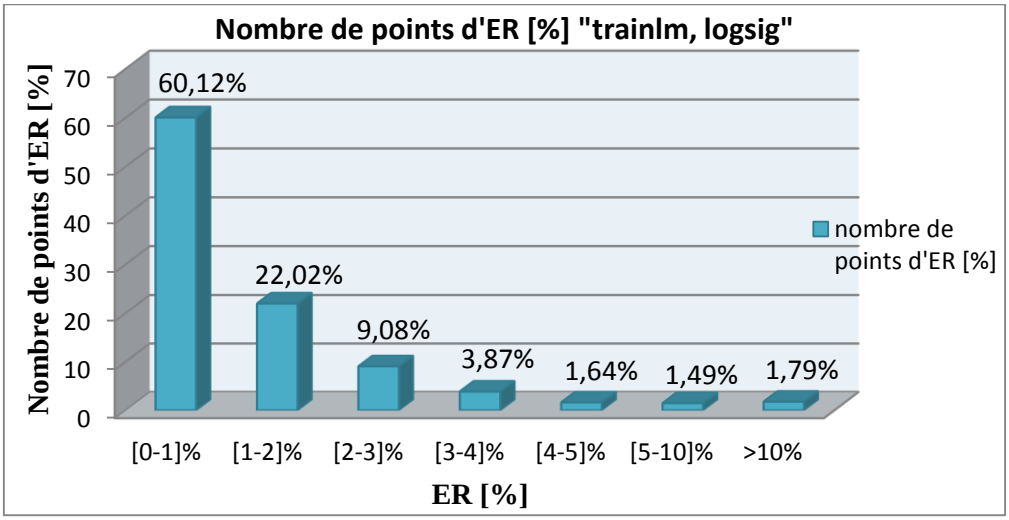


Figure III.11 : Diagrammes de distribution d'ER [%] du modèle RNA « trainlm, logsig ».

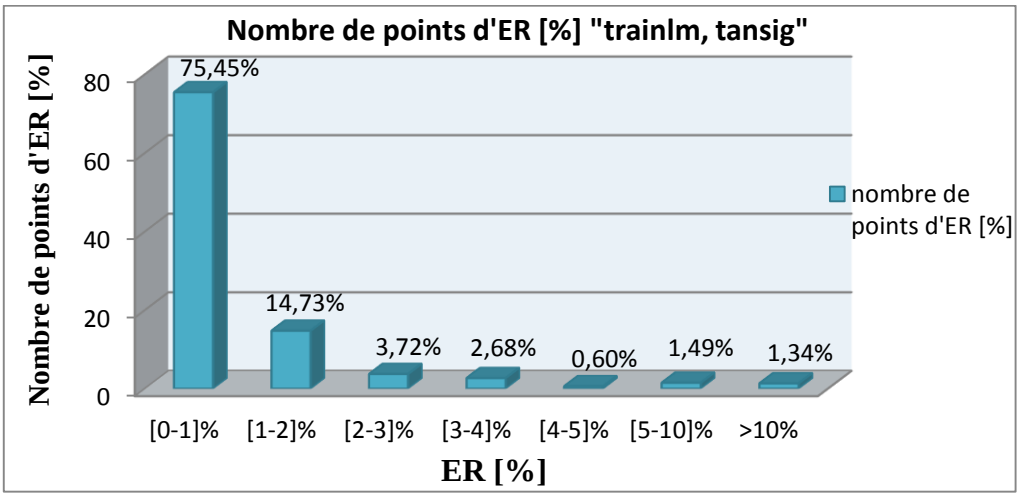


Figure III.12 : Diagrammes de distribution d'ER [%] du modèle RNA « trainlm, tansig ».

Pour confirmer la capacité de généralisation des modèles développés nous avons réalisé des calculs d'interpolation et d'extrapolation pour la température et la composition. Les figures III.13 à III.15 montrent les isothermes de la variation de la tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Dimethyl sulfoxide pour les modèles neuronaux respectivement. Dans les trois cas les points interpolés suivent l'allure des points calculés.

De la même façon les figures III.16 à III.18 représentent le calcul d'interpolation de la température des trois modèles neuronaux. 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.

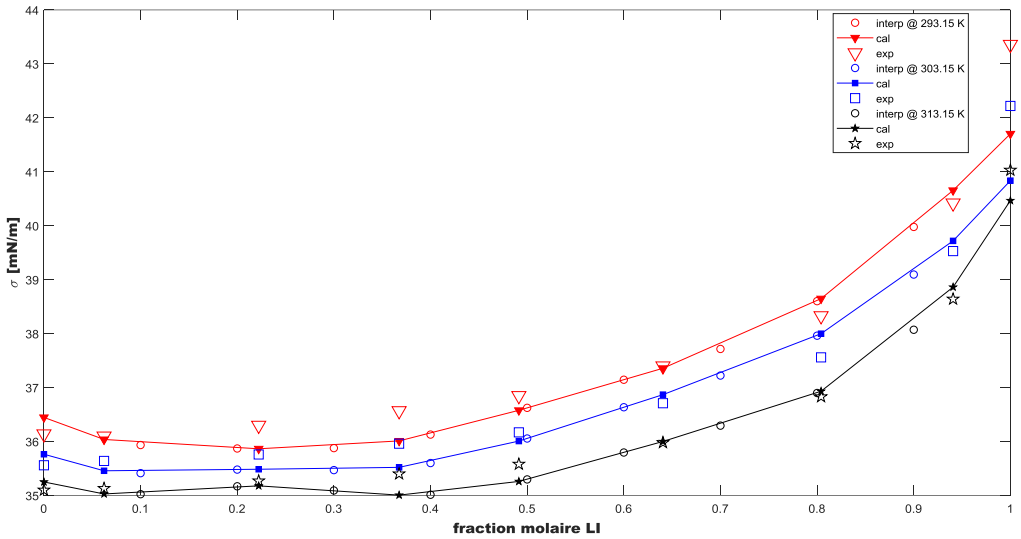


Figure III.13 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainbr,logsig ».

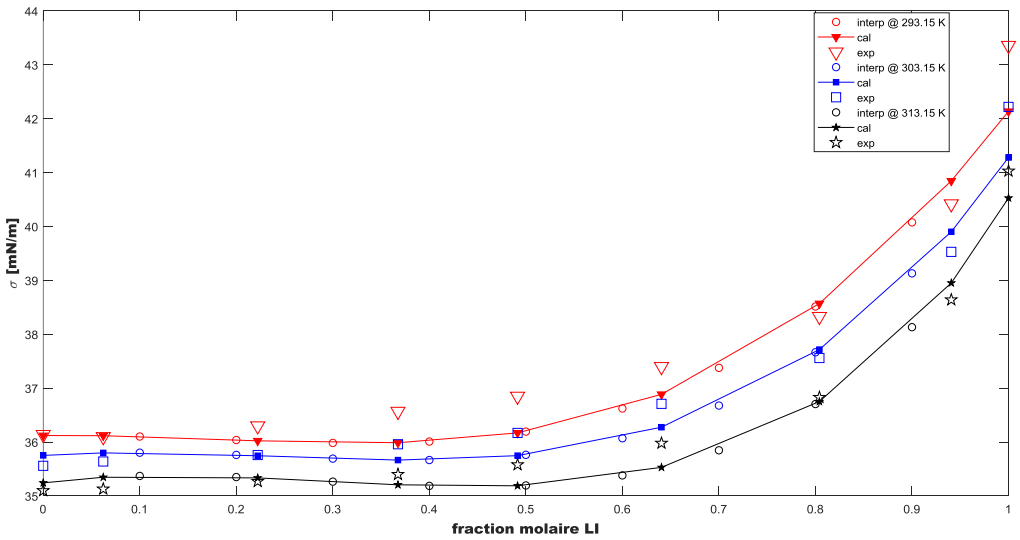


Figure III.14 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,logsig ».

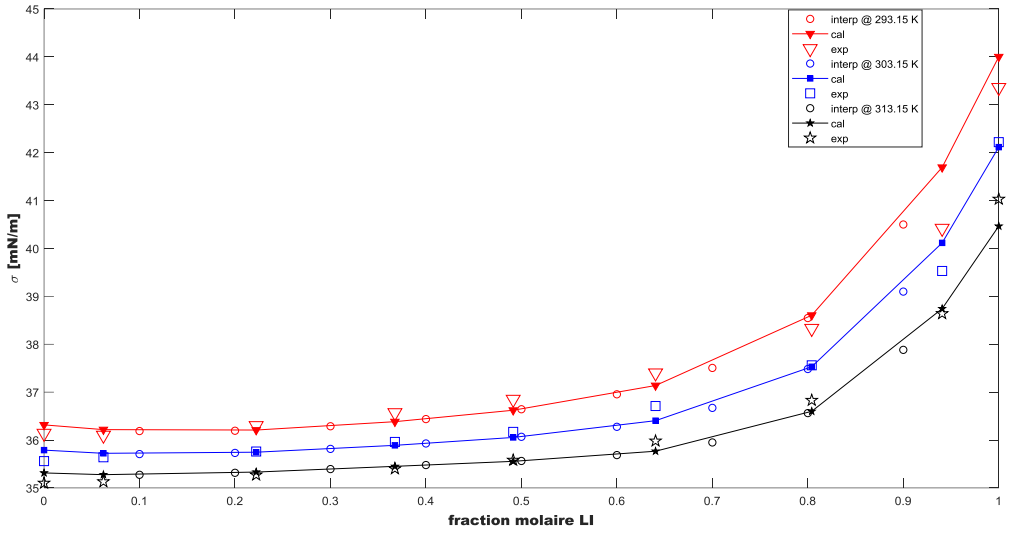


Figure III.15 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,tansig ».

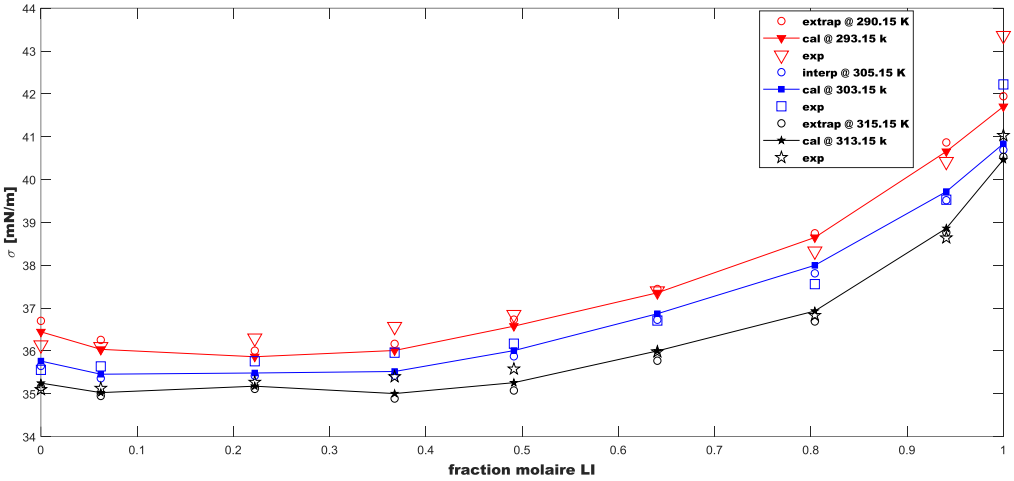


Figure III.16 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainbr,logsig ».

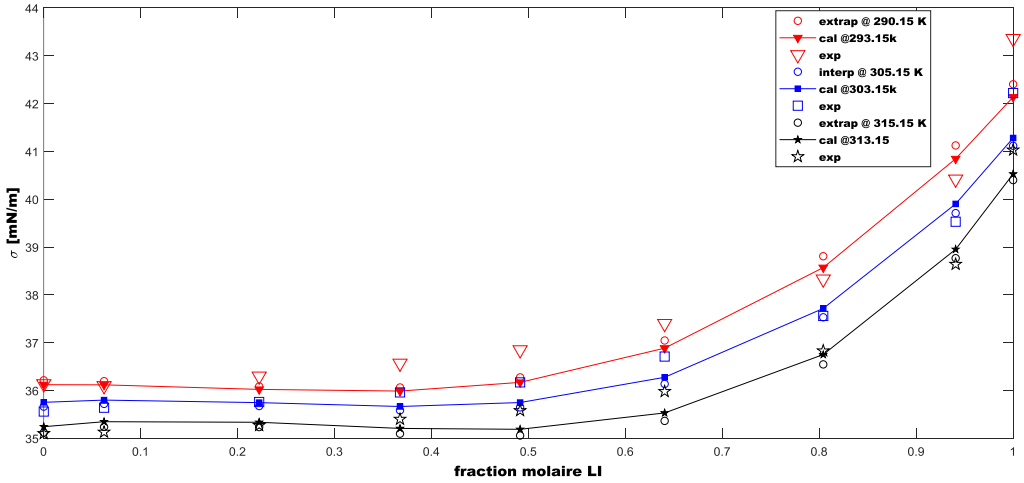


Figure III.17: Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,logsig ».

