N₀ d'ordre : .....

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
University of YAHIA FARES MEDEA



Faculty of Technology
Department of Process and Environmental Engineering

THESIS

Submitted in fulfillment of the Award of

DOCTOR

Of Pharmaceutical chemistry

Speciality : PHARMACEUTICAL ENGINEERING

Prepared & Presented by

EULDJI IMANE

Master in Pharmaceutical Process Engineering at YAHIA FARES University of Medea.

Theme

Contribution To Modeling The Solubility Of Solid Drugs In Different Solvents

Contribution A La Modélisation Dans Le But De Corréler et/ou Prédire La Solubilité De Médicaments Solides Dans Differents Solvants

Defense publicly on : 30/01/2024

Against the jury composed of:

Mm. : KHAOUANE Latifa	Professor at the University of MEDEA	President
Mr. : SI-MOUSSA Chérif	Professor at the University of MEDEA	Supervisor
Mr. : BENKORTBI Othmane	Professor at the University of MEDEA	Co-Supervisor
Mm. : NASRI Loubna	Professor at the University of CONSTANTINE 3	Examiner
Mm. : FERGHANI Zineb	Professor at the University of MEDEA	Examiner
Mr. : HANINI Salah	Professor at the University of MEDEA	Invite

2023/2024

This thesis was prepared within the Laboratory of Biomaterials and Transport Phenomena (LBMPT) at the University of Medea, ALGERIA

Acknowledgments

First and foremost, I would like to thank Allah (SWT) for showering his blessings on me and for giving me the strength and patience to finish my Doctoral studies.

Then, I would like to express my deepest gratitude to the exceptional individuals who supported and guided me throughout this transformative journey, enabling me to successfully obtain my doctorate despite the many unforeseen obstacles that arose.

My most profound appreciation goes to Professor Si-Moussa Chérif and Professor Benkortbi Othmane, my Ph.D. advisors and mentors, for their time, effort, continuous support, patience, motivation, and immense knowledge. Their vast wisdom and wealth of experience have inspired me throughout my studies. I could not have imagined having better advisors and mentors for my Ph.D. study.

I would like to extend my sincere gratitude to Professor Mabrouk HAMADACHE and Professor Abdeltif AMRANE for providing assistance and their insightful feedback

I greatly appreciate the knowledge received through the collaborative work undertaken with Professor paduszynski kamil from the university of the Warsaw University of Technology, Poland.

I would also like to extend my thanks to the members of the jury for having given us the honor and the pleasure of accepting and judging our work.

Also, extend my deepest gratitude to all our teachers for the pieces of knowledge they gave us during our university studies.

Lastly, I want to express my deepest gratitude to my family, who believes in my abilities and support. Your encouragement played an integral role in my accomplishments.

In the end, I thank all the people who contributed, directly or indirectly, to the production of this thesis.

Dedication

I dedicate my dissertation work to my family and many friends.

To my loving parents. Words can hardly describe my thanks and appreciation to you. You have been my source of inspiration, support, and guidance. You have taught me to be unique and determined, to believe in myself, and to always persevere. I am truly thankful and honored to have you as my parents, may Allah protect them and give me strength so that I can give them back a little bit of their goodness and kindness.

To my siblings Rafik, and Riadh who have been by my side since I remember. You have been a listener and a supporter of all my endeavors.

To my dear grandfather and my dear grandmother, I wish them a long life.
To the whole family.

To my dear best friend Widad, I wish you all the best. You have been with me for all the long of this great journey.

To all my friends, and all the teachers of the faculty of technology. Whom have supported me throughout the process. I will always appreciate all they have done.

And finally, to all who wish me only the best.

Imène

ملخص

المساهمة في نمذجة ذوبان الأدوية الصلبة في المذيبات المختلفة

الغرض الأساسي من هذه الأطروحة هو المساهمة في مجال التنبؤ بقدرة الذوبان مع ثلاث تقنيات نمذجة متقدمة تم تصميمها باستخدام معلمات إدخال قوية استنادًا إلى المعلومات الجزيئية؛ يعالج التنبؤ وارتباط قابلية ذوبان الأدوية الصلبة في مذيبات خضراء مختلفة، وهما: ثاني أكسيد الكربون فوق الحرجة والسوائل الأيونية النقية. تغطي هذه الأطروحة مجموعة متنوعة من الأساليب لتوفير حساب للحالة الحالية للفن وتبرر خيارات الأساليب المقترحة للتنبؤ بالذوبان. تقدم المساهمة الأولى استخدام تقنيات العلاقة بين الهيكل الكمي - مع خوارزمية الشبكات العصبية الاصطناعية لربط القابلية للذوبان في 148 مركبة دوائية في ثاني أكسيد الكربون الفائق الحرج. تتعامل المساهمة الثانية مع دراسة مقارنة بين سبعة خوارزميات مختارة من الأرياان: مُحسِّن من اليعسوب، محسن النمل، محسن الذئب الرمادي ، مُحسّن مستعمرة النحل الاصطناعية ، محسن سرب الجسيمات ، محسن الحوت ، وجزيرة هجينة مع محسن وولف رمادي في رمادي ناعم في -تلاشي فرطاميرامترات من خوارزمية الانحدار ناقلات الدعم على قاعدة بيانات تضم 168 مركبًا صيدلانيًا للذوبان في ثاني أكسيد الكربون فوق الحرج. تُظهر المساهمة الثالثة مزيجًا قويًا من تقنيات العلاقة بين الهيكل الكمي-مع نموذج فحص يشبه الموصل للتقنيات القائمة على المذيبات الحقيقية في نمذجة 14 مكونات صيدلانية نشطة في السوائل الأيونية المختلفة. تم تأكيد القوة الجيدة ، المتانة ، والقوة التنبؤية لجميع النماذج المقترحة مع مختلف المقاييس الإحصائية والرسمية.

الكلمات المفتاحية: الذوبان، المركب الدوائي، ثاني أكسيد الكربون فوق الحرج، السوائل الأيونية، العلاقة الكمية بين البنية والملكية، نموذج فحص يشبه الموصل للمذيبات الحقيقية، الشبكات العصبية الاصطناعية، دعم انحدار المتجهات، الاستدلال الفوقي

Abstract

Contribution To Modeling The Solubility Of Solid Drugs In Different Solvents

The primary purpose of this thesis is to contribute to the area of solubility prediction with three advanced modelling techniques built with the use of powerful input parameters based on molecular information; Addresses the prediction and correlation of the solubility of solid drugs in two different green solvents, namely: supercritical carbon dioxide and pure ionic liquids. This thesis covers a variety of approaches to provide an account of the current state of the art and justifies the choices of the proposed methods for solubility prediction. The first contribution introduces the use of Quantitative Structure–Property Relationship techniques with an Artificial neural networks algorithm for correlating the solubility of 148 drug compounds in supercritical carbon dioxide. The second contribution deals with a comparison study between seven selected meta-heuristic algorithms: Dragonfly optimizer, Ant Lion optimizer, Grey Wolf optimizer, Artificial Bee Colony optimizer, Particle Swarm optimizer, Whale optimizer, and a hybrid Particle Swarm with Grey Wolf optimizer in fine-tuning hyperparameters of support vector regression algorithm on a database of 168 pharmaceutical compounds solubilities in supercritical carbon dioxide. The third contribution shows the powerful combination of Quantitative Structure–Property Relationship techniques with a conductor-like screening model for Real Solvents based techniques in modelling 14 active pharmaceutical ingredients solubilities in various ionic liquids. The Good-of-fit, robustness, and predictive power of all proposed models were confirmed with various statistical and graphical metrics.

Keywords: Solubility, Drug Compound, supercritical carbon dioxide, ionic liquids, Quantitative Structure – Property Relationship, conductor-like screening model for Real Solvents, Artificial neural networks, support vector regression, meta-heuristic

Résume

Contribution A La Modélisation De La Solubilité De Médicaments Solides Dans Differents Solvants

L'objectif principal de cette thèse est de contribuer au domaine de la prédiction de la solubilité avec trois techniques de modélisation avancées construites avec l'utilisation de paramètres d'entrée puissants basés sur des informations moléculaires ; Aborde la prédiction et la corrélation de la solubilité des médicaments solides dans deux solvants verts différents, à savoir : le dioxyde de carbone supercritique et les liquides ioniques purs. Cette thèse couvre une variété d'approches pour rendre compte de l'état actuel de l'art et justifie les choix des méthodes proposées pour la prédiction de la solubilité. La première contribution introduit l'utilisation de techniques quantitatives de relation structure-propriété avec un algorithme de réseaux de neurones artificiels pour corréler la solubilité de 148 composés médicamenteux dans le dioxyde de carbone supercritique. La deuxième contribution traite d'une étude comparative entre sept algorithmes méta-heuristiques sélectionnés : l'optimiseur Dragonfly, l'optimiseur Ant Lion, l'optimiseur Grey Wolf, l'optimiseur Artificial Bee Colony, l'optimiseur Particle Swarm, l'optimiseur Whale, et un hybride Particle Swarm avec l'optimiseur Grey Wolf in fine. -réglage des hyperparamètres de l'algorithme de régression vectorielle de support sur une base de données de 168 solubilités de composés pharmaceutiques dans le dioxyde de carbone supercritique. La troisième contribution montre la puissante combinaison de techniques quantitatives de relation structure-propriété avec un modèle de criblage de type conducteur pour les techniques basées sur des solvants réels dans la modélisation de la solubilité de 14 ingrédients pharmaceutiques actifs dans divers liquides ioniques. Le bon ajustement, la robustesse et le pouvoir prédictif de tous les modèles proposés ont été confirmés par diverses mesures statistiques et graphiques.

Mots clés : solubilité, composé médicamenteux, dioxyde de carbone supercritique, liquides ioniques, relation structure quantitative – propriété, modèle de criblage de type conducteur pour les solvants réels, réseaux neuronaux artificiels, régression vectorielle de support, méta-heuristique

List Of Contents

S.N.	Content	Pg. No.
i.	List of Contants……..	i
ii.	List of Abbreviations……..	v
iii.	List of Nomenclature and symbol……..	xiv
iv.	List of Figures…….....	xviii
v.	List of Tables……..	xxi
vi.	List of Publications……	xxiii
	General Introduction	1-9
1.	Background and Motivations…….....	1
2.	Green Solvents as a Solubility Improvement Techniques……..	2
2.1.	Supercritical Carbon Dioxide……..	3
2.2.	Ionic Liquids…….....	5
3.	Experimental Measurement Limitations……..……..	6
4.	Aims and Objectives of this Thesis…….....	8
5.	Thesis Organisation…….....	9
	Chapter I Literature Review	10-84
I.1.	Introduction…….....	10
I.2.	Solubilities of Solid Compounds in Supercritical Carbon Dioxide……	10
I.2.1.	Literature Review on Equation of State Correlations…….....	11
I.2.2.	Literature Review on Empirical and Semi-empirical Equations Correlations	21
I.2.3.	Literature Review on Solution Models……	32
I.2.4.	Literature Review on Group Contibution Techniques…….	35
I.2.5.	Literature Review on Machine Learning Approaches……...	37
I.2.6.	Literature Review on Other Techniques…….	48
I.2.7.	Critical Review and Establish the Significance of QSPR Techniques in our Reported Research……..	53

I.3.	Solubilities of Solute Compounds in Ionic Liquids……	56
I.3.1.	Literature Review on Carbon Dioxide as a Solute……	57
I.3.2.	Literature Review on Hydrogen Sulfide as a Solute…….	66
I.3.3.	Literature Review on Sulfur Dioxide as a Solute……..	71
I.3.4.	Literature Review on Various Gases and Their Mixure as a Solute……	72
I.3.5.	Literature Review on Organic Compounds as a Solute……	77
I.3.6.	Literature Review on Water as a Solute……	78
I.3.7.	Literature Review on Sugar/Sugar Alcohols as a Solute……...	79
I.3.8.	Literature Review on Cellulose as a Solute……...	80
I.3.9.	Literature Review on Other Compounds as a Solutes…….	81
I.3.10.	Critical Review and Establish the Significance of combaing a QSPR model with COSMO-RS Techniques in our Reported Research……	81
I.4.	Conclusions……	84
Chapter II First Contribution: QSPR Modelling of The Solubility of Drug and Drug-Like Compounds in Supercritical Carbon Dioxide		85-110
II.1.	Introduction……	85
II.2.	Procedures……..	85
II.2.1.	Experimental Dataset…….	85
II.2.2.	Descriptor Calculations…….	86
II.2.3.	Selection of Relevant Descriptors…….	87
II.2.4.	Artificial Neural Network……..	93
II.2.5.	ANN-QSPR Modelling……..	93
II.2.6.	Evaluation of the Model's Performance……	95
II.2.7.	Applicability Domain……	98
II.2.8.	Sensitivity Analysis……...	98
II.3.	Results and Discussion…….	99
II.3.1.	Model's Performance…….	99
II.3.2.	William's Plot……	102
II.3.3.	External Test……..	103
II.3.4.	Descriptor's Contribution and Interpretation…….	104
II.3.5.	The use of QSAR for Regulatory Purposes……...	107
II.4.	Conclusions……..	109

Chapter III Second Contribution : A New Hybrid Quantitative Structure Property Relationships-Support Vector Regression (QSPR-SVR) Approach for Predicting the Solubility of Drug compounds in Supercritical Carbon Dioxide 111-147

III.1. Introduction……. 111

III.2. Procedures……... 112

III.2.1. Experimental DataBase…….. 112

III.2.2. Support Vector Regressor……... 112

III.2.3. Meta-heuristic Approaches……..…….. 114

III.2.3.1. Dragonfly Algorithm (DA)…….. 114

III.2.3.2. Ant Lion Optimizer (ALO)……. 116

III.2.3.3. Gray Wolf Optimizer (GWO)…….. 119

III.2.3.4. Artificial Bee Colony (ABC)…….. 121

III.2.3.5. Particle Swarm Optimizer (PSO)…… 123

III.2.3.6. Whale Optimizer Algorithm (WOA)…….. 124

III.2.3.7. Hybrid Particle Swarm Optimization Grey Wolf Optimizer (HPSOGWO).. 126

III.2.4. Hybrid Support Vector Regression Models…… 126

III.2.5. Statistical Criteria for Evaluation of Models Performances……... 129

III.2.6. Applicability Domain……. 130

III.2.6.1. William's plot……. 131

III.2.6.2. Insubria Graph…… 132

III.3. Results and Discussion……... 132

III.3.1. Comparisons of Performances…… 132

III.4. Conclusions……... 147

Chapter IV Third Contribution : A Hybrid Improved Grey Wolf Support Vector Regression Algorithm for Modeling the Solubility of APIs in Pure Ionic Liquids Using σ-profile Descriptors 148-185

IV.1. Introduction……. 148

IV.2. Procedures……... 149

IV.2.1. Overview……. 149

IV.2.2. COSMO-RS and σ-profile…….. 149

IV.2.3. Statistical Data Analysis……. 150

IV.2.4. Dataset Splitting and Output Normalization……... 151

IV.2.5. Model…….. 152

IV.2.6. Model Optimization…… 153

IV.2.7.	Performance Evaluation……..	155
IV.2.8.	Visualization……...	157
IV.2.9.	Sensibility Analyses……	158
IV.2.10.	Model Implementation……	158
IV.3.	Results and Discussion……..	160
IV.3.1.	Performance Comparison……	160
IV.3.2.	External Validation Test for the I-GWO-SVR(r) Model……	178
IV.3.3.	Error Distribution for the I-GWO-SVR(r) Model……	179
IV.3.4.	Outlier Detection for the I-GWO-SVR(r) Model……	181
IV.3.5.	Sensibility Analysis for the I-GWO-SVR(r) Model……	182
IV.3.6.	I-GWO-SVR(r) Model Equation…….....	184
IV.4.	Conclusions…….	184
	General Conclusion	186-188
	Bibliography	189-248
	Annex	249-283

List of Abbreviations

AAD	Average Absolute Deviation
AAED	Average Absolute Relative Deviation
AAPD	Average Absolute Percent Deviation
AAPE	Average Absolute Percentage Error
AAPRE	Average Absolute Percentage Relative Error
AARD	Average Absolute Relative Deviation
AARE	Average Absolute Relative Error
ABC	Artificial Bee Colony
ABC-SVR	Artificial Bee Colony- Support Vector Regression
ACV	Drug Molecule Acyclovir
AD	Applicability Domain
AdaBoost	Adaptive Boosting
ADA-DT	Adaboost Algorithm-Decision Tree
ADA-GRNN	Generative Regression Neural Network
ADA-LR	Adaboost Algorithm-Linear Regression
AIC	Akaike's Information Criterion
AICs	Corrected Akaike's Information Criterion
ALO	Ant Lion Optimizer
ALO-SVR	Ant Lion Optimizer-Support Vector Regression
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
ANN-br-log	Artificial Neural Network(Training Function "Trainbr", Activation Function "Logsig")
ANN-br-tan	Artificial Neural Network(Training Function "Trainbr", Activation Function "Tansig")
ANN-GC	Artificial Neural Network-Group Contribution
ANN-lm-log	Artificial Neural Network(Training Function "Trainlm", Activation Function "Logsig")
ANN-lm-tan	Artificial Neural Network(Training Function "Trainlm", Activation Function "Tansig")

ANN–PSO	Artificial Neural Network-Particle Swarm Optimizer
ANNs	Artificial Neural Networks
ANOVA	Analysis Of Variance
API	Active Pharmaceutical Ingredient
API-IL	Active Pharmaceutical Ingredient-Ionic Liquids System
API-ILs	Active Pharmaceutical Ingredient-Ionic Liquids Systems
APIs	Active Pharmaceutical Ingredients
APSO	Accelerated Particle Swarm Optimization
ARD	Average Relative Deviation
ARM	Association Rule Mining
BA	Bat Optimization Algorithm
BAG	Bagging Regression
BCS	Biopharmaceutics Classification System
BIC	Bayesian Information Criteria
BP	Back Propagation
BP-ANN	Back Propagation-Artificial Neural Network
BP-COSMO-TZVP	Becke-Perdew Functional Conductor-Like Screening Model Triple-Zeta Valence Polarized Basis Set
BR	Bayesian Regularization
BRR	Bayesian Ridge Regression
C_3H_6	Propylene
C_3H_8	Propane
C_4H_{10}	Butane
CDFS	Combined Data Splitting Feature Selection
$CF_3–CFH_2$	Norflurane
CFNN	Cascaded Forward Neural Network
CH_2F_2	Difluoromethane
CH_3F	Methyl-Fluoride
CH_4	Methane
CHF_3	Trifluoromethane
CMIS	Committee Machine Intelligent Systems
CMIS-GP	Committee Machine Intelligent Systems- Genetic Programming
CNN	Convolutional Neural Network
CO	Carbon Monoxide

LIST OF ABBREVIATIONS|

CO_2	Carbon Dioxide
CODESSA	Comprehensive Descriptors For Structural And Statistical Analysis
COSMO	Conductor-Like Screening Model
COSMO-RS	Conductor-Like Screening Model For Real Solvents
COSMO-SAC	Conductor-Like Screening Model For Segment Activity Coefficient
CPA-EOS	Cubic Plus Association-Equation Of State
CP-CNN	Core-Periphery Principle Guided-Convolutional Neural Network
CP-DNN	Core-Periphery Principle Guided-Deep Neural Networks
CP-RNN	Core-Periphery Principle Guided-Recurrent Neural Networks
CSA	Coupled Simulated Annealing
CSA-LSSVM	Coupled Simulated Annealing-Least Square Support Vector Machine
CSP	Cosmo-Sac-Phi
CVD	Co-Volume-Dependent
DA	Dragonfly
DA-SVR	Dragonfly-Support Vector Regression
DBN	Deep Belief Networks
DE	Differential Evolution
DESs	Deep Eutectic Solvents
DJINN	Deep Neural Network Initialization With Decision Trees
DL	Deep Learning
DLH	Dimension Learning-Based Hunting
DMSO	Dimethyl Sulfoxide
DNN	Deep Neural Networks
DT	Decision Tree
EDP	Experimental Data Points
ELM	Extreme Learning Machine
EOS	Equation Of State
EOSs	Equation Of States
ER-EOS	Esmaeilzadeh–Roshanfekr Equation Of State
ERT	Extremely Randomized Tree
ET	Extra Trees

FCNN	Fully Connected Neural Network,
FF-ANN	Feed-Forward Artificial Neural Networks
FIS	Fuzzy Inference System
GA	Genetic Algorithm
GA-LSSVM	Genetic Algorithm- Least Square Support Vector Machine
GA-MLR	Genetic Algorithm-Multi Linear Regression
GA-RBF	Genetic Algorithm-Radial Basis Function
GB	Gradient Boosting
GBTR	Gradient Boosting Trees
GC-EOS	Group Contribution Equation Of State
GC-NLF-EOS	Group Contribution Non-Random Lattice Fluid Equation Of State
GEP	Gene Expression Programming
Gex	Excess Gibbs Energy
GMDH	Group Method Of Data Handling
GP	Genetic Programming
GPR	Gaussian Process Regression
GRAS	Generally Regarded As Safe
GRNN	General Regression Neural Networks
GWO	Grey Wolf Optimizer
GWO-SVM	Grey Wolf Optimizer-Based Support Vector Machine
GWO-SVR	Grey Wolf Optimize-Support Vector Regression
H_2	Hydrogen
H_2S	Hydrogen Sulfide
HFC-RTIL	Hydrofluorocarbons-Room-Temperature Ionic Liquid System
HOMO	Highest Occupied Molecular Orbital
HPSOGWO	Hybrid Particle Swarm With Grey Wolf Optimizer
HPSOGWO-SVR	Hybrid Particle Swarm With Grey Wolf Optimizer-Support Vector Regression
HR	Huber Regression
HSE	Hard-Sphere Expansion
Hybrid-ANFIS	Hybrid-Adaptive Neuro-Fuzzy Inference System
I-GWO	Improved Grey Wolf Optimizer
I-GWO-SVR(g)	Improved Grey Wolf Optimizer- Support Vector Regression(Gaussian Kernal Function)

I-GWO-SVR(r)	Improved Grey Wolf Optimizer- Support Vector Regression(Rbf Kernal Function)
IL	Ionic Liquid
ILs	Ionic Liquids
Iter	Iteration
IUPAC	International Union Of Pure And Applied Chemistry
K–J	Kumar And Johnstone
KK	Krichevsky-Kasarnovsky
KM	Kwak-Mansoori Mixing Rules
KNN	*K*-Nearest Neighbor
KRR	Kernel Ridge Regression
LAR	Least Angle Regression
LCVM	Linear Combination Of Vidal And Michelsen
LK	Lee-Kesler
LM	Levenberg–Marquardt
Logsig	Log-Sigmoid
LR	Linear Regression
LSBoost	Least-Squares Boosting
LSSVM	Least Square Support Vector Machine
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MARS	Multivariate Adaptive Regression Spline
Maxiter	Max Iterations
MCIs	Molecular Connectivity Indices
MD	Molecular Dynamics
m-ER-EOS	Modified Esmaeilzadeh-Roshanfekr Equation Of State
MFO	Moth-Flame Optimization
MFO-SVR(g)	Moth-Flame Optimization-Support Vector Regression (Gaussian Kernal Function)
MFO-SVR(r)	Moth-Flame Optimization-Support Vector Regression (Rbf Kernal Function)
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLP-ANN	Multi-Layer Perceptron-Artificial Neural Network

MLR	Multiple Linear Regression
MMM-EOS	Mansoori-Mohsen Nia-Modarress Equation Of State
MMM-EOS	Mohsennia-Modarress-Mansoori Equation Of State
MQR	Multiple Quadratic Regression
MR	Mukhopadhyay–Rao
mrPR	Panagiotopoulos–Reid
MSE	Mean Squared Error
MST or M–T	Mendez-Santiago And Teja
MVO	Multi-Verse Optimizer
MVO-SVR(g)	Multi-Verse Optimizer-Support Vector Regression(Gaussian Kernal Function)
MVO-SVR(r)	Multi-Verse Optimizer-Support Vector Regression(Rbf Kernal Function)
N_2	Nitrogen
NES	Neuro-Equation-Of-State
NH_3	Azane
NPT-GEMC	Isothermal Isobaric Gibbs Ensemble Monte Carlo
NPT-MC	Isothermal-Isobaric Monte Carlo
NRTL	Nonrandom Two-Liquid
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs
O_2	Oxygen
OECD	Organisation For Economic Cooperation And Development
PC-SAFT	Perturbed-Chain Statistical Associating Fluid Theory
PC-SAFT-EOS	Perturbed-Chain Statistical Associating Fluid Theory Equation Of State
PHSC	Perturbed Hard-Sphere Chain
PHSC-EOS	Perturbed Hard-Sphere Chain- Equation Of State
PLS	Partial Least Square
Pop	Population
PR+MHV1+COSMOSAC	Peng- Robinson+ Modified Huron-Vidal First Order Mixing Rule+ Conductor-Like Screening Model For Segment Activity Coefficient
PR+VDW	Peng- Robinson +Van Der Waals Mixing Rule
PR+WS+Wilson	Peng- Robinson+ Wong-Sandler Mixing Rule+ Wilson

PR-COSMOSAC-EOS	Peng- Robinson Conductor-Like Screening Model For Segment Activity Coefficient Equation Of State
PRCS	Peng- Robinson Conductor-Like Screening Model For Segment Activity Coefficient
PRCS-EOS	Peng- Robinson Conductor-Like Screening Model For Segment Activity Coefficient Equation Of State
PR-cubic-EOS	Peng- Robinson Cubic Equation Of State
PR-EOS	Peng- Robinson-Equation Of State
PRESS	Predictive Residual Sum Of The Squares
PRSV	Peng Robinson Stryjek Vera
PRSV-EOS	Peng Robinson Stryjek Vera- Equation Of State
PR-TS-EOS	Peng Robinson-Two State Equation Of State
PSO	Particle Swarm Optimizer
PSO-ANFIS	Particle Swarm Optimizer-Adaptive Neuro-Fuzzy Inference System
PSO-ANN	Particle Swarm Optimizer-Artificial Neural Network
PSO-SVR	Particle Swarm Optimizer-Support Vector Regression
PSRK	Predictive Soave-Redlich-Kwong
PSRK-EOS	Predictive Soave-Redlich-Kwong-Equation Of State
PTV-EOS	Patel–Teja–Valderrama Equation Of State
PVT	Patel–Teja–Valderrama
qCPA-EOS	Quadrupolar Cubic Plus Association Equation Of State
QSAR	Quantitative Structure–Activity Relationship
QSPR	Quantitative Structure–Property Relationship
QSPR-SVR	Quantitative Structure–Property Relationship-Support Vector Regression
QSTR	Quantitative Structure-Toxicity Relationship
R&D	Results & Discussion
R-1243zf	Trifluoropropene
RAE	Relative Absolute Error
RB	Radial Basis Network
RBF	Gaussian Radial Basis Function
RBF-ANN	Radial Basis Function- Artificial Neural Network
RBFNN	Radial Basis Function Neural Network

RE	Relative Error
RevIL	Reversible Ionic-Liquid
RF	Random Forest
RK-EOS	Redlich-Wong Equation Of State
RM-EOS	Riazi-Mansoori Equation Of State
RMSD	Root Mean Square Deviation
RMSE	Root-Mean Squared Error
RNNs	Recurrent Neural Networks
RR	Ridge Regression
RRSE	Root Relative Squared Error
RST	Regular Solution Theory
RTILs	Room-Temperature Ionic Liquids
SAFT	Statistical Associating Fluid Theory
SAFT-EOS	Statistical Associating Fluid Theory- Equation Of State
SAFT-HR-EOS	Statistical Associating Fluid Theory-Huang-Radosz- Equation Of State
SAFT-VR-EOS	Statistical Associating Fluid Theory-Variable Range- Equation Of State
SCA	Sine Cosine Algorithm
SCA-SVR(g)	Sine Cosine Algorithm-Support Vector Regression(Gaussian Kernal Function)
SCA-SVR(r)	Sine Cosine Algorithm-Support Vector Regression(Rbf Kernal Function)
SC-CO_2	Supercritical Carbon Dioxide
SCF	Supercritical Fluid
SCFs	Supercritical Fluids
SEP	Electrostatic Potential Surface
SI	Swarm Intelligence
SMILES	Simplified Molecular Input Line Entry Specification
SO_2	Sulfur Dioxide
SRK-cubic-EOS	Soave–Redlich–Kwong Cubic Equation Of State
SRK-EOS	Soave–Redlich–Kwong Equation Of State
SRK-MC+vdW	Soave–Redlich–Kwong- Mathias–Copeman +Van Der Waals
SS	Sung And Shim

SSE	Sum Of Square Error
SSY	Sum Of The Squares Of The Response Values Of The Total Observations
SVC	Support Vector Classification
SVM-GC	Support Vector Machine-Group Contribution
SVR	Support Vector Regression
tPC-PSAFT	Truncated Perturbed Chain Polar Statistical Associating Fluid Theory
Trainbr	Bayesian Regularisation Backpropagation
UNIFAC	Uniquac Functional-Group Activity Coefficients
UNIQUAC	Universal Quasi-Chemical
VDW	Van Der Waals
VDW1	Van Der Waals One Mixing Rule
VDW2	Van Der Waals Two Mixing Rule
VDW-EOS	Van Der Waals- Equation Of State
VDWO	Van Der Waals Zero Mixing Rule
VIF	Variation Inflation Factors
VLE	Vapor-Liquid Equilibrium
VPT	Valderrama-Patel-Teja
WAO	Whale Algorithm Optimizer
WAO-SVR	Whale Algorithm Optimizer-Support Vector Regression
WNN	Wide Neural Network
WS	Wong-Sandler Mixing Rules
XGBoost	Extreme Gradient Boosting
ZJ	Zudkevitch-Joffe
βPR-EOS	Modified Peng-Robinson Equation Of State
2DMA2M1P	2-(dimethylamino)-2-methyl-1-propanol

List of Nomenclatures & symbols

Symbol	Name
$\bar{x}^{obs}$	Average Of The Observed Values Liquid Phase
$\bar{x}^{pred}$	Average Of The Predicted Value Liquid Phase
$\bar{Y}_{pred}$	Average Of $\bar{y}^{pred}$ Gas Phase
$\bar{r}_m^2$	Average Of r_m^2 Gas Phase
$\bar{y}^{obs}$	Average Of Y_{obs} Gas Phase
h^*	Threshold Leverage Value
MS_E	Mean Square Residual
MS_R	Mean Square Regression
x_i^{obs}	Observed Value Liquid Phase
x_i^{pred}	Predicted Value Liquid Phase
y_i^{obs}	Observed (Experimental) Value Of Y Gas Phase
y_i^{pred}	Predicted (Calculated) Y-Value Of Training Set, Test Set Or Validation Set Gas Phase
C°	Degree Circus
N_P	Number Of Independent Variables In Each Model,
$P^x(\sigma)$	Peak
Q_{F1}^2, Q_{F2}^2, Q_{F3}^2, & Q_{pred}^2	Different Types Of Cross-Validated Correlation Coefficient
R_m^2 & R'^2_m	Marginal Coefficient Of Determination
R^2	Coefficient Of Determination
R_{adj}	Adjusted Correlation Coefficient
R_{adj}^2	Adjusted Coefficient Of Determination
$S_{\sigma-profile}$	Charge Distribution Area
S_{1-12}	σ-Moment (1-4 Of The Cations, 5-8 Of The Anions, And 9-12 Of The Apis)
Y_{space}	Cutoff Value
a_m	Mixture Energy Parameter Of EOS
b_m	Excluded Volme Paramet Of EOS
k_{12} , k_{ij} and l_{ij}	Binary Interaction Parameter

r_0^2 & r'^2_0	Squared Correlation Coefficient
x^2	Chi-Square Value
α_{12}	Dimensionless Parameter
σ_{11}, σ_{12}	Molecular Size Parameters
h	Leverage Value
ω	Acentric Factor
$ALD - x$	Average Logarithmic Deviation
$ALD - y$	Average Logarithmic Deviation In Mole Fraction
Af	Precision Factor
Bf	Bias Factor
C	Box-Constraint
HE	Excess Enthalpy
Hi	Hat Value
K	Kelvin Degree
KPa	Kilopascal
$-Ln(y2)$	Log-Transformed Property
MPa	Megapascal
Mw	Molecular Weight
N	Number Of Data Points In The Data Set (Training, Test...)
P	Pressure
$P(MPa)$	Pressure In Megapascal
Pc	Critical Pressure
Q^2	Cross-Validated Correlation Coefficient
Q^2ccc	Concordance Correlation Coefficient
R	Correlation Coefficient
T	Temperature
$T\ (K)$	Temperature In Kelvin
Tc	Critical Temperature
b	Bias Term
k & k'	Slopes Of The Corresponding Regression Lines
$k(x_i, x_j)$	Kernel Function
$ln\,\gamma$	Logarithmic Activity Coefficient
$r\ or\ r_k$	Relevance Factor

s	Pooled Standard Deviation
w	Weight Vector
Δr_m^2	Delta Of r_m^2
α	Alpha
β	Beta
δ	Delta
ε	Epsilon
σ	Kernel-Scale
$\varphi\ (x)$	Kernel Function
ω	Omega

Ion designation	Cations
[(CH2)4SO3HMIm]	1-(4-Sulfobutyl)-3-Methylimidazolium
[C2mim] or [EMIM]	1-Ethyl-3-Methylimidazolium
[C4mim] or [BMIM]	1-Butyl-3-Methylimidazolium
[C6mim] or [HMIM]	1-Hexyl-3-Methylimidazolium
[C8mim] or[OMIM]	1-Octyl-3-Methylimidazolium
[Cnmim+]	1-Alkyl-3-Methylimidazolium
[HOemim]	1-(2-Hydroxyethyl)-3- Methylimidazolium
[m-2-HEA]	N-Methyl-2-Hydroxyethylammonium
[P6,6,6,14]	Trihexyl)Tetradecyl-Phosphonium
[PMPY]	3-Methyl-1-Propylpyridinium
[Tf2N]	Bis(Trifluoromethylsulfonyl)Imide
[TFO]	Trifluoromethanesulfonate
[TMA]	Tris(2- Hydroxyethyl)Methyl

	Anions
[Ac]	Acetate
[B(CN)4]	Tetracyanoborate
[BF4]	Tetrafluoroborate
[CF3SO3]	Trifluoromethane-Sulfonate
[eFAP]	Tris(Pentafluoroethyl)Trifluorophospha
[EtSO4]	Ethyl Sulfate
[FAP]	Tris(Perfluoroalkyl)Trifluorophospha

[HSO4]	Hydrogen Sulfate
[LACTATE]	L-(+)-Lactate
[MeSO4]	Methylsulfate
[NTf2] or [Tf2N]	Bis(Trifluoromethylsulfonyl) Imide
[P]	N-Methyl-2-Hydroxyethylammonium Pentanoate
[PF6]	Hexafluorophosphate
[Pr]	N-Methyl-2-Hydroxyethylammoniumpropionate
[SCN]	Thiocyanate
[TF2N]	Bis(Trifluoro-Methanesulfonyl)Imide
[TFO] or [OTf]	Trifluoromethylsulfonate

List of Figures

Figure No.	Caption	Page No.
Figure II.1:	Normality distribution plot of mole fraction of solubility data.	86
Figure II.2:	MLP-ANN structure.	94
Figure II.3:	Scatter plots Predicted vs. Observed values of the solubility for the MLP-ANN model: (a) Global Set, (b) Training Set, (c) Test Set, respectively.	100
Figure II.4:	William's plot for MLP-ANN model.	103
Figure II.5:	Scatter plots Predicted vs. Observed values of the solubility for the MLP-ANN model, external test.	104
Figure II.6:	Results of the relevance factor performed on the QSPR-ANN model.	105
Figure II.7:	Relative contribution of each input variable using the Garson method.	106
Figure III.1:	Flowchart of the hybrid metaheuristic-QSPR-SVR model.	128
Figure III.2:	Convergence curve of the seven algorithms.	133
Figure III.3:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for DA-SVR model.	134
Figure III.4:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for ALO-SVR model.	135
Figure III.5:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for GWO-SVR model.	136
Figure III.6:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for ABC-SVR model.	137
Figure III.7:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for PSO-SVR model.	138
Figure III.8:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for WOA-SVR.	139
Figure III.9:	Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for HPSOGWO-SVR model.	140

Figure III.10:	Bar distribution diagrams of the absolute relative error [%] vs. number of data points in percentage of the training set.	142
Figure III.11:	Bar distribution diagrams of the absolute relative error [%] vs. number of data points in percentage of the test set.	142
Figure III.12:	William's plot for HPSOGWO-SVR model.	143
Figure III.13:	Insubria graph for the HPSOGWO-SVR model.	144
Figure III.14:	Scatter plot Predicted vs. Observed for HPSOGWO-SVR model, validation set.	145
Figure III.15:	Comparison diagram of observed solubility isotherms and those calculated by the HPSOGWO-SVR model, for validation set.	146
Figure. IV.1:	σ-profile distribution of the 14 APIs and their margins limiting each section.	150
Figure. IV.2:	Support vector regression structure.	158
Figure. IV.3:	The flowchart of the proposed I-GWO-SVR(r) model.	159
Figure. IV.4:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-lm-tan : (a) Train set, (b) Test set, (c) Cross-validation, (d) Global.	162
Figure. IV.5:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-lm-log : (a) Train set, (b) Test set, (c) Cross-validation, (d) Global.	163
Figure. IV.6:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-br-tan : (a) Train set, (b) Test set, (c) Global.	164
Figure. IV.7:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-br-log : (a) Train set, (b) Test set, (c) Global.	165
Figure. IV.8:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model I-GWO-SVR(g) : (a) Train set, (b) Test set, (c) Global.	166
Figure. IV.9:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model I-GWO-SVR(r) : (a) Train set, (b) Test set, (c) Global.	167

Figure. IV.10:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MVO-SVR(g) : (a) Train set, (b) Test set, (c) Global.	168
Figure. IV.11:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MVO-SVR(r) : (a) Train set, (b) Test set, (c) Global.	169
Figure. IV.12:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MFO-SVR(g) : (a) Train set, (b) Test set, (c) Global.	170
Figure. IV.13:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MFO-SVR(r) : (a) Train set, (b) Test set, (c) Global.	171
Figure. IV.14:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model SCA-SVR(g) : (a) Train set, (b) Test set, (c) Global.	172
Figure. IV.15:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model SCA-SVR(r) : (a) Train set, (b) Test set, (c) Global.	173
Figure. IV.16:	Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in Ils of the external validation set.	179
Figure. IV.17:	Error of the I-GWO-SVR(r) model in predicting the AIPs solubility in ILs.	180
Figure. IV.18:	Error distribution of the I-GWO-SVR(r) model in predicting the AIPs solubility in ILs.	180
Figure. IV.19:	William's plot of the I-GWO-SVR(r) model.	182
Figure.IV.20:	Insurbia plot of the I-GWO-SVR(r) model.	182
Figure.IV.21:	Results of the relevance factor of the input parameters of of the I-GWO-SVR(r) model.	183

List of Tables

Table No.	Caption	Page No.
Table .II.1:	Statistical parameters for internal validation.	88
Table .II.2:	Statistical parameters for external validation.	89
Table .II.3:	Statistical of multi-collinearity of 13 descriptors.	89
Table .II.4:	Statistical of multi-collinearity of 14 descriptors.	90
Table .II.5:	Statistical of multi-collinearity of 15 descriptors.	90
Table .II.6:	Statistical of multi-collinearity of 17 descriptors.	90
Table .II.7:	Characteristics of the descriptors selected in the optimal model and their VIF.	91
Table .II.8:	Matrices of correlation of the selected features.	92
Table .II.9:	Acceptability criteria (A.C) of a model.	97
Table .II.10:	Validation parameters for the train, test, and global sets for MLP-ANN model.	101
Table II.11:	Different errors for the train, test, and global sets for MLP-ANN model.	102
Table II.12:	AIC values of the four ANN models.	102
Table II.13:	Different errors for the validation set for both cases.	104
Table III.1:	Hyperparameter of the SVR models.	133
Table III.2:	Statistics Parameters of The Seven SVR Models.	141
Table IV.1:	σ-profile descriptors typical cations and anions of the ILs, the APIs and their margins limiting each section.	151
Table IV.2:	Summary of statistics analysis of the dataset.	151
Table IV.3:	The artificial neural network models and their selected abbreviation, transfer functions, and obtained architecture.	160
Table IV.4:	The hybrid support vector regression models and their selected abbreviation, kernel function, and obtained hyperparameters.	161
Table IV.5:	Statistical metrics for the global train, test, and inter-validation set for the 4 ANN models.	174

Table IV.6:	Statistical metrics for the global train, and test set for the 8 hybrid SVR models.	175
Table IV.6:	Statistical metrics for the global train, and test set for the 8 hybrid SVR models (continued).	176
Table IV.7:	AIC & BIC values for the global set of the 12 applied models.	178
Table IV.8:	Statistical metrics for the external validation set for the I-GWO-SVR(r) model.	179

List of Publications

This Thesis Is Based On The Following Published Papers:

- I. Euldji, C. Si-Moussa, M. Hamadache, O. Benkortbi, QSPR Modelling of the Solubility of Drug and Drug-like Compounds in Supercritical Carbon Dioxide, Mol. Inform. 2200026 (2022) 1–17. https://doi.org/10.1002/minf.202200026.

- I. Euldji, A. Belghait, A. Amrane, A new hybrid quantitative structure property relationships-support vector regression (QSPR-SVR) approach for predicting the solubility of drug compounds in supercritical carbon dioxide, (2023) 1–17. https://doi.org/10.1002/aic.18115.

- I. Euldji, W. Benmouloud, K. Paduszynski, C. Si-Moussa, O. Benkortbi, Hybrid Improved Grey Wolf Support Vector Regression Algorithm for Modeling Solubilities of APIs in Pure Ionic Liquids : σ - Profile Descriptors, (2024) 1-16. https://doi.org/10.1021/acs.jcim.3c01876.

General Introduction

General Introduction

1. Background and Motivations

Aqueous solubility is a key parameter in various scientific applications. It is one of basic steps in the successful development cycle of new chemical drug entities [1]. Usually, the low-water-soluble drugs are administered orally either weakly acidic or basic, having low aqueous solubility [2]. Low water solubility, along with permeability, has a direct effect limiting the dissolution rate of an orally administered drug. Thus, it influences the bioavailability and therapeutic action of the drug to reach the target site [3,4].

If a drug does not achieve the desired concentration in the systemic circulation to provide a pharmacological response, a dose increase is often needed. Otherwise, poor water solubility will lead to slow drug absorption, which leads to insufficient gastrointestinal mucosal toxicity and variable bioavailability [2,5]. In fact, increasing doses may lead to increasing cost and development time, and the burden will be shifted to the patient (frequent high-dose administration) [6,7]. As well, increases the side effects of medications.

Recently, more than 40-75% of new chemical entities, categorized as Biopharmaceutics Classification System (BCS) class II or IV drugs, are practically insoluble in water; have low dissolution and bioavailability [8–10]. Additionally, they correspond to 90% of new chemical entities, and 75% of compounds under development. Therefore, it has become paramount a need to introduce advanced enhancement techniques to reduce the risk, time, and capital invested in drug development and overcome this drawback [9].

In the pharmaceutical industry, finding promising ways to improve the bioavailability and solubility of these drugs poses a significant challenge for the orally administered drug delivery system [11]. So far, several approaches are available in the literature to improve the solubility of poorly water-soluble drugs. The techniques are classified into physical and chemical modifications of the drug substance, and other techniques including particle size reduction like micronization and nanosuspension, modification of the crystal habit like polymorphs, amorphous form, and cocrystallization, drug dispersion in carriers like eutectic mixtures, solid dispersions, solid solutions and cryogenic techniques as physical modifications, change of ph, use of a buffer, derivatization, complexation, and salt formation as chemical modification, and supercritical fluid process, use of adjuvant as surfactant, solubilizers,

cosolvency, hydrotrophy, and new excipients as miscellaneous methods [7]. The choice of the enhancement method depends on on the properties of the drug, site of absorption, and the required characteristics of the dosage form, the properties of a drug under consideration, the nature of the excipients to be selected, and nature of the intended dosage form [3,7].

Nevertheless, most of the strategies mentioned above use significant amounts of organic solvents in the formulations, principally to induce the crystallization of a given polymorphic form and particle size, where health and environmental concerns are raised [12,13], since there is a large possibility that molecules of the organic solvents are incorporated into the crystal structure of the low-soluble drugs during the crystallization process[13,14]. In fact, the use of such organic solvents have various other disadvantages, such as being volatile, flammable, and toxic. Therefore, the used solvents must be removed from the pharmaceutical compounds or their levels must be controlled for ensuring human consumption safety [12,13]. Hence, pharmaceutical ingredients must be free from any such traces.

Accordingly, a growing area of research in the development of green technologies is devoted to designing new, environmentally friendly, and tunable solvents, the use of which would meet both technological and economic demands [15]. Indeed, those Green solvents become promising alternatives to conventional solvents in pharmaceutical industries.

2. Green Solvents as a Solubility Improvement Techniques

During the past decades, the pharmaceutical industry has been reported to be one of the largest users of organic solvents [16] such as ethanol, methanol, or dimethyl sulfoxide (DMSO). These solvents are generally applied as solvents or co-solvents in pharmaceutical formulations to improve solubility [13,17]. Therefore, solvents represent an important part of the environmental performance of a process, as well as a major impact on cost, safety and health issues. Consequently, introducing the concept of ‘‘green’’ solvents expresses the goal of minimizing the environmental impact resulting from the use of solvents in chemical production [18].

The physical and chemical barriers of these solvents make them difficult to completely remove from the final manufactured product [19], leading to environmental and safety problems [20]. Therefore, some small amounts of solvents remain in the active pharmaceutical ingredient (API) products. Consequently, the API becomes contaminated during the process of packaging, warehouse storage, or shipping and transportation [16,19] and, of course contains residual

impurities. In fact, numerous organic chemist's solvents were blacklisted, solvent use requires re-thinking and has become a primary focus, mainly in the manufacture of pharmaceuticals [21–23].

Various effort has been made by the pharmaceutical industry to replace toxic solvents with more environmentally friendly alternatives with similar properties, which require more or has looked for innovations. Those solvents are commonly known as green solvents. Grodowska and Parczewski explained [19] study explained the main reasons for replacing organic solvents with environmentally friendly alternatives (e.g., supercritical fluids and ionic liquids) with similar properties in pharmaceutical production processes.

Discovery of new green solvents and improvement of drug solubility have become one of the most important topics for advanced manufacturing as well as process intensification. These features play a vital role and have enormous impacts on the drug effectiveness and the environment [24]. A substance can be recognised as green if it has some basic properties, such as low toxicity, chemically stable, readily available, and easily recyclable...[20].

Water, supercritical and subcritical fluids, room-temperature ionic liquids, and solvents from natural and renewable sources are among the proposed solvents. A full discussion regarding these new solvents is well beyond the scope of this thesis. However, a brief overview of current knowledge regarding the supercritical carbon dioxide and ionic liquids as environmentally harmless solvents and as solubility enhancement methods are presented herein.

2.1.Supercritical Carbon Dioxide

Supercritical fluids (SCFs) are substances whose temperature and pressure are above above the critical condition. A range of intermediate-density states can be obtained by tuning the pressure or temperature [25]. This leads to special properties of the fluid, including high diffusivity, low viscosity, and low surface tension, which allows the fluid to better diffuse through a solid matrix and thus extract natural compounds compared to conventional liquid solvents [15]. SCF have properties intermediate between those of gases and liquids [15,26].

Over the past decades, those fluids have received great attention in various fields, including chemical, biochemical, food, and pharmaceutical, as a clean and safe ‘green solvent’ technology and the most promising technique in pharmaceutical industries due to their suitable density, solvent, transport properties [27,28].

SCF techniques can be used in various applicable methods such as in extraction, separation, and crystallization procedures which are the main steps in API synthesis and particle engineering [29]. Particle design of API is very critical in the manufacturing of solid and liquid dosage forms. They are used to modify and control numerous particles properties, such as various sizes, shapes, crystallinity structures, polymorphism, processability, and solubilities[30–34].

Among the advantages of the SCFs used are their potential regulating properties with temperature and pressure. Also, SCFs are known to be non-toxic, non-flammable, and can achieve high mass and heat transfer rates [30–35]. The current literature on applications of SCF technology applications relates particle engineering in order to improve the solubility of API, bioavailability, and processing capacity. These include, among others, crystal and polymorph modification, complex formation, solid dispersion [29].

Any compound can be examined as a supercritical solvent, the best-suited solvent for a specific application can be determined based on its critical properties, toxicity, safety, flammability, cost and solvating power, phase behaviour, and solubility under the operating conditions [15,36]. Among all available supercritical fluids, supercritical carbon dioxide (SC-CO_2) remains the most widely SCF used; it has been estimated that almost 98% of the reported applications of SCF in the pharmaceutical industry belong to SC-CO_2 [37].

SC-CO_2 provides numerous advantages, including low critical temperature (Tc = 31.10C), critical pressure (Pc = 73.8 bar), non-toxic, non-flammable, non-explosive, readily available, easily removable from the product, and possesses convenient critical properties [27,38], non-corrosive, easily recyclable [28], inert, leaves no traces after the process, and has GRAS (generally regarded as safe) status, high permeability (due to CO_2' small linear structure) [11. These properties made this solvent gain wide acceptance as an effective alternative to toxic solvents [39]. In addition, SC-CO_2 solvent is very sensitive to the variation of temperature and pressure especially near the critical point; this will vary significantly once the temperature or pressure changes slightly. Hence, its ideal properties can be easily obtained by appropriately adjusting the operating temperature and pressure in the supercritical region [28].

SC-CO_2 has been used in various processes, including S-CO_2 power cycle [40], extraction and separation process [41–44] and chemical reaction [45]. In the pharmaceutics and medical chemistry indursties, the solvent is applied for preparation of solid dispersion to enhance the water solubility of drugs. This would be attained by polymorphic control and complex formation behavior of APIs and excipients by CO_2 processing in SCF technique [29].

2.2. Ionic Liquids

Ionic liquids (ILs) are liquid salts comprised entirely of cations and anions ions with melting points lower than 100°C [15]. Those salts have gained interesting attention in both the scientific and industrial community [46,47] as a promising green solvent be sued as an alternative to undesirable conventional solvents (e.g., organic solvents) [48–50]. Various feature of those ion-based materials makes them very attractive, including stability (thermal, chemical, and electrochemical), low melting point (below 100C°) [15,51–53], density, solubility, and miscibility compared with molecular solvents [49], non-flammability and non-volatility of the majority of ILs [54–56]. In addition, ILs possess a low vapor pressure or non-volatility under ambient conditions, tuneable solubility, acidity, basicity, long-range thermal stability, and very low corrosivity relative to mineral acids and bases [46], and wide liquid range, which have tended to several potential applications in a variety of areas [16,57,58].

ILs are made up of ionic species (usually the bulky organic cation and organic or inorganic anion). Therefore, they can be varied and improved, as much as possible to design a solvent with a specific use in mind [59,60]. Due to their adapted characteristics, ILs offer a far better opportunity to establish enhanced size/morphology control compared to organic solvents [61–64].

Usually, ILs are liquids that are only in ionic form. However, there are other classical forms that should be distinguished such as ILs at room temperature or below ≤100C° ILs, fused salt, molten salt, liquids organic salt, and many others [65]. In addition, ILs known as exclusive ions, described as deep eutectic solvents [15], and a few protic ILs including Olah's reagent [66].

Moreover, ILs have been found to be potential candidates for solving some of the major issues society is currently faced with by emerging as a clean, efficient, and eco-friendly alternative resource of volatile organic solvents in various academia and industrial areas, in general [50,66], and in the pharmaceutical industry in particular [13,16,47,53,66,67]. Furthermore, IL properties can be modified based on their application by changing the combination of cations and anions used [66].

Over the past decade, the pharmaceutical fields have witnessed an increasing interest in adding ILs to various applications, to develop environmentally friendly technologies and processes, such as:

(1) the formulation of active pharmaceutical ingredient-ionic liquids systems, abbreviated API-ILs [68],

(2) the use of ILs/IL mixtures as solvents or co-solvents for the synthesis and purification of pharmaceutical compounds [69], and for poorly soluble drug molecules [70,71],

(3) in drug delivery systems[72,73]. This interest has mainly increased due to the faster and easier to preform reactions of ILs, requiring no special apparatus or methodologies, as well as having better reaction conditions, accelerating some more problematic reactions, facilitating the purification and isolation of the desired product compared to volatile organic solvents [67].

The good solvation ability of these solvents has increased the aqueous solubility of APIs enabling them to overcome the disadvantages of the solubility, bioavailability, and polymorphism [13,53] while ILs' brilliant capability has facilitated their use as solvents/co-solvents to improve the solubility/bioavailability of these insoluble APIs [7,13,74]. In addition, the ability to recycle IL has opened a great economic perspective [67].

In comparison, the IL cations have a minor effect on the solvation capability, while anion plays an important role in the solubility of drug molecules. The interactions between the API and IL strongly depend on the combination of cation and anion of the IL [75]. Also, the dissolution of APIs in ILs involves the formation of hydrogen bonds between the IL anions and the drug molecules [75,76]. Furthermore, the possibility of finely adjust the physical chemical properties of IL by simply varying their anion/cation combination makes generating non- toxic and/or biocompatible ILs easy [67]. In fact, ILs are considered promising candidates in the pharmaceutical industries [73]. Theoretically, it is believed that a vast number of ILs could be synthetized with tuned and enhanced properties for the application at hand [77,78].

3. Experimental Measurement Limitations

The experimental determination of drug solubility, in a green solvent or not, is not an isolated event. It must be performed repeatedly under different operation conditions (pressures and temperatures for SC-CO_2 as a solvent, fixed pressures, and varied temperatures for ILs as a solvent) throughout the drug discovery and development process, the assays and their orientation varying depending on the phase [79]. Recently, the knowledge of solute solubility has become a fundamental necessity and must be known before any functional assessment. However, it is not always easy to achieve. Due to several concerns, it is not always possible to calculate this property. The main backdrops are defined below:

- The solubility of a substance fundamentally depends on the solvent used as well as the temperature and pressure. Determining the solubility of all compounds over wide ranges of temperature and pressure is an expensive, challenging, and time-consuming task. Therefore, this is not feasible, mainly for low solubility values.

- The solute solubility in SCF (or SC-CO_2) is a strong function of the temperature and the pressure, which leads to a significant data demand. Calculating the solubility of pharmaceutical compounds in SC-CO_2 at various temperatures and pressures became an essential requirement to assess whether the drug is suitable for treatment by this technology or not. A reasonably good solubility measurement for a single solute in a single SCF requires at least 15 and possibly up to 40 experimental conditions (exp. 5–10 pressures at 3 or 4 different temperatures). These measurements are not a simple task, especially when the solubility is very low [80].

- The selection of cation-anion combinations in ILs has a significant impact on the solubilization mechanism of poorly water-soluble APIs and solvation ability [81,82]. Furthermore, if the calculations rely on the ions' geometries [52,83,84]. Currently, 360,000 possible binary cation-anion combinations are available; for each API, there are 360,000 possible API-IL systems. The number of possible combinations is therefore enormous. There is no way to calculate solubility for all these possibilities. In some cases, laboratory methods are not applicable [85].

This background explains why data on drug solubility in both green solvents (SC-CO_2, and ILs) is still not available for most pharmaceutical compounds. However, the supercritical fluid community has been actively measuring and publishing solubility data since this discovery [80]. In contrast, only a few articles available in the literature have reported the value of drug solubility in ILs (mainly because SC-CO_2 has been used as a solvent for solubility enhancement much earlier than ILs). It is therefore essential to use computational models capable of predicting as efficiently as possible the solubility of drugs in those two solvents.

The development of solubility models has continued in parallel with solubility measurements. Effective models provide the opportunity to reduce measurement demands, the ability to interpolate between measured points and the opportunity to support process design and analysis [80].

Additionally, in many experimental sources for solubility determination, it noticed that significant errors have been repeatedly found in the reporting of solubility values, experimental conditions, and purity information…which limits the use of these data in terms of a database for modeling and model validation. Datasets relying on experimental results often contain high levels of random error or noise from the different employed techniques. Therefore, it is necessary to represent molecules in a way that can be used as inputs. The idea will be adopted in this thesis.

4. Aims and Objectives of this Thesis

The main objective of this thesis is to develop new powerful predictive models, that do not rely on experimental measurement, for the prediction of the solubility of solid drug compounds in two different green solvents, supercritical carbon dioxide, and ionic liquids as a function of the operation conditions (temperature and pressure for SC-CO_2, only the temperature for ILs), and descriptors based on molecular information. To achieve the desired goal, the solubility of pharmaceutical compounds in SC-CO_2, and in ILs under various operations conditions, is collected to build the databases. Powerful algorithms will be investigated, developed, and then validated. Based on the aim targeted, the following objectives were identified for the present research work:

✓ Minimize the use of experiments, to offer the possibility of interpolating between the points measured experimentally, and to provide a tool to support process design and analysis. The ideal target for modeling is purely theoretical predictive models.

✓ Development and validation of suitable models for predicting drugs solubilities in SC-CO_2, and in ILs.

✓ Benefiting from Quantitative Structure-Property Relationship (QSPR) based techniques to obtain powerful inputs that do not rely on experimental measurement, and powerful predictive models for predicting the solubility of drugs in SC-CO_2.

✓ Benefiting from conductor-like screening model for Real Solvents (COSMO-RS) based techniques to propose new molecular descriptors to be combined with QSPR approach to obtain powerful inputs that do not rely too on experimental measurement and a powerful predictive model for estimating pharmaceutical compounds solubilities in pure ILs.

5. Thesis Organisation

The thesis will follow the following plan:

Several authors have studied the prediction of the solubility of pharmaceutical compounds in SC-CO_2 and ILs and therefore many of contribution have been proposed in the literature, which will be briefly explained in **chapter I** so that the proposed models of this thesis will be clarified.

Chapter II addresses the development of a combined ANN-QSPR model for the prediction and correlation of drug and drug-like compounds' solubility in supercritical carbon dioxide. Molecular descriptors that are most correlated to the target, as well as the best-suitable ANN-QSPR architecture, will be determined. The best model will be statistically validated. Sensitivity analysis of the inputs and outliers detection will be also performed.

Chapter III treated a comparison study between seven (7) selected meta-heuristic algorithms: Dragonfly (DA), Ant Lion (ALO), Grey Wolf (GWO), Artificial Bee Colony (ABC), Particle Swarm (PSO), Whale (WAO), and a hybrid Particle Swarm with Grey Wolf (HPSOGWO) optimizers in fine-tuning support vector regression hyper-parameters. The results will be validated through various statistical and graphical appraisals, including an external test.

An advanced artificial intelligence algorithm based on support vector regression (SVR) coupled with an improved version of a well-known nature-inspired method namely grey wolf optimizer (I-GWO) will be implemented in **Chapter IV** for the prediction of solubility of Active Pharmaceutical Ingredient (APIs) in pure Ionic liquids (ILs) at various temperature. In order to validate the model, various statistical and graphical factors will be implied, including calculus of various recommended statistical indexes, analyzes of variance (ANOVA), and residual and outlier detection. In addition, the model will be compared to different machine learning algorithms and an external validation test will be exciting.

A final conclusion of the work will be provided at the end of this thesis.

Chapter 5

Literature Review

Chapter I
Literature Review

I.1. Introduction

This chapter aims to present a literature review of the selected theme of this thesis, which will cover, to the best knowledge of the authors, most of the published research papers related to predicting/modeling the solubilities in green solvents including supercritical carbon dioxide (SC-CO_2) and different ionic liquids (ILs) that have been published by reliable sources over the past four/five decades. It should be noted that this literature review mainly focuses on the different techniques used for the correlation of solubilities of binary solute-solvent systems or ternary systems of solute-solute-solvent only.

This chapter is divided into two main sections. The first section is focused on the detailed presentation of the most theoretical modeling techniques used up for the correlation of the solubility of solid compounds (mainly focused on pharmaceutical compounds) in SC-CO_2, this section is ended with a critical judgment on all reported studies and justified the choice of Quantitative Structure – Property Relationship (QSPR) techniques as a modern solution for the topic.

According to our researchers, until now, there is not much investigation related to the prediction of drug solubility in ILs. Thus, the second section will present the prediction of different solute compounds solubilities in ILs, as well as the section will be ended with an explanation on the selection of the QSPR combined with a on conductor-like screening model for Real Solvents (COSMO-RS) approach to model our property. The importance and usefulness of those methods will be highlighted in each section. A conclusion will close this bibliographical review.

I.2. Solubilities of Solid Compounds in Supercritical Carbon Dioxide

From the literature, it is clear that accurate experimental determination of pharmaceutical compound (or solid compounds) solubilities at various temperature and pressure ranges in SC-CO_2 is a challenge and time-consuming task. Therefore, achieving more

knowledge on the development of promising numerical/computational techniques of solubility calculation to validate the experimental data is essential to improve the quality of research (therefore, the efficacy of novel drugs), which also justifies the increased availability of contributions in this area. A review of most published papers on correlating the solubility of solutes in SC-CO_2 between 1981 and 2023 is reported below:

I.2.1. Literature Review on Equation of State Correlations

Kurnlk et al,. (1981) [86] correlated the solubility of benzoic acid mole fractions in SC-CO_2 as a function of temperature and pressure with the Peng- Robinson equation of state (PR-EOS). The results were found to be satisfying.

Bartle et al., (1989) [87] predicted the solubility of four compounds (naphthalene, phenanthrene, phenol, and 2-naphthol) in SC-CO_2 at six temperatures between 40 °C and 100 °C and pressures up to 240 bar using the PR-EOS. The results were similar to the experimental data.

Bartle et al., (1992) [88] used a PR-EOS to correlate the solubilities of 19 compounds of varied chemical type in SC-CO_2. The binary interaction parameter was used as an adjustable variable. The correlation obtained can be used to obtain rough estimates of the solubility of involatile compounds in SC-CO_2. The results were found to be satisfying.

Chen et al., (1993) [89] correlated the solubilities of cholesterol, stigmasterol, ergosterol, pregosterone, and lauric acid in SC-CO_2. The correlation employs the Patel-Teja cubic EOS. The mixture energy parameter of the EOS, a_m was determined by a UNIFAC group contribution model at an infinite pressure limit. The excluded volme parameter of the EOS, b_m was evaluated by the van der Waals mixing rule with an excluded volume correction term. The accuracy of the mixing model with excluded volwne correction is found to be superior to the traditional van der Waals mixing rules with optimally adjusted binary parameters. The solubilities of heavy steroid molecules in SC-CO_2 were calculated by using cubic Patel-Teja EOS. A modified mixing rule by employing an excluded volume correction term was presented. The excluded volume correction method gave better calculation results than the van der Waals mixing rules with optimally fitted temperature-dependent parameters.

Kwon and Mansoori (1993) [90] incorporated the hard-sphere expansion (HSE) theory in the K~rkwood-Buff fluctuation integral to predict the solubilities of solids in SC-CO_2. Three adjustable parameters (two molecular size parameters σ_{11}, σ_{12}, and one dimensionless

parameter α_{12}) in the HSE model were determined to fit experimental solubility data. The results on the solubihty prediction by the HSE model with the PR-EOS for pure components were better than the PR-EOS with the classical mixing rule in both cases: (1) when the molecular parameters in the HSE model and the binary interaction parameter k_{12} in the mixing rule were assumed to be temperature-dependent; and (2) when α_{12} and k_{12} were considered as temperature-dependent. The results obtained were found to be better than the PR-EOS with the classical mixing rule in both cases.

Ashour (1993) [91] correlated the solubility data of solids obtained at 308 and 313K using a cubic EOS. The effect of the solute-solute interaction parameter has been studied. The equation with three binary interaction parameters gave accurate predictions. The results were similar to the experimental data.

Lee and Kim (1994) [92] applied the PR-EOS and the augmented van der Waals-density-dependent local composition equation to correlate the solubility of the isomers *0-*, *m-*, and *p*-methoxyphenylacetic acid. The results indicate that the latter equation of state correlates better the with Peng-Robinson equation. The results were similar to the experimental data.

Garnier et al., (1999) [93] estimated the sublimation pressure of solids using experimental fusion data and liquid-vapour equilibrium properties obtained from the Peng–Robinson equations of state developed by Coniglio. The results provided by their method are shown to be satisfactory for predicting the solubility of heavy hydrocarbon solids in SC-CO_2 when using a reliable EOS. The results were found to be satisfying.

Chen et al., (2001) [94] estimated the solid solubility of aromatic, fatty acid, and heavy alcohol compounds in SC-CO_2 using the PR-EOS and a modified Huron–Vidal type mixing model with a volume correction term along with the UNIFAC activity coefficient model with its optimally fitted binary interaction parameters. A volume correction term was applied, and its parameters were estimated as functions of the solid molar volume for both non-polar and polar systems. Reasonable sublimation vapor pressures were obtained and their values were well correlated by the Antoine equation. The proposed model gave satisfactory results of the solid solubility in SC-CO_2.

Zhong and Yang (2002) [95] developed a SAFT-EOS combined with eight mixing rules to evaluate the capability of the SAFT approach for modeling the solubility of solids in SC-CO_2. The results were found to be satisfactory when the three-parameter mixing rules are used and the SAFT approach provides good correlative accuracy, where the AARD is normally

smaller than 10%. The present work shows that the SAFT approach can be used to model solid-SCF equilibria, resulting in slightly better correlative accuracy than cubic EOSs.

Madras et al., (2003) [96] combined the PR-EOS with the two-parameter van der Waal's mixing rules for modeling the solubilities of fatty acids in SC-CO_2. The two adjustable interaction parameters in the model are found to vary linearly with the chain length of the fatty acids. Thus this model can be used to predict the solubilities of various fatty acids in SC-CO_2. The results were similar to the experimental data.

Valderrama and Alvarez (2004) [97] applied the temperature-independent mixing rules of KWAK and MANSOORI [98] in cubic EOS to correlate the solubility of solids in SC-CO_2 (eight binary mixtures). Those mixing rules proposed a transformation of standard cubic EOS so that the EOS constants were really constants, and were evaluated and modified. The modification consisted of introducing a volume interaction parameter which only the solute contribution as proposed by one of the authors [Ing. Quim. Spain 34 (394) (2002)380], into the PR-EOS. The proposed model, in which the equation of state constants and the mixing rules are all temperature independent, correlates the solute concentration in the gas phase similarly and better than models that use temperature-dependent constants and temperature-dependent mixing rules.

Housaindokht et al., (2006) [99] performed a comparison between the modified PR-EOS by Danesh et al., (1995) [100], the modified PR-EOS by Gasem et al., (2001) [101], and the Mansoori-Mohsen Nia-Modarress EOS in terms of predicting the solubility of certain solids (Ascorbic acid, Ascorbyl palmitate, Butyl hydroxyl anisole, Dodecyl gallate, and Propyl gallate) in SC-CO_2. From the results, the overall AARD% value predicted by MM-ESO was larger than AARD% values for MPR1 and MPR2 equations of state. Except for Ascorbyl palmitate, as a large asymmetric component, the AARD% value predicted by MMM-EOS was the lowest.

Coimbra et al., (2006) [102] predicted the solubilities of flurbiprofen, ketoprofen, naproxen, and ibuprofen in SC-CO_2 using the Peng–Robinson (PR), Soave–Redlich–Kwong (SRK) and Patel–Teja–Valderrama (PTV) equations-of-state with van der Waals (vdW), Panagiotopoulos–Reid (mrPR) and Mukhopadhyay–Rao (MR) mixing rules. The models were developed in *Mathematica*®, a powerful symbolic computational language. Correlations were compared and discussed on the basis of the employed EOS and mixing/combining rules. The importance of the calculated methods used for the determination of critical and other drug physical properties (namely drug's sublimation pressure) was also highlighted in the discussion

of the quality of correlation results. A potential useful procedure (based on simultaneous correlation of one of the critical or physicochemical properties) was suggested, in order to satisfactorily choose the property estimation methods to be used.

Martinez-Correa et al., (2009) [103] correlated the solubility of squalene in SC-CO_2 with the Chrastil equation. The thermodynamic modeling using the PR-EOS only correlated well with the experimental values obtained when the critical properties and acentric factor reported by Ruivo et al. (2004) [104] were used. The values predicted for the solubility of squalene in SC-CO2 using modeling employing the GC-EOS, were only of the same order of magnitude as the experimental values for pressures below 200 bar, and its performance was influenced by the value of the critical hard sphere diameter.

Esmaeilzadeh et al., (2010) [105] applied the Peng–Robinson (PR) and the Esmaeilzadeh–Roshanfekr (ER) EOS along with several mixing rules including the Wong–Sandler (WS), the co-volume-dependent (CVD) and the van der Waals one (VDW1) and two (VDW2) fluid mixing rules to evaluate the solubilities of 52 mostly used solid compounds in SC-CO_2, from 1776 experimental data points. Moreover, the Van-Laar excess Gibbs energy (Gex) model was used in phase behavior calculations by the WSmixing rule. The optimal values of the proposed thermodynamic model parameters were then evaluated using the DE (differential evolution) optimization strategy. The absolute average deviations of the model results and the optimal values of the adjustable parameters of the model were reported to investigate the capabilities of combinations of each EOS with different mixing rules in calculations of the solubilities. The results indicated that the combination of the ER-EOS with the WS mixing rule led to more accurate results (AAD = 9.0%) compared with other ones.

Esmaeilzadeh et al., (2011) [106] applied various EOS namely, Soave-Redlich-Kwong(SRK), Peng-Robinson (PR), Patel–Teja–Valderrama (PTV), the Esmaeilzadeh-Roshanfekr (ER) and the modified Esmaeilzadeh-Roshanfekr(m-ER) EOS complained with several mixing rules including the Wong–Sandler (WS), the co-volume-dependent (CVD) and the vander Waals one (vdW1) and two (vdW2) fluid mixing rules to calculate the solubilities of 23 mostly used solid compounds in SC-CO_2. In addition, the Van-Laarexcess Gibbs energy(Gex) activity model was also used in phase equilibrium calculations. The probable global optimal values of the proposed thermodynamic model parameters were evaluated by the differential evolution (DE) optimization strategy. The results indicated that the combination of them-ER-EOS with the WS mixing rule led to more accurate predictions (AARD¼7%) in comparison with other ones.

Hosseini et al., (2012) [107] developed a PR-cubic-EOS along with the van der Waals combining rule to model the solubility of cefixime trihydrate, an antibiotic drug, and oxymetholone, a 17alpha-alkylated anabolic-androgenic steroid drug, in SC-CO_2 over a range of 183 to 335 bar, and 121 to 305 bar of pressure, respectively, and temperatures of 308, 318 and 328 K. The results obtained from the semi-empirical models showed that there is good agreement between the experimental data and the results of semi-empirical models with an AARD of 11.92% and 11.74% for cefixime trihydrate and oxymetholone, respectively.

EL HADJ et al., (2013) [108] correlated solubilities of some anti-inflammatory (nabumetone, phenylbutazone and salicylamide), and statin drugs (fluvastatin, atorvastatin, lovastatin, simvastatin and rosuvastatin) in SC-CO_2 at temperature and pressure in ranges of 308–348 K and 100–360 bar, respectively, using the PC-SAFT with one-parameter mixing rule and commonly used cubic EOS Peng-Robinson (PR) and Soave-Redlich-Kwong (SRK) combining with van der Waals 1-parameter (VDW1) and van der Waals 2-parameter (VDW2) mixing rules. The critical properties required for the correlation with PR and SRK were estimated using Gani and Noonalol contribution group methods whereas, PC-SAFT pure-component parameters: segment number (m), segment diameter (σ) and energy parameter (ε/k) were estimated by Tihic's group contribution method for nabumetone. For phenylbutazone and salicylamide, those parameters were determined using a linear correlation. For statin drugs, PC-SAFT parameters were fitted to solubility data, and binary interaction parameters (k_{ij} and l_{ij}) were obtained by fitting the experimental data. The results were found to be in good agreement with the experimental data and showed that the PC-SAFT approach can be used to model solid-SCF equilibrium with better correlation accuracy than cubic EOS.

Rostamian et al., (2014) [109] proposed a modified Redlich–Kwong (βRK) EOS model coupled with the van der Waals zero (VDW0) mixing rule to estimate the solubilities of twenty solids, namly Ascorbic acid, Fluoranthene, Propyl gallate, Acenaphthene, Asprin, Climbazole, Cinnamic acid, Triclocarban, 4-methoxyphenylacetic acid, Phenoxyacetic acid, Cholesterol,Cholesteryl butyrate, Cholestrol acetat,Triphenylene, Ibuprofen, Acetanilide, Propanamide, Butanamide, Chrysene and Dodecyl gallate in SC-CO_2, accounting 440 data points. The results of the proposed model were compared to the results of the PR-EOS model coupled with the van der Waals one (VDW1) and the Wong-Sandler (WS) mixing rules. The obtained results confirmed that the suggested model performed well in reproducing the solubility of these twenty solids in supercritical carbon dioxide with AARD= 5.7 %.

Khansary et al., (2014) [110] carried out an EOS model based on Pitzer correlations for the compressibility factor and virial EOS, as a function of the total number of atoms present in the solute molecule, normal boiling point temperature, and reduced temperature. Thirty dyes and sixty drugs solubility data in SC-CO_2, 2417 data points, were used to build this model, then compared with the PR-EOS plus the two adjustable parameters van der Waals mixing rules. The results showed that the proposed EOS presented more accurate predictions for solubility data in SC-CO_2.

Zabihi et al., (2014) [111] used seven different cubic equations of state (van der Waals, Redlich-Kwong, Soave-Redlich-Kwong, Peng-Robinson, Stryjek-Vera, Patel-Teja-Valderana, and Pazuki) to predict ibuprofen's solubility in SC-CO_2 at the pressure range of 8-13 MPa and the temperatures of 308, 313 and 318 K. The supercritical solution properties were calculated using both single, and twin-parametric van der Walls mixing rules (VDW1, VDW2). The physicochemical properties were also obtained using Joback, Lydersen, and Ambrose methods. Results showed that among all models, the Pazuki equation (AARD = 19.85% using VDW1 and AARD = 8.79% using VDW2) and SRK equation (AARD = 19.20% using VDW1 and AARD = 10.03% using VDW2) predicted the best with the least error.

Hartono et al., (2015) [112] introduced five (5) cubic EOS with respective mixing rules for predicting the solubility of alkanes in SC-CO_2. The five equations of state are the Redlich-wong (RK), Mohsennia-Modarress-Mansoori (MMM), Riazi-Mansoori (RM), Peng-Robinson (PR), and Soave-Redlich-Kwon (SRK). As a result, it was found that the MMM equation is more accurate to predict the property for the 270 data points of the 10 binary systems.

Rostamian and Lotfollahi (2016) [113] used a modified Redlich-Kwong (βRK) and modified Peng-Robinson (βPR) EOSs combined with the van der Waals mixing rule (VDW0) to evaluate the solubilities of ten solid compounds (Methimazole, Ascorbic acid, Ascorbyl palmitate, Propyl gallate, Aspirin, Fluoranthene, Triclocarban, Hinokitiol, Phenol, and Climbazole) in SC-CO_2. The results of the two proposed models were compared to the results of SRK and PR EOS in combination with the VDW1 and the Wong-Sandler (WS) mixing rules and other models reported in the literature. The results showed that the proposed models are well for predicting this dataset, and the βRK-EOS in combination with the vdW0 mixing rule led to better correlation (AARD = 5.1%).

Coelho et al., (2016) [114] correlated seven solid solubilities in SC-CO_2 by applying a Soave-Kwong-Redlich (SRK) EOS with the one-fluid van der Waals mixing rule representing cubic EOSs. 6.7% overall AARD for the SRK-cubic-EOS was obtained. The model was

acceptable and it can be used as a reliable thermodynamic model to predict the solubility of any compound for which there is no sufficient experimental data available.

Sang et al., (2016) [115] performed simplifications and approximations towards three EOS combinations namely, Soave-Redlich-Kwong, Peng-Robinson, and modified Esmaeilzadeh-Roshanfekr, and five mixing rules including, Huron-Vidal, van der Waals one and two, modified covolume-dependent, and Wong-Sandler to overcome some limitation related to sublimation vapor pressure and critical parameter of solutes, such as its difficult to obtain, and the complicity of procedure calculation. The modified models were reduced into two groups, including two and three-parameter models. Then compared with various publication in terms of modeling the solubility of 23 solid solutes in SC-CO_2. The results suggest that the models of that work were better.

Bagheri et al., (2018) [116] carried out five (5) cubic EOS with two mixing rules to predict the solubility of Ibuprofen, Aspirin, Salicylic acid and Griseofulvin in SC-CO_2. Those EOS are van der Waals (VDW), Redlich-Kwong (RK), Riazi-Mansoori (RM), Mohsen-Nia-Moddaress-Mansoori (MMM), and Peng-Robinson (PR) with VDW type I, and Kwak-Mansoori (KM) mixing rules. The results confirmed that the two of the most accurate EOS are PR-KM and PR-vdW with less average absolute percent deviation (AAPD) than the other EOS.

Sodeifian et al., (2018) [30] predicted the solubility of six pharmaceutical compounds (Anti-HIV, Antiinflammatory, and Anti-cancer) in SC-CO_2 using the cubic EOS models (SRK and modified-Pazuki EOSs). Twin-parametric van der Waals (vdW2), and Panagiotopoulos-Reid (mrPR) mixing rules were used for estimating the supercritical solution properties, with three different sets employed for obtaining critical and physicochemical properties of the solid compounds. The capacity of the approach was evaluated based on several statistical criteria, including AARD, R_{adj}, and F-value.

Sodeifian et al., (2018) [117] predicted the solubility of Loratadine in SC-CO_2 at temperatures ranging from 308.15 to 338.15 K and pressures from 120 to 270 bar using three EOS, namely SRK, Qcpa, and PCP-SAFT. The results confirm the superiority of the SRK-EOS model.

Estévez et al., (2020) [118] correlated the solubility of thermolabile solid in SC-CO_2 using an EOS scheme. The solute fugacity coefficient is obtained using the pure-solvent compressibility factor, and empirical solute-to-solvent parameter ratios of cohesion factors and covolumes. The proposed method was formulated with any cubic equation of state where the authors chose to formulate it with the Redlich-Kwong and Peng-Robinson EOS which they

used two adjustable parameters, which are calculated from experimental data for numerous binary systems and used to predict solubilities. The results were favorably compared to different other methods. The advantage of EOS-based models over empirical ones was highlighted in cases where the solubilities are extrapolated beyond the range of experimental data. This contribution presented a simple, yet effective, way to model the solubility of solids in supercritical fluids.

Cai et al., (2020) [119] examined the accuracy of a first-principles thermodynamic model PR-COSMOSAC-EOS in modeling 99 solid solute solubility in SC-CO_2, accounting for 2450 data points, then compared the results to that of a group contribution model PSRK-EOS. The model requested the knowledge of molecular structure, heat of fusion, and melting temperature of the investigated compound which can be obtained from quantum mechanical solvation calculations. The average logarithmic deviation in mole fraction (ALD-y) was found to be 0.79 (or about 510%). This indicated that the PR-COSMOSAC-EOS was able to reasonably estimate solubility order of magnitude without any adjustable model parameter and provided useful solubility information, especially when no experimental solubility data is available.

Ushiki et al., (2020) [120] applied the perturbed-chain statistical associating fluid theory (PC-SAFT) equation of state for predicting the solubilities of metal acetylacetonate (acac) precursors in SC-CO_2. Three PC-SAFT pure-component parameters namely, the segment diameter, segment number, and dispersion energy were determined by adjusting their values to the measured solubilities in organic solvents. Those parameters were then applied to predict the target output over a wide range of pressures and temperatures. The final results provided a successful PC-SAFT description.

Alwi and Garlapati (2022) [121] proposed two forms of simplified EOS models to correlate the solubility data of 11 anticancer drugs in SC-CO_2 where their performance was evaluated with 15 existing solubility models (Garlapati–Madras, Adachi-Lu, Hozhabr et al. Alwi–Garlapati, Keshmiri et al. Khansary et al. Bian et al. Gordillo et al. Chrastil, Jouyban et al. Sung-Shim, Mendez-Santiago–Teja, Jafari et al. Bartle et al. and PR-EOS-VDW2). The correlation ability was quantified in terms of AARD, AIC, and p-test values. The proposed model 1 can be implemented easily with tools that are available on modern personal computers and can estimate better than all models investigated except the Garlapati–Madras model while proposed model 2 predicted the best among all of them.