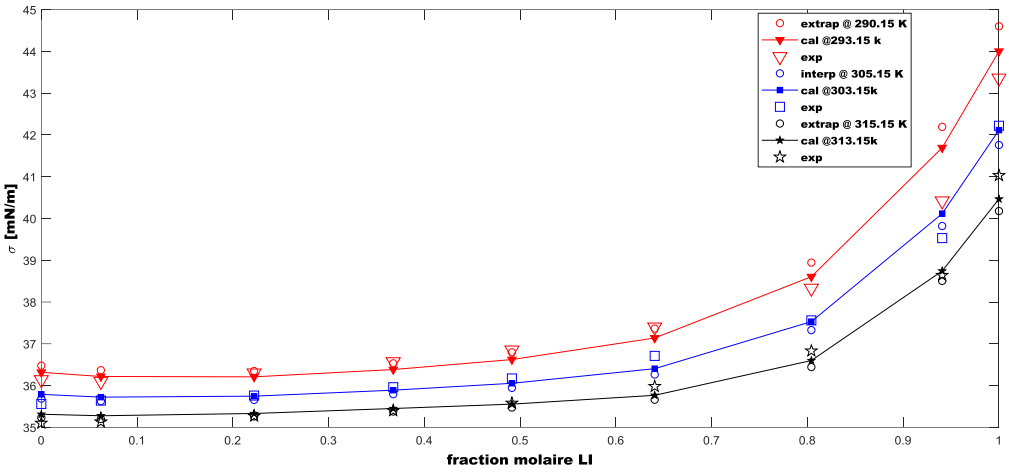


Figure III.18 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle de RNA « trainlm,tansig ».

III.3.4.2 Analyse de sensibilité

Pour déterminer l'importance relative des variables d'entrée, différentes méthodes statistique peut être appliqués (*Olden et al, 2004 et Gevrey et al, 2003)*

Dans ce travail nous avons adoptés la méthode de Garson et la méthode des poids des connections pour calculer l'importance relative de chaque variable.

Les figures III.19 à III.21 montrent l'importance relative calculée par la méthode de Garson pour les trois modèles neuronaux. Pour les trois modèles l'importance relative de la température la plus faible (<5%). Cependant la variable la plus importante varie selon le modèles neuronaux (fraction molaire de LI pour le modèles trainbr,logsig ; sigma profil d'anion pour le modèles trainlm,logsig et sigma profil du solvant pour le modèles trainlm, tansig).

Les figures III.22 à III.24 montrent l'importance relative calculée par la méthode de Con-weight pour les trois modèles neuronaux. Pour les trois modèles l'importance relative les plus importantes sont sigma profil du solvant et la fraction molaire.

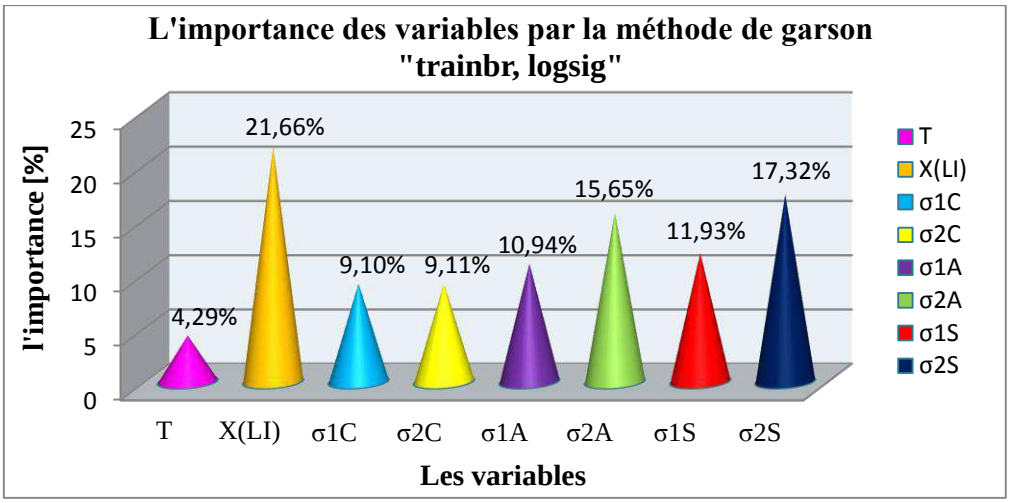


Figure III.19 : L'importance relative des variables par la méthode de Garson du modèle RNA « trainbr,logsig ».

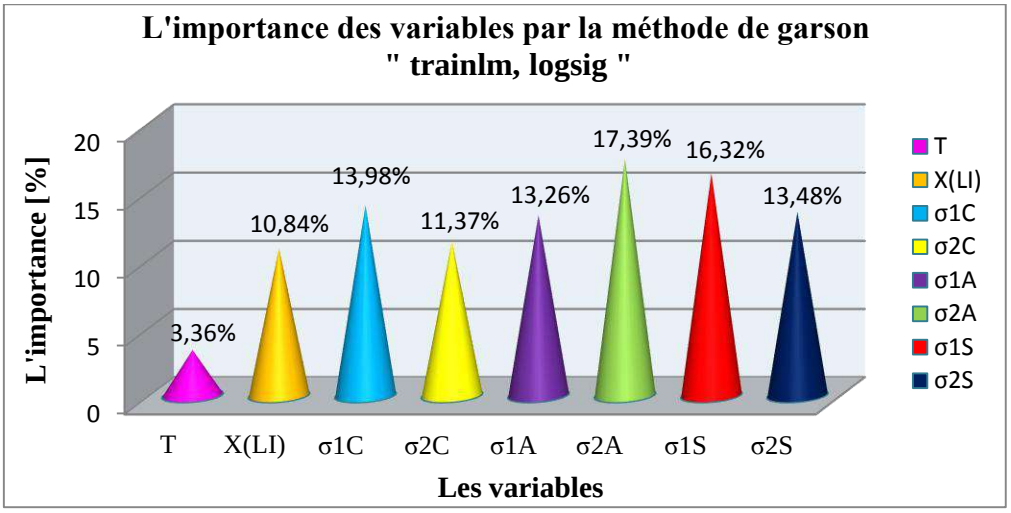


Figure III.20 : L'importance relative des variables par la méthode de Garson du modèle RNA « trainlm,logsig ».

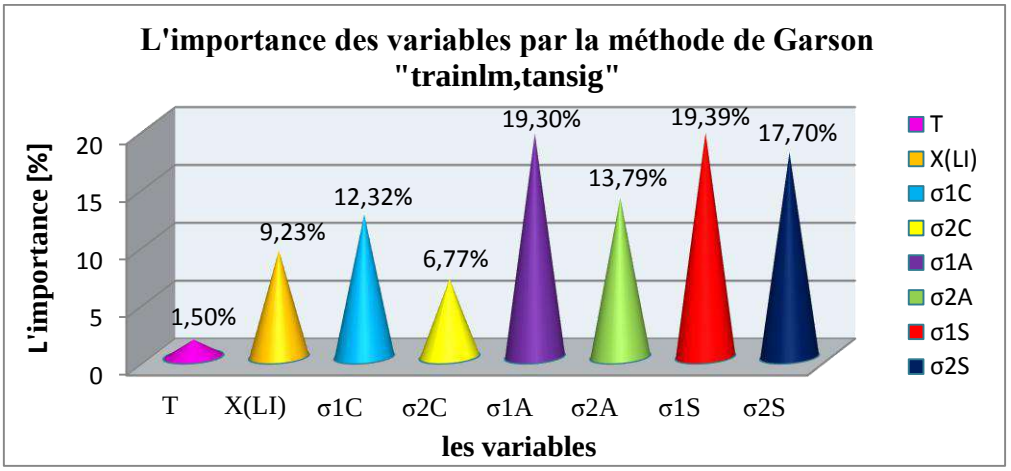


Figure III.21 : L'importance relative des variables par la méthode de Garson du modèle RNA « trainlm,tansig ».

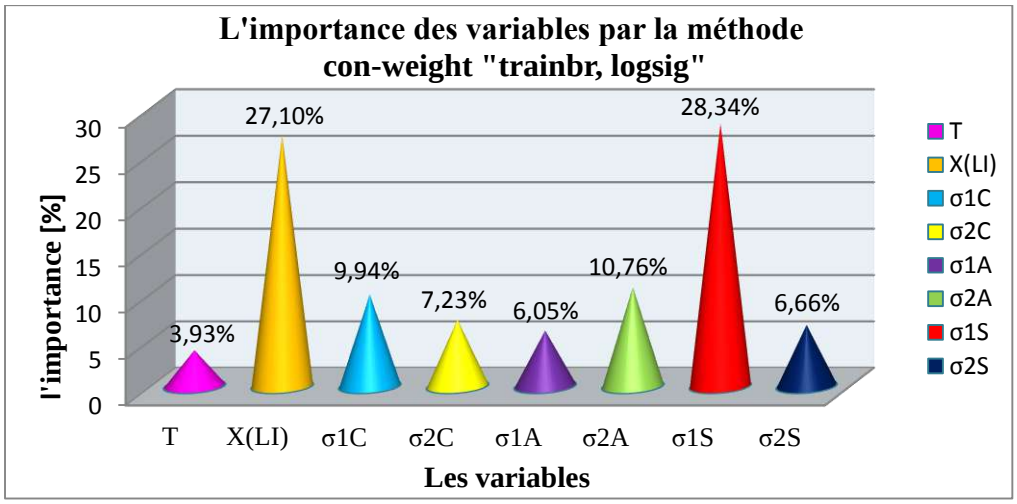


Figure III.22 : L'importance relative des variables par la méthode de con weight du modèle RNA « trainbr,logsig ».

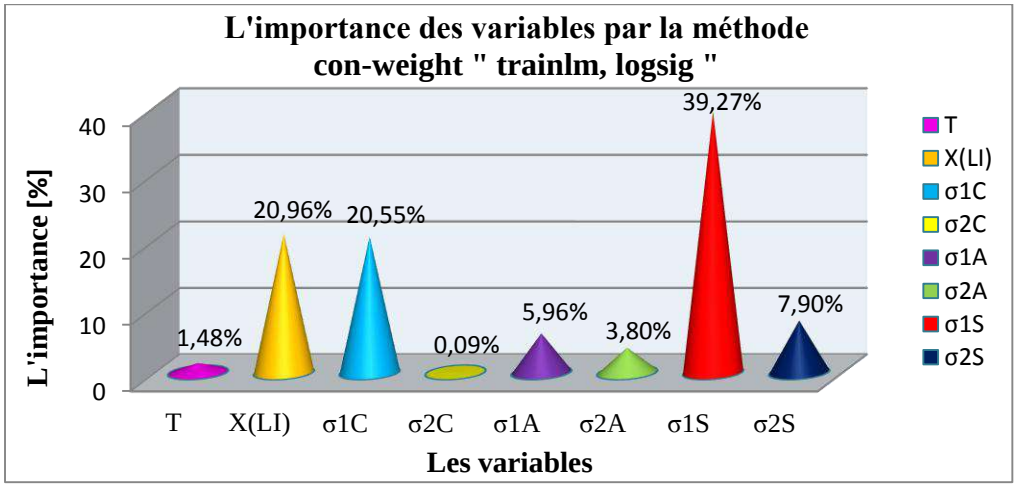


Figure III.23 : L'importance relative des variables par la méthode de con-weight du modèle RNA « trainlm,logsig ».