Ushiki et al., (2022) [122] examined the predictive ability of perturbed-chain statistical associating fluid theory (PC-SAFT) model to estimate the solubilities of two non-steroidal anti-inflammatory drugs (NSAIDs), ibuprofen and ketoprofen in SC-CO_2. The values of the three PC-SAFT pure-component parameters of the drugs were adjusted then used to correlate the target property similer to there previous paper Ushiki et al., (2020) [120]. The predictive ability was found to be almost reproduced the experimental values and better than conventional methods that use cubic-type equations of state and semi-empirical correlation models.

Sodeifian et al., (2022) [123] correlated the solubility of metoclopramide hydrochloride (MCP) in SC-CO_2 using six semi-empirical models three models based on the Peng-Robinson equation of state (PR+VDW, PR+WS+Wilson, and PR+MHV1+COSMOSAC) with different mixing rules and various ways to describe intermolecular interactions as a function of the pressure and the temperature. The results indicated that the PR+MHV1+COSMOSAC is the optimal model with an AARD of 28.9%.

Rahman et al., (2022) [124] developed an artificial intelligence-based predictive model namely, a support vector machine to predict the solubility data of Busulfan drug in SC-CO_2. The temperature and pressure were used as input parameters and the polynomial was selected as the kernel function. Good agreement was observed between the proposed model and the experimental values with an R^2 higher than 0.99.

Wang and Hsieh (2022) [125] proposed combining the PSRK-EOS with sublimation pressures estimated from the Lee-Kesler (LK) equation to predict solubility in SC-CO_2 of 57 solid solutes. Both reported techniques request identical inputs of critical properties and the acentric factor. The average logarithmic deviation (ALD-*x*) of using sublimation pressures from LK with the optimal parameters was found to be 0.84, while using experimental melting properties 0.77. These results demonstrated the feasibility of predicting solubility with only the input of molecular structure.

Shahriari et al., (2022) [126] evaluated a PHSC-EOS model to predict fifteen (15) pharmaceutical compounds' solubility in SC-CO_2 over a wide range of temperatures and pressures. The modeling results were compared to PC-SAFT, COSMO-RS, and COSMO-SAC models. The obtained results indicated good agreement between the PHSC-EOS and the experimental solubility data. The PHSC capabilities were evaluated by estimating the solubility of five (5) pharmaceutical compounds in SC-CO_2 with an adjustable temperature-dependent binary interaction parameter. The results confirmed the essentiality of that parameter to correlate such a property and the applicability of the PHSC EoS in the pharmaceutical industry.

Ushiki et al., (2022) [122] estimated the solubility of salsalate in SC-CO_2 using a perturbed-chain statistical associating fluid theory (PC-SAFT)-based modeling. The PC-SAFT pure-component parameters of the API were first calculated by fitting its solubility data in organic solvents. Then those parameters were employed to estimate the solubilities with the binary interaction parameter (k_{ij}) set once to zero and another time as an adjustable parameter. This type of modeling is more predictive than conventional semi-empirical correlation models and cubic-type equations of state. Thereafter, the results indicated that this method can adequately reproduce the experimental values of the isothermal and isobaric solubilities of salsalate in SC-CO_2 over a wide range of temperatures and pressures. The utilization of readily measurable solubility data of the pharmaceutical compounds in organic solvents combined with PC-SAFT helps with overcoming their unavailability limitation.

Honarvar et al., (2022) [127] used various modeling techniques for the prediction of the solubility of sildenafil citrate in SC-CO_2 at different temperatures (308 to 338 K) and pressure values (12 to 30 MPa). The AARD of the applied techniques was found as follows: SRK-EOS an AARD of 14.33, PC-SAFT-EOS an AARD of 10.26.

Majrashi et al., (2022) [128] predicted the solubility of Chlorothiazide, an antihypertensive drug, in SC-CO_2 over a range of 308–338K and 130–290 bar) with the use of two EOS, namely SRK-EOS, and PR-EOS. The SRK-EOS model outperforms the PR-EOS.

Sodeifian et al., (2023) [129] correlated the solubility of Fexofenadine hydrochloride, an antihistamine drug in SC-CO_2 over a range of different pressures (120–270 bar) and temperatures (308.2–338.2 K) with the use of three cubic EOSs with two mixing rules. The most accurate model was the PR-EOS with Kwak-Mansoori mixing rule with an AARD of 5.48%.

Ardestani et al., (2023) [130] used a non-cubic EOS PC-SAFT, and two versions of the Est´evez et al. model (ECM-PR and ECM-RK) to predict the solubility of famotidine in SC-CO_2 at different temperatures (308-338 K) and pressures (12-30 MPa) ranges. The results indicate that Est´evez et al. model has the highest accuracy with an AARD of 9.92% for the ECM-PR model and 10.70% for the ECM-RK model.

Sodeifian et al., (2023) [131] predicted the solubility of palbociclib in SC-CO_2 at different equilibrium conditions (308-338 K, and 12–27 MPa) using PR-based integrated models. PR-EOS with van der Waals mixing rules (PR + vdW) showed the best accuracy results.

Nateghi et al., (2023) [132] predicted the solubility of Nilotinib hydrochloride monohydrate, an anti-cancer drug, in SC-CO_2 at various temperatures (308–338 K) and pressures (120–270 bar) conditions using PR-EOS (Van der Waals mixing rule). The final model had an *AARD%* of 10.73.

Sodeifian et al., (2023) [133] carried out a PR-EOS and for the prediction of the solubility of buprenorphine hydrochloride, a strong pain reliever, in SC-CO_2. The results were similar to the experimental data.

Venkatesan et al,. (2023) [134] correlated the solubility of Metformin, an anti hyperglycemic, through a Peng-Robinson EOS. The temperature, pressures, and mole solubility fraction range from 308-338K, 140-290 bar and 0.39-1.23 * 10^{-6}, respectively. The average AARD% was found to be 8.24%.

Edrisi and Bakhshi (2023) [135] carried out three thermodynamic model : Soave–Redlich–Kwong equation of state (SRK), cubic plus association equation of state (CPA), and quadrupolar cubic plus association (qCPA) to model the solubility of two drugs, namely ibuprofen and salicylic acid, in SC-CO_2 at various temperatures (308–333 K) and pressures up to 35 MPa. the Adjustable parameters of each model were regressed and reported at the studied temperatures. The results indicate the superiority of the CPA and the qCPA model with AARD% 1.34 and 2.16 for ibuprofen and 1.63 and 1.65 for salicylic acid compared to the other applied methods.

Sodeifian et al., (2023) [136] carried out a theoretical investigation and computational approach to correlate and extrapolate the experimental solubility of Ibrutinib in SC-CO_2 using thermodynamic models based on different approaches; semi-empirical models, activity coefficient (modified model of Margules), fugacity coefficient (*m*PR and *m*RK cubic EOSs) and dilute solution theory by the model of Wang-Tavlarides. The best correlation was found by the modified model of Margules (AARD = 6.68%) and the empirical model of Jouyban*et al*., (9.62%).

I.2.2. Literature Review on Empirical and Semi-empirical Equations Correlations

Kurnlk et al,. (1981) [86] contributed by proposing an experiment equilibrium for modeling the solubility of five different solids in SC-CO_2. A one-pass flow system was used to measure the equilibrium solubilities. The range in temperatures covered was 308-338 K and a range of 80-280 bar for the pressures. The results were similar to the experimental data.

Zlger and Eckert (1983) [137] developed a semi-empirical correlation to predict the solubility of heavy nonpolar solids in SC-CO_2 using only one temperature-independent pure component parameter for each component. The method was based on regular solution theory and the van *der* Waats equation of state. The range of solubility of the 11 solids correlated was 0.00001 to 0.01 mole fraction, 20-80 °C, and 80-500 bar.

Vetere (1997) [138] proposed a simple empirical method to estimate the polar and non-polar solid solubility in SC-CO_2. The model uses two input parameters: the molecular weight, and the melting temperature. A pseudo-sublimation enthalpy was derived for several classes of compounds grouped according to their polarity degree and molecular weight and melting temperature values. The density of carbon dioxide was calculated according to the Vetere EOS from the experimental data of pressure and temperature. The authors applied this model to numerous compounds ranging from high-molecular-weight hydrocarbons to molecules of biological interest, such as vitamins and anti-inflammatory drugs.

Jouyban et al., (2002) [139] correlated the solubility of 106 published drug solubility in SC-CO_2 using an empirical equation with temperature, pressure, and density of pure SC-CO_2 as inputs. The accuracy of the proposed equation was compared to four other empirical equations. The mean AARD for the proposed model was of 21.4 ($\mp$15.9)%, which is an acceptable error when compared with the experimental uncertainty. The mean AARD of the new equation was significantly lower than those obtained from previous empirical models (paired *t*-test, $P<0.0005$). The proposed model presented a more accurate correlation for solubility data in SC-CO_2. It can be employed to speed up the process of SCF applications in the pharmaceutical industry.

Su and Chen (2007) [140] estimated the solubilities for three non-steroidal anti-inflammatory drugs (NSAIDs) of nabumetone, phenylbutazone, and salicylamide in SC-CO_2 using a semi-empirical equation presented by Chrastil, and that presented by Mendez-Santiago and Teja. With optimally adjusted parameters, these two equations yielded satisfactory results where the AARD was less than 7%.

Yamini et al., (2007) [33] correlated solubility data of lovastatin, simvastatin, atorvastatin, rosuvastatin, and flovastatin in SC-CO_2 over a pressure ranges of 121.6–354.6 bar and temperatures in the range of 308–348 K using four semi-empirical density-based models (Chrastil model, Bartle model, Kumar and Johnstone “K–J” model and Mendez-Santiago and Teja “M–T” model). The results showed that by using three temperature-independent parameters these models can be applied for satisfactory solubility predictions at different

pressures and temperatures while the Bartle and M–T models gave much better correlations of the solubility data with an overall AARD% of 6.3 and 8.8%, respectively.

Hernandez et al., (2008) [141] evaluated different, density-based, semi-empirical models in terms of their ability to predict 6 solid solubility in SC-CO_2. The solid compounds used in this study were naphthalene, anthracene, fluorene, hydroquinone, 1,5-naphthalenediamine, and cholesterol. The models used were the methods of Chrastil, del Valle and Aguilera, Adachi and Lu, Méndez-Santiago and Teja, and Bartle. The average error was calculated for all 18 isotherms with each of the models evaluated. Both the Adachi–Lu and del Valle–Aguilera equations provided, in general, lower average error than the other models. According to those results, the authors propose their own equation as a combination of the previous two methods, where the proposed model provided the least error. The obtained model is particularly useful when the reduced density of the solvent is below 1, where previous models tend to fail.

Madras and Garlapati (2009) [142] compared existing solvate complex models to their own set of proposed empirical equation in terms of correlating the solubility of 15 binary systems (solid_SC-CO_2) as a function of temperature, and density of SC-CO_2. The accuracy of the proposed models was evaluated by correlating 15 binary and 18 ternary systems. The proposed model provided the best overall correlations.

Gong and Cao (2009) [143] correlated the solubility of a drug named artemisinin in SC-CO_2 at three temperatures of 313K, 323 K, 333K and pressures from 11 to 31MPa using two density-based methods including, Chrastil and MST. The AARD was 8.32% and 8.33%, respectively. The correlation results presented good agreement with the experimental data.

Nejad et al., (2009) [144] predicted the solubility of 16 solute compounds in SC-CO_2 as a function of the temperature, pressure, and density of pure SC-CO_2 using an empirical equation. The proposed method was compared to three other empirical equations where the proposed model has an average AARD% of (7.46 (±4.54)) which is significantly lower than that obtained from Chrastil et al. model (11.70 (±23.10)) and has the same accuracy as compared with Bartle et al. (6.895 (±3.81)) and Mèndez-Santiago–Teja model (6.39 (±6.41)).

Del Valle et al., (2010) [145] presented various density-based models (Chrastil, Adachi-Lu, del Valle-Aguilera, Sparks, Kumar-Johnston, Bartle, Mendez Santiago-Teja) together with Yu's model and Gordillo's model for predicting solubilities of 27 different pharmaceutical solutes in SC-CO_2. The results showed that, in general, Sparks' equation provided the best fit to the solubility data. However, under certain specific conditions, the best correlation was obtained using Gordillo's equation. By means of a brief comparison with the PR-EOS, semi-

empirical equations presented a more accurate prediction compared to cubic EOS, and presented no drawbacks, such as properties estimation, and computational difficulties.

Kikic et al., (2011) [146] estimated the solubility behavior of pharmaceutical compounds (antioxidants, antibiotics, steroids, and anti-inflammatory) in SC-CO_2 using an activity coefficient model based on linear solvation energy relationships. The parameters of this model were further generalized as a function of the properties of the pharmaceutical compounds. This model can estimate the solubilities within acceptable accuracy for more than 60% of the proposed systems.

Yamini et al., (2011) [147] modeled the solubilities of two antifungal drugs of ketoconazole and clotrimazole in SC-CO_2 at (308 to 348) K and over the pressure range of (12.2 to 35.5) MPa using five density-based models, namely Chrastil, modified Chrastil, Bartle, modified Bartle and Mendez-Santiago and Teja (M–T). The correlation results showed good agreement with the experimental data. A comparison among the five models revealed that the Bartle and its modified models gave much better correlations of the solubility data with an AARD% ranging from 4.8% to 6.2% and from 4.5% to 6.3% for ketoconazole and clotrimazole, respectively.

Nasri et al., (2012) [148] proposed a methodology to correlate and predict the solubility in SC-CO_2 of some polycyclic aromatic hydrocarbons. This methodology used the expanded liquid theory, in which the solid-fluid equilibrium is modeled using the UNIQUAC model. An empirical equation that related the interaction parameters to the solvent density was used to calculate the regressed coefficients of the model. The compounds used were acenaphtene, anthracene, chrysene, fluoranthene, fluorene, naphthalene, phenanthrene, pyrene, perylene, and triphenylene; more than 1400 solubility data points. The results obtained using the proposed technique showed great agreement with the experimental data used.

Hosseini et al., (2012) [107] correlated the solubility of cefixime trihydrate, an antibiotic drug, and oxymetholone, a 17alpha-alkylated anabolic-androgenic steroid drug, in SC-CO_2 over a range of 183 to 335 bar, and 121 to 305 bar of pressure, respectively, and a temperature of 308, 318 and 328 K using four semi-empirical models namely, Bartle, Kumar and Johnstone (K–J), Mendez-Santiago and Teja (M–T) and Chrastil models. The results obtained from the semi-empirical models showed that there is good agreement between the experimental data and the results of semi-empirical models.

Sousa et al., (2013) [149] used density semi-empirical based models namely the Mendez-Santiago-Teja, Bartle, Chrastil, and Spark's models to calculate drugs like norfloxacin,

ofloxacin, and dexamethasone solubilities in SC-CO_2 at different temperature conditions (308.2 K, 318.2 K, and 323.2 K), and in a pressure range between 10 and 35.7 MPa. For all the studied drugs, the models were able to correctly describe the temperature dependence but failed to describe its dependence on density and for all models, the global AARD was lower than 13%.

De Melo et al., (2013) [150] investigated predicting the solubility of solid therapeutic compounds in SC-CO_2 using semi-empirical models. A comparative study between eight models was carried out where all equations provided an average deviation of solubility of approximately 10%. Using the González (with four parameters), and Reddy and Madras equation (with five parameters) was found to be the most practical because it has a smaller number of parameters. With the introduction of a "q exponential function", the authors proposed generalized models based on the Tsallis's statistical non-extensive entropytheory. In fact, the original Chrastil and González equation can be considered as a particular case of the q-Chrastiland q-González models. They presented deviations similar to the other model and showed potential for further studies.

Yamini et al., (2013) [151] correlated the solubility data of desoxycorticosterone acetate (DA), and clobetasole propionate (CP) in SC-CO_2 using four semi-empirical density-based models (Chrastil, Bartle, Kumar and Johnston (K–J) and Mendez-Santiago and Teja (M–T) models), at temperature ranging from (308 to 348) K and pressures from (12.2 to 35.5) MPa. The AARD% ranged from 9.3 to 13.6; 8.9 to 11.9; 6.5 to 10.3 and from 10.4 to 13.4 for Chrastil, Bartle, K–J, and M–T models, respectively. A comparison among the four models revealed that the K–J model gave a much better correlation of the solubilities in comparison with other models.

Haghbakhsh et al., (2013) [152] built a correlation based on the sensitivity analysis and mathematical regression for the prediction of solubility of 31 well-known industrial compounds (1012 solubility data) in the SC-CO_2 at the temperature range of 308–523 K and pressure between 75 and 468 bar. The results were compared with the literature experimental data and different models where the proposed model showed better performance (lower AARD%) and more accurate prediction results with respect to K–J and Chrastil models and confirmed that this correlation can be used as a reliable solubility model in industrial applications.

Tabernero et al., (2013) [153] suggested a generalization of three adjustable parameters Chrastil equation by using Tsallis´ statistical non-extensive entropy theory and the q-exponential function model for modeling the solubility of 58 pharmaceutical compounds in SC-CO_2. The proposed model (q-Chrastil) was compared to Chrastil equations. According to the

results, both models provided deviations of about 15%, slightly higher than those obtained when models with a higher number of adjustable parameters were used. q-Chrastil equation is promising for modeling complex systems at higher temperatures and serves as a starting point for the development of new models.

Hezave And Lashkarbolooki (2013) [154] proposed an empirical correlation with four fitting parameters, optimized with GA, to correlate the solubility of pharmaceuticals in SC-CO_2 at different temperatures and pressures, accounting for 601 data points. The obtained results compared with four widely used density-based correlations including Mendez-Santiago and Teja (MST), Bartle et al., Chrastil, and Kumar and Johnston (KJ) revealed rather good capability of the proposed simple correlation for predicting the solubility of solutes in SC-CO_2. The results were also compared with the results of three EOSs including PR, Pazuki, and modified MPR-EOS's coupled with three different mixing rules, including VDW1, VDW2, and Adachi-Sugie. The comparison revealed that the proposed model was more accurate. Besides, the application of the proposed correlation is easier and faster than EOS's, which needs complex calculations and some physiochemical properties of substances are not always experimentally available.

Sani et al., (2014) [155] applied an empirical model for the mathematical representation of solutes solubility in SC-CO_2. The accuracy of the proposed model was compared to previously published models in the literature using experimental data of dyes and drugs collected from the literature (more than 2400 data points). According to the obtained results, the proposed model presented better accuracy estimation/prediction for in solubility.

Amooey (2014) [34] proposed a correlation to predict drug solubility in SC-CO_2. The main advantages of that correlation are its simplicity and minimal input data, namely temperature and density. The proposed technique showed better performance (lower AARE%) with respect to eight other empirical equations. Due to data sets, the mean AARE for the proposed model was of 6.88%.

Mazloumi et al., (2014) [156] correlated the solute solubility of 24 compounds in SC-CO_2 with an empirical equation, which has four parameters per solute, as a function of the pressure, temperature, and density of pure SC-CO_2. The overall AARD was found 6.54%. as well, comparing those results with the results of a three-parameter and a four-parameter empirical model reveals good accuracy of the proposed empirical model.

Nasri et al., (2014) [157] predicted the solubility of aromatic isomers of hydroxy-benzoic acid and methyl-benzoic acid in SC-CO_2 with a methodology based on the expanded

liquid theory. This theory used the local composition model of UNIQUAC (an activity coefficient semi-empirical model) to model the solid-fluid equilibrium in which the interaction parameters were linked to the solvent-reduced density with empiric exponential form equations. The results achieved with such a model showed to be in good agreement with the experimental data used.

Zabihi et al., (2014) [111] used seven different semi-empirical equations (Bartle, Modified Bartle, Mendez-Teja, Modified Mendez-Teja, Kumar-Johnson, Sung-shim, and Gordillo) to predict ibuprofen's solubility in SC-CO_2 at the pressure range of 8-13 MPa and the temperatures of 308, 313 and 318 K. Results showed that among all models, the most desirable deviation (AARD<10%) was obtained by using Modified Bartle and Modified Mendez-Teja equations in all the studied temperatures.

Bian et al., (2015) [158] proposed a five-parameter empirical model for modeling 45 solute solubility in SC-CO_2 (1130 data points, published data from 2010 to 2015). The results of the proposed model were compared to 14 most common empirical models (Chrastil, Adachi-Lu, del Valle-Aguilera, Sparks, Gordillo, Jouyban, Garlapati-Madras, Ch-Madras, Kumar-Johnston, Méndez Santiago-Teja, Hozhabr, Keshmiri, Khansary) where the lowest overall AARD 55.16%) belongs to the corresponding model.

Yamini et al., (2015) [159] correlated four drug compounds' (azithromycin, erythromycin, clindamycin, and clarithromycin) solubility in SC-CO_2 at temperatures ranging from (308 to 348) K and pressures from (12.2 to 35.5) MPa using four semi-empirical density-based models (Chrastil, Kumar, Johnston, Bartle and Mendez-Santiago, and Teja models). The AARD ranged from 1.7 to 7.4; 2.07 to 8.1; 7.21 to 10.37 and from 5.13 to 9.44 for Chrastil, Bartle, K-J and M-T models, respectively. The results indicated that the use of three temperature-independent parameters allowed to make these models applicable for satisfactory solubility predictions at different pressures and temperatures while the K-J and Chrastil models gave the best correlation of the solubilities in comparison with other models.

Banchero and Manna (2015) [160] correlated the solubility data of niflumic and tolfenamic acid (fenamate NSAIDs with antipyretic, analgesic and anti-inflammatory properties) in SC-CO_2 at temperatures between 313 and 343 K and pressures of 16-40 MPa with fourteen different correlation methods. Results showed that the Chrastil-modified approaches provided the best results. The best fit was obtained with the Adachi-Lu model for niflumic acid (AARD% = 5.8) and the del Valle-Aguilera model for tolfenamic acid (AARD% = 5.5) while the original Chrastil model provided only slightly higher deviations.

Andonova and Sekhar (2016) [161] proposed an empirical model based on the degree-of-freedom analysis to predict the solubility of pharmaceutical compounds (Penicillin G and Penicillin V) in SC-CO_2. The model has three temperature-independent constants. The model was successfully correlated this binary system.

Wang et al., (2016) [162] presented two new semi-empirical models to calculate and predict the solubility data of 40 compounds in SC-CO_2. The 1st model was built based on chrastil's techniques, then the model I was compared with models of chrastil, Adachi-Lu and de Valle. The results showed that the model I can establish better correlation than other models. However; not all mentioned models do not have the ability to predict unknown solubility data. The 2nd model was developed by combining the modified molecular connectivity indices (MCIs) with model I to calculate the solubility data of fatty acids, fatty alcohols, fatty acid esters, and aromatic compounds in SC-CO_2. Noted that, the relation between the solutes' solubility data and their molecular structures was established by a one-off correlation of experimental data and then the inductive relation was used in the model II to predict the solubility data of homogeneous solutes successfully.

Coelho et al., (2016) [114] correlated seven solid solubilities in SC-CO_2 by applying density-based models including Chrastil, Kumar and Johnston, Bartle et al., Méndez-Santiago and Teja, Garlapati and Madras, Nejad et al., and Khansary et al. models. The results obtained revealed that the density-based models recommended more recently (Garlapati and Madras, Nejad et al., and Khansary et al. models) perform better with an over AARD of 6.5, 9.0, and 9.3%, respectively.

Si-Moussa et al., (2016) [163] proposed a density-based model derived by a simple modification of the Jouyban et al. model to estimate the solubility of 100 solid drugs in SC-CO_2, accounting for 2891 experimental data points. The six-parameter model expresses the solubility only as a function of the solvent density and the equilibrium temperature. The comparison in terms of the mean absolute relative deviation for each solid drug and for the entire database between the proposed model and models that were suggested to be mostly more accurate demonstrated that the proposed model presented the best global correlation performance, exhibiting an overall average absolute relative deviation of 8.13%.

Reddy et al., (2018) [164] developed an empirical model based on the degree of freedom analysis to correlate the solubility of twenty-seven industrially important pharmaceutical drug compounds in SC-CO_2. The predictive ability of the proposed model was validated with literature models such as Chrastil, Kumar and Johnston (K–J) (1988), Bartle et al., Mendez-

Santiago and Teja (MST) (1999), Sung and Shim (SS) (1999), Garlapati and Madras (2010), Jafari Nejad et al., Keshmiri (2014), Khansary (2015), Bian et al., (2016). The correlative accuracy of the model was evaluated in terms of AARD, the proposed model was found to correlate better than the existing models in terms of overall AARD.

Sodeifian et al., (2018) [30] predicted the solubility of six pharmaceutical compounds (Anti-HIV, Antiinflammatory, and Anti-cancer) in SC-CO_2 using empirical and semi-empirical models (Chrastil, Mendez-Santiago-Teja, Spark et al. and Bian et al. models) The capacity of the approach was evaluated based on several statistical criteria, including AARD, R_{adj}, and F-value.

Belghait et al., (2018) [165] performed a comparative study between 21 empirical models in terms of predicting the solubility of 210 solid solutes in SC-CO_2, counting 5550 data points as well, the authors proposed an eight-parameter density-based model. According to the results models that related the logarithm of the solubility to the logarithm of solvent density and temperature were more accurate than those that have the pressure as an input, while three-parameter models were the worst. Also, the proposed model was found the best with an AARD range of 0.17-81.99% while the average is 8.88%.

Reddy and Garlapati (2019) [166] used an empirical model developed based on the degree of freedom analysis to estimate the solubilities of 41 pharmaceutical compounds in SC-CO_2. The correlating results were compared with existing solubility models elaborated by other authors and evaluated in terms of global mean ARD, SSE, RMSE, R^2, and R^2_{adji} The proposed model is found to correlate better than existing models.

Vatani et al., (2019) [167] proposed to estimate solute solubility in SC-CO_2 using an empirical equation The model was trained with a dataset of 3,260 experimental data points at different temperatures and pressure. The accuracy of the proposed model was compared with other empirical equations where the AARD was found to be 9.27% which was better than the other models and reasonable accuracy.

Shirazian et al., (2020) [168] investigated the possibility of correlating and extrapolating the measured solubility data of tamoxifen, an anti-cancer drug, in SC-CO_2 as a function of temperature (308-338 K) and pressure (120-400 bar) using five well-known semi-empirical density-based models including Kumar-Johnston (K-J), Garlapati-Madras, Chrastil, Mendez-Santiago and Teja (MST), and Bartle et al., among which K-J model showed the lowest AARD percent of about 8.26%.

Amani et al., (2022) [168] applied well-known empirical models (Adachi and Lu, Ch and Madras, Hozahzbr et al., Bian et al., Mendez-Santiago-Teja) for the prediction of the solubility of six anticancer drugs (Aprepitant, 5-Fluorouracil, Imatinib mesylate, Capecitabine, Letrozole, Docetaxel) in SC-CO_2. The accuracy of these models was evaluated based on various statistical criteria, such as the average absolute relative deviation (AARD%), adjusted correlation coefficient (R_{adj}), and F-value. The best model is Adachi and Lu model with AARD% value of 12.12% and R_{adj} value of 0.97 provided acceptable results for solubility of mentioned drugs in SC-CO_2.

Majrashi et al., (2022) [128] predicted the solubility of Chlorothiazide, an antihypertensive drug, in SC-CO_2 over a range of 308–338K and 130–290 bar) with the use of five empirical models (Chrastil, MST, K-J, Bartle, and Garlapati-Madras). The correlations showed satisfactory accuracy where the K-J model presents the lowest AARD value (3.15%).

Honarvar et al., (2022) [127] used various modeling techniques for the prediction of the solubility of sildenafil citrate in SC-CO_2 at different temperatures (308 to 338 K) and pressure values (12 to 30 MPa). The AARD of the applied techniques was found as follows: Modified Wilson's model an AARD of 7.80, Chrastil an AARD of 7.90), Bartle *et al*. an AARD of 8.75, M´endez-Santiago-Teja an AARD of 6.50, and Kumar-Johnston AARD of 5.30.

Sodeifian et al., (2022) [123] correlated the solubility of metoclopramide hydrochloride (MCP) in SC-CO_2 using six semi-empirical models (Chrastil, Me´ndez-Santiago and Teja, Kumar and Johnston, Bartle, Sodeifian, and Gordillo) as a function of the pressure and the temperature. The results indicated that the Gordillo is the optimal model with an AARD of 11.41%.

Sodeifian et al., (2023) [129] correlated the solubility of Fexofenadine hydrochloride, an antihistamine drug in SC-CO_2 over a range of different pressures (120–270 bar) and temperatures (308.2–338.2 K) with the use of nine density-based correlations with 3–6 adjustable parameters. The most accurate model was the Garlapati and Madras with an AARD of 5.65%.

Ardestani et al., (2023) [130] used the density-based models of Bartle *et al.*, M´endez-Santiago and Teja (MST), Jouyban *et al.*, and Hozhabr *et al.* to predict the solubility of famotidine in SC-CO_2 at different temperatures (308-338 K) and pressures (12-30 MPa) ranges. The model yielded satisfactory results.

Sodeifian et al., (2023) [133] carried out two different approaches, including PR-EOS and a set of semi-empirical models for the prediction of the solubility of buprenorphine hydrochloride, a strong pain reliever, in SC-CO_2. The results of the statistical criteria indicated that the Jouyban and Sodeifian models are better options for correlating solubility data.

Bazaei et al., (2023) [169] calculated the solubility of an anti-cancer drug, Erlotinib hydrochloride, in SC-CO_2 (308-338 K and 12-30 MPa) using PR-EOS and four density-based semi-empirical models including Chrastil, Bartle *et al.*, M´endez-Santiago, Teja, Kumar-Johnston, and modified Wilson model. The results showed that the modified Wilson model exhibited a higher correlation with empirical solubility data in SC-CO_2 (AARD=10.45 %, R_{adj}=0.9876).

Ansari et al., (2023) [170] applied several semi-empirical models (Chrastil, Bartle et al., Kumar & Johnston, Menden-Santiago & Teja, Sodeifian et al., and Jouyban et al.) as well as the modified Wilson model for calculating the solubility of Aripiprazole in SC-CO_2 at various temperatures (308–338 K) and pressures (12–30 MPa). The R^2 score of the Chrastil (0.994), Jouyban et al. (0.993) and Sodeifian et al. (0.992) models showed their highest consistency with the obtained values.

Sodeifian et al., (2023) [131] predicted the solubility of palbociclib in SC-CO_2 at different equilibrium conditions (308-338 K, and 12–27 MPa) using ten semi-empirical equations. Bian's model indicated the best accuracy results.

Sodeifian et al., (2023) [171] calculated the solubility of hydroxychloroquine sulfate in SC-CO_2 at various conditions (P = 12 to 27 MPa and T = 308 to 338 K) using various models (Chrastil, reformulated Chrastil, Méndez-Santiago and Teja (MST), Bartle et al., Reddy-Garlapati, Sodeifian et al., models) and new set of solvate complex models were considered. Among the all models investigated, Reddy-Garlapati and new solvate complex models fitted the best the data with the smallest error.

Nateghi et al., (2023) [132] predicted the solubility of Nilotinib hydrochloride monohydrate, an anti-cancer drug, in SC-CO_2 at various temperatures (308–338 K) and pressures (120–270 bar) conditions using expanded liquid theory (modified Wilson model). The final model presents an *AARD* of 11.28%.

Nateghi et al., (2023) [132] predicted the solubility of Nilotinib hydrochloride monohydrate, an anti-cancer drug, in SC-CO_2 at various temperatures (308–338 K) and pressures (120–270 bar) conditions using ten empirical and semi-empirical models with three

to six adjustable parameters, including Chrastil, Bartle, Sparks, Sodeifian, Mendez-Santiago and Teja (MST), Bian, Jouyban, Garlapati-Madras, Gordillo, and Jafari-Nejad. The final results indicated that all the models showed an acceptable agreement with the experimental data, among them, the Bian model exhibited excellent performance with an *AARD%* of 8.11.

Sodeifian et al., (2018) [117] predicted the solubility of Loratadine in SC-CO_2 at temperatures ranging from 308.15 to 338.15 K and pressures from 120 to 270 bar using seven density-based semi-empirical models, namely Chrastil, Sparks *et al*., Garlapati and Madras, Bartle *et al*., Méndez-Santiago and Teja (MST), Jouyban *et al*. and Bian *et al*. models. The results were found to be satisfying.

Alshahrani et al., (2023) [172] applied eight (8) empirical models via, Kumar- Johnston (K-J), Sung and Shim, Adachi and Lu, Sparks, Bartle, Hozhabr, Khansary, and Jouyban for correlating the solubility of Metoprolol (a beta-blocker agent, widely used to treat hypertension, and some heart diseases) in SC-CO_2 over a pressure and temperature range of 120-300 bar and 308-338K. As a result, it was found that the empirical model proposed by Adachi and Lu has the correlation precision with the lowest AARD% of 13.07.

Venkatesan et al,. (2023) [134] correlated the solubility of Metformin, an anti-hyperglycemic, through eight well-known empirical correlations (Chrastil, Sung and Shim, Adachi and Lu, Si-Moussa, Bartle, Jafari Nejad, Keshmiri, and Jouyban). The temperature, pressures, and mole solubility fraction range from 308-338K, 140-290 bar, and 0.39-1.23 * 10^{-6}, respectively. The overall results showed that the models defined by Jafari Nejad (AARD% of 4.36 and AICc of - 997.07) and Keshmiri (AARD% of 4.73 and AICc of -1048.81) have the highest correlation accuracy.

I.2.3. Literature Review on Solution Models

Chen et al., (2002) [173] employed a solution model approach coupled with the Flory–Huggins equation to predict the solid solubility of biological (steroids, antioxidants, xanthines, drugs, and heavy aromatic) compounds in SC-CO_2 was calculated using the solution model approach. The molar volume of the solute in the supercritical phase was optimally fitted for each solid compound using the experimental solubility data. Satisfactory results were obtained which were also compared to various EOS and semi-empirical equations with more adjustable parameters. The parameters in the solution model showed a regular correlation with the pure component property of the complex solutes.

Su et al., (2007) [174] predicted the solid solubilities of pharmaceutical compounds, including steroids, antioxidants, antibiotics, analgesics, and specific functional drugs, in SC-CO_2 using a regular solution model with the Flory–Huggins equation. Molar volumes of the solids were considered empirical parameters. The logarithms of these parameters were then correlated as a linear function of the logarithms of the densities for SC-CO_2. The results of this equation were found comparable to those from the commonly used semi-empirical equations with more adjustable parameters. The parameters of this model were further generalized as a function of the properties of the pharmaceutical compounds. It was found that the prediction results were within acceptable accuracy for more than 50% of the systems investigated.

Su and Chen (2007) [140] estimated the solubilities of ten NSAIDs in SC-CO_2 by applying the regular solution model coupled with a Flory–Huggins term. The solution model, which has fewer parameters than the semi-empirical equations, yielded comparable correlation results. The parameters in the solution model could be generalized for the specific group of NSAIDs. Finally, the predicted solubilities were demonstrated to be reliable.

Su et al., (2013) [175] predicted the solid solubility data of five fatty acids in SC-CO_2 at different temperatures and pressures using a two-parameter solution model developed from the regular solution model coupled with the Flory–Huggins equation. The developed solution model with fewer parameters yielded correlated results comparable to those of commonly used semi-empirical equations. The solution model parameters can be further generalized based on the fatty acid chain length. According to the results, the proposed predictive solution model provided better-predicted capacities and produced an average deviation in predicted solubilities of 22.1%. In addition, the authors also correlated the solid solubility data of three triglycerides in SC-CO_2 at 313 K. The predictive solution model for triglycerides yields an average deviation in predicted solubilities of 29.8%. These results demonstrated that the solution model used is applicable for the correlation and prediction of solid solubilities of structure-related compounds in supercritical CO_2.

Huang et al., (2013) [176] proposed a regular solution model coupled with the Flory–Huggins equation to model the solid solubility data of 15 pharmaceutical compounds in SC-CO_2 at different temperatures and pressures.The molar volume of the solid solute was considered as a variable composed of two parameters for the model. This volume was then transformed into a reduced form by introducing the critical density of CO_2 and was even simplified to involve only a single parameter. The results of the proposed technique were compared to the correlation results of commonly used semi-empirical models where it was

found that the presented approach was feasible and the results were satisfactory for these complex pharmaceutical compounds. In addition, the only parameter included in the present model can be expressed in terms of the ratio of the molecular weight of the compound to that of CO2, which makes the model more general and allowed it to be applied for prediction.

Madras et al., (2017) [177] correlated the solubility data of ricinoleic acid, methyl ricinoleate, and methyl 10-undecenoate rmined in SC-CO_2, in the range of temperatures of 313–333 K and pressures of 10–18 MPa, using the solution theory and regular solution theory for the activity coefficient. The solubility data were fitted using this model and the AARD for each compound was obtained 1.66%, 5.30%, and 6.43% for ricinoleic acid, methyl ricinoleate, and methyl 10-undecenoate, respectively. These AARD values were comparable to the other three-parameter solution theory models reported in the literature indicating the utility of the developed model.

Sodeifian et al., (2018) [30] predicted the solubility of six pharmaceutical compounds (Anti-HIV, Antiinflammatory, and Anti-cancer) in SC-CO_2 using a regular solution model coupled with the Flory-Huggins equation. The capacity of the approach was evaluated based on several statistical criteria, including AARD, $R_{adj,}$ and F-value.

Sodeifian et al., (2018) [117] predicted the solubility of Loratadine in SC-CO_2 at temperatures ranging from 308.15 to 338.15 K and pressures from 120 to 270 bar using regular solution models. The results were found to be satisfying.

Edrisi and Bakhshi (2023) [135] carried out regular solution theory tuned by the Flory–Huggins equation to model the solubility of two drugs, namely ibuprofen and salicylic acid, in SC-CO_2 at various temperatures (308–333 K) and pressures up to 35 MPa. The Adjustable parameters of the proposed model were regressed and reported at the studied temperatures. The results proved that this method cannot correlate the solubility of the pharmaceutical in SC-CO_2 as well as other models that were compared with (SRK-EOS, CPA-EOS, and qCPA-EOS).

I.2.4. Literature Review on Group Contibution Techniques

Cabral et al., (2001) [178] estimated the solubility of fat compounds (fatty acids, fatty acid esters, and triglycerides) in SC-CO_2 based on the idea of group contribution. This method takes only the knowledge of the chemical structure of a substance to predict this property. Chrastil´s equation was then applied to correlate the solubilities in SC-CO_2. The three-parameters determined from the fitting of the data were calculated. A nonlinear fit routine was

used to determine the chemical group contributions to the three parameters. According to the results, it is possible to correlate the parameters of Chrastil´s equation to molecular groups and The AARD was 8%.

Melo et al., (2004) [179] predicted the solubility of l-Dopa (a drug commonly used for the treatment of Parkinson's disease) in SC-CO_2 using a PR-EOS in a fully predictive way with UNIFAC group contribution activity coefficient model incorporated in its mixing rules using the linear combination of Vidal and Michelsen (LCVM) procedure. Experimental solubility data of caffeic acid in SC–CO_2 was used as the reference basis for this methodology and the results for l-Dopa solubility suggest further experimental studies.

Tomberli et al., (2005) [180] compared the ability of three diffrent group contribution methods in estimating the critical parameters and the Pitzer factor and used the results in correlative predictions of virial coefficients. These virial coefficients were then used to determine the chemical potential of a solute in SC-CO_2 based solely on the solute structure. The group contribution method of Marrero and Gani was shown to best reproduce experimental solubility data with an AARD of 10.4%. The accuracy obtained for several common organic pollutants (naphthalene, benzoic acid, anthracene, phenanthrene, naphthol-2, and pyrene) suggested that the method was applicable for predictions on previously unmeasured systems as long as the molecule has fewer than 20 non-hydrogenic atoms and is relatively rigid.

Del Valle et al., (2013) [181] suggested an empirical model for predicting the solubility of pharmaceuticals in SC-CO_2. The model was performed by using a modified enhancement factor and a group contribution method. According to the results, polar solids with similar sublimation enthalpies, and at similar experimental conditions, showed similar solubility in terms of the proposed modified enhancement factor. Several equations to estimate that modified enhancement factor (depending on the experimental conditions and the sublimation enthalpies) with a deviation less than 5% were determined. These equations can be used as a tool to identify if experimental data follows the common solubility tendency. However, they cannot be used with other types of solids, such as inorganic compounds.

Pazuki and Anvari (2014) [182] applied a version of the SAFT-VR-EOS based on the Yukawa potential for predicting the solubility of diverse organic compounds, mostly drugs, in SC-CO_2. SAFT parameters of pure CO_2 were calculated with the vapor-liquid equilibrium data within the range of 220–300 K. Those parameters were validated by applying them in the prediction of the compressibility factor and density in wide temperature and pressure ranges. The sublimation pressure was obtained by application of a group contribution method prior to

correlating the solubility data, and calculating it in relation to temperature while correlating the solubility data. For each compound, the first isothermal set of the available experimental data was used for correlation and the rest were predicted. Then the superiority and high capability of the proposed method were validated by comparing its results with three empirical models.

Jahromi and Roosta (2018) [183] carried out four group contribution methods to calculate the critical point and acentric factor of eight pharmaceuticals. The calculated properties were evaluated in predicting the vapor pressure and enthalpy of sublimation of the pharmaceuticals by using the PR-EOS. The most accurate critical points were used to estimate the solubility of the pharmaceuticals in the SC-CO_2 by using the PR equation combined with two mixing rules. The AARD of 7.28% indicated the good performance of the Wong Sander mixing rule.

Cai and (2019) [184] employed a PSRK-EOS to estimate the solubility of solid solute solubility in SC-CO_2. The first-principles thermodynamic model PR+COSMOSAC+EOS and four well-known group contribution models were used to estimate critical properties and acentric factors for all solid solutes to overcome the lack of experimental critical properties for most investigated solid solutes. The results confirmed that combining PSRK-EOS with pure substance properties from a certain group contribution model gave the lowest overall deviations, while results obtained from the first-principles thermodynamic model provides greater accuracy for solutes with complicated molecular structures. In general, the PSRK-EOS is able to provided useful solubility information, especially when no experimental data is available.

Mejía et al., (2020) [185] applied the cubic PR and two versions of the SAFT (-HR and -VR Mie) EOS to study the solubility of six vitamin K3 derivatives in SC-CO_2. Group contribution methods were used to compute the thermophysical properties of the six studied molecules so that they will be used in the computation of solubility with the EOS parameters. It was noticed that all the equations of state describe acceptably the target property, but the SAFT-VR Mie model presents the best performance.

Bezaze and Meniai (2022) [186] predicted the solubility of four solid pharmaceutical compounds, benzamide, cetirizine, metaxalone, and niflumic acid in SC-CO_2 at different temperatures and pressures ranging from 100 to 400 bar by using the most appropriate Equation of State (EOS)-Mixing Rule (MR) combination. The group contribution as a function of temperature directly was used to determine the binary interaction parameter k_{ij}. This technique was inserted in the interaction parameters matrix using an optimization technique based on the genetic algorithm. The thermodynamic modeling results using the different EOS and mixing

rules were in good accordance with the experimental data for the studied component conditions. The model provided a better correlation than conventional models that use the Predictive PR-State Equation with the VDW MR mixing rule.

Loreno et al., (2023) [187] used the GC-sPC-SAFT (Group Contribution Simplified Perturbed Chain Statistical Associating Fluid Theory) EOS) to compare two techniques to predict thermophysical properties of ILs ([Cn-mim][BF4], [Cn-mim][PF6], and [Cn-mim][NTf2]), and to correlate the solubility of CO_2 in these ILs. The first strategy which used atmospheric pressure density, speed of sound, and heat capacity as input data (the proposed strategy of Loreno et al. (2019) [188]). The second strategy used high pressure density as input data. The model parameters were obtained with the use of 2B association scheme. The experimental data was better predicted using the 1st strategy.

I.2.5. Literature Review on Machine Learning Approaches

Abbas Karim and Mutlag (2010) [189] correlated the solubility of pharmaceutical compounds in SC-CO_2 using an ANN model. The results were compared to a PR-EOS approach where the ANN technique gave excellent agreement with the experimental data of the systems taken and the percentage of AARD ranges from 5.668×10-5 % to 2.089×10-3 % meanwhile the model when using PR-EOS gave good agreement with (%AARD) ranges from 3.34% to 8.67%.

Mehdizadeh and Movagharnejad (2011) [190] presented a genetic algorithm based least square support vector machine approach to calculate the solubility of 25 different solutes in SC-CO_2. The model consisted of three inputs including temperature, pressure, and density of SC-CO_2, a single output (solubility). The prediction results were compared with seven well-known semi-empirical correlations where the present technique showed an AARD of about 4.92% while the best semi empirical-equation has an AARD of about 13.60%.

Mehdizadeh and Movagharnejad (2011) [191] suggested a ANN model to estimate the solubility of 30 different compounds in SC-CO_2. The model accuracy was compared with seven semi-empirical equations. The results confirmed the ability of the ANN model in predicting with an AARD (5.3%) less than the best semi-empirical equation (15.96%).

Mohammadi et al., (2011) [192] calculated the solubilities of 21 of the commonly used industrial solid compounds in SC-CO_2, accounting for 795 data points, using an optimized three-layer FF-ANN algorithm with the presence of critical properties of the solute and

operating temperature and pressure. The results were an R^2 of 0.9533 and an AARD of about 14% from the experimental values.

Üstündağ et al., (2012) [193] investigated the application of a feed-forward artificial neural networks consisting of one hidden layer for modeling the solubility of binary mixtures of lipids and SC-CO_2. The database used contained the solubility data of minor and major solid lipid compounds (lauric acid, myristic acid, palmitic acid, stearic acid, behenic acid, trilaurin, trimyristin, tripalmitin, stigmasterol, and β-carotene) in SC-CO_2 containing 437 data points, and used the temperature, pressure, and solute properties (acentric factor, critical pressure, critical temperature) as inputs. The network was trained using the Bayesian regularization combined with the Levenberg-Marquardt backpropagation algorithm and used in order to improve network generalization and estimate the effective number of network parameters (weights and biases). The average cross-validation values of AARD and R^2 were 5.78% and 0.9681, respectively. These results confirmed the predictive power of ANN for the modeling of solubility in supercritical carbon dioxide.

Hamid et al., [194] predicted the solubilities of two natural compounds found in mangosteen peel (squalene and α-cubebene); the compounds exhibit antimicrobial, antibacterial, and anticancer activity for cancer disease treatment, in SC-CO_2 using a multi-layer feedforward back-propagation ANN model (48 solubility data). The input variables were temperature and pressure. The optimal ANN model consisted of one hidden layer and five neurons was obtained with a minimum value of MSE (0.011), while the values of R for training, validating, and testing obtained were of 0.933, 0.982, and 0.927 proceedings.

Vaferi et al., (2013) [195] carried out an FF-ANN model to estimate the solubilities of solid aromatic compounds in SC-CO_2 as a function of the temperature, pressure, critical properties, and acentric factors of each solute. The backpropagation learning algorithm was applied to adjust the parameters of the multi-layer perceptron (MLP) network. After various training and examinations, a single hidden layer network composed of nineteen hidden neurons was selected as an optimal topology. The capability of the proposed ANN model was compared with those published results which have obtained by SAFT combined with eight different mixing rules (one, two, and three parameters mixing rules) and PRSV equation of state (EOS). The best accurate results were obtained by the proposed MLP model with an AARD% = 4.99, MSE = 7.08×10-7 and R^2 = 0.99699 (overall AARD% for SAFT approach with one, two, and three parameters mixing rules were 16.15, 12.32% and 7.65%, respectively while PRSV-EOS showed AARD% of 21.10%).

Tatar et al., (2016) [196] developed an Adaptive Neuro-Fuzzy Inference System (ANFIS) for modeling the solubility of several solid species in SC-CO_2, a total of 795 data points. The Genetic Algorithm (GA) was implemented to optimize the radius of influence of initial Fuzzy Inference System (FIS). A technique based on the conjugation of hybrid and Particle Swarm Optimization (PSO) algorithm namely, CHPSO was used to optimize the tuning parameters of the FIS structure. Results confirmed that the developed conjugate ANFIS model was robust and precise in the prediction of experimental data and is superior to other literature models.

Lashkenari and KhazaiePoul (2016) [197] proposed a *K*-nearest neighbor (KNN) as a nonparametric and easy learning model to predict the solubility of 11 polycyclic aromatic hydrocarbons in SC-CO_2. The model was constructed with optimum values of K for different compounds and then compared with several well-known semi-empirical models (Chrastil, Adachi and Lu, Li, Garlapati and Madras, Sparks and Bian). The results showed that the KNN model has an AARD of about 12.47% for 11 polycyclic aromatic hydrocarbons modified density-based correlations (Sparks and Bian) while the best semi-empirical equation (Sparks) results an AARD of about 11.50%.

El hadj et al., (2017) [198] carried out a nonlinear hybrid neural network with a particle swarm optimization (PSO) to build a model that can predict with precision the solubility of binary systems (solid drugs_SC-CO_2). The obtained model consisted of five neurons in the input and hidden layers, and one neuron in the output layer (5-5-1). Statistical analysis of the predictability of the optimized neural network model trained with trainpso algorithm showed excellent agreement with experimental data. Furthermore, the comparison in terms of AARD% for each binary showed that the ANN–PSO model correlated far better with the solubility of the solid drugs in SC-CO_2. A control strategy was also developed by using the inverse ANN method. The sensitivity analysis showed that all studied inputs have a strong effect on the solubility and allowed the estimation of some solid properties from the solubility data with good accuracy without the need for the group contribution methods.

Baghban et al., (2017) [199] applied a supervised technique, namely, the least square support vector machine (LSSVM) to estimate the solubility of 33 different drug compounds in SC-CO_2. The corresponding output was calculated as a function of temperature, pressure, density of SC-CO_2, and two physical properties of the drugs, i.e. molecular weight and melting point. The employed LSSVM model was compared with eight correlations. The obtained results were an AARD of 5.61%, and a square regression coefficient of 0.9975 for the testing set.

Therefore, the model developed in this work can be reliably applied to the studied drugs by only knowing their physical properties.

Aminian el al., (2017) [200] proposed four neural network-based models for estimating fifteen solutes' solubility in SC-CO_2 at different temperatures and pressures. The generalize-ability of each network was tested and then compared to the results from different semi-empirical models, namely Chrastil, Kumar and Johnston, Bartle, Mendez-Santiago and Teja, and two EOS, namely the Peng-Robinson, and Soave-Redlich-Kwong. The average absolute deviation (AAD) of the best proposed ANN was found to be 5.42%, which was found to be a reasonable accuracy compared to those of density-based and EOS models.

Bian et al., (2017) [31] carried out a grey wolf optimizer-based support vector machine (GWO-SVM) for correlating 18 solute solubility (1148 data points including 814 training data points and 334 testing data points) in SC-CO_2 as a function of the temperature, pressure, and the density of SC-CO_2. The model successfully correlated this property better than 27 commonly used empirical models with the lowest AARD (3.20%). It was also found that the overall AARD is less dependent on material type for the proposed GWO-SVM model.

Baghban et al., (2018) [201] employed a least-square support vector machine (LSSVM) approach in order to predict 33 different drug solubilities in SC-CO_2. The property was estimated as a function of the temperature and pressure, SC-CO_2 density, and 20 different chemical substructures. The proposed model was compared to 8 previously reported semi-empirical correlations. The results showed satisfaction with an AARD of 4.92% and R^2 of 0.998 for the testing dataset. Therefore, the proposed LSSVM can be applied as a reliable predictive tool to estimate the drugs' solubility, if drugs' chemical structures are available.

Sodeifian et al., (2018) [30] predicted the solubility of six pharmaceutical compounds (Anti-HIV, Anti-inflammatory, and Anti-cancer) in SC-CO_2 using an artificial neural network-based (ANN-based) model. The capacity of the approach was evaluated based on several statistical criteria, including AARD, R_{adj}, and F-value. Results of the analysis of variance (ANOVA) indicated that the ANN-based model provided the best results in terms of correlating the experimental solubility of the pharmaceutical compounds in SC-CO_2 compared to three different approaches.

Vatani et al., (2019) [167] proposed to estimate solute solubility in SC-CO_2 using a fuzzy-genetic model. The model was trained with a dataset of 3,260 experimental data points at different temperatures and pressure. The accuracy of the proposed model was compared with

other empirical equations where the AARD was found to be 7.74% which was better than the other models and reasonable accuracy.

Abdi-Khanghah et al. (2018) [202] investigated in the performance of modeling the solubility of alkanes in SC-CO_2 using two types of artificial neural networks, i.e., Radial Basis Function (RBF) and Networks Multi-layer Perceptron (MLP). The final results showed that the MLP-ANN has better performance than RBF-ANN in the prediction.

Li et al., (2020) [203] developed an improved double-population chaotic accelerated particle swarm optimization (APSO) algorithm combined diffusion theory based on the particle evolution utilizing diffusion energy. This technique was applied in the training of parameters of the radial basis function ANN and then a model for predicting the solubility of 8 polymers in SC-CO_2 was built. The results indicated high prediction accuracy, an ARD of 0.0036, an R^2 of 0.9970, and an RMSE of 0.0152.

Wang and Su (2021) [204] preformed various machine learning approaches to model the solubility of Lornoxicam in SC-CO_2 at different levels of temperature and pressure (as inputs). The three methods used were: Medium Gaussian SVM, Wide Neural Network, and Stepwise Linear. The computational outcomes and fitting show that the SVM algorithm and the stepwise technique precisely predicted the target property, while the approach of Wide Neural Network (WNN) over-predicted the solubility values due to the small amount of data points in the datasets in the training step of simulations. The final obtained result was that ML methods were robust techniques for such a property but requested to have larger datasets to provide a model with high accuracy.

Zhu et al., (2021) [205] implemented a machine learning method based on a Neuro-fuzzy system, i.e., ANFIS (Adaptive neuro-fuzzy inference system) for simulating the solubility of an anticancer drug (Busulfan) in SC-CO_2 (32 data points). The processing parameters, including pressure and temperature, were evaluated in the process and were taken into account in the modeling of the supercritical system. Both training and testing steps in developing the ANFIS model indicated great accuracy and were very close to the experimental values. In developing the ANFIS model, Trimf membership function was used for the generation of the FIS structure. The Fuzzy system was generated using the Grid Partitioning technique, which indicated great accuracy in the prediction of busulfan solubility data. The model developed in this work indicated that the ANFIS model can be used to accurately predict the solubility of drugs in supercritical solvents, which can be consequently used for the production of drugs with improved efficacy.

Cao et al,. (2021) [24] carried out a neural-based modeling method for the prediction of drug (Chloroquine, used for treating malaria and Coronavirus) solubility in SC-CO_2, 32 experiments. The neural network was developed considering temperature and pressure as the inputs, while the solubility was set as the output. The results of training and validation indicated that the developed neural network is robust in predicting solubility, and high accuracy with R^2 higher than 0.99 was obtained for both the training and validation set.

Ayari et al., (2022) [206] constructs a stacked/ensemble model by combining three up-to-date machine learning tools (i.e., extra tree, gradient boosting, and random forest) to predict the solubility of twelve anticancer drugs in SC-CO_2 (311 phase equilibrium samples), as a function of solute and solvent properties and operating conditions. The statistical metrics indicate, an AARD of 8.62%, an MAE of 2.86E-6, an RAE of 2.42%, an MSE of 1.26E-10, and an R^2 of 0.99809. The results confirmed that the suggested stacked model demonstrates excellent performance for correlating and predicting the output.

An et al., (2022) [207] simulated the pharmaceutical compound (Fenoprofen) solubilities in SC-CO_2 on the basis of machine learning techniques. The pressure and temperature were the inputs of the system, whereas the response was the solubility in mole fraction, a dataset of 32 data points. The applied models were CNN, DNN, and GRNN where the hyper-parameters of these methods were fine-tuned with the Bat algorithm. All the three models presented a R^2 higher than 0.9 while the RMSE of 7.57E-05, 7.25E-05, and 3.94E-05 for DNN, CNN, and GRNN, respectively. The results confirmed the over-performance of GRNN and the significant impact of the Bat algorithm in improving the accuracy of the employed models.

Amani et al., (2022) [168] applied a well-known machine learning algorithm namely, ANN for prediction of the solubility of six anticancer drugs (Aprepitant, 5-Fluorouracil, Imatinib mesylate, Capecitabine, Letrozole, Docetaxel) in SC-CO_2. The final results were an AARD% value of 1.65% and a R_{adj} value of 0.9960 where the model was found to be better than six known empirical models, i.e., the most appropriate model for correlation of drugs solubility data.

Liu et al., (2022) [208] utilized three different decision tree-based ensemble models, namely, Random forest (RF), Gradient Boosting Trees (GBRT), and Extremely Randomized Tree (ERT) to predict the solubility of busulfan as an anticancer drug in SC-CO_2. The model takes two inputs temperature and pressure. The overall statistic evaluation showed that all models have an R^2 score of more than 0.9, RMSE error rates of 1.80E-04, 1.72E-04, and 1.03E-

04, and a MAPE metrics 4.51E-01, 4.87E-01, and 3.62E-01 for RF, ERT, and GBRT models, respectively. The results indicated that the GBRT was the best predictive model.

Nguyen et al., (2022) [209] predicted the solubility of two APIs, namely, salsalate and decitabine in SO-CO_2 using an artificial neural network having one hidden layer, two inputs (temperature and pressure), and one output (solubility), and 32 experimental data points. The model was developed and tested at various pressures (120 to 400 bars) and temperatures (308 to 338K). the model's accuracy was found to be better than previous correlations. The results of the training and validation sets indicated great accuracy with R^2 higher than 0.99 for both drugs.

Kostyrin et al., (2022) [210] implemented an ANN model for predicting the solubility of Oxaprozin in SC-CO_2 at different operational conditions by variation of temperature and pressure of the process. The best model was obtained with R^2 greater than 0.99 for the both training and validation sets. The developed ANN model was indicated to be robust in predicting the values of drug solubility in a wide range of conditions, which can save time, and cost of the measurements.

Alzhrani et al., (2022) [11] used three different artificial intelligence algorithms; Kernel Ridge Regression (KRR), least angle regression (LAR), and Multilayer Perceptron (MLP), for predicting the solubility of Lenalidomide, an anti-cancer agent, inside the SC-CO_2 in a wide range of pressure and temperature. The available data includes 28 rows of data with two inputs, including temperature and pressure, and two outputs, density, and solubility. The final results selected MLP as the primary model of that research. The models illustrated R^2 scores of 0.999 and 0.994 for density and solubility while the Max error was 2.92E+00 and 6.44E-02 for these outputs, which shows the accuracy and significant generality of the model.

Zhao et al., (2022) [211] carried out two support vector-based models namely, linear SVM and Nu-SVM to correlate the solubility of Busulfan in SC-CO_2 over a wide pressure range of 120–400 bar and at various temperatures. The pressure and temperature serve as the model's inputs. The Nu-SVM model showed an R^2 score of 0.993, an MAE of 1.4453E-05, and an MAPE of 1.23755E-01while the linear SVM had an R^2 score of 0.856, an MAE of 7.3430E-05, and an MAPE of 5.12260E-01.

Byabani-Givo et al., (2022) [212] developed a radial basis function neural network (RBF-NN) to correlate the solubility of some 1,4-dihydropyridine derivative drugs in SC-CO_2 based on pressure, temperature, molecular weight, melting point, density, carbon number, and hydrogen numberThe results were statistically validated with various statistics where the R^2

was obtained between 0.981 and 0.99 for each and overall compound and the model developed can be reliably used to investigate the solubility of drugs by known physical properties.

Rezaei et al., (2022) [213] performed different machine learning scenarios to simulate the solubility of twelve non-steroidal anti-inflammatory drugs (NSAIDs) in the SC-CO_2 as a function of Physical characteristics of the drugs (molecular weight and melting temperature), operating conditions (pressure and temperature), and solvent properties (SC-CO_2 density). A database of 254 laboratory-measured solubility data. The considered NSAIDs are Fenoprofen, Flurbiprofen, Ibuprofen, Ketoprofen, Loxoprofen, Nabumetone, Naproxen, Nimesulide, Phenylbutazone, Piroxicam, Salicylamide, and Tolmetin. The authors compared twelve intelligent models from three categories (artificial neural networks, support vector regression, and hybrid neuro-fuzzy) in terms of prediction accuracies and approved that ANFIS is the best tool for the proposed task. The hybrid optimization strategy adjusts the cluster radius of the subtractive clustering membership function to 0.6111. This model estimated with the AAPRE = 3.13%, MSE = 2.58 × 10-9, and R^2 = 0.99919. In addition, the proposed hybrid approach was found to be more reliable than the EOS and available correlations of the literature.

Tianhao et al., (2022) [214] applied three common techniques including, Multi-layer Perceptron (MLP), decision tree, and random forest on a database of two inputs (T, and P), one output, and 32 rows of experimental data for solubility of busulfan in SC-CO_2. The MSE error of the optimized models was found to be 1.77E-05, 6.72E-05, and 3.28E-05 for RF, MLP and DT, respectively. The results obtained were good and the models can be used as reliable methods in the pharmaceutical area.

Liu et al., (2022) [215] developed three computational models for the estimation of drug (Salsalate) solubility in SC-CO_2. The dataset contains two P and T inputs, one output, and 32 data points. The models used were three Gaussian Process Regression (GPR) models: simple (raw) GPR, Ada-boosted GPR, and Bagged GPR. The statistic metrics of the three models were as follows: the R^2 of 0.9779, 09992, and 0.9795, the RMSE error ratings of 1.25E-04, 1.20E-04, and 1.29E-04, for the GPR, Adaboost + GPR, and Bagging + GPR, respectively. Furthermore, the MAPE of 7.2389E-02, 7.2831E-02, and 7.329E-02, respectively. Finally, the boosted GPR model was chosen as the best accurate model.

Abdelbasset et al., (2022) [216] estimated the solubility of a drug model, namely salsalate in SC-CO_2 at different values of temperature and pressure using several machine learning to include, Linear SVR (Support Vector Regression), Nu SVR, and Bayesian Ridge Regression (BRR) models with two inputs (T and P). The simulation results indicated that linear

SVR, Nu SVR, and BRR have R^2 scores of 0.784, 0.998, and 0.876, an exhibit MAE error rates of 3.71E-04, 8.41E-05, and 3.04E-04, and an RMSE of 4.21E-04, 1.39E-04, and 3.60E-04, respectively. Indeed, the model of Nu SVR was chosen as the best model based on these statistical metrics and some visual examination.

Li et al., (2022) [217] carried out Tree ensemble models, including Extremely randomized Trees (Extra Tree), Adaptive Boosting (AdaBoost), and Gradient Boosting models for modeling the solubility of Chlorothiazide, an antihypertensive, in SC-CO_2 as a function of two input features, namely temperature and pressure, and one output, namely solubility. The hyperparameters of the models were optimized using GA. According to the evaluations, the gradient boosting model showed the best results (R^2=0.9820, MAE=1.51E-02, RMSE= 2.51E-02, and MAPE=1.59E-02).

Namazi et al., (2022) [218] estimated Decitabine, chemotherapeutic agent, solubility in SC-CO_2 by employing different machine learning-based mathematical models namely, AdaBoost (Adaptive Boosting) to boost three base models including Linear Regression (LR), Decision Tree (DT), and GRNN on a database of 32 sample points, two input features: P (bar) and the T (k). ADA-DT (Adaboost Algorithm-Decision Tree), ADA-LR (Adaboost Algorithm-Linear Regression), and ADA-GRNN (Generative Regression Neural Network) models showed MAE of 6.54E-05, 4.66E-05, and 8.35E-05, an R^2 score of 0.986, 0.983, and 0.911 scores, respectively. The results indicated that the ADA-LR was selected as the primary model.

Huwaimel and Alharby (2023) [219] carried out a variety of multiple tree-based machine-learning techniques to predict the solubility of Lenalidomide in SC-CO_2. The models take the temperature and the pressure as input features. The selected models were Decision Tree (DT), Extra Trees (ET), and Gradient Boosting (GB) coupled with the SCA algorithm to obtain better robustness models. The models were abbreviated as SCA-DT, SCA-ET, and SCA-GB, and have R^2-scores of 0.932, 0.951, and 0.997, a RMSE of 0.0948, 0.0822, 0.0203, respectively. The results indicated that the SCA-GB is the best model for predicting the solubility of this drug compound in SC-CO_2.

Begum (2023) [220] applied bagging regression (BAG) to three core models, namely SVM, KNN, and linear regression (LR) optimized with the WOA meta-heuristic to estimate the solubility of the Hyoscine drug as a function of the temperature and pressure. The performance of these models was compared with R^2, RMSE, MAE, and AARD% metrics. The results confirmed that the BAG-KNN model outperformed the remaining models with an R^2 of 0.989 and a low RMSE of 0.041, while the BAG-LR model had the highest AARD% value of

7.50974. Generally, the bagging approach used by the authors proved to be effective in improving the accuracy of the core models, while the WOA optimization algorithm was successful in fine-tuning their hyperparameters.

Huwaimel and Abouzied (2023) [221] employed three machine learning models, the Multilayer Perceptron (MLP), the Kernel Ridge Regression (KRR), and the Gaussian Process Regression (GPR), fine-tuned their hyperparameters with the Bat optimization algorithm (BA) to estimate the solubility of a drug in SC-CO_2 using a small dataset. The pressure and temperature were selected as input features and solubility as the output. In the end, it was found that the GPR model was the most effective with an R^2 of 0.9889, MAPE error of 2.52E-2, RMSE of 1.96E-2, and a Max error value of 2.90E-2.

Waskita et al., (2023) [222] applied four ML approaches, including Ridge regression (RR), Huber regression (HR), Random forest (RF), and Gaussian process regression (GPR) to correlate the solubility of Azathioprine in SC-CO_2. The statistical metrics for the models were found as follows the R^2 equal to 0.974, 0.6518, 0.966, and 1.0, the RMSE error rates of 2.843 ×10-13, 7.036 ×10-12, 5.673 ×10-13, and 1.054 ×10-30, MAE metrics of 1.205 ×10-6, 2.151 ×10-6, 5.997 ×10-7 and 9.419 ×10-16 for RR, HR, RF, and GPR, respectively. AIn conclusion, it was confirmed that RR, RF, and GPR algorithms were good candidates for predicting the solubility of this anticancer drug in SC-CO_2.

Hani et al., (2023) [223] applied various ML techniques for predicting the solubility of nanomedicine preparation of solid dosage oral medications in SC-CO_2 as a function of two input parameters: the pressure (120–300 bar) and temperature (308–338 K). The ML algorithms used (MLP, *v*-SVR, and MLR) were coupled with the GWO optimizer. All models were found to have acceptable results, while MLP regression has shown the best results, with an RMSE error rate of 2.98×10^{-2} and an R^2 score of 0.9797.

Obaidullah et al., (2023) [224] presented a comparative study of three different models named: Deep Neural Network (DNN), Quantile Regression (QR), and K-Nearest Neighbors (KNN) coupled with a Fireworks Algorithm (FWA) optimizer in terms of correlating the solubility of rivaroxaban in SC-CO_2. The database contained two inputs temperature (T), pressure (P), and one output, mole fractions of rivaroxaban. The results indicated that the DNN model exhibited outstanding performance (R^2=0.99831, MAE=2.49870E-02) compared to the KNN model (R^2= 0.98029, MAE= 8.88397E-02), and the QR model (R^2= 0.87362, MAE= 1.55840E-01).

Zhang and Mahdi (2023) [225] investigated the application of an SVR algorithm with Quadratic Kernel (QSVR) for the prediction of the solubility of lornoxicam in SC-CO_2 using the temperature T(K) and pressure P(bar) as input variables. The Genetic Algorithm (GA), Tabu Search (TS), and Bayesian Hyperparameter Optimization (BHO) optimization methods were also employed to fine-tune the hyper-parameters of the QSVR model and enhance its predictive accuracy. All results of the three models, BHO-QSVR (R^2=0.96725, MAE=1.75666E-05), GA-QSVR (R^2=0.95346, MAE= 1.56725E-05), and TS-QSVR (R^2=0.95882, MAE=1.84075E-05), were statistically valid.

Li et al., (2023) [226] predicted the solubility of Chloroquine (antimalaria medication with anti-SARS-CoV activity) in SC-CO_2 with the use of various non-mechanistic based models and a small dataset of 32 laboratory data points. The implemented models were Random Forest (RF), KNN, and Extra Tree (ET). The final results based on R^2 are 0.999, 0.978, and 0.972 for ET, RF, and KNN, respectively.

Hao et al., (2023) [227] correlated the solubility of drug desoxycorticosterone acetate in SC-CO_2 by employing Decision Tree (DT), Theil-Sen (TS), and Gaussian Process Regression (GPR) core models coupled with Adaboost ensemble method and EPO for model tuning. The database used contains 45 data rows, two inputs temperature (Kelvin unit) and pressure (Megapascal unit), and one output: drug solubility in mole fraction. The best results belong to the EPA-GPR model with an R^2=0.997, RMSE=1.59E-01, and MAE=1.16E-01.

Ghazwani et al., (2023) [228] used three regression models, namely Gaussian Process Regression (GPR), Gamma Regression (GR), and K-nearest neighbor (KNN), optimized by the Directional Bat Algorithm (dBA), for modeling the solubility of phenytoin drug in SC-CO_2 as functions of temperature (T) and pressure (P). The highest accuracy was exhibited by GPR (R^2 value of 0.999 for the solubility, RMSE error of approximately 0.087, and an AARD% of 2.29%).

Ghazwani and Begum (2023) [229] predicted the solubility of phenytoin in SC-CO_2 with the use of three different models: Multilayer Perceptron (MLP), NU-SVR, and Poisson Regressor (POR) coupled with the BA algorithm as a function of T(K) and P(bar). The prediction results revealed that the MLP model provides the highest reliable forecasting with an R^2 score of 0.998, an MSE of 0.051, an MAE of 0.200, and a MAPE of 0.080.

Nateghi et al., (2023) [132] predicted the solubility of Nilotinib hydrochloride monohydrate, an anti-cancer drug, in SC-CO_2 at various temperatures (308–338 K) and pressures (120–270 bar) conditions using ML algorithms (random forest, decision trees,

multilayer perceptron, and deep neural network. The final models had a respective R^2 values of 0.9933, 0.9799, 0.9724 and 0.9701.

I.2.6. Literature Review on Other Techniques

Agrawal et al., (1999) [230] used an Isothermal-isobaric Monte Carlo (NPT-MC) simulations with the Widom test particle technique to predict the solubility of the explosive 1,3,5-trinitro-1,3,5-s-triazine (RDX) in SC-CO_2 over a temperature ranging from 308K to 353K and a pressures range of 10.4 to 48.3 MPA. The interaction between RDX and CO_2 was described a Lennard–Jones potential energy function. The results obtained were in good agreement with the experimental values.

Garcı´a et al., (2000) [231] applied a Response Surface Methodology approach to correlate the solubility data of solids (2,6-dimethylnaphthalene, 2,3-dimethylnaphthalene, phenanthrene, and benzoic acid) in SC-CO_2 directly to the pressure and temperature. Data points were correlated by means of quadratic models, with an R^2 greater than 0.99, and with an AAPE error of less than 5%. This study provided a possible guide for saving experimental effort and contributes with a simple model that could be used for industrial design purposes.

Jouyban et al., (2007) [232] used the descriptors computed by HyperChem® software along with T and P as inputs to represent the solubility of 40 drug molecules in SC-CO_2 carbon dioxide using an ANN with an architecture of 15-4-1. The average of absolute error AE(± SD) of data sets was 0.4 (±0.3) when all data points were used as training set and the solubilities were back-calculated. The AE for predicted solubilities using a trained network employing 1/3 of data points from each set was 0.4 (± 0.3) and this finding reveals that the network was well-trained using a limited number of experimental data. To provide a full predictive method, data sets were divided into two sets, and the network was trained using 20 data sets and the next 20 sets were used as prediction sets. The produced average AEs (± SD) were 1.7 (±1.1) and 1.6 (±1.5), for two sets of analyses.

Shimoyama and Iwai (2009) [233] carried out a COSMO base activity coefficient model to predict the solubilities of 16 pharmaceuticals solid solutes in SC-CO_2. This method describes that the system is composed of the surface segments on the solvent molecule and vacancy, unlike the conventional model based on the COSMO method. The density change of SCF can be represented by the change in the surface area of the vacancy. This prediction model was referred to “COSMO-vac (vacancy)” model. The averaged deviations between the logarithmic experimental and predicted results were smaller than the unity. Furthermore, the predicted

results for the solutes composed of only C, H and O atoms were better than those for the solutes, including the other atoms. The percentage of the predicted results within the order of the experimental data at the pressure over 15MPa was higher than that at the pressure below 15MPa.

Wang and Lin (2013) [234] carried out a predictive approach for 46 drug solubilities estimation in SC-CO_2 (T = 293.15-473 K, P = 8.5-50 MPa, and 1160 solubility data ranging from 10−7 to 10−2). A quantum mechanics-based, COSMO-SAC solvation model was used to obtain the temperature and composition dependence of the interaction parameters a(T,x) and b(x) of the PR-EOS approach. As a result, this technique does not necessitate the input of the experimental data of the mixture. The average logarithmic deviation (ALD-x) was found to be 0.81 (a factor of 5.3). The proposed method, the PR + COSMOSAC+EOS, was thus a useful tool for a priori prediction of solid solubility in SC-CO_2, as well as it can be used for other solvents.

Li et al., (2013) [235] proposed an optimization method, the pattern search method, to optimize binary interaction parameters for correlating solute solubilities in SC-CO_2 (10 binary systems) with cubic EOS models. The results showed great agreement with the corresponding experimental data and that the AARD values were mostly less than 10% (less than 1% in some cases). Meanwhile, a comparison between the pattern search method and the genetic algorithm reflected that the former gave better correlation performance and a more significant reduction in computing times than the latter.

Shimoyama et al., (2014) [236] suggested an EOS based on the hole theory with molecular surface charge density for modeling eighteen drug solubility in SC-CO_2. The density change of SC-CO_2 can be represented by the number of holes in the system. The molecular interaction energy parameter was calculated by applying the interactions of segments on the molecular surface, given by a quantum calculation using a conductor–like screening model. Only one parameter, the coordination number around a molecule, was fitted to the experiment. The effect of the molecular pair on the coordination number on the correlated results was investigated. It was found that the results using the fitted parameter of the solute–solute pair coordination number resulted in the modeling performance better than those of the carbon dioxide–solute coordination number. The AARD for the drug composed of C, H, and O atoms acetylsalicylic acid, benzoic acid, ferulic acid, (S)-naproxen, p-benzoquinone, propyl gallate, salicylic acid, and vanillic acid is 0.38 smaller than those for c 1 compounds including N, F, I

and S atoms, 2 amical-48, benzocaine, caffeine, carbamazepine, (±)-flurbiprofen, methimazole, 3 phenazopyridine, theobromine, theophylline and uracil, 0.59.

Li et al., (2014) [237] proposed an expanded liquid model, and a modified solution model for predicting solid solubilities in SC-CO_2. The results indicate that in the expanded liquid model, the accuracy of the solubility parameters of SC-CO_2 can affect the change of the variables; if the parameters are accurate, the model's variables are nearly temperature-independent, while inaccurate solubility parameters result in variables slightly temperature-dependent. When using the solubility parameter of the solute as a variable and calculating the solubility parameters of SC-CO_2 with an accurate technique in the expanded liquid model, means that the solubility parameter of the solute was a linear function of the ratio of the solubility parameter and molar volume of SC-CO_2. Aconrding to this linear function, the expanded liquid model was improved. In the modified model, the slope and the intercept of the linear function were used as the fitting parameters. The modified model was compared with the other two-parameter expanded liquid models and the results showed the modified model can provide better correlation results with the AARD being 13.31%. A group contribution method was developed to estimate the parameters in the modified model. Based on the modified model and the group contribution method, the solubilities of most of the solutes in SC-CO_2 can be estimated in the order of magnitude.

Li et al., (2015) [238] estimated the solubility of three fenamate drug compounds namely, niflumic acid, clofenamic acid, and tolfenamic acid in SC-CO_2 with the semi-empirical models (Teja model, Chrastil model, Bartle model, and Kumar–Johnston (K–J)), a compressed gas model combined with the reference solubilities, and the modified expanded liquid model combined with the group contribution method. The results demonstrated good consistency with the measured solubilities. The solubilities of the three medicines in SC-CO_2 were calculated with the compressed gas model. With the experimental reference solubilities known, their solubilities can be correlated and predicted satisfactorily. The AARDs are 5.0% and 11.9% respectively while using the modified expanded liquid model, the average AARD% is 9.6%. Combined with the group contribution method, this model can estimate the solubilities of those compounds in the order of magnitude.

Bian et al., (2015) [239] investigated a combined model, based on the optimization theory, to correlate 48 different compounds' solubility in SC-CO_2, a total of 1116 data points. The capability of the proposed model was evaluated by comparing its results with mostly more accurate empirical models (Ch-Madras, Garlapati-Madras, Adachi-Ju, Jouyban, Sparks, and

Bian) in calculations of the solubilities and seven other combined models. The results confirm that the Sparks-Jouyban-Bian combined model produced more accurate results (AARD=4.37%).

Shimoyama et al., (2015) [240] proposed a mixed solid phase model using an EOS based on the hole theory for solubility calculation in SC-CO_2. A quantum calculation based on a conductor-like screening model (COSMO) was applied to calculate EOS parameters from the molecular information of the twenty-three species of pharmaceutical compounds. The correlation results of the mixed solid phase model were compared with those of the conventional pure solid phase model, where it was noticed that pharmaceutical compounds with S, N, F, Cl, and I atoms using the mixed solid phase model have the best calculation performance compared to those of the pure solid phase model. The mixed solid phase model was also applied for the calculation of the CO_2 solubility in the solid phase.

Reveco-Chilla et al, (2016) [241] utilized molecular dynamics (MD) simulations to study the binary system of menadione in explicit SC-CO_2. A force field was developed for menadione and tested by comparing computed and experimental monoclinic crystal structures at 283 K and 0.01 MPa. Lattice parameters obtained from anisotropic isothermal-isobaric MD simulations agreed reasonably well with experimental values, with an *AARD*% less than 7%. A previously validated force field for SC-CO_2 was used, and simple mixing rules were used to describe cross-interactions. Canonical ensemble MD simulations were used to estimate the association number for CO_2 about menadione and the enthalpy required to form an SC-CO_2 solvate complex with menadione as a function of temperature and CO_2 density. Spatial distribution functions were computed to better understand the nature of the molecular-level interactions between menadione and SC-CO_2, as well as between associating menadione molecules.

Sodeifian et al., (2021) [242] estimated the equilibrium solubilities of sulfabenzamide compound in SC-CO_2, at various pressures and temperatures, by applying two thermodynamic models based on solid-liquid equilibrium criteria. Moreover, PR-EOS Couple with Van der Walls mixing rules (VDWmr) and a few other empirical solubility models have been explored. The models were found quite well at predicting the data.

Soares et al., (2022) [243] compared the performance of COSMO-SAC-Phi (CSP) with other predictive models from literature in terms of predicting the solubility of solids in SC-CO_2, for 2450 experimental data points with pressures ranging from 45 to 500 bar and temperatures varying from 298 to 473 K. For the calibration of the pure compound CSP model parameters,

pure compound saturation pressure, and liquid volume were used, and no binary parameters were adjusted. The average logarithm deviation in mole fraction was found to be 0.46 and 0.51 for the CSP (2021) and CSP (2019) versions, respectively. The CSP (2021) results represent an overall improvement of 27 %, 48 %, and 64 % when compared to PSRK, PRCS, and SRK-MC+vdW results, respectively. This indicates that CSP can give useful solubility estimates, especially when no experimental solubility data are available.

Faress et al., (2022) [243] suggested a universal correlation based on the modified version of the Arrhenius equation (a combination of an Arrhenius-shaped term and a departure function) to predict the solubility of anti-cancer drugs (316 solubility data) in SC-CO_2. The model takes as inputs the pressure, temperature, and CO_2 density. The reliability and prediction ability of the proposed correlation was validated and compared ten available empirical correlations in the literature. The proposed model has the lowest AARD (9.54%).

El Hadj et al., (2022) [244] used a Neuro-equation-of-state approach (NES), an approach based on the association of ANN and PC-SAFT-ESO to conceive a model for estimating the solubility of solid drugs in SC-CO_2. This technique can be divided into three main stages, first, the optimization of direct ANN for predicting solids_SC-CO_2 phase equilibria (15 binary systems), then the ANN inverse was performed to be an alternative to group contribution methods (GCMs) for estimating the pure components and physical properties and enhance the PC-SAFT-EOS to estimate phase equilibria parameters and finally, ANN-PC-SAFT approach is used to estimate the solubility of 213 solid solutes in SC-CO_2. The proposed method was found successfully applied to the phase equilibria modeling the data with high accuracy (AARD showed a variation from 2 to 6%) and allowed to enhance the phase equilibria modeling by reducing the number of optimized parameters and surpassing the main drawbacks faced in this area mainly the non-availability of physical properties and EOS pure component properties and the limitation of the equation of state.