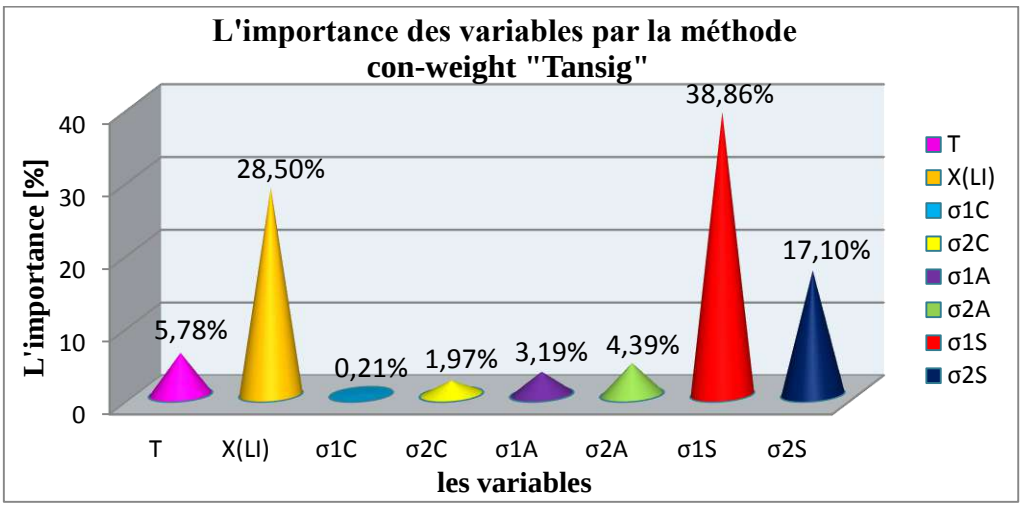


Figure III.24 : L'importance relative des variables par la méthode de con-weight du modèle RNA « trainlm,tansig ».

III.3.5 Modélisation par ANFIS

III.3.5.1 Développement de modèle

La base de données expérimentale a été divisée de manière aléatoire en trois ensembles «apprentissage», «test» et « validation » selon un ratio de 70%, 15 % et 15%. Bien qu'il existe plusieurs méthodes de développement du système d'inférence flou (FIS), dans cette étude, le système Takagi Sugeno FIS est développé par une méthode de partitionnement sur grille. La fonction genfis a été choisie comme fonction de traitement dans le programme MATLAB, qui utilise la partition de grille pour développer le système FIS TSK. Les fonctions d'appartenance (Membership Functions MF) ***linear*** de type ***gbellmf*** ont été utilisées dans la modélisation. Il est essentiel de former le FIS initial pour obtenir de meilleures performances. Dans cette étude, le FIS initial a été formée selon une méthode prédéfinie dans la boîte à outils de la logique floue MATLAB, appelée méthode hybride. Le tableau III.9 présente les paramètres détaillés pour l'entraînement du FIS principal.

Tableau III.9 : Détails des paramètres d'apprentissage du système d'inférence floue (FIS)

Paramètre	Description/Valeur
Structure Flou	Type Sugeno
Nom	ANFIS
FIS initial pour l'apprentissage	genfis
Fonctions d'appartenance des entrées	Generalized bell-shaped membership function (gbellmf)
Fonctions d'appartenance de sortie	Linear
Nombre de Fonctions d'appartenance associées avec chaque sortie	2
Nombre d'entrées	8
Nombre de sorties	1
Méthode d'optimisation	Hybride (Moindre carrés et back-propagation)
Nombre maximum d'époques d'apprentissage	30
Taille du pas initial	0.01
Taux de diminution de la taille des pas	0.9
Taux d'augmentation de taille de pas	1.1
Nombre de règles	256

L'évaluation du modèle est réalisée à l'aide d'outils statistiques et graphiques. Les tableaux III.10 et III.11 résument respectivement le calcul des différents critères statistiques

et les différentes erreurs du modèle ANFIS. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité.

Tableau III.10 : Critères statistiques du modèle ANFIS.

Critère statistique	Train	Test	Global
Q2F1	0.9860	0.8672	0.9444
Q2F2	0.9861	0.8658	0.9444
Q2F3	0.9860	0.8340	0.9403
Q2CCC	0.9929	0.9285	0.9714
R2	0.9860	0.8658	0.9444
R	0.9930	0.9312	0.9719
k	1.0001	1.0100	1.0032
K'	0.9988	0.9779	0.9922
R_0^2	1.0000	0.9988	0.9999
$R_0'^2$	1.0000	0.9947	0.9993
M	-0.0142	-0.1536	-0.0588
N	-0.0142	-0.1489	-0.0581
R_m	0.8691	0.5501	0.7219
A_f	1.0181	1.0447	1.0261
B_f	1.0004	0.9932	0.9982

Tableau III.11 : Différentes erreurs par le modèles ANFIS.

Erreurs	Train	Test	Global
ERAM%	1.8074	4.2138	2.5308
REQM	1.2633	4.3426	2.6048
EQM	1.5959	18.8582	6.7848

L'évaluation graphique est réalisée à l'aide de diagramme de dispersion (pour les ensembles d'apprentissage, de test, de validation et global), diagramme de ER% et le diagramme de distribution de ER%. La figure III.25 montre le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale. Ce diagramme montre clairement la concordance entre les valeurs calculées et les valeurs expérimentales à travers une dispersion serrée des points sur la première bissectrice. Cependant on peut remarquer l'existence d'un certain nombre de points légèrement éloignés qui seront élucidés par la suite à travers une analyse de diagnostic des points aberrants.

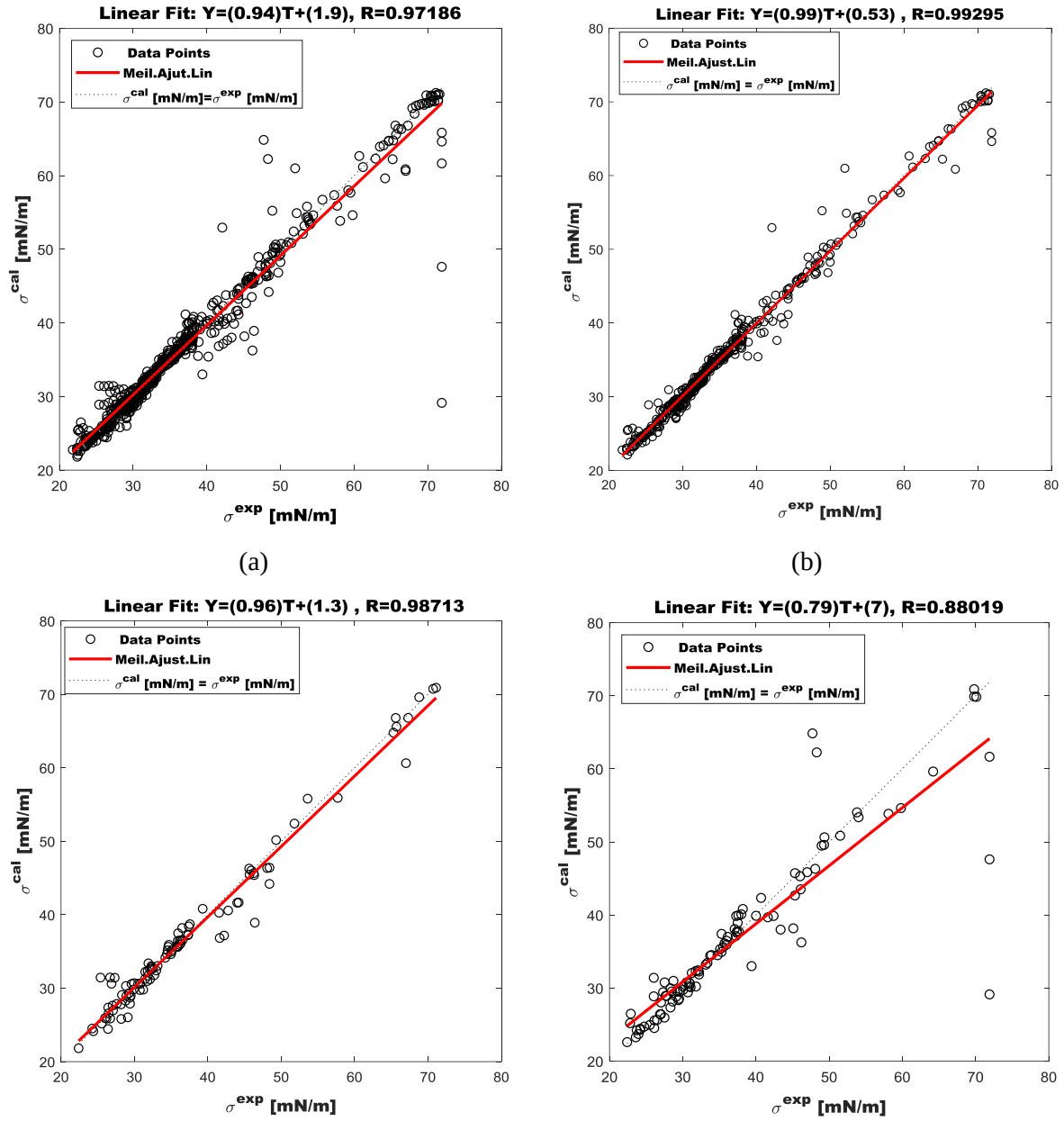


Figure III.25 : Diagramme de dispersion de la tension superficielle calculé en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble de la validation, (d) Ensemble du test par le modèle ANFIS.

Afin de visualiser quantitativement la distribution des erreurs relatives nous avons adoptés deux types de diagrammes. Le premier diagramme montre l'erreur relative en fonction de la tension superficielle et le deuxième montre la distribution de cette erreur. La figure III.26 représente l'erreur relative en fonction de la tension superficielle. Pour L'erreur maximale la quasi totalité des points sont situés entre ±25 %

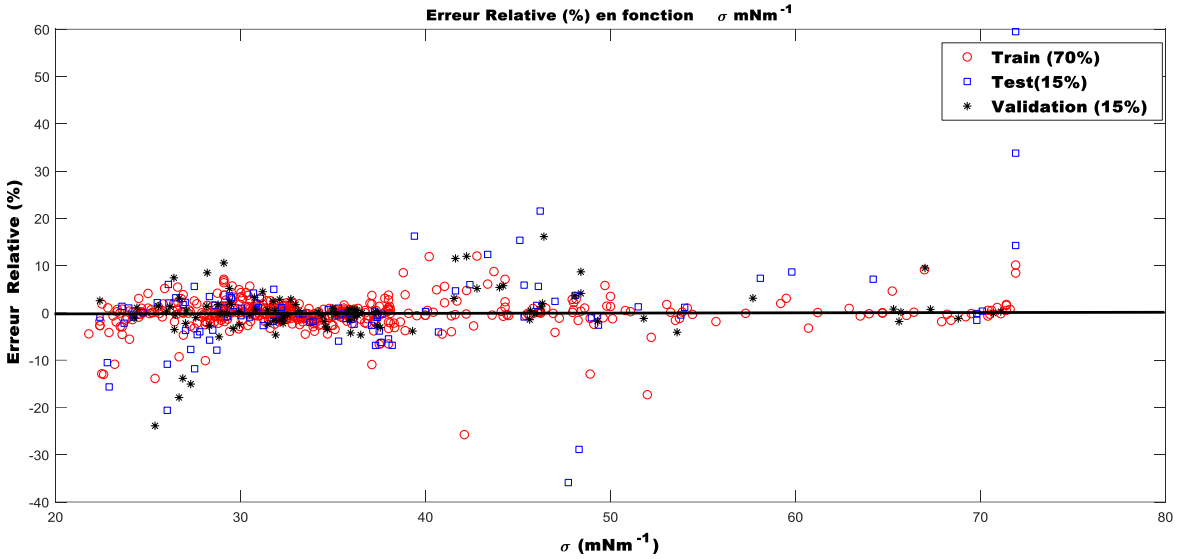


Figure III.26 : L'erreur relative (%) de la tension superficielle par le modèle ANFIS.