Abadian et al., (2022) [245] predicted the solubility of Riluzole, a neuroprotective agent, in SC-CO_2 at different temperatures and pressures with the use of three thermodynamic models of (1) eight density-based correlations, (2) two EOSs, and (3) COSMO-Vacancy (Vac) activity coefficient. The results of the empirical models, Jouyban and Bian, show the best correlation. The combined SRK-EOS and PR-EOS with Wilson activity coefficient by means of Wong Sandler (WS) mixing rule with three curve-fitting parameters, presented an excellent fit with the experimental data with an AARD% of 5.0 and 7.0% for SRK-EOS and PR-EOS, respectively.

Alwi et al., (2023) [246] applied three methods: (1) Eight models that are based on density (Chrastil, K-J, Bartle et al., MST, Alwi–Garlapati, Del Valle–Aguilera, Adachi–Lu, and Sparks et al.); (2) Two EOS, namely PC-SAFT and the SRK-EOS; and (3) Association theory based models developed by Reddy-Madras, Rajasekhar-Madras model, Garlapati-Madras model, and a new modified association theory model to correlate the solubility of Clonazepam in SC-CO_2 at different temperatures (308 to 338 K) and pressures (12–30 MPa). According to the results, all examined models could produce acceptable fits.

I.2.7. Critical Review and Establish the Significance of QSPR Techniques in our Reported Research

The idea of predicting the solubility of pharmaceutical compounds (or solid compounds) in SC-CO_2 gained so much attention over more than the last 40 years. Yet various researchers still contributing to this field with different techniques and numerous datasets only with the hope that one day to build an ideal model able to predict any compound's solubility in the supercritical carbon dioxide solvent with no need for experimental data, as well to minimize the use of experiments, to offer the possibility of interpolating between the points measured experimentally, and to provide a tool to support process design and analysis. So far, this plan is still out of the picture. The models are far form being effective, although, a large amount of experimental data on the solubility of solids (including solid drugs) in SC-CO_2 is now available, which keeps increasing every year, allowing further attempts to develop models analogously to the measurement of solubility. This will allow one day reaching the optimal proposed goal.

As highlighted in the abovementioned section, over the year, authors developed various methods and techniques to correlate such as property. They started with the use of the thermodynamic equation of state (PR-EOS) as one of the early published papers, still ongoing with the introduction of new machine learning approaches, as it can be seen that most of the latter published papers belong to section I.2.5. to furthermore understand this transition a brief explanation should be addressed.

The solubility of a solid in a gas is a thermodynamic solid-gas equilibrium problem, thus, in the early years' authors focused on predicting this binary system with thermodynamics-based techniques such as the equation of state (EOS). This method can be divided into three main classes: cubic EOS, complex non-cubic EOS, and conductor-like screening model (COSMO) based on quantum mechanical calculation results to determine the interaction

between molecules in a mixture, used to predict the solubilities. Then the focus passes to empirical or semi-empirical correlation, solution model mostly the Flory–Huggins equation or Group Contribution Methods.

These techniques can well represent the solid compounds' solubility based on the influences of operating conditions. Despite that, their limitations must not be overlooked. Particularly, those models cannot provide qualitatively correct predictions. In addition, they require either sublimation pressures or melting properties to determine the solid phase fugacity of the studied solute, also, the knowledge of critical and physicochemical properties, and the acentric factor. Previous studies have shown that it is viable to estimate critical properties and the acentric factor from group contribution methods or PRCS-EOS. However, lacking experimental melting properties and sublimation pressures remains unsolved.

Unlike thermodynamically based models, models based on soft computing techniques such as machine learning algorithms have been developed and seemed more robust. Operating conditions are basically considered inputs for the models. Still, despite their fast computation ability, these techniques require a large amount of data and typically do not broaden well outside the ranges and classes of the systems from which they were recognized, and the accuracy is conditional on the availability and validity of experimental data because they are correlative methods.

Particularly, all mentioned modeling approaches rely on experimental data. Therefore, it is necessary to improve the abovementioned aspect by replacing the experimental input parameters with molecular structure information, which makes the new model being independent of experimental data. Viz, to represent molecules in a way that can be used as an input. An interesting approach that is suitable for fulfilling such a request, is the quantitative structure-property relationship (QSPR) techniques (QSAR descriptors).

The field of QSAR, QSPR, and QSTR modelling is the heading robust cheminformatics approach that can combine the Activity, Property, and Toxicity to the molecular structure, respectively, to explore and exploit the dependence of chemical, biological, toxicological, or other types of activities or properties on their molecular characteristics. QSAR modeling was commonly used as an advanced optimization technique in drug discovery research. However, in recent years QSAR modeling has grown, diversified and evolved from application to small series of homogeneous compounds using comparatively simple regression approaches to the analysis of very large datasets including thousands of diverse molecular structures adopting a

wide variety of machine learning and statistical techniques. Additionally, the approach focuses on the field of drug-like property prediction and chemical hazard assessment.

Khayamian and Esteki (2004) [247] carried out a wavelet neural network (WNN) model for estimating solubilities of naphthalene, biphenyl, fluorene, phenanthrene, and triphenylene in SC-CO_2 over a temperature range of 308–333K and a pressure range of 80–135 bar. The proposed model was constructed in QSPR using six descriptors namely, temperature, pressure, volume of the molecule, highest occupied molecular orbital (HOMO), dipole moment, and number of double bonds in the molecules. These descriptors were chosen by a stepwise-MLR method. The capability of the model was evaluated and found an R^2 of 0.996, an *F-value* of 1947, and a small standard error of 0.087.

Hemmateenejad et al., (2008) [248] conducted a QSPR study on predicting the solubility in SC-CO_2 of some recently synthesized anthraquinone, anthrone, and xanthone derivatives with a dataset of 29 molecules in various temperatures and pressures (1190 solubility data). The feature selection and model development were performed with the combined data splitting feature selection (CDFS). The selected techniques for modeling were a linear (multiple linear regression; MLR) and nonlinear (artificial neural network; ANN) methods and were validated by cross-validation and an external test (three compounds, which did not have a contribution in model development steps). Both the proposed models resulted in accurate prediction whereas more accurate results were obtained by the ANN model with an RMSE of 0.095 (RMSE=0.284 for the MLR model) and that features selected by the CDFS method showed better predictive ability than the descriptors of the conventional stepwise feature selection method.

Valenzuela et al., (2014) [249] built a QSPRs technique to predict the parameters of the solubility equation of Chrastil using a small set of descriptors obtained from their equilibrated 3D structure after molecular dynamic simulations in implicit CO_2. In order to predict the solubility in SC-CO_2 of 30 compounds as a function of system pressure and temperature aids the selection of process conditions for extraction processes. The best set of descriptors was chosen by the leave-one-out strategy while the leave-10%-out was used for cross-validation. The obtained ANN models with very good correlation (R^2>0.91 for the test set) were then tested with five new compounds for validation. In three of them, the models had gave great accurate predictions of the solubility as a function of pressure and temperature. In an alternative approach, accurate predictions were obtained for the five compounds if the experimental value of the solubility at a reference temperature was known and only the effect of pressure and temperature was obtained through the model.

Gurikov et al., (2016) [250] employed two models, the semi-empirical Chrastil model and the regular solution model to correlate a solubility database for high-boiling point compounds in SC-CO_2 and its mixtures with modifiers where 600 compounds (more than 20,000 experimental points) were used. Both proposed approaches basically required empirical parameters which were linked to the molecular structure of the solute (QSPR). The presented hybrid thermodynamic/QSPR models provided exceed the limits of the experimental data and predicted the solubility at a given pressure and temperature fully entirely on molecular structure.

Moussaoui et al., (2020) [251] carried out two computational intelligence techniques based on QSPR models to predict the solubility of 145 solid solutes in SC-CO_2 using a database of 3637 values. Dragon software was used to calculate the descriptors while the genetic algorithm (GA) was implemented to optimize the subset of the significantly contributed descriptors. The results showed that the SVR-QSPR model was more accurate than the ANN-QSPR with An overall AARD of 1.345% and 2.772%, respectively. The proposed SVR-QSPR was compared with other models reported in the literature.

This background explained the idea behind carrying a QSPR model for the prediction and correlation of drug and drug-like compounds solubility in supercritical carbon dioxide which will represented in Chapter II. Then developed in Chapter III with the application of more effective intelligence algorithms instead of traditional methods.

I.3. Solubilities of Solute Compounds in Ionic Liquids

Compared to the SC-CO_2, the utilization of ionic liquids (ILs) as a green solvent come much later. Despite that, those relatively novel types of compounds have gained so much attention in recent years with great importance from academic, industrial, and environmental perspectives and have the potential to be recognized as a substitution for conventional volatile organic solvents in reaction and separation processes. No flammability, high thermal stability, a wide liquid range, and electric conductivity are some properties of ILs, which make them impressive more and more. ILs can be linked freely with wide possible combinations of cations and anions, possible candidates are numerous. Therefore, it is time-consuming and expensive to carry out experimental measurements, the enormous possibilities, also make it a challenging task the design and search for the best IL for a specific application, which requires an efficient method to screen ionic liquids regarding their solvent power is mandatory to find an optimal combination and amazing models to predict the solubilities. Various types of solute solubilities in ILs were correlated recently. According to the solute used a review of most published papers

on correlating the solubility of solutes in ILs from the year 2004 until the end of the year 2023 is reported below:

I.3.1. Literature Review on Carbon Dioxide as a Solute

Kim et al., (2004) [252] used a group contribution form of a non-random lattice–fluid model was applied to predict solubilities of CO_2 in ILs; [bmim][PF6], [C6mim][PF6], [emim][BF4], [C6mim][BF4], [emim][Tf2N] and [C6mim][Tf2N], at 298.15K and up to 1MPa. The results show the agreements between data and calculated solubilities to be generally good for ionic liquids with 1-alkyl-3-methylimidazolium cations and [PF6], [BF4], and [Tf2N] anions with an estimated accuracy less than 0.002 in mole fraction.

Kroon et al., (2006) [253] estimated accurately the phase behavior of ionic liquid + CO_2 systems based on the truncated perturbed chain polar statistical associating fluid theory (tPC-PSAFT) EOS. The method accounts precisely for the dipolar interactions between IL molecules, the quadrupolar interactions between CO_2 molecules, and the Lewis acid-base type of association between the IL and the CO_2 molecules. The experimental vapor-liquid equilibrium of all data was calculated using a single linearly temperature-dependent binary interaction parameter. The results were in good agreement with the experimental data.

Torrecilla et al., (2008) [254] applied various AI techniques namely, Multiple Linear Regression (MLR), Multiple Quadratic Regression (MQR), Radial Basis Network (RB), and MultiLayer Perceptron (MLP) Neural Network (NN) models for the estimation of the CO_2 solubility in 1-n-ethyl-3-methylimidazolium hexafluorophosphate, 1-n-hexyl-3-methylimidazolium hexafluorophosphate, 1-n-butyl-3-methylimidazolium tetrafluoroborate, 1-n-hexyl-3-methylimidazolium tetrafluoroborate, and 1-n-octyl-3-methylimidazolium tetrafluoroborate ILs at sub and supercritical conditions. The fitting performance of the models was validated with different statistical and with analyses of residuals from the interpolative models by graphical and numerical methods and central tendency and statistical dispersion tests were applied. The MPE values in all five studied systems were less than 12, 8, 2.2, and 2.1% for MLR, MQR, MLP, and RB models, respectively. The numerical analysis of residuals from non-linear models was considered into account. The optimized interpolated models were tested and the MPE values calculated by the non-linear models were less than 3.3%.

Kim et al., (2011) [255] correlated the solubility of CO_2 in ILs: 1-butyl-3-methylimidazolium hexafluorophosphate and 1-nonyl-3-methylimidazolium

hexafluorophosphate at temperatures of 293.15 and 298.15Kand pressure up to 4MPa using the group contribution non-random lattice fluid equation of state (GC-NLF-EOS) proposed by Lee and co-workers. The group parameters for CG-NLF-EOS were slightly modified at a limited range to accommodate recent experimental data and better prediction at high pressure and long alkyl chains.

Mortazavi-Manesh et al., (2011) [256] developed a simple model to estimate the CO_2 solubilities in ILs under normal processing and conditioning pressures for natural gas fluids. The correlation was developed with the use of a PR-EOS for the calculation of gas fugacity coefficients, the COSMO-RS-based techniques for the calculation of unsymmetrical activity coefficients, and the molecular weight of ILs. According to the results, the proposed model correlated the available saturation pressure data within 13% with an average deviation of 261 kPa and a bias of 8 kPa. The results were also compared favorably against another method available in the literature where the average error and deviation for the same data are within 47% and 823 kPa and the bias of −822 kPa. The method was tested against IL and CO_2 mixtures hidden from the development model with an average error under 14.9% with an average deviation of approximately 71 kPa.

Ashrafmansouri and Raeissi (2011) [257] used a group contribution approach to correlate the phase behaviour of CO_2 + IL systems, within a temperature range of 313.15–353.15K and pressures up to 100 bars, based on the statistical associating fluid theory (SAFT) which was proposed by Lymperiadis et al [258]. The ability of the model to describe the phase behaviour of these systems was confirmed and examined by using the optimized group parameters to predict equilibrium data of some CO2 + IL systems not included in the optimization database. The results show good agreement with the experimental solubility data.

Eslamimanesh et al., (2011) [259] carried out an ANN procedure to represent the solubility of SC-CO_2 in 24 most lyused ILs, counting 1128 data points. The model developed was a three-Layer Feed Forward Neural Network having critical properties of ILs and operational temperature and pressure as inputs. The results showed an R^2 of 0.993 and an AAD of 3.6% confirming the prediction capability of the presented model.

Lei et al., (2012) [260] used a PR-EOS to correlate the solubility data of CO_2 in a series of ILs namely, 1-ethyl-3-methylimidazolium tetrafluoroborate ([EMIM]+[BF4]−), 1-butyl-3-methylimidazolium tetrafluoroborate ([BMIM]+[BF4]−), 1-octyl-3-methylimidazolium bis(trifluoromethylsulfonyl)-imide ([OMIM]+[Tf2N]−), and their binary mixtures (i.e.,

[EMIM]+[BF4]− +[OMIM]+[Tf2N]− and [BMIM]+[BF4]− + [OMIM]+[Tf2N]−) at temperatures (313.2 and 333.2) K and pressures up to 6 MPa.

Gonzalez-Miquel et al., (2012) [261] used a quantum-chemical approach COSMO-RS to develop structure−property relationships of reversible ionic-liquid (RevIL) solvents for CO_2 capture. A variety range of structures were included in this study: silylamines with varying alkyl chain lengths attached to the silicon and amine functionality, silylamines with fluorinated alkyl chains, sterically hindered silylamines, and carbon-based analogues. Based on the results the simple COSMO-RS computational approaches built can be used as a guide for designing new RevILs, and estimations well the experimental data. Hence, theoretical tools as a guide can dramatically reduce the time and expense associated with the synthesis of large numbers of potential structures for CO_2 capture.

Chen et al., (2012) [262] confirmed the ability of the PC-SAFT equation in representing the solubility of CO_2 in ILs where the parameters of pure imidazolium based ILs were calculated by applying experimental densities and then correlated regarding the molecular weight and structure of the solvents; The density was predicted with high accuracy. The k_{ij} parameter was fitted to experimental liquid−vapor equilibria data. then, a correlation based on the group contribution concept was conducted to determine this temperature-dependent parameter. Finally, the ability of the model to describe accurately the solubility in imidazolium-based ILs was verified.

Manesh et al., (2012) [263] calculated the solubility of CO_2-ILs systems with a semi-empirical correlation. The theoretical COSMO–RS calculation method was used to calculate the chemical potential of CO_2 in a wide variety of ILs and the SRK-EOS was used to calculate the fugacity coefficient of the CO_2 vapor phase. The model was correlated with available literature data, yielding an average error of AAR=23% and a small bias.

Macías-Salinas et al., (2013) [264] proposed a formal modeling approach of the CO_2 solubility prediction in ILs, which has cations of the imidazolium type together with anions of the [TF2N] type. The applied method was a cubic EOS (Soave or Peng−Robinson) coupled with modern mixing rules of the Wong−Sandler type. The results indicate that the proposed model can correlate and/or predict the CO_2 solubility in ionic liquids over diverse conditions of temperature, pressure, and solute composition.

Haghbakhsh et al., (2013) [265] proposed a correlation to predict CO_2 solubility in 27 common ILs as a function of the temperature and pressure, for 3073 data. The ILs studied ranged within a variety of families, having various anions and cations. The results were

compared with the results of popular engineering models of PR-EOS and SRK-EOS. The proposed correlation was found to be more accurate with an AARD OF 4.9% (AARD% values of 43.5%, 44.3% for PR, SRK, respectively). In fact, when binary interaction coefficients of PR and SRK were optimized to experimental data, the model kept having better results (AARD% values of 17.2%, 16.9%, and 4.9% for PR, SRK, and the proposed model, respectively).

Sedghamiz et al., (2014) [266] predicted the solubility of carbon dioxide (CO_2) and hydrogen sulfide (H_2S) in numerical ILs by applying an ANN approach and a dataset of 2930 and 664 solubility data of CO_2 and H_2S, respectively. The databases were divided into three sets of train (70%), test (15%), and validation (15%) set. The model was built with one hidden layer through a hyperbolic tangent sigmoid transfer function. Optimum neurons were 23 and 14 for CO_2 and H_2S solubility, respectively. The AAD% and R^2 were 3.58% and 0.9947 for CO_2 and 2.07% and 0.9987 for H_2S system. The final results were also compared to the results of a PR-EOS with an optimized k_{ij}, as well as with an empirical correlation with different constants where the ANN model was found to have higher accuracy and lower error among the three approaches.

Baghban et al., (2015) [267] preposed a multi-layer perceptron artificial neural network (MLP-ANN) and an adaptive neuro-fuzzy interference system (ANFIS) based on the pressure, temperature of the system, critical temperature (Tc), critical pressure (Pc), acentric factor (ω), and molecular weight (Mw) of pure ionic liquids to estimate CO_2 solubility in presence of various ILs, a total of 728 experimental data points. The results of the models were compared to the results of thermodynamic models based on PR-EOS and SRK-EOS. The R^2 values were 0.9972, 0.9336, 0.7323, and 0.6989, while the MSE was 0.000133, 0.002942, 0.002702, and 0.005578 for MLP-ANN, ANFIS, PR-EOS, and SPK-EOS models, respectively. Therefore, the MLP-ANN model has better ability and better performance in estimating.

Tatar et al., (2015) [268] carried out four different methods based on artificial intelligence namely, RBF, MLP, LSSVM, and CMIS to predict CO_2 solubility in 14 different ILs, 728 data points (582 data for constructing the model, and 146 for validation). The chosen inputs were pressure, temperature, critical pressure, critical temperature, molecular weight, and acentric factor. Based on the results, the predicted values were in great agreement with the experimental data and the CMIS predicted the best with an AAED of 7%. A comparison between developed models and previously published ones, such as EOS methods, revealed the

superiority of the proposed models (without the problems of the presented thermodynamic models).

Hizaddin et al., (2015) [269] proposed a SRK-EOS technique for predicting CO_2 solubility in imidazolium-based ionic liquids containing bis(trifluoromethylsulfonyl)imide [Tf2N], tetrafluoroborate [BF4], and hexafluorophosphate [PF6] anions. The UNIFAC PSRK matrix was extended to include these new groups. The model predictions showed satisfactory agreement with the experimental data at pressures up to 5 MPa, with an overall root mean square deviation (RMSD) of 0.31. The prediction ability of the model was tested for ILs containing long alkyl chains, exhibiting an overall RMSD of 0.40. The results indicated that the proposed model provides a reasonable prediction.

Ali et al., (2016) [270] applied a PR-EOS model to predict the solubility of CO_2 in specific ILs at high pressure by tuning the interaction parameters to fit the experimental data reported in the literature. The model demonstrated excellent accuracy in predicting the solubility of CO_2 at pressure values less than the critical pressure of CO_2. However, at higher pressures; the calculated solubility diverged from the experimental values. Furthermore, the type of anion and cation used in the IL affected the performance of the PR-EOS.

Mehraein and Riahi (2016) [271] proposed an efficient approach based on the combination of genetic algorithm-multi linear regression (GA-MLR) and least-squares support vector machines (LS-SVM) to build a nonlinear QSPR model for predicting CO_2 solubilities in ILs utilizing a set of 21 rows (CO_2 solubility in 21 different ILs at 303K and 10 bar). The results were very satisfactory prediction results; R^2=0.9620, and $RMSE$=0.015, and were compared to the results of a QSPR-MLR model (R^2=0.876, and $RMSE$=0.027). The LS-SVM model drastically enhanced the ability of prediction in QSPR studies and was superior to the MLR model.

Tatar et al., (2016) [272] carried out different intelligent approaches such as MLP-ANN, GA-RBF, CSA-LSSVM, Hybrid-ANFIS, PSO-ANFIS, and CMIS for the calculation of CO_2 solubility in methyldiethanolamine, and N-methylpyrrolidone in a wide range of temperature and pressure. The input data of the models were temperature, pressure, and saturation pressure, while the output parameter was the solubility of CO_2. The models' reliability was investigated through different graphical and statistical methods where it was found that the CMIS approach has better results compared to other intelligent-approaches and that intelligent based algorithms are powerful alternatives for time-consuming and difficult experimental processes of solubility measurement.

Zoubeik et al., (2016) [273] correlated using PR-EOS with a quadratic mixing rule and the nonrandom two-liquid (NRTL) model the CO_2 solubility for four different ILs (1-ethyl-3-methylimidazolium L-(+)-lactate ([EMIM][LACTATE]), 3-methyl-1-propylpyridinium bis[(trifluoro-methylsulfonyl]imide ([PMPY] [TF2N]), 1-(4-sulfobutyl)-3-methylimidazolium bis(trifluoro-methanesulfonyl)imide ([(CH2)4SO3HMIm][TF2N]), 1-(4-sulfobutyl)-3-methylimidazolium hydrog-en sulfate ([(CH2)4SO3HMIm][HSO4])). The studies systems were at several fixed temperatures (313.15, 323.15, 333.15 K) and at pressures up to 20 bar. The model correlated well with all the ILs but could not correlated the [PMPY][Tf2N], as it behaved like [BMIM][Ac], known to react chemically with CO_2.

Ali et al., (2016) [270] used a COSMO-RS model to predict the solubility of CO_2 at atmospheric pressure in eight imidazolium-based ILs resulting from the combination of ethyl, butyl, hexyl, and octyl-imidazolium cations with two anions: bis(trifluoromethylsulfonyl)imide ([Tf2N]) and Trifluoromethanesulfonate ([TFO]). The results indicated relatively acceptable qualitative consistency between the experimental and predicted values.

Venkatraman and Alsberg (2017) [274] used ML-based approaches to establish QSPR between molecular structures of cations and anions and their CO_2 solubilities with a database of 10,000 IL-CO_2 solubility data of 185 ILs measured at different operating temperatures and pressures. The results were in good agreement with the available experimental measurements for both single decision tree (MAE= 0.10) and ensemble random forest (MAE=0.04) models. Also, their results were more accurate than those obtained with the quantum chemistry-based COSMOtherm predictions.

Panah (2017) [275] estimated the solubility of CO_2 in the fourteen ionic liquids ([emim][eFAP], [emim][EtSO4], [emim][OTf], [emim] [Tf2N], [bmim][BF4], [bmim][PF6], [hmim][PF6], [hmim][Tf2N], [omim][PF6],[omim][Tf2N], [HOemim][BF4], [HOemim][OTf], [HOemim][PF6] and [HOemim][Tf2N]) using the CPA-EOS. The critical temperature of ILs was calculated with the modified Lydersen-Joback-Reid method, and pure parameters of ILs were calculated with CPA-EOS. Also, the interaction parameters between the ILs and CO_2 were calculated. The proposed systems were modeled by the SRK-EOS and the obtained results were compared with the obtained ones by the CPA-EOS. For all the studied systems, the AAD % was lower than 10%.

Al-fnaish and Lue (2017) [276] investigated the solubility of CO_2 in several methylimidazolium bis (trifluoromethylsulfonyl) imide ILs or [Cnmim][NTf2] where n = 2, 4, 6, and 8 using the PC-SAFT. The pure component parameters of the ILs were calculated and

two strategies were tested to model the ILs. In one cas, ILs were treated as neutral molecules, while in the second one, ILs were modeled as two charged ions: imidazolium cation [Cnmim]+ and bis (trifluoromethylsulfonyl) imide anion $[NTf2]^-$. For both cases, four different self-association schemes are examined: non-associating, 2-site, 3-site, and 4-site schemes. The 4-site association scheme with two donors and two acceptors provided the best results for almost all the investigated binary systems, with an AARD of 1.54%–4.98%.

Deng et al., (2019) [277] combined critical properties of ionic liquids with deep learning models (CP-DNN, CP-CNN, CP-RNN) to establish theoretical prediction models of CO_2 solubility in ILs. The predictive performance of these frameworks was similar (or exceed) to the predicted effects of the techniques based on thermodynamic models (PR, SRK) and ML method (XGBoost). For CP-RNN, the R^2 value is 0.988, CP-CNN is 0.999, and CP-DNN is 0.984. This research can avoid complex computational characterization, it is to provide a theoretical method to further enrich and improve the data information analysis of this database.

Zhang and Xu (2020) [278] developed a least-squares boosting (LSBoost) model to predict the solubility of SC-CO_2 in 24 ILs with the use of critical properties of ILs and biphasic system parameters as descriptors. The model required a few inputs, including include the critical temperature, the critical pressure, the acentric factor of the ILs, and the temperature and pressure of biphasic systems. The model was highly accurate, stable, and promising as a fast, robust, and low-cost tool for solubility predictions and better performance than other models reported in previous studies.

Moosanezhad-Kermani et al., (2020) [279] investigated the application of the group method of data handling (GMDH) intelligent method as a powerful computational approach for predicting CO_2 solubility in 60 ILs with a temperature lower and upper than 324 K and 4726 data points while seven different ILs were chosen to perform an external test. The selected inputs were temperature (T), pressure (P), critical temperature (Tc), critical pressure (Pc), and acentric factor (ω). The results showed a proper fit between the experimental and the predicted data. Noted that the predicted negative values of the proposed models were considered zero and that the established models were compared to other proposed models that exist in the literature.

Song et al., (2020) [280] predicted the solubility of CO_2 in various ILs. The relationship between CO_2 solubility and IL structure, temperature, and pressure was calculated with a group contribution (GC) method. Two different non-linear regression machine learning algorithms, namely artificial neural network-Group contribution (ANN-GC) and support vector machine-Group contribution (SVM-GC) were employed to develop the models. The models were

accomplished using a database of 10,116 solubility data points at different temperature and pressure ranges. The results obtained indicated that the ANN-GC performs better than the SVM-GC model with an MAE of 0.0202 and an R^2 of 0.9836 for the test-set data.

Aghaie and Zendehboudiand (2020) [48] applied four ML algorithms including Least Squares Support Vector Machine (LSSVM), Decision Tree (DT), Random Forest (RF), and Multilinear Regression (MLR) to correlate CO_2 solubilities in a structurally diverse set of ILs based on thermodynamic properties and QSPR-based model. The authors used two sets of databases; the first considerd critical properties, molecular weight, and acentric factor of pure ILs as the input information, and the second one included structural descriptors of cations and anions as the input parameters. The significant inputs (e.g., Homo-Lumo fraction and Disps (cation)) were selected with the Genetic Algorithm (GA)-MLR algorithm. The predictive ability of the models was analyzed by a four-fold external cross-validation procedure. The results indicated that a great match between the predicted values and experimental measurements was attained for RF and DT techniques for both descriptors and thermodynamics properties database, and the structural descriptors-based models were more accurate and robust than those built on the critical properties.

Ouaer et al., (2020) [281] evaluated the performance of two data-driven techniques, namely multilayer perceptron (MLP), and gene expression programming (GEP), in estimating the solubility of CO_2 in 13 ILs as the function of pressure, temperature, and four thermodynamical parameters of the IL using 744 experimental data points (80% of the points for training and 20% for validation). The MLP algorithm was optimized with two backpropagation-based techniques: Levenberg–Marquardt (LM) and Bayesian Regularization (BR) and validated with numerous statistical and graphical assessments. Then the results were compared to the results of PR-EOS and SRK-EOS. The comparison confirmed the slightly better performance of the MLP-LMA model (R^2=0.9965, and RMSE=0.0116).

Valeh-e-Sheyda et al., (2021) [282] estimated the CO_2 solubilities in pyridinium-based IL mixtures using a dataset of 430 solubility data and two different feed-forward back-propagation neural network models with three (molecular weight, absorption temperature, and equilibrium pressure) and five (three molecular descriptors, the presence of ether groups as a categoric factor, and the number of carbon in hydrocarbon chain and\or ether groups in functionalized pyridinium-based ILs) input variables. The statistical analysis of the obtained results proved that the developed molecular descriptor-based model, with the estimated

(*RMSE=9.06-E03*, and R^2=0.995 for the test data) can be employed for the detailed assessment of the CO_2 loading capacities of pyridinium-based IL solutions.

Wang et al., (2021) [283] developed a molecular-based mathematical model for the calculation of CO_2 solubility in different deep eutectic solvents (DESs). The database used includes 1011 data points at various temperatures and pressures. The COSMO-RS-derived descriptors of involved hydrogen bond acceptors and hydrogen bond donors of DESs were calculated. The efficiency of the input variables was explored by a qualitative analysis using a simple MLR model. Then a random forest algorithm was employed to develop a more accurate nonlinear QSPR model. The proposed QSPR model with COSMO-RS-derived parameters as molecular descriptors showed better results than the COSMO-RS model, traditional thermodynamic models, and equations of state methods.

Zhang et al., (2022) [284] predicted the CO_2 equilibrium solubility in aqueous 2-(dimethylamino)-2-methyl-1-propanol (2DMA2M1P) solution, at various temperatures conditions, and CO_2 partial pressure and amine concentration, with an extended Cf model. The predicted results show a good agreement with the experimentally measured data with an AAD of 2.08%. Furthermore, the proposed model was validated with the published data of tertiary amines, and the results indicated that the developed model was in good universality in predicting CO_2 equilibrium solubility.

Feng et al., (2022) [285] carried out 12 different ML algorithms to train models for estimating CO_2 solubility in different ILs as function of temperature (T), pressure (P), molecular weight (Mw), critical temperature (Tc), and critical pressure (Pc) of solutions. The statistic evaluation indicated that the Stacking algorithm presented the most accurate model in the IL-CO_2 system, with an average MSE, RMSE, MAE, and R^2 of 0.001, 0.025, 0.018, and 0.969 respectively. In fact, the combination of Voting and Stacking algorithms can balance out each model's weaknesses and produce a more accurate model. Those algorithms can be considered as promising candidates for the estimation of acid gases solubility in ILs.

Tian et al., (2023) [286] carried out an QSPR-based model for the calculation of CO_2 solubilities in ILs, a total of 13,055 CO_2 solubility data in 164 kinds of ILs with different temperature and pressure ranges. The proposed model was built by combining ionic fragments contribution (IFC) with SVM and ANN to create the relationship between CO_2 solubility and ILs structure. The results confirmed that both IFC-SVM (R^2=0.9855) and IFC-ANN (R^2=0.9732) models can accurately predict.

Heidari et al., (2023) [287] carried out two ANFIS models to correlate the solubility of CO_2 in 26 different ILs with various ionic structures under subcritical and supercritical CO_2 conditions at a wide range of pressure (0.34–14596 kPa) and temperature (278.1–368.4 K). The 1st model toke the critical temperature (Tc), melting temperature (Tm) and molecular weight (Mw) of ILs as input variables, while the 2nd was a function of critical temperature (Tc), critical pressure (Pc), and acentric factor (ω) of ILs. Both models predicted well the laboratory data of training and test data, mainly in critical conditions, while the 2nd model (R^2=0.9950) presented better predictive accuracy than the 1st model (R^2=0.9500).

Ali et al., (2023) [288] applied a PR-EOS model associated with three mixing rules, including modified Van der Waal (MR1), the Quadratic (MR2), and the Wong Sandler (MR3) for calculating the solubility data of CO_2 in seven ILs at various pressure (0.119-65.2 bar) and temperatures (298.2-353 K) ranges. The modified Lydersen-Joback-Reid method was used to determine the critical properties of ILs. The final results showed very similar results to the literature.

I.3.2. Literature Review on Hydrogen Sulfide as a Solute

Jalili et al., (2009) [289] correlated the H_2S solubilities in three ILs, namely, 1-butyl-3-methylimidazolium hexafluorophosphate ([bmim][PF6]), 1-butyl-3-methylimidazolium tetrafluoroborate ([bmim][BF4]), and 1-butyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide ([bmim][Tf2N]), at temperatures ranging from (303.15 to 343.15) K and pressures up to 1 MPa using the Krichevsky-Kasarnovsky equation.

Sakhaeinia et al., (2010) [290] correlated the solubility of H_2S in ILs, 1-ethyl-3-methylimidazolium hexafluorophosphate ([emim][PF6]), at temperatures from (333.15 to 363.15) K, and 1-ethyl-3-methylimidazolium bis(trifluoromethyl) sulfonylimide ([emim][Tf2N]) at temperatures ranging from (303.15 to 353.15) K and pressures up to about 2.0 MPa using two models: the Krichevsky-Kasarnovsky (KK) equation, and the extended Henry's law combined with the Pitzer's virial expansion for the excess Gibbs energy. Henry's law constants (at zero pressure) were obtained at different temperatures from the obtained experimental solubility data.

Ahmadi et al., (2014) [291] investigated the capability of a three-layer feed-forward ANN algorithm trained with Back Propagation (BP) and particle swarm optimization (PSO) in estimating the solubility of H_2S in 11 different ILs (465 data points) as a function of acentric factor (ω), critical temperature (Tc), critical pressure (Pc) of Ils accompanied by pressure (P)

and temperature (T). the optimized architecture was found to be 5-9-1. The results showed an R^2 of 0.99218 and an MSE of 0.00025 for the PSO-ANN predictive model while the values of R^2= 0.95151 and MSE=0.00335 for the BP-ANN predictive model. Therefore, the PSO-ANN model could attain significantly better results than the BP-ANN model.

Ahmadi et al., (2014) [292] applied an LSSVM model optimized with a GA algorithm for forecasting H_2S solubility in various ILs utilizing 465 data samples (e.g., where 11 ionic liquids are included). The results demonstrated great predictive performance (R^2=0.997594, and MSE=6.6507E-05).

Ahmadi et al., (2014) [293] appraised the effectiveness of one of the new generation soft computing methodologies, called gene-expression programming (GEP) for estimating H_2S solubility in 11 ILs with a database of 465 experimental data. The temperature, pressure, acentric factors, critical temperature, and pressure of ILs were considered independent input variables, while H_2S solubility was the dependent output variable. The modeling results showed an R^2= 0.9902, and MARE=0.0438% confirming that the proposed model was comprehensive and accurate enough to be used for predicting such data. As well, the results were compared to the results of two well-known engineering approaches named SRK-EOS, and PR-EOS, and indicated that the proposed technique was more accurate than the widely used aforementioned thermodynamic models.

Faúndez et al., (2015) [294] carried out a three-layer ANN with 6, 10, and 1 neuron, respectively (6, 10, 1) architecture for the correlation of H_2S solubility in ILs using twelve binary hydrogen sulfide+ionic liquids mixtures counting 392 data points for training and 104 data points for testing. The selected inputs were the temperature and pressure of the binary systems. An AAD lower than 4.0% and the maximum individual absolute deviation in solubility of 12.6% were found. The proposed neural network model was a good alternative method for the estimation of such a dataset.

Zhao et al., (2016) [295] carried out two QSPR models based on charge distribution area (*S*σ-profile) descriptors and extreme learning machine (ELM) algorithm for predicting H_2S solubility in ILs, a total of 1282 pieces of data belonging to 27 ILs. The models end up with an AARD% and R^2 of 3.73 % and 0.998, as well as 3.80 % and 0.997, respectively.

Amedi et al., (2016) [296] used a feed-forward MLP-ANN, an ANFIS, and an RBF-ANN approach for correlating the H_2S solubilities in various ILs based on the critical temperature (Tc), critical pressure (Pc), and molecular weight (Mw) of pure ionic liquids, and 664 experimental data points. Different statistical criteria were used to evaluate the models'

performance. The values of R^2 and MSE achieved for MLP-ANN, ANFIS, and RBF-ANN were 0.9951 and 0.000117, 0.901 and 0.002268, and 0.9679 and 0.000787, respectively. Therefore, the ability and acceptable performance of using the MLP-ANN as an accurate model for estimating Hydrogen Sulfide solubility in ILs was shown versus other computational intelligence models.

Kamgar and Esmaeilzadeh (2016) [297] predicted the solubility of H_2S in [hmim][Pf6], [hmim][Bf4], and [hmim][Tf2N], in the temperatures of 303.15 to 343.15 K and pressure up to 11 bar, by UNIQUAC, NRTL and COSMO-RS models. The interaction parameters of UNIQUAC and NRTL were calculated as a function of temperature. The absolute relative deviations for the UNIQUAC and NRTL were found similar to some extent in the temperature domain, while the COSMO-RS model did not need experimental data and can compete with the other models.

Zhao et al., (2016) [298] carried out an extreme learning machine (ELM) intelligence algorithm on a dataset of H_2S solubility in 27 ILs (including 23 imidazolium-based and four ammonium-based ILs), 1282 pieces of data were used to build and test the model. The R^2 and the RMSE for the test set were found 0.990 and 0.0301, respectively. The results showed that the established ELM model is applicable for predicting the solubility of H_2S in ILs, which is important for the design, simulation, and analysis of new gas-sweetening processes.

Shojaeian (2016) [299] developed a Peng Robinson-Two state (PR-TS) association model to model the PVT behaviour of pure ILs and solubility of H_2S in 12 various imidazolium-based ILs. At first, the experimental liquid density of different pure ILs at various temperatures and pressures was correlated in order to obtain the parameters of the PR-TS-EOS for pure components, an AAD% of 0.24% was found. Two association schemes were considered in this study (presence or not of association behaviour) and estimated by simultaneously optimization of vapor pressure and liquid density data. For binary systems in the first model, only one interaction parameter was used, while in the second model, two binary interaction parameters were used. The results of the VLE calculation of the H_2S+IL systems showed that the best accuracy with AAD=3.40% is obtained when H_2S was considered as a non-association component and used as two binary interaction parameters for modeling the binary systems. Overall, the results of the present model for pure and binary systems comprising H_2S and ILs were in very good agreement with the experiments.

Panah (2017) [275] estimated the solubility hydrogen sulfide in the fourteen ionic liquids ([emim][eFAP], [emim][EtSO4], [emim][OTf], [emim] [Tf2N], [bmim][BF4],

[bmim][PF6], [hmim][PF6], [hmim][Tf2N], [omim][PF6],[omim][Tf2N], [HOemim][BF4], [HOemim][OTf], [HOemim][PF6] and [HOemim][Tf2N]) using the CPA-EOS. The critical temperature of ILs was calculated with the modified Lydersen-Joback-Reid method, and pure parameters of ILs were calculated with CPA-EOS. Also, the interaction parameters between the ILs and the gas were calculated. The proposed systems were modeled by the SRK-EOS and the obtained results were compared with the obtained ones by the CPA-EOS. For all the studied systems, the AAD % was lower than 10%.

Al-fnaish and Lue (2017) [276] investigated the solubility of H_2S in several methylimidazolium bis (trifluoromethylsulfonyl) imide ILs or [Cnmim][NTf2] where n = 2, 4, 6, and 8 using the PC-SAFT. The pure component parameters of the ILs were calculated and two strategies were tested to model the ILs. In one cas, ILs were treated as neutral molecules, while in the second one, ILs were modeled as two charged ions: imidazolium cation [Cnmim]+ and bis (trifluoromethylsulfonyl) imide anion $[NTf2]^-$. For both cases, four different self-association schemes were examined: non-associating, 2-site, 3-site, and 4-site schemes. The 4-site association scheme with two donors and two acceptors provided the best results for almost all the investigated binary systems, with an AARD of 2.76%–6.62%.