Cette remarque relative à l'ER est mieux illustrée sur le diagramme de distribution de la figure III.27

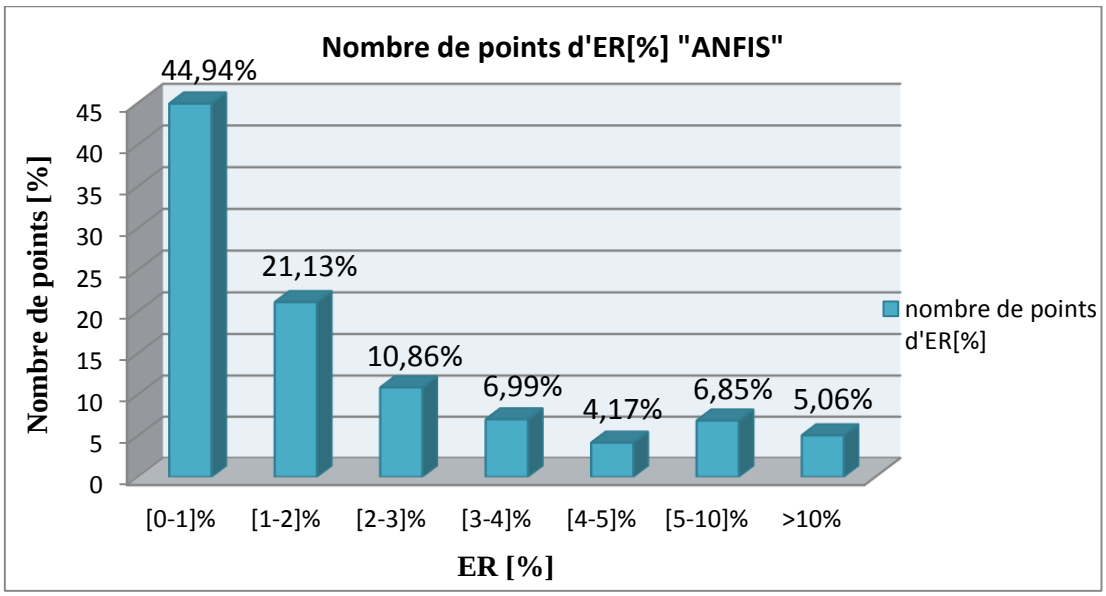


Figure III.27 : Diagrammes de distribution d'ER [%] du modèle ANFIS.

La figure III.28 représente les isothermes de la variation de la tension superficielle en fonction de la composition du mélange [C2MIm] [BTI]-Dimethyl sulfoxide. Cette figure illustre la capacité d'interpolation du modèle ANFIS. Les points interpolés suivent l'allure des points calculés approximativement.

La figure III.29 représente le calcul d'interpolation de la température. L'isotherme d'interpolation (305.15 K) semble suivre l'allure prévisible. Cependant pour les isothermes d'extrapolation (290.15 K et 315.15 K), on remarque que isotherme à 290.15K présente une

meilleure extrapolation par apport à l'isotherme de 315.15K. Pour la première les points sont bien situés au dessus de ceux de l'isotherme de 293.15 K bien que l'écart entre ces deux isothermes est faible. Alors que pour l'isotherme de 315.15K les points correspondant à faible concentration de LI au dessous de points de l'isotherme 313.15K.

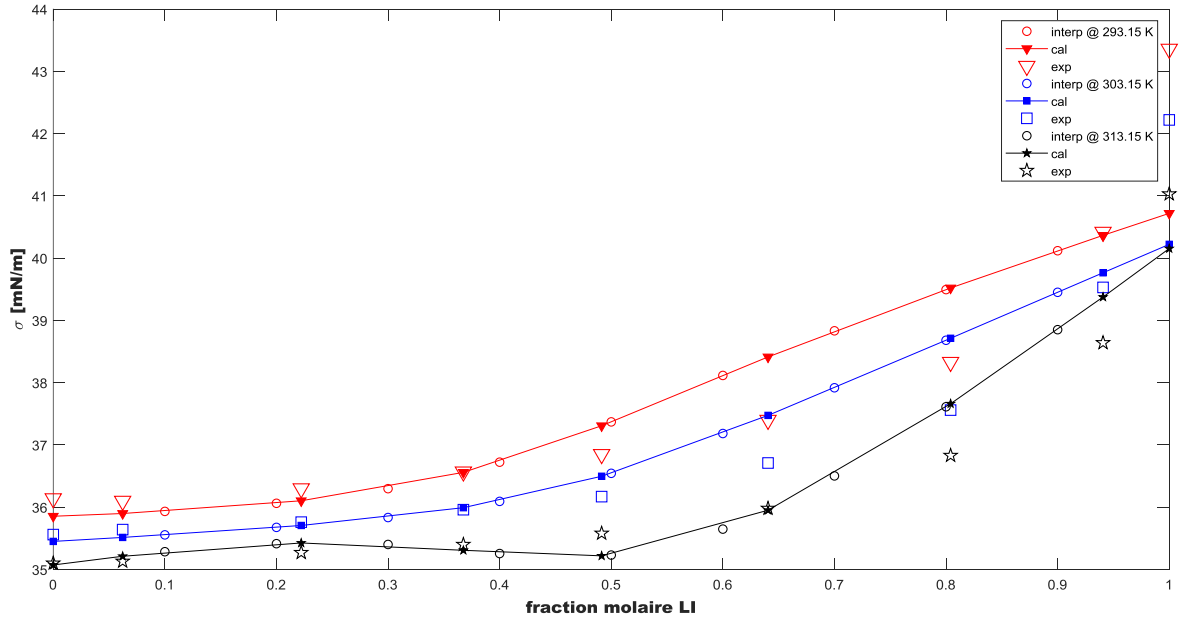


Figure III.28: Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle ANFIS.

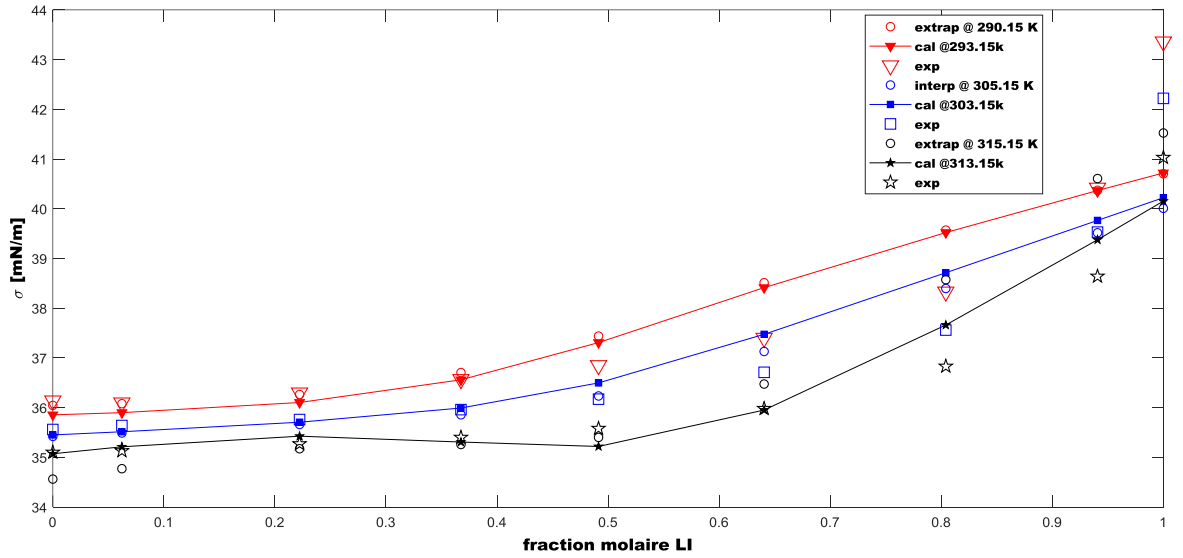


Figure III.29 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle ANFIS.

III.3.6 Modélisation par SVM

Le développement du modèle SVM est réalisé par la fonction ***fitrsvm*** de MATLAB. Cette fonction compte plusieurs options de calculs relatives au choix de la méthode de la validation croisée, la fonction noyaux (gaussian, rbf et polynomial), le solver (SMO, ISDA, L1QP). Elle vise à déterminer les valeurs optimales de la constante C (boxConstraint), epsilon et le paramètre de la fonction noyaux (KernelScale) du modèle. Dans cette étude nous avons choix la méthode de Cross-validation « holdout » que nous avons fixé à 0.3 qui signifie de 70% des données sont utilisées pour l'apprentissage et 30% pour la validation.

Les paramètres optimisés du modèles SVM sont présentés dans le tableau III.12

Tableau III.12 : Les paramètres optimisés du modèle SVM.

C	**ε**	**K**	**Fonction kernel**
124	1.8	0.9	gaussian
Quantité de vecteurs de support : 471			

Tableau III.13 : Critères statistiques du modèle SVM.

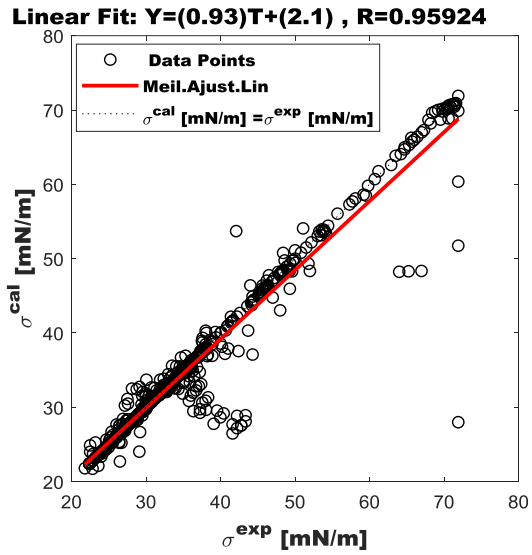
Critère statique	**Train**	**Test**	**Global**
Q2F1	0.9579	0.8514	0.9181
Q2F2	0.9608	0.8433	0.9177
Q2F3	0.9579	0.7923	0.9084
Q2CCC	0.9785	0.9186	0.9575
R2	0.9579	0.8433	0.9177
R	0.9793	0.9229	0.9592
k	1.0091	1.0218	1.0133
K'	0.9877	0.9654	0.9802
R_0^2	0.9989	0.9946	0.9978
$R_0'^2$	0.9981	0.9866	0.9954
M	-0.0428	-0.1794	-0.0874
N	-0.0419	-0.1700	-0.0847
R_m	0.7640	0.5153	0.6578
A_f	1.0152	1.0496	1.0254
B_f	0.9918	0.9750	0.9867

Tableau III.14 : Différentes erreurs du modèles SVM.

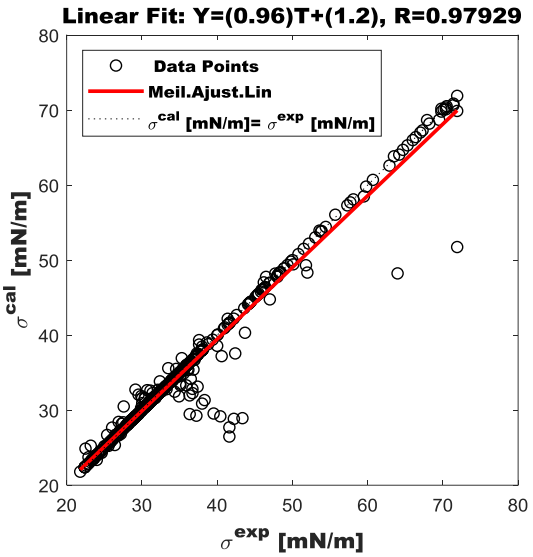
Erreurs	Train	Test	Global
ERAM%	1.3914	4.3660	2.2811
REQM	2.1476	4.7723	3.1693
EQM	4.6120	22.7745	10.0445

Les tableaux III.13 et III.14 résument respectivement le calcul des différents critères statistiques et les différentes erreurs de modèle SVM optimisé. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité. Cependant il faut noter un écart notable entre l'erreur de ensemble d'apprentissage et l'ensemble de test, qui peut être explique une tendance de sur apprentissage.

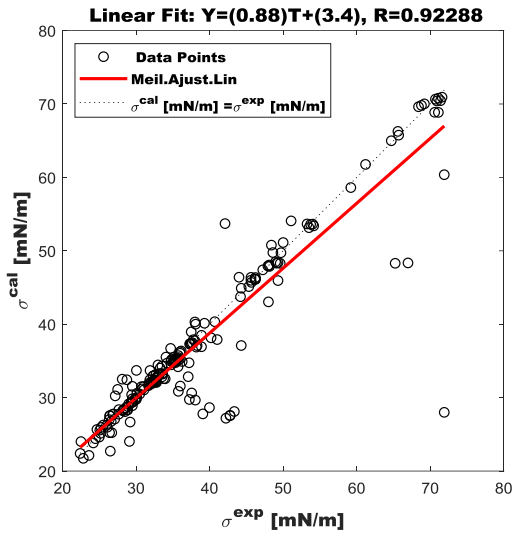
La figure III.30 montre le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale. Ces diagrammes montrent clairement la concordance entre les valeurs calculées et les valeurs expérimentales à travers une dispersion serrée des points sur la première bissectrice. Cependant on peut remarquer l'existence d'un certain nombre de points éloignés qui peuvent être attribué à l'existence de points aberrants et la qualité de généralisation du modèle.



(a)



(b)



(c)

Figure III.30 : Diagramme de dispersion de la tension superficielle calcule en fonction de la tension superficielle expérimental : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble du test par le modèle SVM.

Afin de visualiser quantitativement la distribution des erreurs relatives nous avons adoptés deux types de diagrammes. Le premier montre l'erreur relative en fonction de la tension superficielle et le deuxième montre la distribution de cette erreur. La figure III.31 représente erreur relative en fonction de la tension superficielle et la figure III.32 montre La distribution des erreurs relatives.

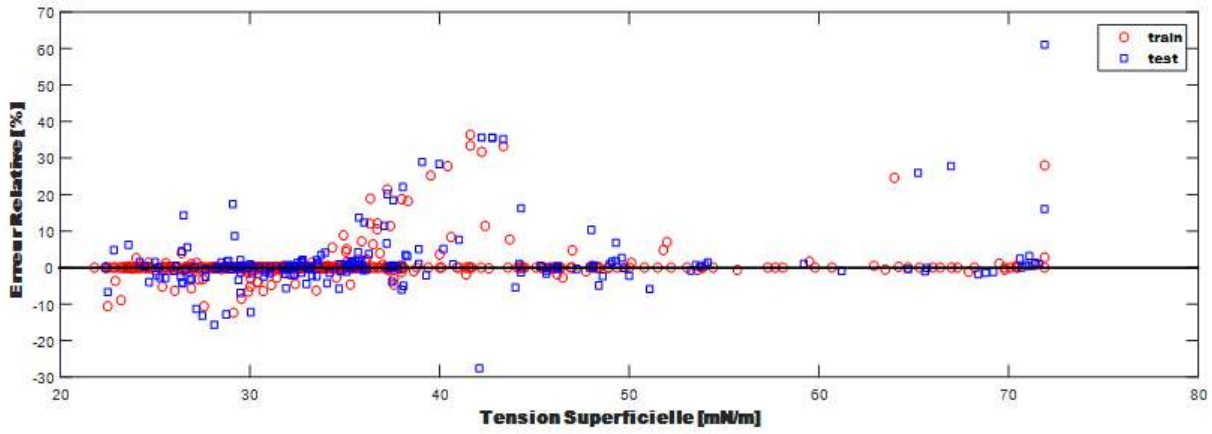


Figure III.31 : L'erreur relative (%) de la tension superficielle par le modèle de SVM.

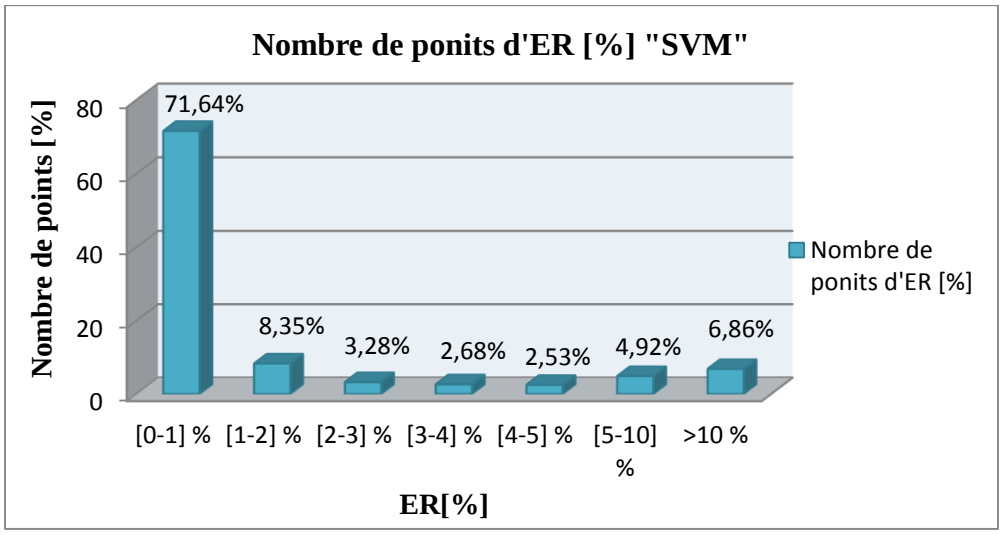


Figure III.32 : Diagrammes de distribution d'ER [%] du modèle SVM.

III.3.7 Modélisation par LSSVM

Dans cette étude la modélisation par LSSVM est réalisée à l'aide de la boite à outils LS-SVMlab . Publié en ligne. Accédé le 15 janvier 2018,

(URL: http: // www.esat.kuleuven.be/sista/lssvmlab/.)

Tout d'abord, un algorithme approprié implémenté dans LS-SVMlab détecte et redimensionne les variables continues, catégorielles et binaires. Ensuite, les paramètres de LSSVM (constante de régularisation γ et le paramètre kernel σ^2) sont déterminés par minimisation d'une mesure de performance sélectionnée; cette étape est appelée tuning. L'étape d'apprentissage suivante utilise la constante de régularisation γ et le paramètre kernel σ^2 ajustés dans la première étape pour déterminer les paramètres α et b sont du modèle LSSVM.

La dernière étape consiste à tester la capacité de généralisation du modèle sur ensemble des données réservées au test. Les paramètres optimisés sont situés au tableau III.15

Tableau III.15 : Les paramètres d'optimisation du modèles LSSVM.

γ	σ^2
3714.1	1.2178

Les tableaux III.16 et III.17 résument respectivement le calcul des différents critères statistiques et les différentes erreurs de modèle LSSVM optimisé. Il est clair que tous les critères statistiques vérifient les conditions d'acceptabilité. Cependant il faut noter un écart

notable entre l'erreur de ensemble d'apprentissage et l'ensemble de test, qui peut être explique une tendance de sur apprentissage.

Tableau III.16 : Critères statistiques du modèles LSSVM.

Critère statique	Train	Test	global
Q2F1	0.9943	0.9370	0.9753
Q2F2	0.9943	0.9369	0.9753
Q2F3	0.9943	0.9275	0.9742
Q2CCC	0.9971	0.9680	0.9876
R2	0.9943	0.9369	0.9753
R	0.9971	0.9682	0.9876
k	1.0001	0.9955	0.9987
K'	0.9994	0.9988	0.9992
R_0^2	1.0000	0.9998	1.0000
$R_0'^2$	1.0000	1.0000	1.0000
M	-0.0058	-0.0671	-0.0253
N	-0.0058	-0.0673	-0.0253
R_m	0.9189	0.7021	0.8221
A_f	1.0068	1.0282	1.0132
B_f	1.0002	1.0038	1.0013

Tableau III.17 : Différentes erreurs de modèles LSSVM.

Erreurs	Train	Test	Global
ERAM%	0.6840	2.9266	1.3581
REQM	0.8189	2.9087	1.7356
EQM	0.6706	8.4604	3.0122

La figure III.33 montre le diagramme de dispersion de la tension superficielle calculée en fonction de la tension superficielle expérimentale. Ces diagrammes montrent clairement la concordance entre les valeurs calculées et les valeurs expérimentales à travers une dispersion serrée des points sur la première bissectrice. Cependant, comme pour les modèles précédant on remarque l'existence d'un certain nombre de points légèrement éloignés qui peuvent être attribués à l'existence des points aberrants.

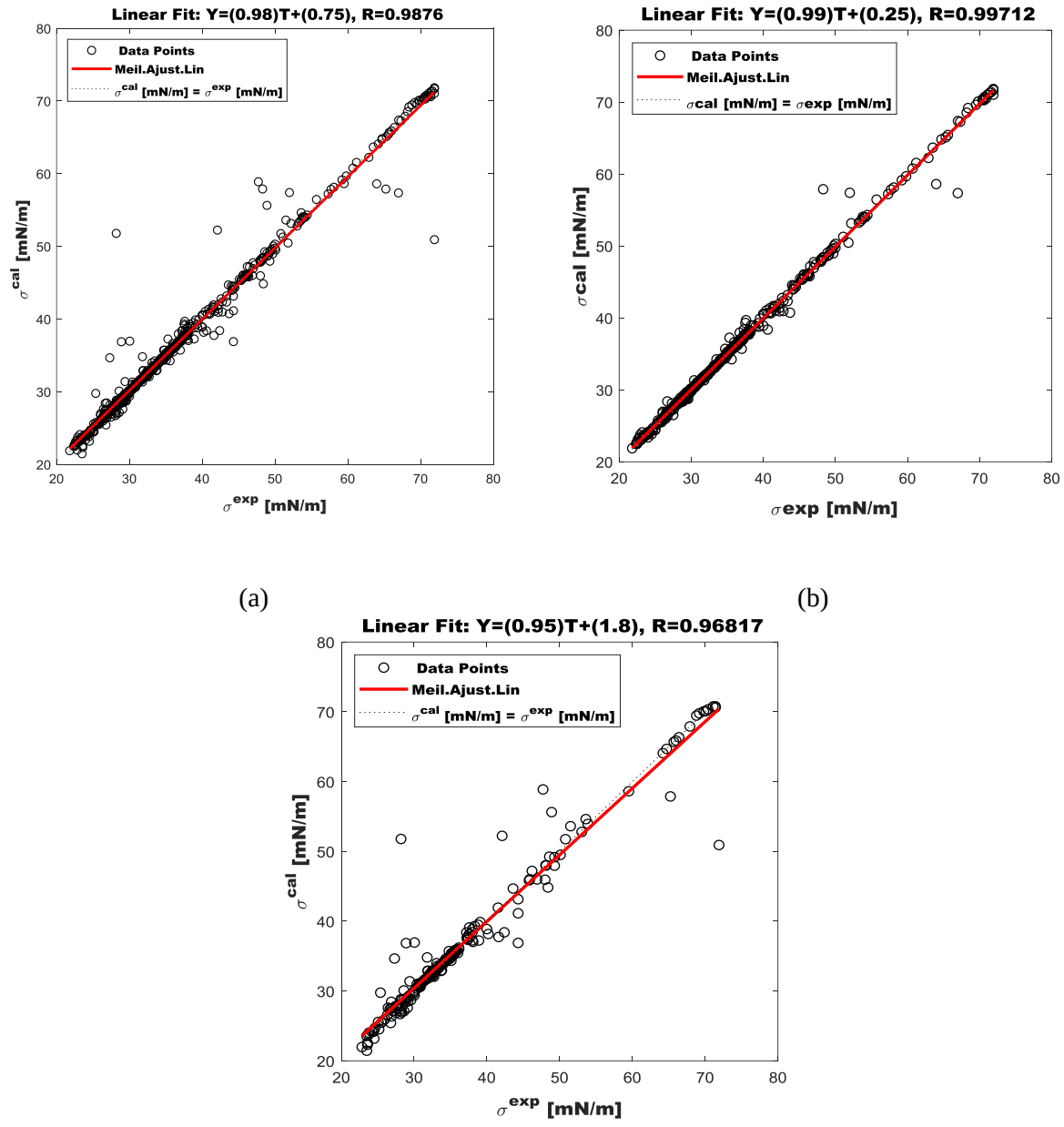


Figure III.33 : diagramme de dispersion de la tension superficielle calculé en fonction de la tension superficielle expérimentale : (a) Base globale, (b) Ensemble d'apprentissage, (c) Ensemble du test par le modèle LSSVM.