Kang et al., (2018) [300] proposed a class of molecular descriptors, namely, electrostatic potential surface (*S*EP) to correlate the H_2S solubility in ILs with the use of 1318 experimental data points of 28 ILs, including 7 cations and 12 anions covering diverse temperatures and pressures. Two quantitative models (ELM1 based on the isolated ions and ELM2 based on the ion pairs) were established based on combining the *S*EP descriptors with the extreme learning machine (ELM) algorithm. The AARD% for the total set of ELM1 and ELM2 models was found to be 5.87% and 3.84%, respectively. The results showed that the SEP descriptors can extensively be employed to predict properties of ILs due to their rich information at the electron level.

Mesbah et al., (2020) [301] suggested a group contribution correlation and an empirical relation model to estimate the solubility of H_2S in methylimidazolium-based ILs. The temperature ranges from 303.15 to 363.15K while the pressure range is 60.8-2016.8Kpa. The Nelder-Mead simplex algorithm was used to find the constants of the functionality relations. Both the correlations were trained with 407 data points of H_2S solubility in nine methylimidazolium-based ILs and tested through 121 H_2S solubility data points of three different methylimidazolium-based ILs to ensure generality. The results indicated that both suggested correlations are robust and have satisfactory error trends.

Amar et al., (2021) [302] introduced an advanced committee machine intelligent systems (CMIS) for predicting the solubility of H_2S in 33 ILs (1243 data points). The implemented CMIS models were gained by linking robust data-driven techniques including multilayer perceptron (MLP) and cascaded forward neural network (CFNN) beneath rigorous schemes using a group method of data handling (GMDH) and genetic programming (GP). The performed comprehensive error investigation revealed that the implemented paradigms yielded very satisfactory prediction performance and that the CMIS-GP provides better accurate estimations, and the best-prior paradigms compared with the other intelligent models. Besides, the proposed model exhibited overall AARD and R^2 values of 2.3767% and 0.9990, respectively. In conclusion, the trend analysis demonstrated that the tendencies of CMIS-GP predictions were in excellent accordance with the real data regarding pressure and temperature.

Mousavi et al., (2022) [303] used four deep learning (DL) approaches for the estimation of H_2S solubility in Ils, namely, convolutional neural network (CNN), recurrent neural networks (RNNs), deep belief networks (DBN), and deep neural network initialization with decision trees (DJINN). The dataset used contains 1516 data points of H_2S solubility in 37 ILs where the chemical structure of ILs, temperature, and pressure were considered as the model's inputs. The statistical results were compared to various proposed techniques of the literature where it was revealed that the CNN model provided more accurate, rapid, flexible, and inexpensive estimations with an AAPRE of 2.92% and an R^2 of 0.99. Taylor's diagram proved that the proposed method was very efficient, accurate, and realistic for predicting the solubilities as a function of chemical structure, pressure, and temperature, while trend analysis revealed that the trends of CNN predictions are in great agreement with the measured data based on the pressure.

Abdi et al., (2022) [304] carried out six robust machine learning methods, including adaptive neuro-fuzzy inference system, leasts quares support vector machine (LS-SVM), radial basis function, cascade, multilayer perceptron, and generalized regression neural networks for predicting H_2S solubilities in fifteen various ILs with an experimental databank comprising 792 datasets as a function of the temperature, pressure, acentric factor, critical pressure, and critical temperature of investigated ILs. The final results illustrated that the LS-SVM accurately predicts the solubility, and possesses R^2, RMSE, MSE, RRSE, RAE, MAE, and AARD of 0.99798, 0.01079, 0.00012, 6.35%, 4.35%, 0.0060, and 4.03, respectively.

Feng et al., (2022) [285] carried out 12 different ML algorithms to train models for estimating H_2S solubility in different ILs as function of temperature (T), pressure (P), molecular weight (Mw), critical temperature (Tc), and critical pressure (Pc) of solutions. The statistic

evaluation indicate that the Voting algorithm has the most accurate model in IL-H_2S system, with an average MSE, RMSE, MAE, and R^2 of 0.002, 0.032, 0.024, and 0.934 accordingly. The combination of Voting and Stacking algorithms can balance out each model's weakness and produce a more accurate model. Those algorithms can be considered as a promising candidate for the estimation of acid gases solubility in ILs.

Mousavi et al., (2023) [305] carried out a variety of ML techniques, including white-box ML (GMDH, and GP), deep learning (DBN), and ensemble learning (XGBoost), to determine the solubility of H_2S in ILs. The models were constituted within an extensive database of 1516 data points on the H_2S solubility in 37 ILs and as a function of seven input variables, such that temperature (T), pressure (P), critical temperature (Tc), critical pressure (Pc), acentric factor (ω), boiling temperature (Tb), and molecular weight (Mw). The most precise and accurate model was found to be the XGBoost model, with an AAPRE of 1.14%, RMSE of 0.002, SD of 0.01, and R^2 of 0.99.

Ali et al., (2023) [288] applied a PR-EOS model associated with three mixing rules, including modified Van der Waal (MR1), the Quadratic (MR2), and the Wong Sandler (MR3) for calculating the solubility data of H_2S in seven ILs at various pressure (0.119-65.2 bar) and temperatures (298.2-353 K) ranges. The modified Lydersen-Joback-Reid method was used to determine the critical properties of ILs. The final results showed very similar results to the literature. The MR2 mixing rule was the best compared with MR1 and MR3.

I.3.3. Literature Review on Sulfur Dioxide as a Solute

Bahmani et al., (2015) [306] estimated sulfur dioxide (SO_2) solubilities in many ILs with an ANN model as a function of temperature, pressure, critical temperature, critical pressure, and acentric factor. The uses ILs are consisting of 1-butyl-3-methylimidazolium acetate [BMIM][Ac], 1-nbutyl-3-methylimidazolium methylsulfate [BMIM][MeSO4], 1-Butyl-3-methylimidazolium tetrafluoroborate ([BMIM][BF4]), 1-butyl-3-methylimidazolium hexafluorophosphate ([BMIM][PF6]), 1-butyl-3-methylimidazolium bis(trifluoromethylsulfonyl) imide [BMIM][Tf2N], 1-ethyl-3-methylimidazolium tetrafluoroborate ([EMIM][BF4]), 1-Hexyl-3-methylimidazolium tetrafluoroborate ([HMIM][BF4]) and 1-Hexyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide [HMIM][Tf2N]. The network was built with 70% of the database (30% for the test) and ended up having one hidden layer with 27 neurons. The results confirmed that the ANN model was a successful method for estimating the data.

Mokarizadeh et al., (2020) [307] suggested two non-linear machine learning algorithms namely, the least square support vector machine couples with the genetic algorithm (GA-LSSVM) and an artificial neural network (ANN), to calculate the solubility of SO_2 in eight (8) ILs, accounting for 232 data points and seven (7) inputs including the pressure, temperature, critical temperature, boiling temperature, critical pressure, critical compressibility factor, and acentric factor The statistical and the graphical results showed the superiority of GA-LSSVM model with an AARD of 4.6%, illustrating good accuracy and validity.

Mohammadi et al., (2022) [308] presented four soft computing methods including deep belief network (DBN), group method of data handling (GMDH), genetic programming (GP), and K-nearest neighbour (KNN), and five EOS namely, Valderrama-Patel-Teja (VPT), Zudkevitch-Joffe (ZJ), Peng-Robinson (PR), Redlich-Kwong (RK), and Soave-Redlich-Kwong (SRK) EOSs for correlating the solubility of SO_2 in ILs with a database of 374 experimental data points of SO_2 solubility in 15 types of Ils. The results showed that the soft computational models outperformed the EOSs models, while the DBN model was the most reliable by having an average absolute percent relative error (AAPRE) of 3.56%. Moreover, the proposed simple-to-use GMDH mathematical correlation provided good calculations (AAPRE= 8.05%). Despite the low performance of the EOSs compared to the intelligent models, the PR-EOS gave better estimations among other EOSs.

I.3.4. Literature Review on Various Gases and Their Mixure as a Solute

Camper et al., (2004) [309] used a regular solution theory to model the solubility of gases (CO_2, C_2H_4) in room-temperature ionic liquids, RTILs, at low pressures. The results confirm the ability of the proposed model in correlating the solubilities without taking the intermolecular interactions between the gas and the anion into account, but instead using empirical constants that depend only on the gas. The model is limited to low pressure.

Wang et al., (2006) [310] used a molecular thermodynamic model (SWCF-EOS) to correlate the solubilities of gases such as CO_2, C_3H_6, C_3H_8, C_4H_{10} in 21 ILs. The relative deviation between the calculated molar volume and experimental data was found to be less than 0.2%. the results confirm that one temperature-independent adjustable interaction parameter is acceptable to correlate the data, but the accuracy was further improved using two temperature-independent adjustable parameters. For a reasonable description of the solubility of C_3H_6, C_3H_8, C_4H_{10}, two temperature-dependent binary interaction parameters were required. Based on the results, the presence of water in ILs has a large influence on the calculations. For systems with

water content lower than 0.1%, the ARD of bubble point pressure was 3.14 and 4.90% using two parameters and one parameter, respectively. For systems containing C_3H_6, C_3H_8, C_4H_{10} two temperature-dependent adjustable parameters were needed to obtain a good fit, and the corresponding deviation of the gas solubility was less than 2%, except for C_3H_8.

Kim et al., (2007) [311] carried out a group-contribution lattice fluid EOS for the correlation of the solubility of pure and mixed gases (carbon dioxide, propane, carbon dioxide + nitrogen, and carbon dioxide + propane) with ILs namely, 1-hexyl-3-methylimidazolium tri(bisfluoromethylsulfonyl)imide [C6mim][Tf2N] at 298.15K and up to 1MPa. The results were in reasonably good agreement with experimental data.

Carlisle et al., (2008) [312] carried out a group contribution technique to calculate the solubility parameters of the functionalized RTILs, then uses these values with a regular solution theory to predict the solubility and selectivity of the three gases namely, CO_2, N_2, and CH_4. These predicted data values were found to be in good agreement with those measured experimentally. The group contribution method offers a valuable guide for systematically designing functionalized RTILs with specific gas solubility and selectivity performance.

Martin et al., (2009) [313] carried out a group contribution equation of state (GC-EOS) for calculating the phase behavior of systems with ILs {(CO_2, O_2, or SO_2) + ionic liquids} with Tf2N anion and cations of the families 2,3-dimethyl-imidazolium, 1-alkyl-1-methyl-pyrrolidinium, and 1-alkyl-3-methyl-pyridinium. The GC-EOS was able to reproduce experimental data with deviations of the same order of experimental uncertainty. With the correlated parameters, it will be possible to predict the phase behavior of systems with ILs of the families considered in this work.

Manan et al., (2009) [314] evaluated the predictive capability of COSMOthermX in predicting the solubility data of 15 gases in 27 ILs (IUPAC database). First, gas solubilities estimations were performed on selected molecular solvents at constant temperature and pressure. Then, performed on 14 gas solubility at temperatures ranging from (278 to 368) K at 0.1 MPa in water. The study has demonstrated that COSMOthermX was capable of predicting, qualitatively, gas solubilities in ILs and, hence, reducing the amount of unnecessary experimental measurements prior to specific applications using ILs.

Sharma et al., (2009) [315] predicted the hydrogen (H_2) and carbon monoxide (CO) in [Bmim][PF6] at different temperatures and pressures up to 25 bar using correlation with various dimensionless parameters including modified Reynolds or Froude numbers to account for the

critical impeller speed. The best fit was obtained for models of the form $Sh \propto Re^b(Fr - Fr_{cr})^c Sc^{0.5}$ or $Sh \propto (Re - Re_{cr})^b Fr^c Sc^{0.5}$.

Yokozeki and Shiflett (2010) [316] applied a generic van der Waals equations of state (VDW-EOS) with their well-tested mixing rule for multi-component mixtures for phase equilibrium calculations mainly, the solubility of gases (CO_2, CF_3–CFH_2, SO_2, and NH_3) in room-temperature ILs. The results were compared to the results of the PR-EOS and the RK-EOS. All the models could present equally well the phase behaviors (temperature–pressure–composition diagrams). Also, the prediction results were satisfactory.

Blath et al., (2011) [317] predicted the solubility of CO_2, CH_4, N_2, and CO in a broad number of ILs (anions like [NTf2]−, [B(CN)4]−, [BF4]−, [CF3SO3]−, and [FAP]−, as well as cations like imidazolium, pyridinium, piperidinium, pyrrolidinium, sulfonium, and phosphonium) using two empirical models (Camper et al., (2005) [318] and Moganty and Baltus (2010) [319]) with the proposing of their own empirical model, which describes Henry's law constant as a function of the molar mass of the IL, as long as physical interactions between gas and IL are dominant. In this study, the absorption process was determined by physical phenomena. The Henry's law constant data was then implemented to the model of Camper et al., (2005) [318] as well as Moganty and Baltus (2010) [319] and expanded to non-imidazolium-based ILs and gases like methane and nitrogen. The final results indicated that Henry's law constant for CO_2, CH_4, and even N_2 can be described as a function of the reciprocal molar mass, whatever sort of ions respectively ion pairs the ionic liquid consists of, as long as physical interactions are dominant.

Oliferenko et al., (2011) [320] carried out multiple regression- and neural network-based models for the calculation of solubilities in 48 ILs of 23 industrially important gases as a function of molecular polarizabilities, special Lewis acidity, basicity descriptors calculated for the IL cations and anions, and gaseous solutes. The results were particularly good for the neural network model. Validation was established with an external dataset, again with a high-quality fit.

Shahriari et al., (2011) [321] applied a modified perturbed hard-sphere chain (PHSC) EOS for the calculation of acid gas (CO_2, H_2S) solubility in ILs over wide ranges of temperature and at pressures ranging from 0.1 to 100 MPa. The proposed model is a combination of the PHSC-EOS and a polar term that uses just one temperature-dependent binary interaction parameter. The polar contribution was based on the segment approach. The specific interactions between molecules were taken into account. The result showed an average deviation of less

than 0.14%, an absolute average deviation of less than 2.4%. The capability of the proposed model was evaluated by the prediction of solubility behaviour for the ternary H_2S/CO_2/[C4mim][PF6] system.

Sousa et al., (2013) [322] applied two modeling approaches for predicting the gas solubility of hydrofluorocarbons (CHF_3, CH_2F_2 and CH_3F) in four RTILs namely, 1-ethyl-3-methylimidazolium bis(trifluoromethanesulfonyl) imide([C2mim][NTf2]), (trihexyl)tetradecyl-phosphoniumbis(trifluoro-methylsulfonyl) imide ([P6,6,6,14][NTf2]), and N-methyl-2-hydroxyethylammoniumpropionate ([m-2-HEA][Pr]) and pentanoate ([m-2-HEA][P]), within the temperature range 288 K to 308 K and at atmospheric pressure. The first model was a cubic plus association equation of state (CPA-EOS) using only pure components' physical properties, while the second was a semi-empirical model based on RST whose parameters are obtained through least-squares fitting of experimental values. The models could predict qualitatively the solubilities between HFC-RTIL systems.

Singh et al., (2014) [323] correlated the solubility of CO_2, H_2, and their mixture in the ILs 1-alkyl-3-methylimidazolium bis(trifluoromethanesulfonyl)amide ([Cnmim+][Tf2N-], n=4,6) using the isothermal isobaric Gibbs ensemble Monte Carlo (NPT-GEMC) method. It is found that the predicted absorption isotherms for pure gases in the studied ionic liquids agree reasonably well with the available experimental results.

Amedi et al., (2016) [296] used a feed-forward MLP-ANN for correlating the solubilities of a ternary mixture containing CO_2, H_2S, and ILs based on the critical temperature (Tc), critical pressure (Pc), and molecular weight (Mw) of pure ionic liquids, and 89 experimental data points. Different statistical criteria were used to evaluate the model's performance. The values of R^2 and MSE achieved for MLP-ANN were 0.9955 and 0.000082 correspondingly. Therefore, the ability and acceptable performance of using the MLP-ANN as an accurate model for estimating was confirmed.

Al-fnaish and Lue (2017) [276] investigated the solubility of a ternary mixture including CO_2, H_2S, and C8mimNTf2 IL using the PC-SAFT. The pure component parameters of the ILs were calculated and two strategies were tested to model the ILs. In one cas, ILs were treated as neutral molecules, while in the second one, ILs were modeled as two charged ions: imidazolium cation $[Cnmim]^+$ and bis (trifluoromethylsulfonyl) imide anion $[NTf2]^-$. For both cases, four different self-association schemes were examined: non-associating, 2-site, 3-site, and 4-site schemes. The 4-site association scheme with two donors and two acceptors provided the best

results for the investigated ternary systems was successfully represented, with an AARD of 6.24% for CO_2 and 7.99% for H_2S, without the need for binary interaction parameters.

Kamgar et al., (2017) [324] estimated the solubility of different gases including CO_2, CO, CH_4, H_2, and N_2 in Choline Chloride/Urea as a deep eutectic solvent at a temperature ranging from 308 and 333 K and pressure up to 5 MPa using two thermodynamic models (NRTL and COSMO-RS). The performance of the proposed models indicated that the NRTL model predicted the solubility data better than the COSMO-RS model. This was justified by that the COSMO-RS model does not use any adjustable parameters in solubility calculation.

Amar et al., (2021) [325] established rigorous models, viz. cascaded forward neural network (CFNN), radial basis function neural network (RBFNN), and gene expression programming (GEP), for modeling the solubility of N_2O in various ILs. The obtained results verified that the employed models can predict solubilities with high accuracy. Moreover, the CFNN model optimized using Levenberg-Marquardt (LM) algorithm was the best predictive paradigm (R^2=0.9994 and RMSE=0.0047).

Urtiaga et al., (2022) [326] developed a prescreening tool based on the use of ANNs to correlate the solubility of F-gases in ILs from easily accessible properties of the pure compounds; the critical properties of the gases or the molar mass and volume of the IL. The UC-RAIL dataset that contained more than 4300 solubility data of 24 F-gases in 52 ILs was applied for this study which ends up with obtaining an ANN model capable to predict the fed dataset with an AARD of 10.93%, and an MAE of 0.014. Furthermore, the proposed model showed great prediction ability in estimating R-1243zf, an F-gas that was not present in the training set. Finally, the developed ANN was implemented in an easy-to-use spreadsheet that will allow it to extend its use in the prescreening of ILs towards the abatement and recovery of high environmental impact refrigerant gases.

Lv et al., (2022) [327] carried out two simple models, named Gene Expression Programming (GEP) and Group Method of Data Handling (GMDH), and four tuned EOSs, including PR-EOS, SRK-EOS, ZJ-EOS, and RK-ESO, to predict the solubility of CO_2-N_2 on a database of 289 experimental data. This study indicated that tuned EOSs outperform untuned EOSs while SRK-EOS outperform all EOSs (RMSE=0.00128, R^2=0.9561), and that the GMDH model has the best predictive performance (RMSE=0.000564, R^2=0.9792).

Tian et al., (2023) [286] carried out an QSPR-based model for the calculation of N_2 solubilities in ILs, a total of 415 N_2 solubility data in 38 kinds of ILs with different temperature and pressure ranges. The proposed model was built by combining ionic fragments contribution

(IFC) with SVM and ANN to create the relationship between N_2 solubility and ILs structure. The results confirmed that both IFC-SVM (R^2=0.9966) and IFC-ANN (R^2=0.9909) models can accurately predict.

Nakhaei-Kohani et al., (2023) [328] studied a database of H_2 solubilities in 13 Ils (580 data points) in a vast range of temperature (278.2-453.15 K) and pressure (0.433-552 bar) for building four intelligent models, including RF, AdaBoost-SVR, DBN, and Multivariate Adaptive Regression Splines (MARS) models, based on two different inputs: chemical substructures based inputs and thermodynamic attributes of ILs based inputs, along with the temperature and pressure for both sets. The results indicated that the DEN gives the best proficiency for both input sets with a RMSE of 0.00106 and 0.00066 for the first and second input sets, respectively, and an R^2 value of 0.9991 and 0.9996 for the first and second input sets, respectively. Also, the results of all four models were compared with EOSs, and it was found that the intelligent models have higher efficiency better performance than EOSs.

I.3.5. Literature Review on Organic Compounds as a Solute

Katritzky et al., (2008) [329] investigated by a QSPR methodology the Ostwald solubility coefficients and partition coefficients of organic solutes in eight different ILs by molecular descriptors calculated solely from their structures. Multi-linear regression approaches that have two to four-parameter were selected and have an R^2 ranges from 0.913 to 0.992. In addition, the QSPR models were validated using the leave-one-out cross-validation method.

Shan-Fei et al., (2010) [330] performed a QSPR study on the solubility of 84 organic compounds in 4 different ILs based on VolSurf parameters and with the use of the partial least square (PLS) statistical method where good results were obtained for both training and test sets. MLR analysis indicated that hydrophilic parameterW1 was the most significant parameter, which showed that hydrophilicity is a key factor that effects the solubility of organic compounds in ILs.

AlSaleem et al., (2015) [331] estimated the solubility of three halogenated hydrocarbons, namely, carbon tetrachloride, chloroform, and bromoform in 12 hydrophobic ILs for temperatures ranging between 25 and 45 °C with a COSMO-RS model. The predictive capability of the proposed model is in excellent qualitative agreement with experimental data both for temperature dependence and for the cations structure effect.

I.3.6. Literature Review on Water as a Solute

Freire et al., (2008) [332] applied a predictive method based on unimolecular quantum chemistry calculations (COSMO-RS) for the prediction of water solubilities in tetradecyltrihexylphosphonium-based ILs with the bromide, bis(trifluoromethylsulfonyl)imide, bis(2,4,4-trimethylpentyl)phosphinate, chloride, decanoate, and dicyanimide anions at temperatures between (288.15 and 318.15) K and atmospheric pressure. The results of correlating this binary system were found to provide fine qualitative and quantitative predictions of the experimental data for extremely hydrophobic ILs. Less accurate predictions were observed with the increase of the anion hydrophilic character.

Can et al., (2021) [333] calculated the solubility data for water in 16,137 ILs (a combination of 163 cations and 99 anions) using machine learning models namely, Association rule mining (ARM), decision tree (DT), and multilayer fully connected neural network, a deep learning model, FCNN. COSMO-RS-based methods were applied to compute water's solubilities in ILs while the molecular descriptors for the individual cations and anions were calculated with a semi-empirical PM3 approach. The ARM was used to identify the descriptors leading to low water solubility in ILs, while the DT analysis provided heuristic rules for the selection of cations and anions to form ILs with low water capacity. The prediction results of the FCNN showed high accuracy, indicating that the anionic and cationic descriptors used were sufficient to represent the performance of ILs. The classification ability of the DT was confirmed with the use of experimental water solubility data for 49 ILs. The DT model suitably classified the experimental solubility of 46 of these ILs. The deep learning predictions of solubility were also in agreement with the experimental data within the accuracy level of COSMO-RS calculations. In general, It was found that the anionic descriptors were more influential to predict the water capacity of ILs, while the cationic descriptors made limited contributions.

I.3.7. Literature Review on Sugar/Sugar Alcohols as a Solute

Carneiro et al., (2013) [334] modeled the solubility of two sugar alcohols (xylitol and sorbitol) in five different Ils including 1-ethyl-3-methylimidazolium dicyanamide, 1-butyl-3-methylimidazolium dicyanamide, AliquatR dicyanamide, trihexyltetradecylphosphonium dicyanamide, and 1-ethyl-3-methylimidazolium trifluoroacetate between 288 and 339 K with a PC-SAFT using a molecular-based associative approach for ILs. The three pure-IL parameters

were fitted to pure-IL density, activity coefficients of 1-propanol at infinite dilution in ILs, and/or xylitol solubility in ILs. This method allowed accurate modeling of the pure-IL data and the mixture data with only one binary interaction parameter k_{ij} between sugar and IL or sugar alcohol and IL. In cases where only pure-IL density and activity coefficients of 1-propanol at infinite dilution in ILs were used for the IL-parameter estimation, the solubility of sugars and sugar alcohols in ILs could be predicted (k_{ij}=0 between sugar and IL or sugar alcohol and IL) with reasonable accuracy.

Paduszyński et al., (2013) [335] presented theoretical thermodynamic studies of binary mixtures composed of sugars (D-(+)-glucose, D-(-)-fructose, and sucrose) and two low-viscous ILs based on 1-butyl-3-methylimidazolium cation and dicyanamide or trifluoroacetate anion as a function of temperature. The obtained solid−liquid equilibrium phase diagrams of the experimental data were analyzed in terms of perturbed-chain statistical associating fluid theory (PC-SAFT). The results showed that the PC-SAFT-EOS model was capable of accurately representing the phase behaviour of the studied systems quantitatively (The measured data were in very good agreement with the literature data). According to the results, the predictive performance of the PC-SAFT approach shows to be good if binary interaction parameters were incorporated into the combining rules. However, the values of the corrections were relatively small. As a conclusion, the proposed PC-SAFT (and possibly other SAFT-based ideas) coupled with simple molecular schemes can be perceived as an interesting and reliable thermodynamic tool for modeling systems of a high degree of complexity.

Mohan et al., (2015) [336] performed a COSMO-RS-based screening for the prediction of the solubility of monosaccharides (D-(+)-Glucose, D-(+)-Xylose, D-(-)-Fructose, and D-(+)-Galactose) in two ILs over the temperature range 303.15–373.5 K. The IL data set consisted of 31 cations and 23 anions comprising cations such as imidazolium, ammonium, pyridinium, pyrrolidinium, phosphonium, sulfonium, and piperidinium. The RMSE values of mono-saccharides such as D-(+)-Glucose, D-(+)-Xylose, D-(-)-Fructose and D-(+)-Galactose in 1-ethyl-3-methylimidazolium thiocyanate [EMIM][SCN] were 3.11%, 3.12%, 2.32% and 1.83%, respectively while the corresponding figures in Tris(2- hydroxyethyl)methyl ammonium methylsulfate [TMA][MeSO4] is 1.47%, 1.20%, 2.94% and 0.70%. The local thermodynamic models NRTL and UNIQUAC were used to calculate the experimental solubility data. The binary interaction parameters for both models gave an RMSE of less than 2%.

Bakhtyari et al., (2023) [337] carried out ML approaches for modeling the solubility data of sugar alcohols in ILs. The applied methods ANN, LSSVR, and ANFIS were computed

utilizing a large databank (647 data points of 19 sugar alcohols and 21 ILs). According to the results, ANFIS offered the best accuracy (AARD%=f 7.43%, and R^2= 0.98359). the model was trained with 85% of the dataset. Consequently, this model can be effectively employed in predicting this databank.

I.3.8. Literature Review on Cellulose as a Solute

Kahlen et al., (2010) [338] computed the relative solubilities of cellulose in ILs with a COSMO-RS-based technique combined with different types of models describing the combinatorial contribution to the chemical potential. The results were in good qualitative agreement with the experimental data and it was confirmed that the application of COSMO-RS, is well suited for a relatively quick, qualitative screening of a large number of ILs.

Mai et al., (2016) [339] predicted the cellulose dissolution in ILs with a QSAR modeling technique and the molecular descriptors of ionic liquid's constitutional ions derived from the CODESSA program. The proposed model based on cellulose molar solubility exhibited better correlation (R^2 of 0.93 vs 0.88) and predictability (R^2 of 0.89 vs 0.83) than those based on mass percentage solubility. while, models developed by the multivariate adaptive regression spline (MARS) method requested fewer variables (13 vs 57–59) with better predictability (R^2 of 0.83–0.89 vs 0.45–0.51) compared to those developed by multiple linear regression (MLR) techniques. The results showed that the molecular descriptor of ILs could be successfully used to develop QSAR models for easing the in silico and a priori screening/selection of ILs customized for specific applications.

Chu et al., (2018) [340] proposed a computational approach based on COSMO-RS for the prediction of cellulose solubilities in ILs. The COSMO-RS-based method was also applied for the calculation of the thermodynamic property of logarithmic activity coefficient (ln γ) and excess enthalpy (HE). Then, four sets of ILs were utilized to calculate lng, and HE based on four different cellulose models. The goodness-of-fit of linear regressions between the experimental cellulose solubilities and the calculated ln γ and HE shows that lng is more reliable than HE for predicting the solubility data. However, HE is more suitable for the prediction of the dissolution ability of halogen-based ILs. The overall results show that the cellulose models gave comparably good prediction results regarding the dissolving power of ILs based on the estimated lng, but the cellobiose model was acknowledged as the optimal model due to the relatively higher prediction ability (R^2) across different IL datasets.

I.3.9. Literature Review on Other Compounds as a Solutes

Shojaeian and Fatoorehchib (2019) [341] investigated the phase equilibrium behaviour of refrigerants with 37 various imidazolium-based ILs binary solutions at different temperatures and pressures with a total of 969 data points using a Peng Robinson-Two State equation of state (PR-TS-EOS). At first, the proposed model was applied to pure components to determine the parameters of the model. The AARD% was found to be 0.58% for the vapor pressures and 1.59% for liquid densities. The model was applied to estimate the vapor-liquid equilibrium of refrigerant+IL binary systems. The results showed a AARD% of 3.54% and were found to be better than the results of the Peng-Robinson EOS and Peng Robinson-Kwak Mansoori EOS.

Chu and He (2019) [342] developed an automated computational framework (MoDoop) for the calculation of biopolymer solubilities in ILs on the basis of COSMO-RS calculations of two thermodynamic properties: logarithmic activity coefficient (ln γ) at infinite dilution and excess enthalpy (HE) of the mixture. The calculation was based on the optimized two-dimensional structures of biopolymer models and ILs by searching for the lowest-energy conformer and optimizing molecular geometry. The evaluation results showed that ln γ is a more reliable property to predict lignin solubilities in ILs and the p-coumaryl alcohol model was considered the best model to represent lignin molecules.

I.3.10. Critical Review and Establish the Significance of combaing a QSPR model with COSMO-RS Techniques in our Reported Research

According to the literature review, the majority of contribution belongs to predicting gas solubility in various ILs. Begin with carbon dioxide (or SC-CO_2), Hydrogen Sulfide, sulfur dioxide, then the remanding gases. This can be mainly explained by the importance of adequate knowledge of the solubility of acid gases in ILs at different thermodynamic conditions in gas processing, carbon sequestration, and mostly, elimination of CO_2 from gas mixtures... Therefore, a precise prediction of this is a key parameter in the design perspective of OIL-based separation processes. In addition to various other published papers related to estimating several different types of solutes solubilities in ILs including water, organic compounds, sugar, cellulose, and the list goes on… each solute has various industrial and environmental impacts which make its prediction very essential.

Another interesting use of ionic liquids is as a replacement for conventional solvents (e.g. organic solvents) in pharmaceutical industries, to increase the solubility of drug

compounds with no traces compared to excipients polar organic solvents. This aspect is one of the most important recherche fields, which is the focus of our recherche. To the best of our knowledge, there is not yet much investigation into modeling the solubility of active pharmaceutical ingredients (APIs) compounds in pure ILs. In fact, very few papers can be cited.

Lotfi et al., (2017) [343] correlated the solubility of the sparingly soluble drug molecule acyclovir (ACV) in a wide variety of ILs with a COSMO-RS technique. The results were validated by good agreement between the predicted and experimental results. The solubility of ACV was greatly affected by the structure of ILs, particularly the anionic moiety.

In addition, too few other authors, proposed a thermodynamic-based correlation of an API in a single ILs (in some contributions, the temperature was chosen as the output instead of the solubility), or papers that study drugs' solubility in mixed ILs, such as:

Lee and Lin (2016) [344] investigated predicting drug molecules' solubilities in single- and mixed IL solutions using the COSMO-SAC activity coefficient model, 442 solubility data points ranging from 0.93 to 2.84×10-4 in mole fraction and temperature ranging from 248.92 K to 488.3 K. The root-mean-square deviation was found to be 91% (0.65% for ln" property").

Mahmoudabadi and Pazuki (2021) [345] established a predictive type of PC-SAFT equation of state by incorporating the COSMO computations to predict the solubility data of pharmaceutical compounds in pure and mixed solvents. The model was compared with the COSMO-SAC and COSMO-RS models. The results showed that the RMSE of predictive PC-SAFT-EOS is similar to the COSMO-RS model but lower than the RMSE for the COSMO-SAC model.

From a general point of view, there is not yet a paper that focuses on predicting different APIs solubilities in various pure ILs with great agreement with the experimental measurement. This can be clarified by the difficulty of modeling solid-liquid equilibrium. As it is well known, solubility can be affected by various parameters. In this particular case, the equilibrium is affected by four variables, namely, the temperature, and the natural behaviour of the drug compounds, cation, and anion used. Therefore, to model such a complex database, it is quite significant to have powerful inputs.

As explained in the previous section. It is important to use molecular structure information as input parameters to make a predictive model independent of experimental data, as well as to find a way to represent that information so that they can be used as inputs/ as a set of features for any type of approach.

So far there are various molecular representations such as SMILES, fingerprints, molecular graphs… Despite the achievements obtained using these molecular representations, they present several shortcomings. So the search for a universal representation of molecular structures is a necessity.

Conductor-like Screening Model for Real Solvents (COSMO-RS) method [346] is one of the widely used methodologies that has recently attracted so much attention. This method is one of the new advanced computational techniques based on quantum mechanical calculations and statistical thermodynamics [347]. The COSMO-RS was first introduced by Klamt et al. [346], which is a novel and an *a priori* predictive approach that uses molecular structure as the input parameter, namely, Charge distribution area ($S_{\sigma\text{-profile}}$). The σ-moments of ions $S_{\sigma\text{-profile}}$ is an a priori quantum chemical descriptor, which can represent rich information such as electronic structure and molecular size which is calculated by the COSMO, has been proved to be an effective descriptor for the QSPR model [295,348–352]. Mainly, the independence of experimental data for calculations makes this technique a powerful predictive of various circumventing challenges and limitations encountered in correlative modelling [347,352].

This full background inspired us the most, to contribute by applying the concept of the σ-profile (σ-moment) as a new molecular descriptor for predicting the solubility of APIs in ILs and aiming at mitigating the disadvantages of classical methodologies. This idea will be shown in the 4th chapter. For more details Colo´n et al. [353] explained how σ-profiles are better candidates to be used as inputs, such as their ability to describe molecules of any size. The authors prove that those in charge distributions are promising candidates for a universal molecular descriptor.

I.4. Conclusions.

This chapter has covered a wide spectrum of techniques that was used for the prediction of drugs' solubility in two different green solvents, namely, supercritical carbon dioxide and ionic liquids. The literature review involves an extensive study of research publications, the status of the defined problem, databases, and techniques that were applied by the researchers to investigate the problem, the obtained results, as well as other related details.

The first section reported, to the best of our knowledge, most of the published papers related to estimating solid compounds solubilities in SC-CO_2 from the year 1981 until now. The contributions of this section were described based on the used techniques, including

thermodynamic-based techniques and machine-learning methods. It was found that all mentioned modeling approaches rely on experimental data. Thus, their accuracy is conditional on the availability and validity of those experimental data. Therefore, as a novelty, a quantitative structure-property relationship-based technique (QSAR descriptors) was proposed to compute the solubility of drug (or drug-like) compounds in SC-CO_2. Furthermore, details are available in chapters II and III.

The second section, section reported, to the best of our knowledge, most of the published papers related to correlating solute solubility in ILs from the year 2004 until 2023. Various types of solutes were declared such as Carbon Dioxide, Hydrogen Sulfide, Sulfur Dioxide, water, sugar, cellulose, organic compounds… Reviewing the literature has implied that there are several authors that have contributed to predictive models for estimating/correlating various types of solutes while few papers almost no papers performed predictive models for modeling the solubility amount of active pharmaceutical ingredients (APIs) in ILs. Therefore, this background inspired us to contribute by employing the QSPR model combined with the Conductor-like Screening Model for Real Solvents (COSMO-RS) theory for the prediction of APIs solubilities in pure ILs using σ-moment ($S_{\sigma\text{-profile}}$) as input parameters. This study addressed the calculation of the σ-profile of APIs for the first time. More details will be held in Chapter IV.

Chapter 07

First Contribution

Chapter II

First Contribution: QSPR Modelling of The Solubility of Drug and Drug-Like Compounds in Supercritical Carbon Dioxide

II.1. Introduction

In this chapter, a quantitative structure-property relationship (QSPR) modeling was investigated to predict drug and drug-like compounds solubility in supercritical carbon dioxide. To do this, this chapter will carry on three main sections. The first section 'Procedures' will describe in detail all the necessary stages of the development of the database, the calculation and the selection of the descriptors, the applied learning technique, as well as the statistical and graphical metrics used to validate the built model. The second section, 'Results & Discussion', will discuss the results obtained. The chapter will end with a 'Conclusion' as the third section, which will summarize all the important remarques and results achieved from this study.

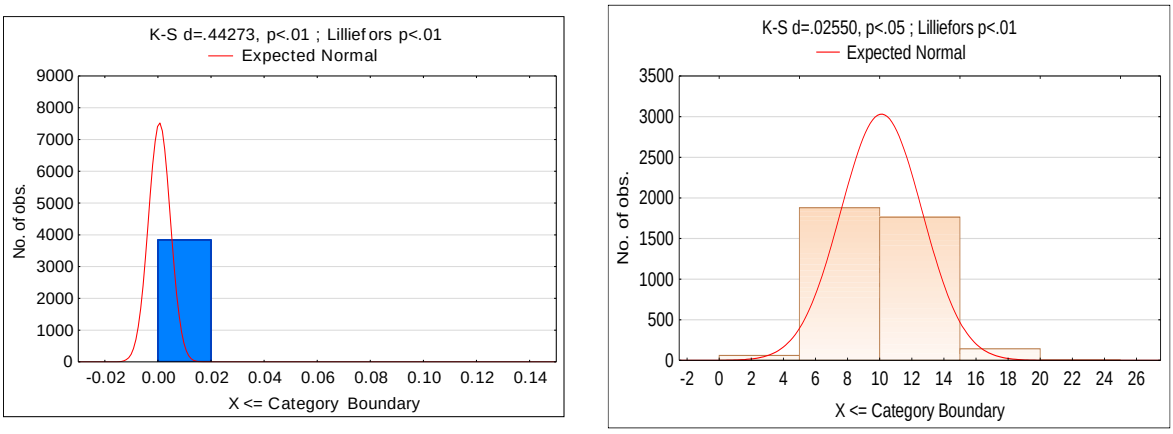
II.2. Procedures

II.2.1. Experimental Dataset

Data collection, known as the 'heart of modeling', is one of the main steps for the development of high-quality models. For this, it is necessary to select cautiously the data points that will be then used for building a reliable model. In this study, a total of 3,971 experimental data points (EDP) of drug and drug-like compounds solubilities in supercritical carbon dioxide have been collected from the literature.

A whole of 148 pharmaceutical compounds solubility (mole fraction 0.000000001–0.131) measured in SC-CO_2 from 298 K to 373.15 K and from 80 bar up to 500 bars were used for the model development. The dataset is an update from a previously performed database by Si-Moussa et al., (2017) [354] by increasing the number of drug compounds from 99 (except for Azadirachtin) to 148 (**Table A1**), accounting from 2,905 to 3,971 EDP and increasing the number of inputs by taking the two inputs of the previous study (the temperature and the pressure) and adding the molecule.

The output property was converted to the corresponding pLn property "-Ln(y2)" to ensure a normality distribution. The normality distribution was checked using statistical tests. The normality distribution plots are presented in **Figure II.1**.



(a) log-normal. (b) log-transformed.

Figure II.1 Normality distribution plot of mole fraction of solubility data.

II.2.2. Descriptor Calculations

The SMILES notation of the 148 pharmaceutical compounds was determined with the Chembioa drawing software. 1544 molecular descriptors were obtained by using an online computer software program called PaDEL [355]. The 1544 descriptors were then reduced to 124 using various variable reduction methods, beginning with removing features with zero values. Then, the invariant descriptors, namely those with absent values or those with almost the same values "descriptor with all values ~ constant", were manually removed. Any descriptor that have identical values for 75% of the samples and all descriptors with a relative standard deviation <0.05 were removed. Finally, half of the descriptors showing an absolute value of the Pearson correlation coefficient >0.95 were also removed. The objective of this step is to avoid the phenomenon of overfitting, and elimination of the insignificant descriptors.

II.2.3. Selection of Relevant Descriptors

Among all 124 descriptors, only a few will be chosen to define the model. In this step, the database will be horizontally reduced by applying the genetic algorithm method (GA-MLR); GeneticAlgorithm_v4.1 (https://sites.google.com/site/dtclabgatool/), to the smallest significant combination that can estimate the output property with an acceptable error. The algorithm takes as inputs the training set file (80% of the database) with an inter-correlation of 0.9 (various inter-correlation has been tested 0.75, 0.8, 0.85, and 0.9 where the results of the comparison indicate a negligible difference between them, hence an inter-correlation of 0.9 has been eventually selected). The remained inputs were kept by-default. Its should be mentioned that a database of 133 pharmaceutical compounds and 3590 EDP were used in this study.