Afin de visualiser quantitativement la distribution des erreurs relatives nous avons adoptés deux types de diagrammes. Le premier montre l'erreur relative en fonction de la tension superficielle et le deuxième montre la distribution de cette erreur. La figure III.34 représente l'erreur relative en fonction de la tension superficielle. Pour L'erreur maximale les quasis totalité des points sont situés entre ±20 %

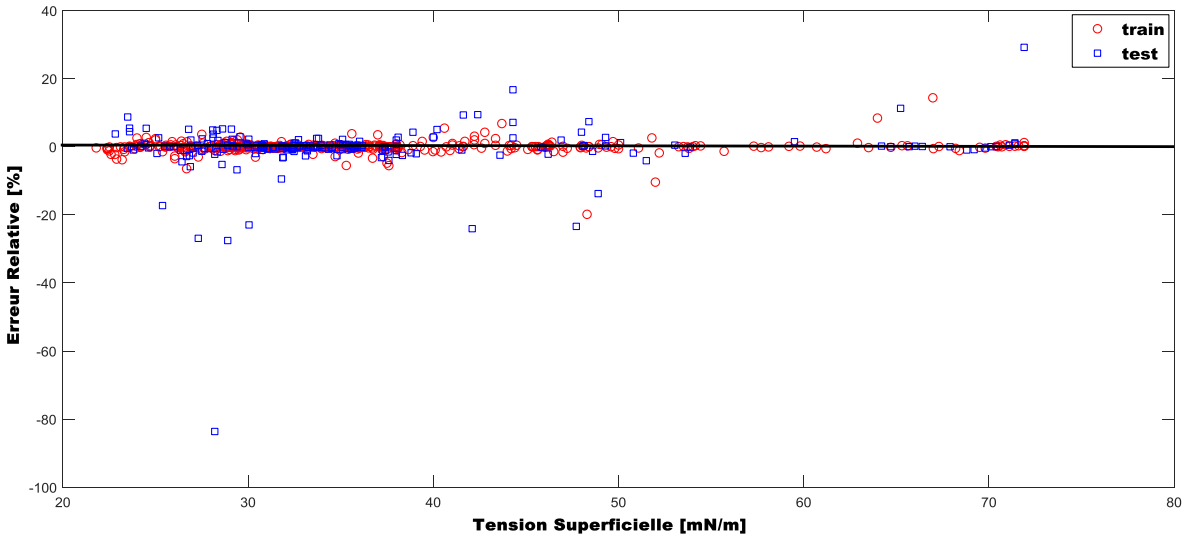


Figure III.34 : L'erreur relative (%) de la tension superficielle par le modèle LSSVM.

Ces remarques relatives à l'ER sont mieux illustrées sur le diagramme de distribution de figure III.35

Pour confirmer la capacité de généralisation du modèle développé nous avons réalisé des calculs d'interpolation et d'extrapolation pour la température et la composition. La figure III.36 montre les isothermes de la variation de la tension superficielle en fonction de la composition du mélange [C2MIm][BTI]-Dimethyl sulfoxide pour le modèle LSSVM. Les points interpolés suivent l'allure des points calculés.

De la même façon la figure III.37 représente le calcul d'interpolation et d'extrapolation de la température. Si la capacité d'interpolation (305.15 K) est acceptable par contre la capacité d'extrapolation (290.15 K et 315.15 K) est mauvaise. Ceci signifie que le modèle LSSVM ne peut pas être utilisé en dehors du domaine de température étudié.

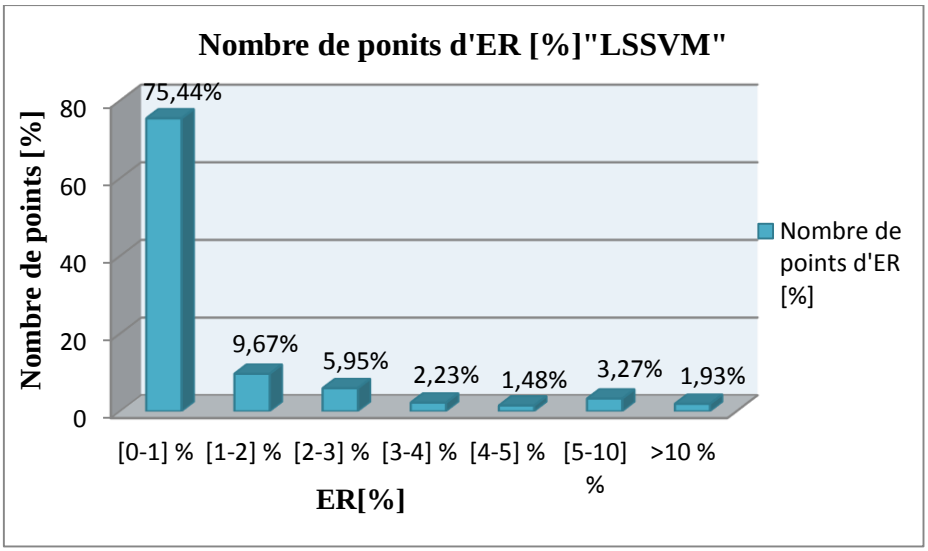


Figure III.35: Diagramme de distribution d'ER[%] du modèle LSSVM.

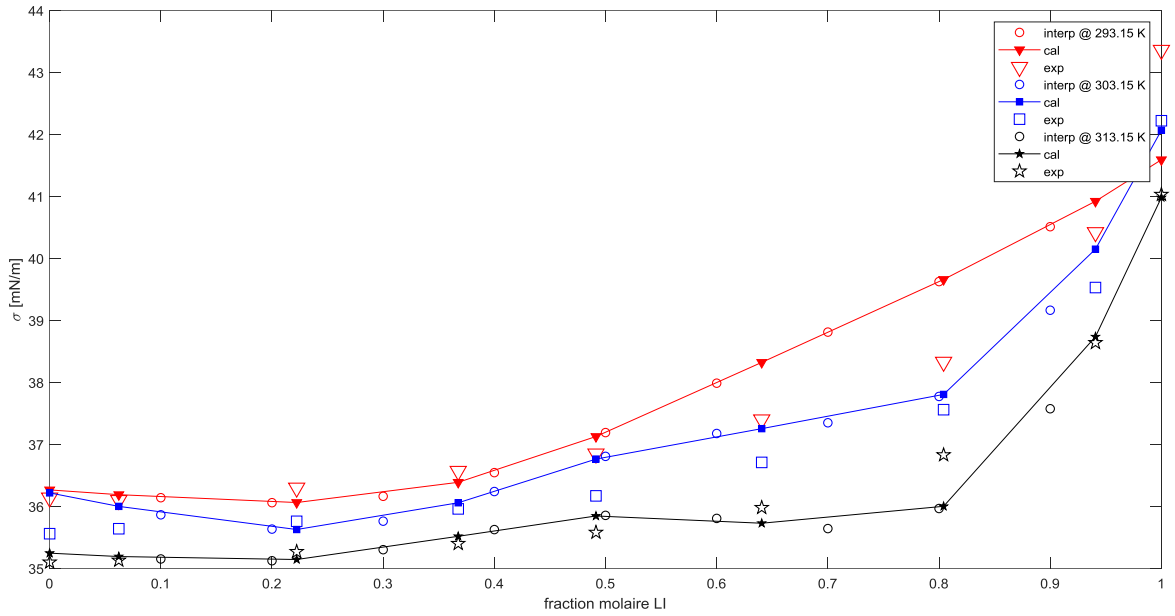


Figure III.36 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle LSSVM.

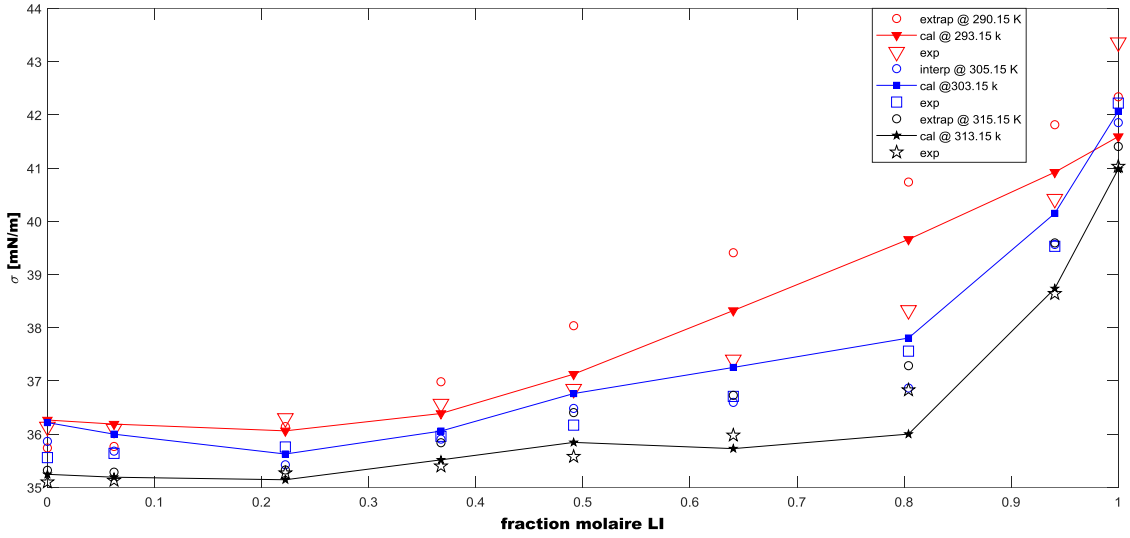


Figure III.37 : Isothermes de la variation de tension superficielle en fonction de la composition des mélanges binaires de [C2MIm][BTI]-Dimethyl sulfoxide du modèle LSSVM.

III.4 Analyse de sensibilité des modèles ANFIS, SVM et LSSVM

Pour étudier en profondeur la tension superficielle du mélange à l'aide du modèle ANFIS, SVM et LSSVM, une analyse de sensibilité a été réalisée pour évaluer l'effet des paramètres d'entrée (la température, la fraction molaire du LI ainsi que les profiles σ pour les cations, les anions et les solvants) sur la tension superficielle. L'entrée la plus pertinente peut être reconnue par le facteur de pertinence r, compris entre -1 et +1. Le facteur de pertinence (r) a été appliqué pour évaluer le degré d'influence de chaque paramètre sur la tension

superficielle du mélange, les valeurs ont été calculées à l'aide de la formule suivante *(Baghban et al, 2018)* :

$$r_k = \frac{\sum_{i=1}^{n}(X_{k,i}-\overline{X_k})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(X_{k,i}-\overline{X_k})^2 \sum_{i=1}^{n}(y_i-\bar{y})^2}}, k = 1, \dots, 8 \qquad \text{III.2}$$

Où $X_{k,i}$ est l'ième valeur de l'entrée k, Yi est l' ième valeur de sortie, $\overline{X_k}$ est la valeur moyenne de l'entrée k, $\bar{y}$ est la valeur moyenne du paramètre de sortie, La valeur absolue de r_k a une relation directe sur la sortie, comme on peut le voir sur la figure III.38:

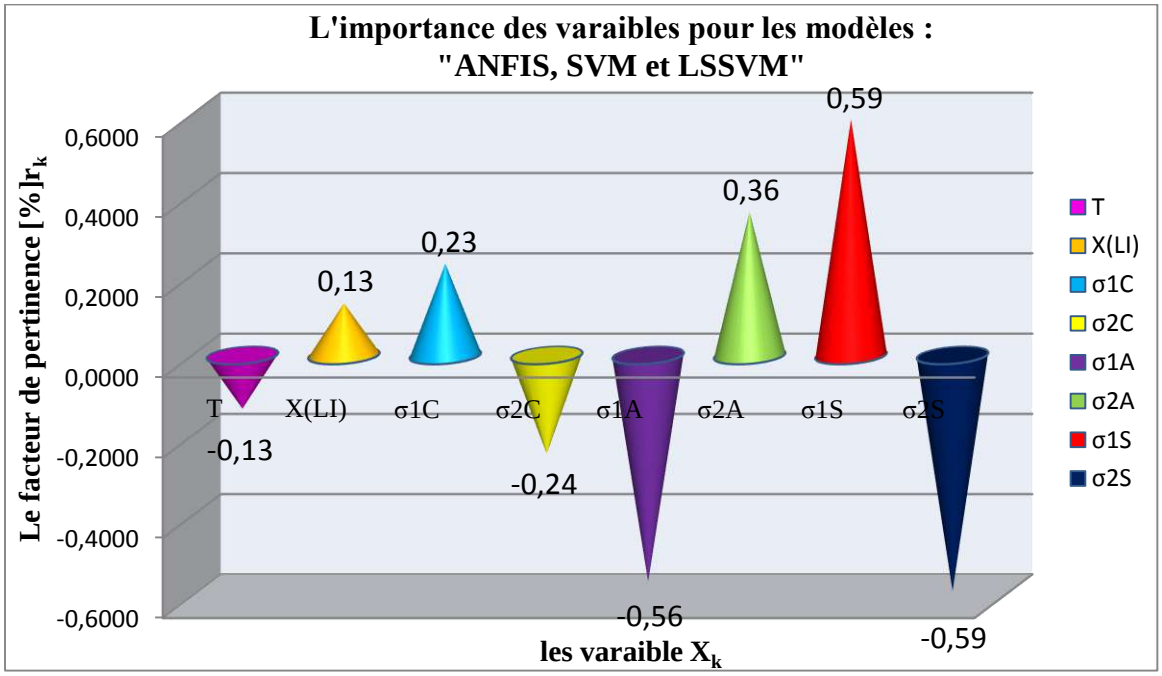


Figure III.38 : Résultats de l'analyse de sensibilité réalisée sur les modèles ANFIS, SVM et LSSVM.

Comme on peut le voir sur la figure III.38, la tension superficielle montre une dépendance directe des variables (σ1S, σ2A, σ1C, X(LI)) et une dépendance opposée des variables (σ2S, σ1A, σ2C, T). De plus, σ1S et σ2S sont les variables d'entrées les plus pertinentes avec un facteur de pertinence de 0,59 et de -0.59 respectivement par contre T et X(LI) sont les moins pertinents avec un facteur de pertinence de 0.13 et de -0.13 respectivement.

III.5 Domaine d'applicabilité

Les diagrammes de dispersions des modèles optimisé ont montré un certains nombres des points éloignés de la première bissectrice qu'on a attribué à l'existence probable de points aberrants. Ceci nous a motivé à réaliser un diagnostic des données potentiellement suspectes

selon la méthode décrite dans les deux références suivantes (Hosseinzadeh et Hemmati-Sarapardeh, 2014 et Baghban et al, 2017).

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 (Hosseinzadeh et Hemmati-Sarapardeh, 2014, Baghban et al. 2017). 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 est créée sur la base de la formule suivante :

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

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 qui représente les résidus normalisés en fonction du ‘’Leverage’’ levier (Hat) peut détecter graphiquement les valeurs aberrantes. Dans ce travail le modèle RNA ayant une meilleure capacité prédictive que les systèmes ANFIS, SVM et LSSVM a été choisi pour l’analyse de détection des valeurs potentiellement suspectes. La figure III.39 montre le graphique de William établit à partir des résultats du modèle de RNA. Dans ce graphique, le levier critique (H^*) (seuil) est généralement fixé à la valeur donnée par l'équation suivante:

$$H^* = \frac{3(n+1)}{m} = \frac{3(8+1)}{672} = 0.0402 \qquad \text{III.5}$$

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

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

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

Le résidu normalisé égal à trois est normalement 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’’.

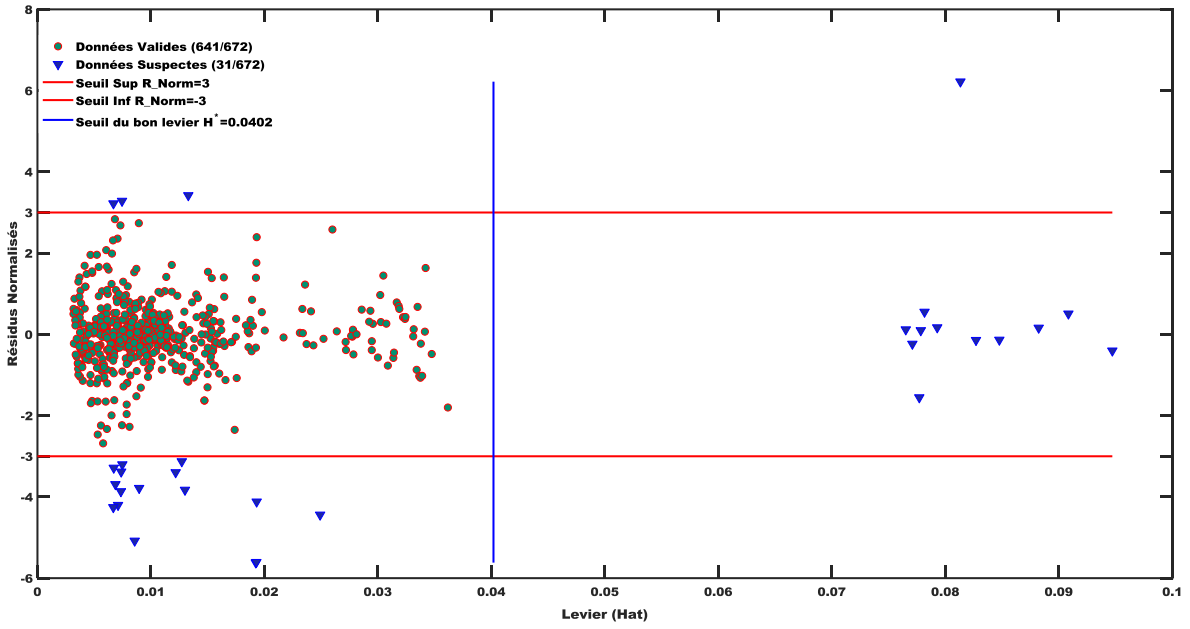


Figure III.39 : Diagnostic des données potentiellement suspectes et domaine d'applicabilité du modèle.

Comme le montre la figure III.39, on remarque l’existence de 31 points aberrants sur les 672 points de la base de données. Ceci confirme donc les constatations faites lors des interprétations des diagrammes de dispersion et les diagrammes de l’erreur relative.

III.6 Comparaison des résultats des modèles optimisés

Afin de visualiser les résultats de modélisation par les méthodes d'intelligence computationnelle, considérées toutes comme acceptables, on termine ce chapitre par un récapitulatif de comparaison

La figure III.40 montre une comparaison de l'ERAM% obtenue par les différents modèles.

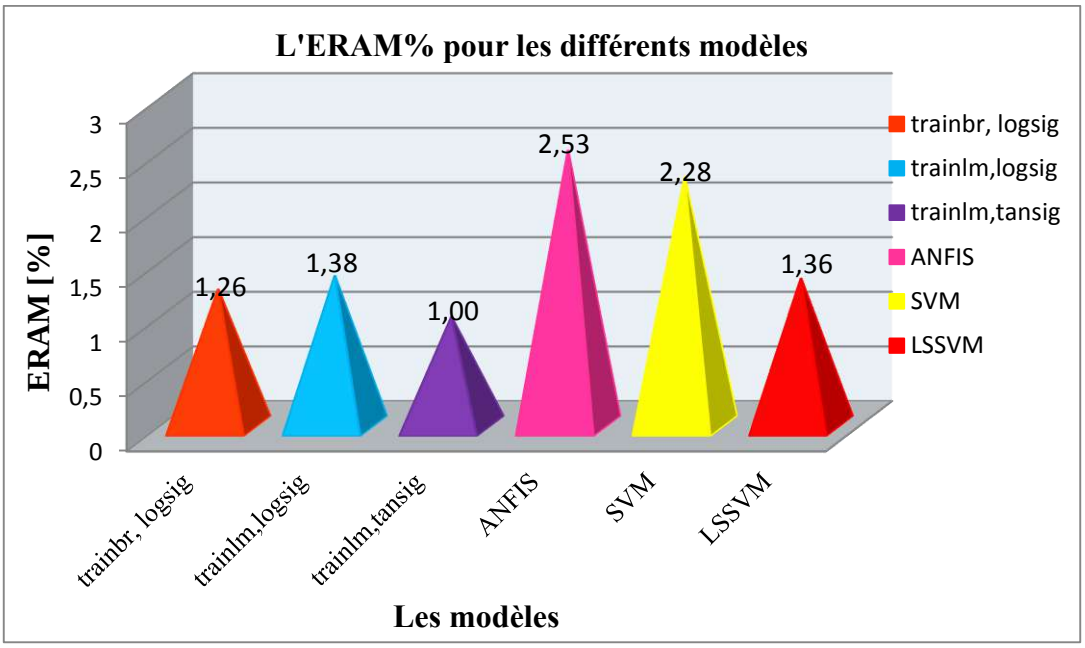


Figure III.40 : Comparaison des ERAM% globales.

La comparaison de la capacité de généralisation et la tendance de sur-apprentissage des différents modèles est illustrées par les figures de distribution de l'ER des ensembles d'apprentissage et de test (figure III.41 et figure III.42). Ces figures montrent que le modèles SVM possède la capacité de généralisation la plus faible et une tendance de sur-apprentissage plus élevé. Alors que le modèle de RNA (trainlm, tansig) possède la capacité de généralisation la plus élevée. Ces remarques sont confirmées par le tableau de comparaison des ERAM% des ensembles d'apprentissage et de test (tableau III.18). Les modèles de RNA présentent les écarts les plus faibles par rapports aux autres méthodes, par conséquent des capacités de généralisation plus élevées.

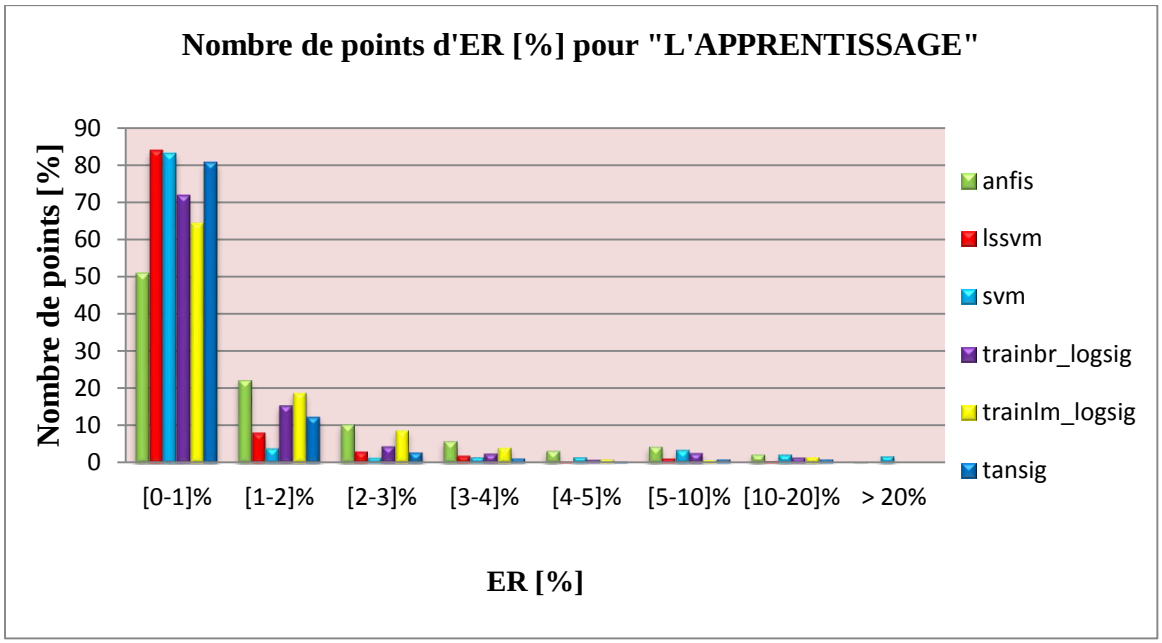


Figure III.41 : Figure de distribution de l'ER des ensembles d'apprentissage

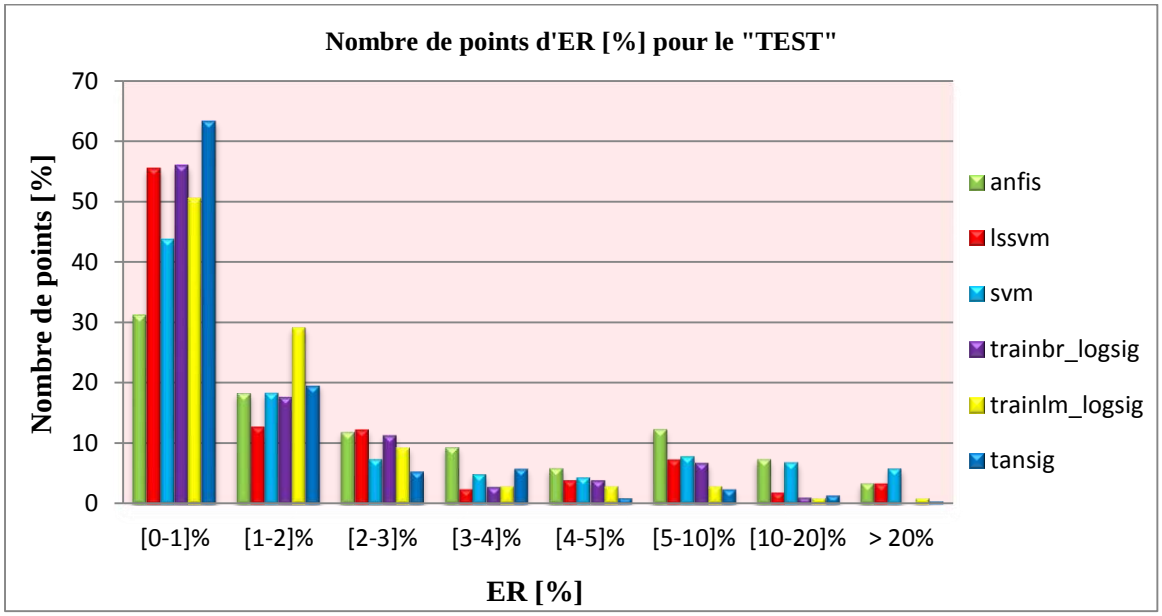


Figure III.42 : Figure de distribution de l'ER des ensembles de test

Tableau III.18 : Comparaison des écarts entre l'ERAM% de l'ensemble de test et de l'ensemble d'apprentissage pour les différents modèles développés.