The selection of the fitting number of descriptors was carried out over an interval of 5 to 19 features, establishing 14 combinations where each combination contained the data of the temperature, pressure, and from 3 to 17 descriptors. In order to select the best combination, the 14 models obtained were statistically examined using various metrics including, R, R^2, RMSE, Q^2, Q^2_{pred}, k, k', $\bar{r}^2_m$, and Δr^2_m (their formula are available in **section II.2.6**). The statistical results of 14 models are represented in the **Table II.1** and **Table II.3** for the training and test set, respectively.

The results of below Tables prove the acceptability criteria of all databases (or models). Note that a high R^2 means a more accurate model, and smaller RMSE values yield a more reliable prediction. Some interesting remarks that are worth mentioning are given hereof:

- The value of RMSE, Δr^2_m shows a negative correlation with the number of descriptors; The more the number of descriptors increases, the more those values decrease.

- The values of k and k' are technically constant (≈1) which means that the number of descriptors has no influence on those two parameters. i.e., no correlation.

- For the remaining statistics; R, R^2, Q^2, Q^2_{pred}, and $\bar{r}^2_m$ their correlation with the number of descriptors is a positive correlation; the more the number of descriptors increases, the more those values parameters increase.

Databases with 13 to 17 descriptors are more likely to be great candidates for predicting. For better judgments, more tests should be introduced.

Table II.1 Statistical parameters for internal validation

Training set						
Number of descriptors	**R**	**R^2**	**RMSE**	**Q^2**	$\bar{r}_m^2$	Δr_m^2
5	0.9395	0.8827	0.8640	0.8827	0.8290	0.1067
6	0.9690	0.9389	0.6235	0.9389	0.9091	0.0588
7	0.9726	0.9460	0.5865	0.9460	0.9202	0.0511
8	0.9829	0.9661	0.4644	0.9661	0.9498	0.0322
9	0.9851	0.9704	0.4340	0.9704	0.9546	0.0291
10	0.9869	0.9739	0.4076	0.9739	0.9615	0.0246
11	0.9882	0.9765	0.3867	0.9765	0.9649	0.0225
12	0.9892	0.9784	0.3706	0.9784	0.9670	0.0214
13	0.9905	0.9811	0.3477	0.9810	0.9716	0.0176
14	0.9926	0.9852	0.3069	0.9852	0.9779	0.0143
15	0.9906	0.9814	0.3443	0.9814	0.9724	0.0177
17	0.9928	0.9856	0.3004	0.9856	0.9785	0.0137
19	0.9939	0.9878	0.2791	0.9878	0.9821	0.0112

Multi-collinearity between the descriptors of the models 13, 14, 15, and 17 descriptors was identified by determining their variation inflation factors (VIF) using XLSTAT® (Addinsoft, New York, NY, USA). A model is known as acceptable if the VIF falls into the range of 1–5. The results are denoted in **Table II.3.** to **II.6.** The results of the Table indicate that all models have at least one VIF higher than 5, except the model of 13 inputs, which has all VIF values lower than 3.265, which confirms the acceptability of that combination compared to the rest and proves the over-performance of the model that has a combination of 13 features, best-fit combination, indicating that the obtained model has statistical significance, and the descriptors were found to be reasonably orthogonal.

Table II.2 Statistical parameters for external validation

Test set								
Number of descriptors	**R**	**R^2**	**RMSE**	**Q^2_{pred}**	**k**	**k'**	**$\bar{r}^2_m$**	**Δr^2_m**
5	0.9400	0.8836	0.9050	0.9823	1.0004	0.9921	0.8312	0.1019
6	0.9705	0.9418	0.6402	0.9911	0.9990	0.9972	0.9146	0.0471
7	0.9723	0.9453	0.6202	0.9917	0.9976	0.9988	0.9156	0.0419
8	0.9854	0.9709	0.4528	0.9956	0.9969	1.0012	0.9525	0.0133
9	0.9874	0.9750	0.4199	0.9962	1.0002	0.9982	0.9604	0.0273
10	0.9875	0.9750	0.4194	0.9962	0.9980	1.0004	0.9572	0.0203
11	0.9891	0.9783	0.3906	0.9967	0.9983	1.0003	0.9648	0.0139
12	0.9897	0.9795	0.3800	0.9969	0.9990	0.9997	0.9672	0.0169
13	0.9907	0.9815	0.3614	0.9972	0.9980	1.0008	0.9687	0.0103
14	0.9924	0.9849	0.3264	0.9977	1.0008	0.9983	0.9742	0.0155
15	0.9899	0.9799	0.3757	0.9969	0.9990	0.9997	0.9698	0.0128
17	0.9901	0.9801	0.3856	0.9967	0.9992	0.9994	0.9635	0.0275
19	0.9931	0.9863	0.3105	0.9979	0.9991	1.0000	0.9778	0.0100

Table II.3 Statistical of multi-collinearity of 13 descriptors

Descriptors	**T(k)**	**P(Mpa)**	**AATS3v**	**MATS2e**	**GATS4c**	**GATS3v**	**GATS4e**
Tolerance	0.983	0.972	0.532	0.822	0.349	0.769	0.316
VIF	1.017	1.029	1.881	1.216	2.861	1.300	3.164
Descriptors	**GATS3s**	**nBondsM**		**AVP-0**	**SHBd**	**MLogP**	**MLFER_S**
Tolerance	0.741	0.401		0.471	0.568	0.427	0.490
VIF	1.349	2.494		2.125	1.761	2.341	2.041

Table II.4 Statistical of multi-collinearity of 14 descriptors

Descriptors	**T(k)**	**P(Mpa)**	**AATS3v**	**MATS2e**	**GATS4c**	**GATS3v**	**GATS4e**
Tolerance	0.971	0.963	0.368	0.822	0.350	0.767	0.315
VIF	1.030	1.038	2.719	1.217	2.858	1.303	3.173
Descriptors	**GATS3s**	**nBondsM**	**AVP-0**	**SHBd**	**ETA_dEpsilon_B**	**MLogP**	**MLFER_S**
Tolerance	0.741	0.246	0.425	0.531	0.184	0.380	0.404
VIF	1.350	4.059	2.354	1.884	**5.449**	2.630	2.474

Table II.5 Statistical of multi-collinearity of 15 descriptors

Descriptors	**T(k)**	**P(Mpa)**	**AATS4v**	**ATSC2m**	**AATSC2v**	**GATS3v**	**GATS1e**	**GATS3s**
Tolerance	0.967	0.949	0.118	0.318	0.290	0.290	0.466	0.630
VIF	1.034	1.053	**8.452**	3.144	3.452	3.425	2.309	1.588
Descriptors	**AVP-0**	**Mp**	**ETA_dEpsilon_B**		**ETA_BetaP**	**MLogP**	**MLFER_E**	**JGT**
Tolerance	0.145	0.106	0.092		0.070	0.353	0.293	0.202
VIF	**6.878**	**9.390**	**10.897**		**14.317**	2.831	3.416	4.953

Table II.6 Statistical of multi-collinearity of 17 descriptors

Descriptors	**T(k)**	**P(Mpa)**	**AATS5v**	**ATSC7c**	**AATSC0p**	**GATS4e**	**GATS2s**	**SpMax3_bhm**	**ASP-7**
Tolerance	0.979	0.953	0.200	0.903	0.709	0.733	0.730	0.089	0.378
VIF	1.022	1.050	**5.009**	1.108	1.410	1.364	1.370	**11.286**	2.648
Descriptors	**nssCH2**	**ETA_dEpsilon_D**		**SIC1**	**CIC0**	**ZMIC5**	**piPC9**	**N5HeteroRing**	**JGl3**
Tolerance	0.280	0.471		0.257	0.124	0.310	0.111	0.670	0.584
VIF	3.576	2.125		3.894	**8.061**	3.230	**9.031**	1.492	1.713

Therefore, the selected descriptors were as follows: AATS3v, MATS2e, GATS4c, GATS3v, GATS4e, GATS3 s, nBondsM, AVP-0, SHBd, MLogP, and MLFER_S, with the addition of T (K) and P (MPa). More details on the characterization of those selected descriptors are available in **Table II.7**. In order to get more details on the correlation between the selected descriptors, the correlation matrix has been established. The result is shown in **Table II.8.** The value of the correlation coefficient of each pair of selected descriptors was ≤ 0.590, with one exception value equal to 0.758. This result proves that the selected descriptors were independent.

Table II.7 Characteristics of the descriptors selected in the optimal model and their VIF

Descriptors Name	Type	Class	Description "Physical-chemical meaning"	VIF
AATS3v	2D	Auto correlation	Average Broto-Moreau autocorrelation - lag 3 / weighted by van der Waals volumes	1.782
MATS2e	2D	Auto correlations	Moran autocorrelation of lag 2 weighted by Sanderson electro-negativity	1.226
GATS4c	2D	Auto correlations	Geary autocorrelation - lag 4 / weighted by charges	2.875
GATS3v	2D	Auto correlations	Geary autocorrelation of lag 3 weighted by van der Waals volume	1.292
GATS4e	2D	Auto correlations	Geary autocorrelation of lag 4 weighted by Sanderson electro-negativity	3.265
GATS3s	2D	Auto correlations	Geary autocorrelation of lag 3 weighted by I-state	1.247
nBondsM	2D	Bond Count	Total number of bonds that have bond order greater than one (aromatic bonds have bond order 1.5).	2.472
AVP-0	2D	Chi Path	Average valence path, order 0	2.090
SHBd	2D	Electro-topological State Atom	Sum of E-States for (strong) hydrogen bond donors	1.823
MlogP	2D	MannholdLogP	Moriguchi octanol-water partition coeff. (logP)	2.207
MLFER_S	2D	Molecular linear free energy relation "MLFER"	Combined dipolarity/polarizability	1.947
T(K)	-	-	Temperature	1.024
P(MPa)	-	-	Pressure	1.030

Table II.8 Matrices of correlation of the selected features

	T(k)	**P(Mpa)**	**AATS3v**	**MATS2e**	**GATS4c**	**GATS3v**	**GATS4e**	**GATS3s**	**nBondsM**	**AVP-0**	**SHBd**	**MLogP**	**MLFER_S**	**pLn Property**
T(k)	**1**	0.004	0.040	-0.063	-0.083	-0.035	-0.057	0.023	0.022	0.037	-0.060	0.034	0.002	-0.039
P(Mpa)	0.004	**1**	0.061	-0.018	-0.013	-0.047	-0.031	-0.111	0.076	0.003	-0.023	0.096	0.102	-0.225
AATS3v	0.040	0.061	**1**	0.016	0.023	-0.266	0.046	-0.128	0.528	-0.321	-0.138	-0.061	0.365	0.217
MATS2e	-0.063	-0.018	0.016	**1**	0.080	0.148	0.100	-0.283	-0.058	-0.202	0.175	-0.209	-0.094	-0.062
GATS4c	-0.083	-0.013	0.023	0.080	**1**	-0.022	**0.758**	0.172	-0.015	0.013	-0.080	0.168	0.057	-0.202
GATS3v	-0.035	-0.047	-0.266	0.148	-0.022	**1**	0.098	-0.056	-0.426	0.139	-0.073	-0.109	-0.278	-0.146
GATS4e	-0.057	-0.031	0.046	0.100	**0.758**	0.098	**1**	0.168	-0.122	-0.112	-0.029	-0.384	-0.199	-0.126
GATS3s	0.023	-0.111	-0.128	-0.283	0.172	-0.056	0.168	**1**	0.017	0.210	-0.015	-0.099	-0.023	-0.242
nBondsM	0.022	0.076	0.528	-0.058	-0.015	-0.426	-0.122	0.017	**1**	-0.313	0.003	0.252	0.590	0.008
AVP-0	0.037	0.003	-0.321	-0.202	0.013	0.139	-0.112	0.210	-0.313	**1**	-0.456	0.433	-0.046	-0.216
SHBd	-0.060	-0.023	-0.138	0.175	-0.080	-0.073	-0.029	-0.015	0.003	-0.456	**1**	-0.407	0.016	0.180
MLogP	0.034	0.096	-0.061	-0.209	0.168	-0.109	-0.384	-0.099	0.252	0.433	-0.407	**1**	0.377	0.039
MLFER_S	0.002	0.102	0.365	-0.094	0.057	-0.278	-0.199	-0.023	0.590	-0.046	0.016	0.377	**1**	0.250
pLn Property	-0.039	-0.225	0.217	-0.062	-0.202	-0.146	-0.126	-0.242	0.008	-0.216	0.180	0.039	0.250	**1**

II.2.4 Artificial Neural Network

Artificial neural networks (ANNs) are supervised learning computation systems consisting of a number of simple, extremely interconnected processing units, which were introduced by McCulloch and Pitts [356–359]. Such a system processes information by its dynamic state response to external inputs. These knowledge processing patterns were inspired by mimicking the neural network of the human brain [357]. The ANN method offers efficient performances for dynamic and nonlinear regression and classification problems [358]. The algorithm contains input, hidden, and output layers. Each layer consists of a set of interconnected processor components (neurons or nodes). Their output is the input layer of the next layer.

The output signal from one layer is attached to the next layer through means of weight factors by an intermediate that amplifies or weakens the signals [357]. A transfer function (activation function) suchlike a linear or sigmoid function was exploited to compute the outputs of nodes in the hidden and output layers. The number of neurons in the input and output layers is the number of input and output variables, respectively. There is no special way to select the number of neurons in the hidden layer; thus, the number of hidden layers and the nodes in each hidden layer is usually related to the complexity of the problem and the trial-and-error method [358,360]. There are various neural networks with different training algorithms. A multilayer perceptron (MLP) trained with Bayesian regularisation backpropagation (trainbr) was considered in this study. MLP are feed-forward multilayer networks that supply flexible frameworks for non-linear function prediction [361]. The ANN modelling process can be summarised in two main phases: (1) assigning network structure and (2) adjusting the weight of connections between neurons, as explained in [357]. More details can be found elsewhere [357,360–362].

II.2.5 ANN-QSPR Modelling

In this study, the dataset of 3971 EDP were randomly divided into two major sets. The first set have 3828 EDP that were applied to develop the model, while the second set of 143 was hidden from the model and kept as an external test (validation set). Using the selected 13^{th} features, the first set of 3828 EDP was taken to build the feed-forward MLP model, which has a single input, hidden and output layer.

The implementation of the algorithm was carried out under MATLAB® R2019b environment. The dataset was divided using MATLAB® *dividerand* function into two random

sets as follows: 70% (2695 EDP) for the training set and 30% (1155 EDP) for the test set. The net was trained using the training function trainbr "Bayesian regularisation back-propagation". The logsig "Log-sigmoid" function was selected as a transfer function for hidden layer and purelin "Linear transfer" function for the output layer. The number of hidden neurons was selected to 10 "13 nodes of the input×0.75~10". One output neuron was used to represent the observed solubility fraction (y_2). Several 25–30 iterations were carried out. The model was optimized based on their average absolute relative error (AARD%) as an objective function. Therefore, the MLP-ANN has the following architecture {13-10-1}.

Figure II.2 illustrates the architecture of MLP-ANN used in this study.

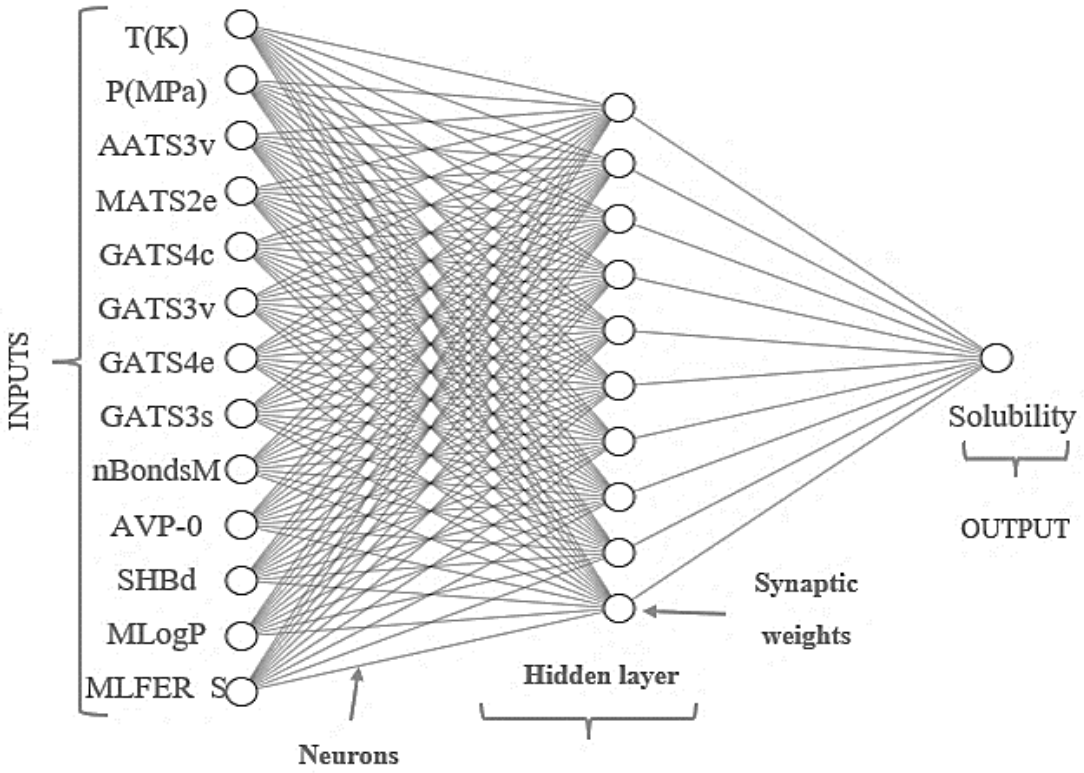


Figure II.2 MLP-ANN structure

II.2.6 Evaluation of the Model's Performance

Numerous statistics metrics were implemented to evaluate the performance of the QSAR model, as defined [363–366] by:

$$Q^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{i\,(training)}^{obs} - y_{i\,(training)}^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_{i\,(training)}^{obs} - \bar{y}_{(training)}^{obs}\right)^2} \quad \text{(II.1)}$$

$$R^2 = r^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{obs} - y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{obs} - \bar{y}^{obs}\right)^2} \quad \text{(II.2)}$$

$$R = r = \frac{\sum_{i=1}^{N}\left(y_i^{obs} - \bar{y}^{obs}\right)\cdot\left(y_i^{pred} - \bar{y}^{pred}\right)}{\sqrt{\sum_{i=1}^{N}\left(y_i^{obs} - \bar{y}^{obs}\right)^2}\cdot\sqrt{\sum_{i=1}^{N}\left(y_i^{pred} - \bar{y}^{pred}\right)^2}} \quad \text{(II.3)}$$

$$k = \frac{\sum_{i=1}^{N}\left(y_i^{obs}\cdot y_i^{pred}\right)}{\sum_{i=1}^{N}\left(y_i^{pred}\right)^2} \quad \text{(II.4)}$$

$$k' = \frac{\sum_{i=1}^{N}\left(y_i^{obs}\cdot y_i^{pred}\right)}{\sum_{i=1}^{N}\left(y_i^{obs}\right)^2} \quad \text{(II.5)}$$

$$r_0^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{obs} - k y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{obs} - \bar{y}^{obs}\right)^2} \quad \text{(II.6)}$$

$$r'^2_0 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{obs} - k' y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{obs} - \bar{y}^{obs}\right)^2} \quad \text{(II.7)}$$

$$m = \frac{\left(r_0^2 - r'^2_0\right)}{r_2} \quad \text{(II.8)}$$

$$n = \frac{\left(r_2 - r_0^2\right)}{r_2} \quad \text{(II.9)}$$

$$r_m^2 = r^2\left(1 - \sqrt{r^2 - r_0^2}\right) \quad \text{(II.10)}$$

$$r'^2_m = r^2\left(1 - \sqrt{r^2 - r'^2_0}\right) \quad \text{(II.11)}$$

$$\bar{r}_m^2 = \frac{\left(r_m^2 + r'^2_m\right)}{2} \quad \text{(II.12)}$$

$$\Delta r_m^2 = \left[r_m^2 - r'^2_m\right] \quad \text{(II.13)}$$

$$\bar{y}^{obs} = \frac{\sum_{i=1}^{N} y_i^{obs}}{N} \quad \text{(II.14)}$$

$$\bar{y}^{pred} = \frac{\sum_{i=1}^{N} y_i^{pred}}{N} \quad \text{(II.15)}$$

$$PRESS = \sum_{i=1}^{N}\left(y_i^{obs} - y_i^{pred}\right)^2 \quad \text{(II.16)}$$

$$SSY = \sum_{i=1}^{N}\left(y_i^{pred}\right)^2 \quad \text{(II.17)}$$

$$\mathrm{AARD\%} = 100 \cdot \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left| \frac{\mathrm{y_i^{obs}} - \mathrm{y_i^{pred}}}{\mathrm{y_i^{obs}}} \right| \quad \text{(II.18)}$$

$$\mathrm{MSE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \left(\mathrm{y_i^{obs}} - \mathrm{y_i^{pred}} \right)^2 \quad \text{(II.19)}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum \left(\mathrm{y_i^{obs}} - \mathrm{y_i^{pred}} \right)^2}{\mathrm{N}}} \quad \text{(II.20)}$$

$$MAE = \frac{1}{N} X \sum_{i=1}^{N} \left| y_i^{obs} - y_i^{pred} \right| \quad \text{(II.21)}$$

$$Bf = 10^{\left(\frac{\sum_{\mathrm{i}=1}^{\mathrm{N}} \log \frac{\mathrm{y_i^{obs}}}{\mathrm{y_i^{pred}}}}{\mathrm{N}} \right)} \quad \text{(II.22)}$$

$$Af = 10^{\left(\frac{\sum_{\mathrm{i}=1}^{\mathrm{N}} \left| \log \frac{\mathrm{y_i^{obs}}}{\mathrm{y_i^{pred}}} \right|}{\mathrm{N}} \right)} \quad \text{(II.23)}$$

Where, Q^2 is the cross-validated correlation coefficient (also notated CV, q^2), R^2 is the coefficient of determination, r is the correlation coefficient between the observed and predicted value of compounds with interception, k & k' are the slopes of the corresponding regression lines, r_0^2 is the squared correlation coefficient between the observed and predicted value of compounds without intercept, r'^2_0 it bears the same meaning as r_0^2 , but uses the reversed axes, $\bar{r}_m^2 \& \Delta r_m^2$ are theaverage and delta of r_m^2. $\mathrm{y_i^{obs}}$ is the observed (experimental) value of Y, $\mathrm{y_i^{pred}}$is the predicted (calculated) Y-value of training set, test set or validation set, N is the Number of data points in the data set (training, test...), $\bar{y}^{obs}$ is the Average of Y_{obs}, $\overline{\mathrm{Y}}_{\mathrm{pred}}$ is the Average of $\bar{y}^{pred}$, PRESS is the Predictive Residual Sum of the Squares, SSY is the Sum of the Squares of the response values of the total observations. AARD is the Average Absolute Relative Deviation, MSE is the Mean Squared Error, $RMSE$ is the Root-Mean Squared Error, MAE is the mean absolute error, Bf is the bias factor, and Af is the precision factor.

These statistical metrics were used to decide whether the model is sound and unbiased ("good model") or not. These metrics were used as acceptability criteria, as shown in **Table II.9**.

Table II.9 Acceptability criteria (A.C) of a model

Parameter	**m**	**n**	**R^2**	**Q^2**	**Q^2_{pred}**	**K**	**k'**	**$\bar{r}^2_m$**	**Δr^2_m**	**B_f**	**A_f**
A.C value	<0.1		>0.6	>0.5	>0.6	0.85< <1.15		>0.5	<0.2	0.9-1.05	

Moreover,

- R^2 -Q^2 show whether a model is overfitted or not. It must be less than 0.3.
- A PRESS/SSY ratio smaller than 0.4 shows that the QSAR model is reasonable. A ratio below 0.1, shows an excellent model, and confirms that the developed models predict better than chance.
- Perfect agreement between the observed values and the predicted values give a *Bf* (or A_f, after ignoring whether the difference between the predicted value and the observed value is positive or negative) of 1. A value greater or less than 1 shows an intrinsically safe. More detail can be found elsewhere [363–366].

Another interesting statistic parameter, namely, Akaike's information criterion (AIC) [367,368] was also calculated:

$$AIC = Nln\left(\frac{SSE}{N}\right) + 2n_p + \frac{2n_p(n_p+1)}{N-(n_p+1)} \quad \text{(II.24)}$$

$$SSE = \sum_{i=0}^{N}(y_i^{obs} - y_i^{pred})^2 \quad \text{(II.25)}$$

Where, N is the number of experimental samples, Np is the number of the parameters in the model and SSE is the sum of square error.

This statistical metric is used to compare different models and identify the best model that represents the experimental results. A small AIC value means that the model fitted well the experimental results. Normally, a single AIC value is meaningless; the AIC value of the model studied in this chapter was compared to three ANN models.

II.2.7 Applicability Domain

The applicability domain of this study was verified with the adaptation of William's plot. William's plot is a plot of standardized residuals versus leverage values, Hat diagonal values (hi) are usually used to visualize the applicability domain [369–371] and to identify outliers' responses. i.e., compounds with standardized residuals greater than two standard deviation units (Y outliers) and structurally influential chemicals (X outliers) in the model, i.e., hi>h*. In this plot, the applicability of the domain was recognized inside a squared area within ± 3 standard deviations and a threshold leverage value (h*).

The leverage (h) value of a compound in the original independent variable space and the warning leverage (h*) are defined as follows [372].

$$\mathbf{hi} = \mathbf{x_i}(\mathbf{X^T X})^{-1}\mathbf{x_i} \quad \text{(II.26)}$$

$$\mathbf{h}^* = \frac{\mathbf{3(p+1)}}{\mathbf{n}} \quad \text{(II.27)}$$

Where, x_i is the descriptor-row vector of the query compound, X k x n is a matrix containing the k descriptor values for each one of the n training compounds, n is the total number of samples in the training set, and p is the number of descriptors involved in the correlation.

II.2.8 Sensitivity Analysis

ANNs are powerful tools that can be used to predict and analyze easily non-linear relationships. Various authors highlighted the interest in adopting ANNs instead of linear statistical models. In fact, starting from input variables, ANNs can predict the output variable, but they often ignore the mechanisms that occur within the network are often ignored. ANNs are often recognized as black boxes. Several authors explored these problems and proposed algorithms to illustrate the role of variables in ANN models and the contribution of features. Determination of the contribution of each variable is of prime importance. This explanatory aspect of ANNs was measured in the present study. These methods were used to detect the

influence of each input variable and its contribution to the output. They are procedures used to estimate the relative contribution of each input variable.

Various different techniques that allow contribution analysis were used. The Garson method to calculate the relative importance of each input variable [373] was considered in this study, as well as the relevance factor *r*, which is in the range of −1 to +1 and is stated by the following equation [374].

$$r_k = \frac{\sum_{i=1}^{n}(X_{k,i}-\bar{X}_k)(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_{k,i}-\bar{X}_k)^2 \sum_{i}^{n}(Y_i-\bar{Y})^2}} \quad \text{(II.28)}$$

Where, $X_{k,I}$ is the i^{th} importing parameter, Y_i is the i^{th} exporting value, $\bar{X}_k$ is the average value of the k^{th} input,$\bar{Y}$ is the average value of exporting parameter, and *n* is the number of sets.

II.3 Results and Discussion

II.3.1 Model's Performance

Figure II.3 shows the scatter plots of the proposed model. i.e., predicted vs. observed values for the training, test, and global set. A close correlation between the values predicted by the ANN model and the observed values of the mole solubility fraction was determined, except for few outliers points, considered as accepted out of range. **Table II.10** represents the statistical metrics results of the three sets, **Tables II.11** and **II.12** shows the error parameter for the sets and the AIC values, respectively.

In summary, the results above are in agreement with the acceptability criteria. Where the non-linear ANN model gave good results with the highest global correlation coefficients (R) 0.976, as well as the best robustness (Q^2) 0.953, and the coefficients of determination (R^2) 0.953. In addition, to the error parameters that showed highly acceptable results such as the global RMSE (0.516) and AARD% (3.775). The calculation of the coefficient of determination minus the cross-validated correlation coefficient was lower than 0.3, indicating non-over-fitting

for this model. The calculated global PRESS value divided on the SSY were lower than 0.1. This proved the efficiency of the developed model, which predicts better than chance.

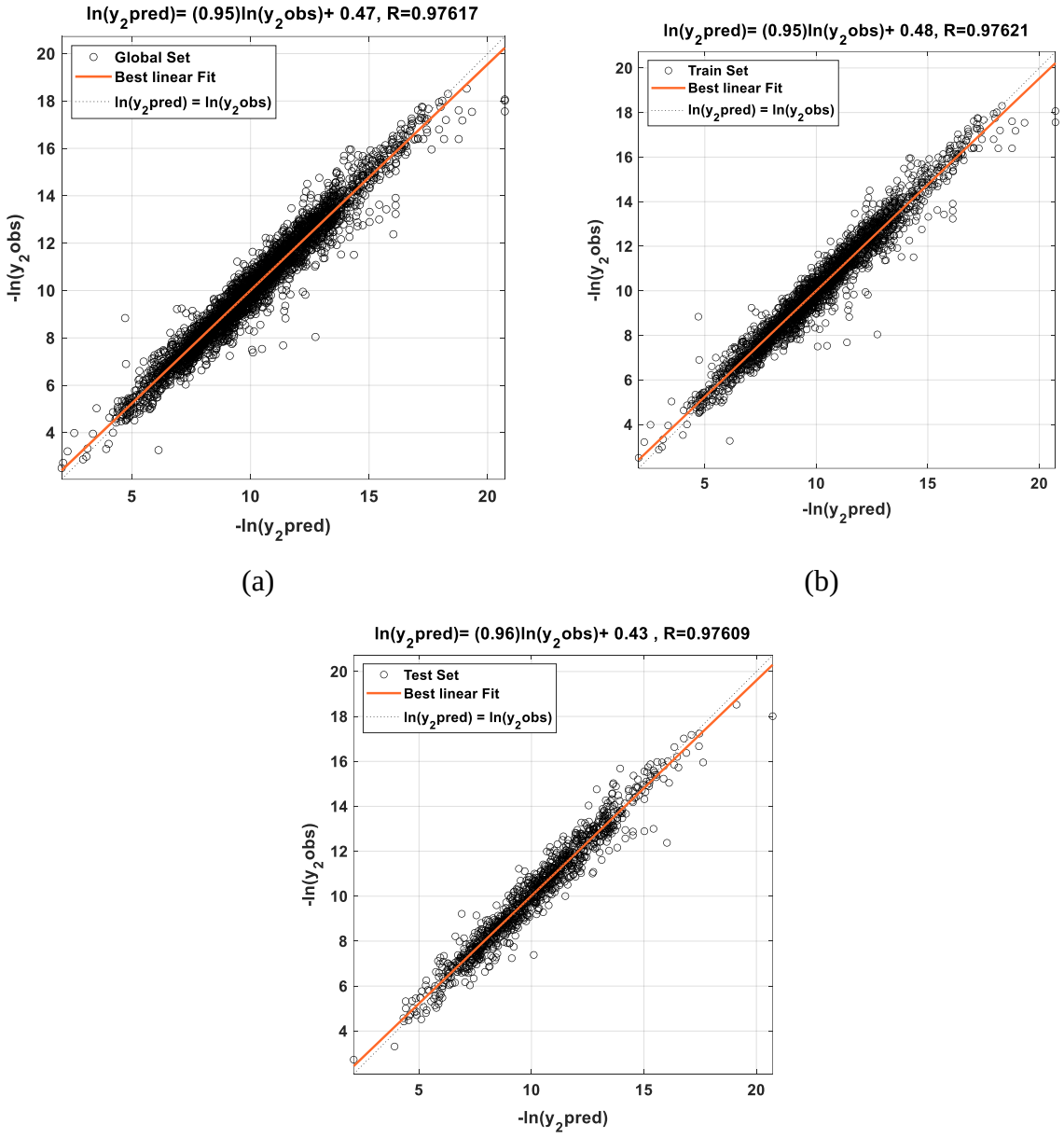


Figure II.3 Scatter plots Predicted vs. Observed values of the solubility for the MLP-ANN model: (a) Global Set, (b) Training Set, (c) Test Set, respectively.

Table II.10 Validation parameters for the train, test, and global sets for MLP-ANN model.

Statistical values	Train	Test	Global
Q^2	0.9529	0.9528	0.9528
R^2	0.9529	0.9528	0.9528
R	0.9762	0.9761	0.9761
k	1.0000	0.9975	0.9993
k'	0.9976	0.9998	0.9983
R_0^2	1.0000	0.9999	1.0000
$R'_0{}^2$	0.9999	1.0000	0.9999
m	-0.0415	-0.0472	-0.0432
n	-0.0414	-0.0473	-0.0432
R_m^2	0.7683	0.7594	0.7687
$R_m'^2$	0.7685	0.7695	0.7688
Bf	1.0014	1.0034	1.0020
Af	1.0379	1.0391	1.0383
$\bar{r}_m^2$	0.7684	0.7694	0.7687
Δr_m^2	0.0002	0.0001	0.0001

These results indicated that the ANN not only performed well in model development but also had excellent prediction capacity, suggesting a non-linear correlation between the solubility fraction and the relevant descriptors.

$$-\boldsymbol{ln(solubility)} = \boldsymbol{w_2} \cdot \left(\frac{\boldsymbol{Inputs}}{\boldsymbol{1-e^{-w_1 \cdot Inputs}}} + \boldsymbol{b_1}\right) + \boldsymbol{b_2} \quad \text{(II.29)}$$

The obtained w_1, w_2, b_1 and b_2 are provided in the annex.

In **Table II.12**, the values of AIC for the four ANN models clearly denote that the experimental data is well fitted by corresponding model of this work.

Table II.11 Different errors for the train, test, and global sets for MLP-ANN model.

Error	Train	Test	Global
AARD%	3.7338	3.8704	3.7748
RMSE	0.5084	0.5339	0.5162
MSE	0.2585	0.2851	0.2665
MAE	0.3651	0.3690	0.3621

Table II.12. AIC values of the four ANN models.

Model	Training function	Activation function		AIC
		Hidden layer	Output layer	
Corresponding Model	Trainbr	logsing	purelin	-5034.59
Model 1	Trainbr	tansing	purelin	-4895.65
Model 2	Trainlm	logsig	purelin	-4926.54
Model 3	Trainlm	tansing	purelin	-4527.42

II.3.2 William's Plot

The applicability domain (AD) of the proposed model was examined with William's plot [372]. The vertical line is the critical leverage value (h*), and the horizontal lines ±3 are the cutoff value for the cutoff value for Y_{space}. In **Figure II.4** it can observe that none of the drug-like compounds in the whole dataset have leverage higher than the warning h* value of 0.016. 2633 data points belonging to the train set were in-between the two horizontal lines,

which means 98.25% of the train set was covered. Out of 1148 data points of the test set, only 16 data point showed a distribution outside the two horizontal lines (98.6% of the test set was covered). 98.35% of the whole database was covered. It should be noted that only 63 EDP showed a distribution out of the two horizontal lines (1.65% outranges). Thus, these results show that the MLP-ANN model complied with the third principle of the OECD. Accordingly, the model developed in this study provided excellent predictions for 148 drug-like compounds (3828 EDP). It can be used to predict the molar solubility fraction of those compounds, particularly for those that have not been tested as well as new drugs.

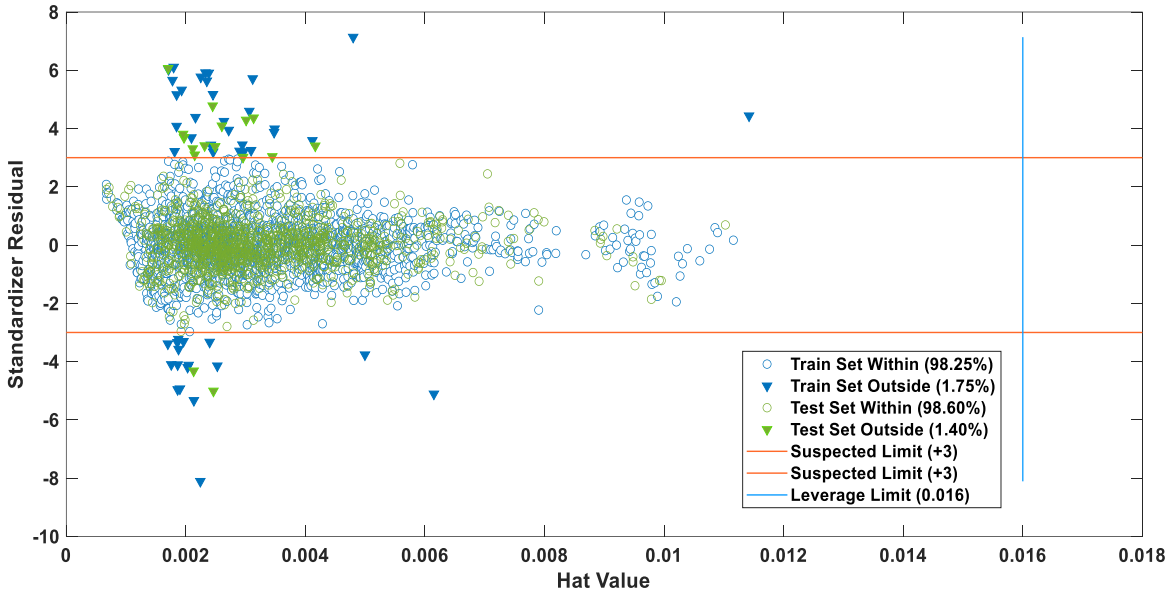
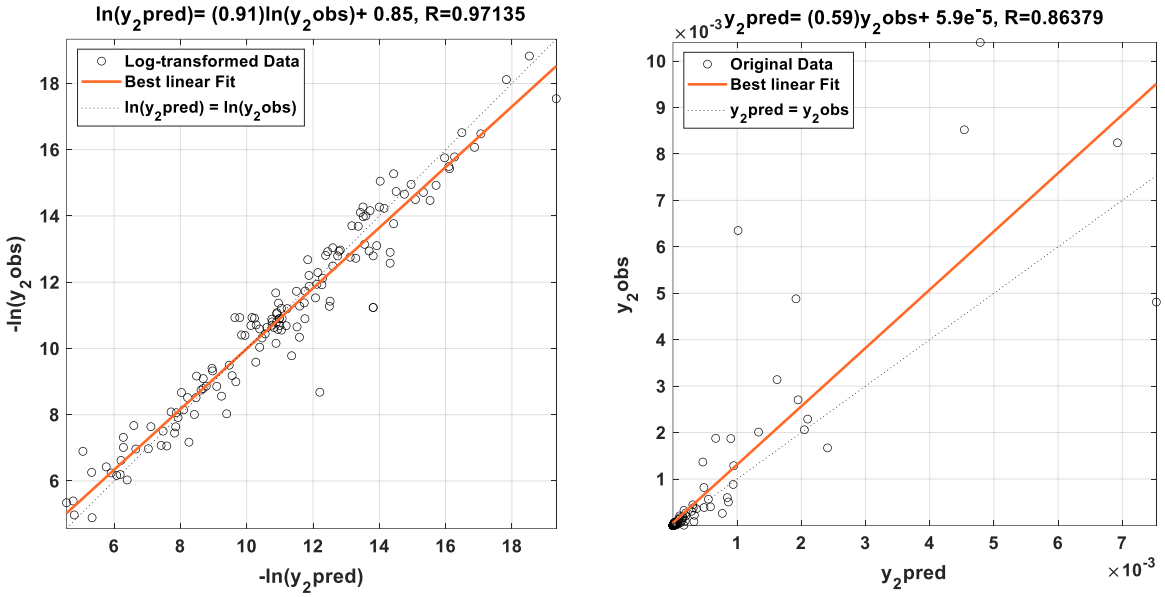


Figure II.4. William's plot for MLP-ANN model.

II.3.3 External Test

The hidden dataset from the model, which included 143 EDP, for the same selected number of features, was used for a further test of the models. i.e., prediction of the solubility fraction of a non-used dataset, known as the validation set or external test. The correlation between both log-transformed and original data is shown in **Figure II.5**, proving that this model can represent the predicted and the observed solubility fraction of validation set with a close positive correlation; while the **Table II.13** shows an acceptable error for both cases. To this end, the model is valid for prediction of future data.