	Train	**Test**	**Ecart (Test-Train)**
RNA (trainbr, logsig)	1.1516	1.5602	0.4086
RNA(trainlm, logsig)	1.2495	1.6928	0.4433
RNA(trainlm, tansig)	0.8457	1.3742	0.5285
ANFIS	1.8074	4.2138	2.4064
SVM	1.3914	4.3660	2.9446
LSSVM	0.6840	2.9266	2.2426

III.7 Conclusion

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 des 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, ANFIS, SVM et LSSVM sont tous statistiquement valides avec une ERAM globale comprise entre 1% et 2.53%. Cependant, les modèles obtenus par la méthode de RNA ont des performances de corrélation plus précises et des capacités de généralisation et d'interpolation nettement meilleures. Bien que les techniques ANFIS, SVM et LSSVM aient à peu près le même niveau de précision pour l'ensemble des données de test, LSSVM a un développement mathématique plus simple, ce qui conduit à une vitesse de convergence plus rapide, et par conséquent, il est moins coûteux en calcul. L'analyse des données a révélé l'existence d'un certain nombre de données suspectes et a permis de définir le domaine d'applicabilité des modèles développés. En résumé, l'élaboration réussie de modèles pour la corrélation de la tension superficielle des mélanges de LI dans la présente étude nous permet de promouvoir ces méthodes pour d'autres propriétés physicochimiques des mélanges d'IL dans le cadre d'études futures visant à surmonter les limitations du développement des industries et des technologies basées sur les LI.

Conclusion générale

Conclusion générale et perspectives

Cette étude nous a permis d'apprécier l'importance accordée actuellement aux : (i) liquides ioniques comme de nouveaux classes de matériaux dont les applications potentielles ne cessent de se révéler ; (ii) modèle 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) méthodes d'intelligence computationnelle comme outils de modélisation ayant la capacité d'apprendre une tâche spécifique à partir de 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.

La tension superficielle des liquides ioniques 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 ceux de séparation, d'adsorption, de distillation et de performances de membranes biologiques, pour ne citer que quelques exemples. Par conséquent le développement de modèles pour prédire ou corréler cette propriété est fondamental pour la conception et l'exploitation de ces procédés.

Dans ce travail l'utilité et la capacité des modèles basés sur les réseaux de neurones artificiels (RNA), les système d'inférence neuro-flou adaptatif (ANFIS), les machines à vecteur de support (SVM) et la version à moindres carrés des machines à vecteurs de support (LSSVM) optimisés par divers algorithmes d'optimisation, pour corréler la tension superficielle de systèmes binaires contenant divers LI avec l'eau et LI avec divers solvants organiques ont été évaluées à l'aide des données expérimentales tirées de la littérature. L'ensemble de données contient 672 points de données expérimentaux. 70% des données, choisies aléatoirement, de l'ensemble de données ont été utilisées comme ensemble d'apprentissage et les données restantes ont été utilisées comme ensemble test. La température (T), la composition du LI (X_{IL}), 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 utilisés comme variables 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 40 mélanges binaires LI+Eau et LI+différents solvants organiques par les méthodes d'intelligence computationnelle donne des résultats très satisfaisants. Compte tenu de leur précision et de leur simplicité, ces modèles peuvent être utilisés efficacement dans la pratique.
2. Les modèles développés par les méthodes de RNA, ANFIS, SVM et LSSVM sont tous statistiquement valides avec une ERAM globale comprise entre 1% et 2.53%. Cependant, les modèles obtenus par la méthode de RNA ont des performances de corrélation plus précises et des capacités de généralisation et d'interpolation nettement meilleures.
3. Bien que les techniques ANFIS, SVM et LSSVM aient à peu près le même niveau de précision pour l'ensemble des données de test, LSSVM a un développement mathématique plus simple, ce qui conduit à une vitesse de convergence plus rapide, et par conséquent, il est moins coûteux en calcul.
4. 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.

En résumé, l'élaboration réussie de modèles pour la corrélation de la tension superficielle des mélanges de LI dans la présente étude nous permet de promouvoir ces méthodes pour d'autres propriétés physicochimiques des mélanges de LI dans le cadre d'études futures visant à surmonter les limitations du développement des industries et des technologies basées sur les LI.

Références bibliographiques

Références bibliographiques

Almeida H.F.D., Lopes-da-Silva J.A., Freire M.G. et Coutinho J.A.P., J. Chem. Thermodyn, 2013.

Amich.G et Benrouissi A., Modélisation neuronale des coefficients d'activité à la dilution infinie de l'eau dans les liquides ioniques en utilisant des profils sigma comme descripteurs moléculaires, Memoire de master, universite yahia fares de medea, Algerie 2016.

Ayad A., Contribution à l'étude de la réduction des émissions de cov par les liquides ioniques, Thèse de doctorat, Universite Abou-Bekr Belkaid – Tlemcen, Algerie 2017.

Baghban J.A., **Shafiee M., Ahmadi M.H. et Chau K.W.**, Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids, Engineering Applications of Computational Fluid Mechanics, 2018.

Baghbana A., Ahmadib M.A. et Shahrakia B.H., Prediction carbon dioxide solubility in presence of various ionicliquids using computational intelligence approaches, J. of Supercritical Fluids, 2015.

Bouchardy L., Elaboration de liquides ioniques (chiraux) réversibles et applications en catalyse organique et en glycochimie. Carbènes N hétérocycliques chiraux : synthèse et application dans la réaction d'addition conjuguée, thèse de doctorat de l'université paris –saclay préparée a l'université paris-sud ,2016 .

Chouih KH., Modélisation du coefficient d'Activité à dilution infinie des liquides ioniques dans différents solutés (eau, n-alcanes, cycloalcanes) en utilisant les machines à vecteur de support mémoire de master université Dr yahia fares Médéa, Algérie 2018.

Diedenhofen M. et klamt A.,COSKO6RS as a tool for property prediction of IL mixtures-A review ,Fluid Phase Equilibria ,2010.

Diop A., Extraction, Depol ymerisation et valorisation de la lignine kraft de la liqueur noire, Thèse de doctorat, L'université du Québec À Trois-Rivières,Canada 2014.

Domanska U., Zawadzki M. et Lewandrowska A., 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. Thermodynamics 48,2012.

Domaoska U., **Pobudkowska A. et Rogalski M.**, J. Colloid Interf. Sci., 2008.

Geppert-Rybczyoska M., Lehmann J. K. et Heintz A., J. Chem. Eng. Data, 2011.

Geppert-Rybczyoska M., Lehmann J. K., Safarov J. et Heintz A., J. Chem. Thermodyn., 2013.

Gevrey M., Dimopoulos I. et Lek S., Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecological Modelling, 2003.

Gutel T., Les liquides ioniques, leur utilisation et leur role comme solvants de réaction catalytique, Thèse de doctorat, Université Claude Bernard - Lyon I, Français 2007.

Hashemkhani M., Soleimani R., Fazeli H., Lee M., Bahadori A. et Tavalaeian M., Prediction of the binary surface tension of mixtures containing ionic liquids using Support Vector Machine algorithms , Journal of Molecular Liquids 211,2015.

Huang Y., Zhang X., Zhao. Y., Zeng S., Dong H., Zhang S., Physical Chemistry Chemical Physics, 2015.

KASRI S., Etude et Modélisation de Déclin de Potentiel de Surface par les Réseaux de Neurones, de magister, Faculté des Sciences de l'ingénieur université badji mokhtar, Annaba, 2010.

Kharroubi J., Etude de techniques de classement "Machines à vecteurs supports" pour la vérification automatique du locuteur. domain_other. Télécom Paris Tech, 2002.

Mohamadally H. et Fomani B., SVM : Machines à Vecteurs de Support ou Separateurs a Vastes Marges**,** BD Web, ISTY3**,** Versailles St Quentin, France 2006.

Morton M.D. et Hamer C.K., Ionic liquids – the beginning of the end or the end of the beginning?- A look at the life of ionic liquids through patent claims, Separation and Purification Technology, 2017.

Moutiers G. et Billard I., Les liquides ioniques : des solvants pour l'industrie, Technique de L'ingénieur, 2011.

Olden J.D., Joy M.K. et Death R.G., An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, Ecological Modelling, 2004.

Paduszynski K., 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., 2016.

Quijano G., Couvert A., Amrane A., Ionic liquids: Applications and future trends in bioreactor technology, Bioresource technology, 2010.

Razavi R., Bemani A., Baghban A., Mohammadi A. H.et Habibzadeh S., An insight into the estimation of fatty acid methyl ester based biodiesel properties using a LSSVM model fuel 243 ,2019.

Revelli A.L., 'Etude thermodynamique des liquides ioniques: Applications à la protection de l'Environnement'. Thèse de doctorat : Génie des Procédés et des Produits. Inst. Nat. Polytech, Lorraine, 2010.

Rilo E., Pico J., García-Garabal S., Varela L. M. et Cabeza O., Fluid Phase Equilibr., 2009.

Rilo E., Domínguez-Pérez M., Vila J., Varela L. M. et Cabeza O., J. Chem. Thermodyn., 2012.

Rizkalla N., Nanoparticules et réseaux de neurones artîfîciels: de la préparation à la modélisation, Thèse de doctorat, Faculté de Pharmacie, Unïversité de Montréal, canada 2005.

Russo J.W. **et Hoffmann M.M.,** 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, 2011.

Shabri A. et Suhartono, Stearmflow forecasting using leat-squares support vector machines, Journal des Sciences Hydrologiques, 2012.

Si-Moussa C., Modélisation des Équilibres de Phases à Pressions Élevées par les Réseaux de Neurones Artificiels, Thèse de Doctorat, École Nationale Polytechnique, Alger 2010.

Soleimani R., Dehaghani A.H.S., Shoushtari N.A., Yaghoubi P., et Bahadori A., Toward an intelligent approach for predicting surface tension of binary mixtures containing ionic liquids, Korean J. Chem. Eng, 2017.

Suparta w. et Maulana Alhasa K. ,Modeling of Tropospheric Delays Using ANFIS, Springer International Publishing , 2016.

Tariq M., Freire M. G., Saramago B., Coutinho J. A. P., Lopes J. N. C. et Rebelo L. P. N., Surface tension of ionic liquids and ionic liquid solutions, JChem. Soc. Rev, 2012.

Todeschini R., Ballabio D., et Grisoni F., Beware of unreliable Q2! A comparative study of regression metrics for predictivity assessment of QSAR models, J. Chem. Inf. Model., 2016.

Torrecilla J. S., Rafione T., García J. et Rodríguez F., J. Chem. Eng. Data, 2008.

Wan Normazlan W.M.D., Sairi N.A., AliasY., Udaiyappan A.F., Jouyban A., et Khoubnasabjafari M., Composition and Temperature Dependence of Density, Surface Tension, and Viscosity of EMIM DEP/MMIM DMP + Water + 1Propanol/2-Propanol Ternary Mixtures and Their Mathematical Representation Using the Jouyban−Acree Model, J. Chem. Eng. Data, 2014.

Wandschneider A., Lehmann J. K. et Heintz A., J. Chem. Eng. Data, 2008.

Wang J.Y., Zhao F.Y., Liu Y.M., Wang X.L. et Hu Y.Q., Fluid Phase Equilibr, 2011.

Wang J.Y., Zhang X.J., Hu Y.Q., Qi G.D. et Liang L.Y., J. Chem. Thermodyn., 2012.

Zhang S., Zhou Q., Lu X., Song Y. et Wang X., Physicochemical Properties of Ionic Liquid Mixtures-Springer Netherlands, 2016.

https://fr.wikipedia.org/wiki/Machine_%C3%A0_vecteurs_de_support#Historiquelearningtorankchallenge.yahoo.com

Http://www.perso.univrennes1.fr/mathieu.pucheault/_PDF/Publis/2010_Pucheault_Dunod.pdf (Consulté le 03/05/2017).

Http://www.perso.univrennes1.fr/mathieu.pucheault/_PDF/Publis/2010_Pucheault_Dunod.pdf (Consulté le 03/05/2017).