(a) log-transformed data. (b) original data.

Figure II.5 Scatter plots Predicted vs. Observed values of the solubility for the MLP-ANN model, external test

Table II.13 Different errors for the validation set for both cases

Error	Log-transformed Data	original Data
AARD%	5.1982	99.3617
RMSE	0.7507	8.4365e-04
MSE	0.5635	7.1175e-07
MAE	0.5395	0.0013

II.3.4 Descriptor's Contribution and Interpretation

After the selection of the molecule descriptors, 11 pertinent descriptors with two intensive state variables (pressure and temperature) were used to develop the ANN-QSPR model, which were nBondsM, GATS4c, AVP-0, MLFER_S, AATS3v, SHBd, MLogP, GATS4e, GATS3s, GATS3v, MATS2e, P (MPa) and T (K). To evaluate the scope of each

descriptor for the prediction of solid drugs solubility, the relevance and the relative contributions of the 13 descriptors to the ANN model were determined and a description of these descriptors is shown:

The absolute value of the relevance factor has a direct relation to the output. As can be seen in **Figure II.6**, the solubility showed a straight dependency on the descriptors named: AATS3v, nBondsM, SHBd, MLogP, and MLFER_S, and an opposite dependency on the temperature and the pressure, and the rest of descriptors. In addition, MLFER_S, and pressure were the most relevant input variables, with a relevance factor of +0.25 and -0.24 respectively. Contrariety, the temperature and nBondsM were the least relevant, with a relevance factor of -0.03 and + 0.01, respectively.

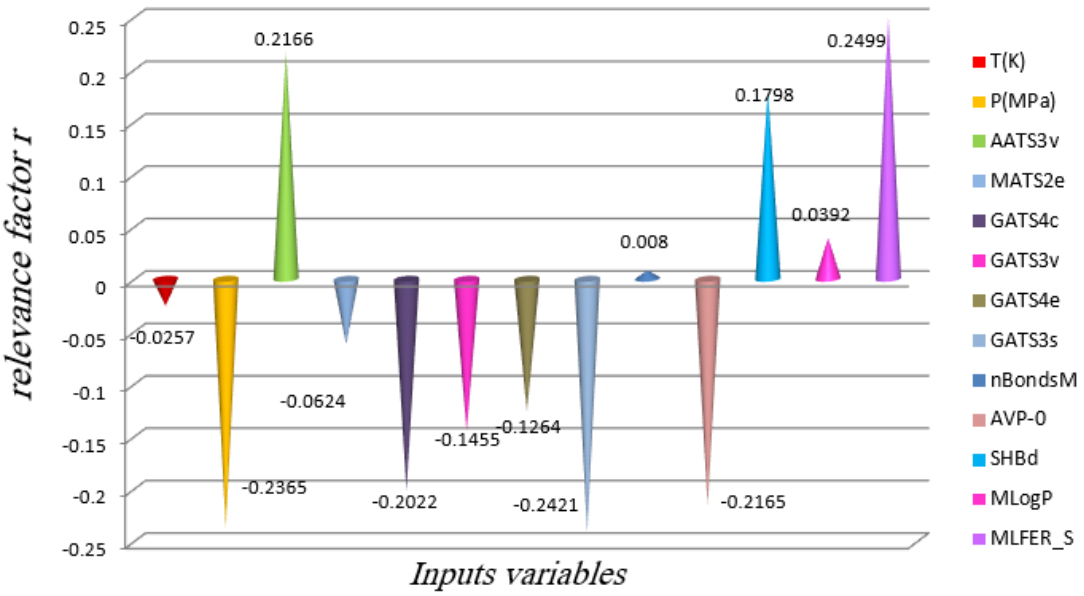


Figure II.6 Results of the relevance factor performed on the QSPR-ANN model

Figure II.7 depicts sensitivity analysis for ANN-QSPR model and shows that the GATS4e and GATS3s share the highest contribution (11.59% and 11.57, respectively) where the lowest contribution was by T (K) (1.94%). The rest of the inputs had an importance between 4.43% and 11.06%. The relative contribution level of the descriptors was found to be in the following order: GATS4e, GATS3s, AATS3v, GATS4c, SHBd,MLogP, MLFER_S, AVP-

0,nBondsM, GATS3v, MATS2e, P(MPa) and T(K) (**Figure II.7**). In this case, we can see that the variables that have the highest relative contribution were successively GATS4e, GATS3s, AATS3v.

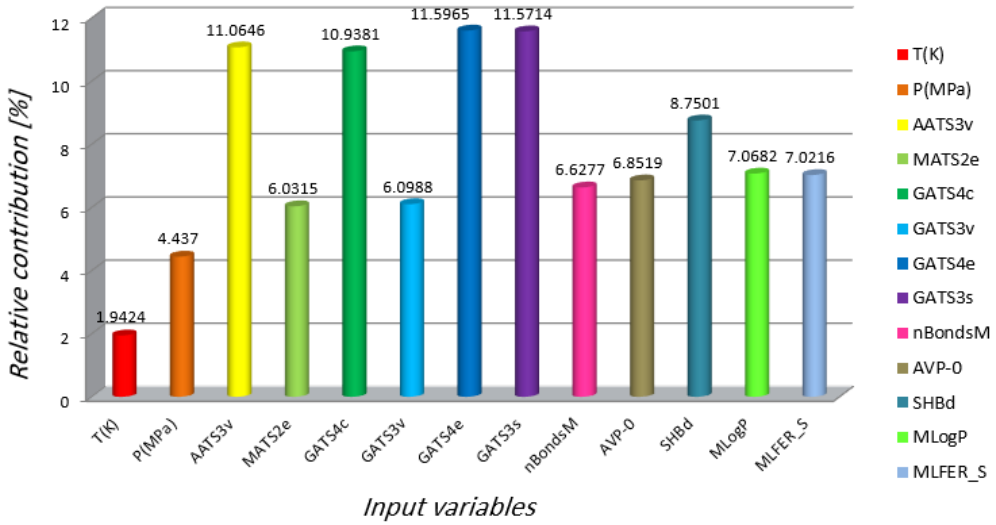


Figure II.7 Relative contribution of each input variable using the Garson method

GATS4c, AATS3v, GATS4e, are a 2D auto correlation molecular descriptor, which is known respectively as the geary autocorrelation-lag 4/weighted by charges, autocorrelation of lag 3 weighted by vdw volume, geary autocorrelation of lag 4 weighted by atomic Sanderson electro-negativity. SHBd describes the electropological state (E-state) in different aspects. SHBd represents the sum of E-states for (strong) hydrogen bond donors. An increase in the value of this descriptor means that the molecule is more and more hydrophobic. AVP-0 is a topological descriptor that gives information regarding the connectivity of various atoms in the molecule. This descriptor is able to take into account the presence of hetero-atoms in a molecule, as well as double and triple bonds. GATS3s is 2D auto correlation molecular descriptors, which describes how the considered property (atomic masses, polarizability, charges and electro-negativities) is distributed along a topological molecular structure. On the other hand, GATS3v and MATS2e are defined as Geary autocorrelation of lag 3 weighted by van der Waals volume and Moran autocorrelation of lag 2 weighted by Sanderson electro-negativity, respectively. The molecular descriptor MLOGP (which is the square root of MLOGP2) is considered as a measure of lipophilicity of a molecule. The MLFER_S descriptor is a descriptor for combined dipolarity/polarizability. It is one of the molecular linear free energy relations. The 2D descriptor “nBondsM” is defined as the dimensionless total number

of bonds that have bond order greater than one (for example aromatic bonds have bond of 1.5) and is encoded to an unsaturation. It referred to as the number of non-kekulized structures. In summary, in this model, it is concluded that the lipophilicity, electro-topology and electronic properties (polarizability, electro-negativity, atomic masses, van der Waals volume, unsaturation, and charges) of a molecule are of major importance in the solid drugs solubility, whereas temperature and pressure does not appear to be a major factor. Although both are very important in experiments. The mathematical formula for predicting the molar solubility fraction of drug-like compounds obtained using XLSTAT® (**Addinsoft**, New York, NY,USA) is given by:

$$\begin{aligned} Ln(y_2) = -(&22.4581 - 7.2613 \cdot 10^{-3} \cdot T(k) - 9.0914 \cdot 10^{-2} \cdot p(MPa) + 0.0184 \\ &\cdot AATS3v - 2.1674 \cdot MATS2e - 2.9553 \cdot GATS4c - 3.6541 \\ &\cdot GATS3v + 1.6024 \cdot GATS4e - \ 2.3767 \cdot GATS3s - 0.1746 \\ &\cdot nBondsM - 11.0664 \cdot AVP0 + 0.6300 \cdot SHBd + 0.5326 \cdot MLogP \\ &+ 1.0102 \cdot MLFER_S \end{aligned} \qquad \text{(II.30)}$$

II.3.5 The use of QSAR for Regulatory Purposes

This study aims to use QSARs combined with an ANN as an alternative for experimental data to predict and correlate drug solubilisation in supercritical carbon dioxide at specific temperature and pressure values. To facilitate the consideration of a QSAR model for regulatory purposes, the development of QSPR models should comply to the five principles for the QSAR validation drawn by the Organisation for Economic Cooperation and Development (OECD) [375], namely: (1) a defined endpoint; (2) an unambiguous algorithm; (3) a defined domain of applicability; (4) appropriate measures of goodness of fit, robustness, predictability; (5) a mechanistic interpretation, if possible. These criteria were applied to the newly ANN-QSPR model developed in the present work:

1. *OECD criterion 1:* the first step of the QSAR development was the selection of the appropriate endpoint (s). The selected endpoints solubility for drug and drug-like compounds was further defined according to the following endpoints criteria:

- Target: Molar solubility fraction.
- Solid compounds: drug or drug-like compounds.
- Compounds solubility in SC-CO_2.
- Temperature range: 298-373.15 k.
- Pressure range: 80-500 bars.
- Solubility mole fraction range: 0.000000001-0.131.
- Molecular descriptor: PaDEL Descriptors of the solid compounds.
- Molecular descriptor range: AATS3v (114.80-411.05), MATS2e (-0.25-0.45), GATS4c (0.46-1.66), GATS3v (0.66-1.39), GATS4e (0.38-1.91),GATS3s (0-1.91), nBondsM (1-27), AVP-0 (0.47-0.73), SHBd (0-2.36), MLogP (1.35-4.98), and MLFER_S (0.22-6.19).

The QSPR-ANN model predicts a defined endpoint.

2. *OECD criterion 2:* the MLP-ANN model developed in this present work is a clear and an easily applicable algorithm carried out under MATLAB® R2019b environment. Both training and activation functions are defined, the number of layers and the number of nodes in each layer is already given, which fulfil the criterion on an unambiguous algorithm. The transparency is ensured.

3. *OECD criterion 3:* QSAR models are only applicable within a specific set of conditions known as applicability domain, which identifies limitations and scope of the model. The applicability domain of the ANN-QSPR model developed was analyzed using a Williams plot, and it is well documented.

4. *OECD criterion 4:* The goodness of the model was reached by verifying the model fitting and the model robustness by calculating various statistical criteria, including the determination coefficient (R^2 >0.9) and the cross-validation coefficient (Q^2>0.9). Executing Y-scrambling to confirm the existence of a correlation between the selected modelling features and the studied response. The internal performance ends up indicating a well-founded (not by chance) original model. Moreover, the predictability (as determined by external validation) was reached by characterizing the correlation between predicted and experimental values in the validation set for both log-transformed and original data. In addition, the AARD, RMSE, MSE

that summarizes the overall errors of model, were used to measure and compare prediction accuracy in the training, test and validation sets. The error values were close and there were no significantly large residual values for the validation set, demonstrating that the model had not over-fitting. The model exhibited a satisfactory goodness-of-fit, robustness and predictability.

5. *OECD criterion 5:* the mechanistic interpretation of solid drug solubility in SC-CO_2 shows to be mostly influenced by the pressure, lipophilicity, electro-topology and electronic properties of a molecule, while the temperature has the lowest influence. However, experimental measurements indicate that temperature has a remarkable influence on the solubility.

For this, the developed ANN-QSPR model for drug and drug-like compounds solubility in SC-CO_2 meets all the OECD criteria for use in a regulatory context.

II.4 Conclusions

In this chapter, supercritical carbon dioxide solubility data of 148 drug compounds collected from the literature were applied to develop a combined ANN-QSPR approach for the prediction of drugs' solubilities in SC-CO_2. 11 molecular descriptors were recognized as relevant after reducing the number of molecular descriptors using numerous reduction methods. Then, a total of 3971 experimental data points, representing the solubility (mole fraction) of the drug in SC-CO_2 at different temperatures (K) and pressure (MPa), were used for training and validation of the proposed model. The architecture of the optimized network was achieved by varying different network parameters. The best model was obtained using the training function trainbr "Bayesian regularization back-propagation", logsig "Log-sigmoid transfer function" activation function for the hidden layer, network architecture of {13-10-1}.

Statistical studies suggested were tested to assess the predictive performance of the developed model. The results indicated that all the metrics fell within the domain of acceptability, therefore confirming the model's validity. All validations proved that the built QSAR model was robust, satisfactory, meets all the OECD principles for QSAR validation, and can be used to correlate reliably the solubility of drug-like compounds, for interpolation purposes for drugs of the data set within the model space and for extrapolation purposes for drugs not included in the data set...

The results of this study were found very satisfactory. However, this work can be furthermore improved with the application of new modeling techniques, like hybrid approaches, for correlating such a property. More details are exposed in the upcoming chapter.

Chapter ৩৩৩

Second Contribution

Chapter III

Second Contribution: A New Hybrid Quantitative Structure Property Relationships-Support Vector Regression (QSPR-SVR) Approach for Predicting the Solubility of Drug Compounds in Supercritical Carbon Dioxide

III.1 Introduction

The ability to predict and to correlate the molar fraction solubilities of drug compounds in supercritical carbon dioxide (SC-CO_2) is a key strength that should be useful in the pharmaceutical area. The support vector regression (SVR) algorithm is one of the main supervise-learning techniques used for predicting. However, this technique reported difficulty in defining its hyper-parameters. Classical techniques for optimizing are usually time-consuming and do not provide an exact or feasible solution. Therefore, there is a necessity to implement meta-heuristic algorithms. The term meta-heuristics refers to a global search technique or all modern nature-inspired optimization algorithms, used in various problems where the solution methods are inexact and near-optimal. The algorithms are highly effective and have general applications with high performance and a great driving force for optimality which makes them more suitable to be used than traditional mathematical methods for parameter estimation for such nonlinear systems.

The aim of this chapter is to select the best optimization algorithm for fine-tuning the support vector regression algorithm, in the form of a comparison study between seven (7) selected meta-heuristic algorithms: Dragonfly optimizer (DA), Ant Lion optimizer (ALO), Grey Wolf optimizer (GWO), Artificial Bee Colony optimizer (ABC), Particle Swarm optimizer (PSO), Whale optimizer (WOA) and a hybrid Particle Swarm with Grey Wolf optimizer (HPSOGWO). Note that the same structure of the previous chapter will be adopted in this chapter.

III.2 Procedures

III.2.1 Experimental DataBase

A database of 168 drug compounds accounting for 4490 solubility data points (mole fraction 0.000000001–0.131) measured in supercritical carbon dioxide at a temperature ranges from 298 K to 373.15 K and a pressure range of 80 to 500 bars were collected from the literature. These data were updated from previous a study [39].

As explained in Chapter II, a QSPR\QSAR modelling was performed on 148 drug compounds (3971 EDP) in order to determine the best-relevant combinations of features that can predict and correlate this property (see **Table A2**). According to the results, the proposed model meets all the OECD principles for QSAR validation and indicated that a combination of 13 descriptors is satisfactory for modelling the solubility.

Therefore, the selected PaDEL-descriptors were as follows: AATS3v, MATS2e, GATS4c, GATS3v, GATS4e, GATS3s, nBondsM, AVP-0, SHBd, MLogP, and MLFER_S, with the addition of T (K) and P (MPa). The output property was converted to the corresponding log-transformed property "-Ln(y2)" to guarantee the linear distribution as reviewed in [39] (see Chapter II). The final dataset was divided into two main sets. The first set of 4330 EDP was retained for modelling; while the second set of 160 EDP was kept hidden from the models, i.e., an external test.

III.2.2 Support Vector Regressor

The support vector machines are classification (SVC) and regression (SVR) techniques that use the machine learning concepts to increase the predictive accuracy. It was first suggested by Vladimir Vapnik in 1995 for the purpose of surpassing the traditional neural network algorithms that had suffered from severe difficulties with generalisation and producing models [376,377]. For regression modelling, the SVR algorithm is based on mapping the input vector x into a higher-dimensional feature space F using a kernel function and achieving the optimum hyper-plane that separates them [378–383] via, a nonlinear projecting function φ (*x*), and then to conduct a linear regression in F space [381–383]. The linear regression function is given as follows [272,380,384]:

$$f(x) = \omega \cdot \varphi(x) + b \tag{III.1}$$

Where φ (*x*) is the kernel function, *w* and *b* denote the weight vector and the bias term, respectively. That can be obtained by minimising the cost function [380,385]:

$$cost\ function = \frac{1}{2}\,\omega^2 + C\sum_{i=1}^{k}(\zeta_i^- + \zeta_i^+) \tag{III.2}$$

$$Subject\ to \begin{cases} y_i - (w \cdot \varphi(x_i) + b) \le \varepsilon + \xi_i^+ \\ (w \cdot \varphi(x_i) + b) - y_i \le \varepsilon + \xi_i^- \\ \xi_i^+, \xi_i^- \ge 0, \quad i = 1,2,3,\dots,n \end{cases}$$

A standard dualisation technique using Lagrangian multipliers was emploied for easily solving the optimisation problem in Equation. (III.2). The dual formulation of this problem is represented as follows [386]:

$$Maximize \left\{ -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i, \alpha_i^*)\left(\alpha_j, \alpha_j^*\right)(\varphi(x_i) \cdot \varphi\left(x_j\right)) - \varepsilon\sum_{i=1}^{n}(\alpha_i, \alpha_i^*) + \sum_{i=1}^{n} y_i(\alpha_i, \alpha_i^*) \right\} \tag{III.5}$$

$$Subject\ to \begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i \le C \quad i = 1,2,3,\dots,n \\ 0 \le \alpha_i^* \le C \quad i = 1,2,3,\dots,n \end{cases}$$

Here α_i, α_i^* are nonlinear Lagrangian multipliers. Resolving the dual maximisation problem in Equation. (III.3) provides the SVR function:

$$F(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)(\ \varphi(x_i) \cdot \varphi\left(x_j\right)) + b \tag{III.4}$$

The vector inner-product $(\varphi(x_i) \cdot \varphi\left(x_j\right))$ signifies the mapping function. The function can be replaced by a kernel function $k\left(x_i, x_j\right)$ as shown in Equation. (III.4):

$$F(x, \alpha_i, \alpha_i^*) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) k(x_i, x_j) + b \quad \text{(III.5)}$$

In this work, the Gaussian kernel function was selected Equation. (III.6):

$$k(x_i, x_j) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{(III.6)}$$

III.2.3 Meta-heuristic Approaches

III.2.3.1 Dragonfly Algorithm

The dragonfly algorithm (DA) is an optimiawzion technique that was first introduced by Seyedali Mirjalili in 2015 [387]. This approach is inspired by the static and dynamic behaviour of dragonflies in nature. In the static swarm process, the dragonflies hunt prey; the dragonflies behave equally to the exploitation phase in meta-heuristic algorithms in which they fly in small groups frequently over a well-determined small area and much closer to the land. While the dynamic behaviour is similar to the exploration phase, migratory swarms form as large groups (hundreds of thousands of dragonflies) and fly in a single direction for long distances to find food resources [356,386,387]. The life cycle of agents is divided into two different phases: the nymph phase and the adult phase [388]. The natural behaviour of each dragonfly in the swarm obliges the move toward nurturing sources and distract outward enemies [356,386]. Based on Reynolds' paper [389], the behavior of swarm follows five different principles that are essential in finding the weights solution [356,386,387,389]:

1. *Separation*

This step aims to avoid the collision of individuals with their neighbours that is close to its position, the static swarm. It is given as follows:

$$S_i = \sum_{j=1}^{n} X - X_j \quad \text{(III.7)}$$

Where X and X_j are the positions of the current individual and the j^{th} neighbouring individual, respectively. n is the number of neighbouring individuals.

2. *Alignment*

This step denotes the velocity matching between dragonflies of the same group. It is defined by:

$$A_i = \frac{\sum_{j=1}^{n} V_j}{n} - X \quad \text{(III.8)}$$

Where v_j is the velocity of neighbouring individual j.

3. *Cohesion*

The cohesion refers to the tendency of members towards the center of the swarm's group, the neighborhood mass. It can be computed as:

$$C_i = \frac{\sum_{j=1}^{n} X_j}{n} - X \quad \text{(III.9)}$$

4. *Attraction to food*

Typically, an individual's goal is to survive. The purpose of attraction to food is that all dragonflies must move toward the food. As shown following this mathematical formula:

$$F_i = X^+ - X \quad \text{(III.10)}$$

Where X^+ shows the position of the food source.

5. *Distraction from enemy*

The final step refers to when individuals move far away from the enemy's sources in order to survive; it is calculated as follows:

$$E_i = X^- - X \quad \text{(III.11)}$$

Where X^- shows the position of the enemy.

Each dragonfly updated their position with the aim of testing other weight solutions and getting another fitness value is attained by calculating ΔX and X using Equation. (III.12) and Equation. (III.13) [356,387]:

$$\Delta X_i = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta X_t \quad \text{(III.12)}$$

$$X_{t+1} = X_t + \Delta X_{t+1} \quad \text{(III.13)}$$

$$e = 0.1 - 1 * (\frac{0.1}{\frac{I}{2}}) \quad \text{(III.14)}$$

$$w = 0.9 - i * (\frac{0.9-0.4}{I}) \quad \text{(III.15)}$$

Where s, a, c, f, e, and w are the weights of their corresponding element. w is calculated using Equation. (III.15). i is the current iteration and I is the number of iterations, and e is calculated in Equation. (III.14). s, a, and c are three different random numbers between 0 and 2e; f is a random number between 0 and 2. More details can be found at [387].

III.2.3.2. Ant Lion Optimizer

This method is inspired by the hunting mechanism of antlions in nature and is introduced by Mirjalili [390] in 2015. The antlions (doodlebugs) belong to the *Myrmeleontidae* family of insects and *Neuroptera* order. They have a medium lifespan of approximately three years that it is spent as larvae except for 3-5 weeks of that period is spent in adulthood [391]. At the larvae stage, ant lions are recognised for their unique process of hunting and preferably of hunting ants. Using their massive jaw, antlions dig conical pits in the sand. Then, the larvae hide and wait at the bottom of that pit, sharp enough for the prey to be trapped in (most cases, ants). Once the antlion observed that the prey are inside the trap, they started to throw sands out of the hole across the prey. Therefore, the ants will fail to escape from the pit and slide to the bottom of that pit. After that, antlions would consume the prey and thrown away the leftovers. With this strategy, individuals prepare the pit for the next hunt. As well as the chance of survival increases [391,392].

The mathematical representation of this method is defined as follows [390–392]:

1. Random walks of Ants

$$X(t) = [0, cumsum(2r(t_1 - 1), cumsum(2r(t_2) - 1, \cdots, cumsum(2r(t_n) - 1)] \quad \text{(III.16)}$$

Where n is the maximum number of iterations, cumsum calculates the cumulative sum, and t is the step of the random walk. Hence,

$$r(t) = \begin{cases} 1 & if\ rand > 0.5 \\ 0 & if\ rand < 0.5 \end{cases} \quad \text{(III.17)}$$

Here, $r(t)$ is a stochastic function and rand is a random number generated with uniform distribution at interval of [0, 1]. The positions of ants are saved and considered during optimisation in the matrix:

$$M_{ant} = \begin{bmatrix} A_{1,1} & A_{1,2} & \cdots & \cdots & A_{1,d} \\ A_{2,1} & A_{2,2} & \cdots & \cdots & A_{2,d} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{n,1} & A_{n,2} & \cdots & \cdots & A_{n,d} \end{bmatrix} \quad \text{(III.18)}$$

Where M_{Ant} is the matrix for saving the position of each ant, $A_{i,j}$ shows the value of the j^{th} variable of i^{th} ant, n is the number of ants, and d is the number of variables. Based on the random walk, ants update their position with each step of optimization. The min-max normalisation equation is used to normalise the random walks:

$$X_i^t = \frac{(X_i^t - a_i) \times (d_i - c_i^t)}{(d_i^t - a_i)} + c_i \quad \text{(III.19)}$$

where a_i is the minimum of random walk of i^{th} variable, d_i is the maximum of random walk of i^{th} variable, c_i^t is the minimum of i^{th} variable at t^{th} iteration, and d_i^t is the maximum of i^{th} variable at t^{th} iteration.

2. Trapping in Antlion's pit

Antlion's trap affects the random walks of ants. It is defined as follows:

$$c_i^t = Antlion_j^t + c^t$$
$$d_i^t = Antlion_j^t + d^t \quad \text{(III.20)}$$

Where c^t represents the minimum of all variables at t^{th} iteration, d^t indicates the vector including the maximum of all variables at t^{th} iteration, c_i^t is the minimum of all variables for i^{th} ant, d_i^t is the maximum of all variables for i^{th} ant, and $Antlion_j^t$ shows the position of the selected j^{th} antlion at t^{th} iteration.

3. Building trap

In this phase, the ALO involved a roulette wheel operator to choice antlions according to their fitness during optimization. This algorithm offers high chances of catching ants by the fitter antlions.

4. Sliding Ants toward Antlion

At this point, the radius of the random walks hypersphere of ants is decreased adaptively in the mathematical modelling by applying the following Equation.(III.21):

$$c^t = \frac{c^t}{I}$$
$$d^t = \frac{d^t}{I} \quad \text{(III.21)}$$

Where I is a ratio, c^t is the minimum of all variables at t^{th} iteration, and d^t indicates the vector, including the maximum of all variables at t^{th} iteration.

5. Catching Pray and Rebuilding the Pit

This is the final stage of hunting. Based on the position of the latest hunted ant, the antlions update their position to improve the possibility of catching new prey. It is mathematically formulated as follows:

$$Antlion_j^t = Ant_i^t \; if \; f(Ant_i^t) > f\left(Antlion_j^t\right) \quad \text{(III.22)}$$

Where t represents the current iteration, $Antlion_j^t$ is the position of the selected j^{th} antlion at t^{th} iteration, and Ant_i^t represents the position of i^{th} ant at t^{th} iteration.

6. Elitism

In the ALO algorithm, an elite represents the best-obtained antlion over the whole iteration. This elite affects the movement of the rest of the ants during iteration. Hence, it is implicit that every prey walks randomly around a selected antlion by the roulette wheel and the elite altogether as follows:

$$Ant_i^t = \frac{R_A^t + R_E^t}{2} \quad \text{(III.23)}$$

Where R_A^t is the random walk around the antlion selected by the roulette wheel at t^{th} iteration, R_E^t is the random walk, and Ant_i^t represents the position of i^{th} ant at t^{th} iteration.

III.2.3.3 Gray Wolf Optimizer

The Grey wolf optimizer is also a meta-heuristic algorithm proposed by Mirjalili et al., in 2014 [393] for solving optimized problems. This algorithm is primarily inspired by the social leadership and hunting behaviour of grey wolves in nature [394]. Grey wolves are social animals that live in a pack of 5-12 members. In each path, there are four types of wolves divided based on their responsibilities and decisions making roles, during the process of prey hunting,

to alpha (α) wolf as the group leader, beta (β) wolves as the second ones in command, then delta (δ) wolves the subordinate ones, those three are considered as the best solution to lead the rest wolves known as omega (ω) wolves, the fourth type, toward promising areas for the aim of retching the global solution [380,394,395]. Those wolves have a special mechanism of hunting, considering three main stages. In the beginning, grey wolves approach the prey using a process of tracking and chasing. Then, they encircle the prey and harass it until the prey stops moving. The final step is to attack them, as mentioned in [396]. The GWO algorithm follows the following mathematical technique of hunting [380,393–396]:

1. Encircling

The Encircling behaviour of prey by grey wolves is mathematically computed as shown:

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)\right| \quad \text{(III.24)}$$

$$\vec{X}(t+1) = \vec{X}_P(t) - \vec{A} \cdot \vec{D} \quad \text{(III.25)}$$

Here, t is the current iteration, $\vec{X}_P$ is the vector of the prey position, and $\vec{X}$ indicates the vector of the grey wolf position, $\vec{A}$, $\vec{C}$ are coefficient vectors which can be obtained by:

$$\vec{A} = \overrightarrow{2a} \cdot \vec{r}_1 - \vec{a} \quad \text{(III.26)}$$

$$\vec{C} = 2 \cdot \overrightarrow{r_2} \quad \text{(III.27)}$$

Where components of $\vec{a}$ are linearly decreased from 2 to 0, over the course of iterations and $\vec{r}_1$ &$\vec{r}_2$ are random vectors in [0, 1].

2. Hunting

The mathematical modelling, α (the best candidate solution), β, and δ are assumed to have better knowledge about the position of the prey. As a consequence, it is mandatory to save the first three best solutions attained so far and to update the positions of the remaining wolves (ω) according to those solutions. The formulas of hunting mechanism are set as follows:

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1} \cdot \overrightarrow{X_\alpha} - \vec{X}\right|, \overrightarrow{D_\beta} = \left|\overrightarrow{C_2} \cdot \overrightarrow{X_\beta} - \vec{X}\right|, \overrightarrow{D_\delta} = \left|\overrightarrow{C_3} \cdot \overrightarrow{X_\delta} - \vec{X}\right| \quad \text{(III.28)}$$

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1} \cdot \left(\overrightarrow{D_\alpha}\right), \overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2} \cdot \left(\overrightarrow{D_\beta}\right), \overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_3} \cdot \left(\overrightarrow{D_\delta}\right) \quad \text{(III.29)}$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \quad \text{(III.30)}$$

Where C_1, C_2, and C_3 are calculated by Equation. (III.27), Xα, Xβ, and Xδ are the first three best solutions at iteration t. A1, A2, and A3 are calculated as in Equation. (III.26), and D_α, D_β, and D_δ are defined as in Equation. (III.28).

Using the objective function each time the wolves are repositioned, an iteration occurred, and the solutions are re-evaluated. Then, the alpha, beta, and delta wolves become the three best solutions, respectively. On the other hand, the omega wolves update their positions in the next iteration. These moves are repeated until the stopping criterion is met.

3. Attacking

The hunting mode ends with the moment that wolves start the attack. This can be mathematically represented by the value of $\vec{a}$ which is linearly decreased over the course of iterations controlling the exploration and exploitation:

$$\vec{a}(t) = 2 - (2 \times t)/MaxIter \quad \text{(III.31)}$$

Between the prey position and their current position, wolves change their positions randomly. More details can be found at [393].

III.2.3.4 Artificial Bee Colony

The artificial bee colony algorithm is suggested by karaboga [397] in 2005, then its performance was investigated in 2007 [398]. The inspiration comes from the behaviours of honeybees in finding food sources and sharing that information with the others in the nest. In the algorithm, individual bees are divided into three types (employed, onlooker, and scout)

where each agent plays different roles in the process, as illuminated by [356]. The process of the ABC algorithm is shown below [399]:

1. Initialization

At the start, a percentage (%) of the population is randomly sprayed into the solution space. Then, the fitness value, the nectar amount, is determined. The percentage (%) represents the ratio of employed bees compared to the total population.

2. Move the Onlookers

The probability of selecting a food source is obtained by the following Equation.(III.32):

$$P_i = \frac{F(\theta_i)}{\sum_{k=1}^{S} F(\theta_k)} \quad \text{(III.32)}$$

Where θ_i denotes the position of the ith employed bee, $\mathcal{S}$ represents the number of employed bees, and P_i is the probability of selecting the ith employed bee. After selecting a food source for the onlooker bees to move to by the roulette wheel selection, the amount of nectar is determined. This movement follows Equation. (III.33):

$$x_{ij}(t+1) = \theta_{ij} + \phi(\theta_{ij}(t) - \theta_{kj}(t)) \quad \text{(III.33)}$$

Where x_i denotes the position of the ith onlooker bee, t denotes the iteration number, θ_k is the randomly chosen employed bee, j represents the dimension of the solution and $\phi(\cdot)$ produces a series of a random variable in the range $[-1, 1]$.

3. Move the Scouts

The scouts are mathematically moved by the given Equation. (III.34):

$$\theta_{ij} = \theta_{ijmin} + r \cdot (\theta_{ijmax} - \theta_{ijmin}) \quad \text{(III.34)}$$

Where r is a random number and $r \in [0, 1]$.

The upcoming steps aim to update the fitness value and the position of the best food source obtained so far as well to check if the termination condition is satisfied or not. Based on the response, the algorithm outputs the results or repeats the program from step 2.

III.2.3.5 Particle Swarm Optimizer

The particle swarm optimizer is a well-known population-based meta-heuristic optimization techniques. This algorithm was first introduced by Kennedy and Eberhart [400–402] in 1995. Its strategy was inspired by the social behavior of bird flocking, fish schooling, a swarm of bees, and sometimes from the social behaviour of humans when it searches for food [402,403].

Firstly, the PSO used the search domain to generate a random initial population. The position information of the best particle and the best location of each particle in the swarm are permanently saved in the memory. By applying the Equations. (III.35) and (III.36) all particles in the swarm can update their positions [403]:

$$\bar{x}_{n+1}^{i} = \bar{x}_{n}^{i} + \bar{v}_{n+1}^{i} \quad \text{(III.35)}$$

$$\bar{v}_{n+1}^{i} = \omega \bar{v}_{n}^{i} + c_1 r_1 \left(\bar{p}_{n}^{i} - \bar{x}_{n}^{i}\right) + c_2 r_2 \left(\bar{p}_{n}^{g} - \bar{x}_{n}^{i}\right) \quad \text{(III.36)}$$

Where i refers to the particle in the swarm. n is the iteration step carried out, and r_1 and r_2 values represent random numbers in the range [0, 1]. ω is the inertia weight parameter. The coefficients c_1 and c_2 represent the optimisation parameters, $\bar{x}$ is the position vector, $\bar{v}$ is the velocity vector, and $\bar{p}^{i}$ is the best position information that the i^{th} particle has achieved and finally, $\bar{p}^{g}$ represents the best position information available in the swarm.

For the purpose of escaping the local minimums, the new position and new velocity of a particle, in the PSO algorithm, are replaced with a random position inside the search space. The processor is kept until an optimum result, or the maximum of iterations number is achieved.

III.2.3.6 Whale Optimizer Algorithm

The whale optimizer algorithm is another meta-heuristic algorithm introduced by Mirjalili & Lewi [404], in 2016. The inspiration is obtained from observing the hunting behavior of humpback whales. Those humpbacks whales have a special hunting behavior common as the bubble-net feeding method, used for hunting schools of fish. Hence, whales create a 9-shaped net with bubbles that encircles the prey and simplifies it for the whales to eat them during hunting [405–407]. This approach is described in more detail in [404,406]. The mathematical model for performing the optimization follows the following three phases:

1. Encircling prey

In this step, humpback whales encircle the prey that they have found. Then, change their position towards the best search agent through a course of iterations. This method considers that the position of the target prey is the best or near to the optimum solution. The mathematical representation is defined by [404,406,407]:

$$\vec{D} = \left|\vec{C} \cdot \overrightarrow{X^*}(n) - \vec{X}(n)\right| \quad \text{(III.37)}$$

$$\vec{X}(n+1) = \overrightarrow{X^*}(n) - \vec{A} \cdot \vec{D} \quad \text{(III.38)}$$

Where n indicates the current iteration, X^* is the position vector of the best solution were obtained so far iteration n, $\vec{X}$ is the position vector of each agent, $\vec{A}$, $\vec{C}$ are computes as follows:

$$\vec{A} = \overrightarrow{2a} \cdot \vec{r}_1 - \vec{a} \quad \text{(III.39)}$$

$$\vec{C} = 2 \cdot \overrightarrow{r_2} \quad \text{(III.40)}$$

Where components of $\vec{a}$ are linearly decreasing from 2 to 0, over the course of iterations and $\vec{r}_1$ & $\vec{r}_2$ are random vectors in [0, 1].

2. Bubble-net attacking method

This attack combines two strategies that can be mathematically defined as follows:

2.1. Shrinking encircling mechanism: This method is simulated by decreasing the value of a in equation (III.39) which is clarified in [404,406].

2.2. Spiral updating position: the helix-shaped movement of humpback whales is calculated using Equations. (III.41) and (III.42):

$$\vec{X}(n+1) = \vec{D}' \cdot e^{bv} \cdot \cos(2\pi v) + \vec{X}^*(n) \quad \text{(III.41)}$$

$$\vec{D}' = \left|\vec{X}^*(n) - \overrightarrow{X(n)}\right| \quad \text{(III.42)}$$

Here, v is a random number in [-1,1], b is constant for defining the shape of the logarithmic spiral, $\vec{X}^*$ is the position vector of the prey position, and $\vec{X}$ is the position vector of the humpback whale. Note that there is a probability of 50% that the humpback whales swim either around the prey within a shrinking circle or along a helical path. Below is the its mathematical representation:

$$\vec{X}(n+1) = \begin{cases} \vec{X}^*(n) - \vec{A} \cdot \vec{D} & \text{if } p < 0.5 \\ \vec{D}' \cdot e^{bv} \cdot \cos(2\pi v) + \vec{X}^*(n) & \text{if } p > 0.5 \end{cases} \quad \text{(III.43)}$$

Where p is a random number in [0,1].

3. Search for prey

In the bubble net technique, humpback whales search randomly for prey to find the position of the optimal agent. The mathematical formula is defined as follwos:

$$\vec{D} = \left|\vec{C} \cdot \vec{X}_{rand} - \vec{X}\right| \quad \text{(III.44)}$$

$$\vec{X}(n+1) = \vec{X}_{rand} - \vec{A} \cdot \vec{D} \quad \text{(III.45)}$$

Where X_{rand} is a random position vector. More detail can be found elsewhere [404–406].

III.2.3.7 Hybrid Particle Swarm Optimization Grey Wolf Optimizer

The hybridised version of particle swarm optimization with the grey wolf algorithm was established by applying a low-level co-evolutionary mixed hybrid, as illuminated by [403]. The inspiration came to improve the ability of exploitation in PSO with the ability of exploration in GWO to benefit from both variant's strengths in order to reduce the possibility of trapping into a local minimum [403,408].

The HPSOGWO approach uses the following improved equations. The Equation. (III.46) is used to update the search space of the first three agent positions. The inertia constant is used to control the exploration and exploitation of an individual wolf in the search space. For combining PSO and GWO variants, Equation. (III.47.a) is proposed for the velocity, while Equation. (III.47.b) for updating [408]:

$$\begin{aligned} \vec{d}_\alpha &= |\vec{c}_1 \cdot \vec{x}_\alpha - \omega * \vec{x}| \\ \vec{d}_\beta &= |\vec{c}_1 \cdot \vec{x}_\beta - \omega * \vec{x}| \\ \vec{d}_\delta &= |\vec{c}_1 \cdot \vec{x}_\delta - \omega * \vec{x}| \end{aligned} \quad \text{(III.46)}$$

$$v_i^{k+1} = \omega * \left(v_i^k + c_1 r_1\left(x_1 - x_1^k\right) + c_2 r_2\left(x_2 - x_i^k\right) + c_3 r_3\left(x_3 - x_i^k\right)\right) \quad \text{(III.47.a)}$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad \text{(III.47.b)}$$

III.2.4 Hybrid Support Vector Regression Models

SVR is a robust supervised learning approach to solve nonlinear problems of regression type. The major concern for researchers, with this technique, is the struggle of defining their best hyper-parameters values for a specified database that can ensure sufficient accuracy with minimal error. Notice that unsuitable SVR parameters values may cause over-fitting or under-fitting issues. Usually, the grid search technique and gradient descent method were the most used algorithms to determine SVM parameters. Those algorithms reported many difficulties such as convergence to local minima point, computational complexity, and height

computational time requirement led to the development of meta-heuristic optimization algorithms. Recently, some of them were adopted to determine the SVR parameters as a replacement to the traditional algorithms. Genetic algorithm (GA) and particle swarm optimisation (PSO) can be distinguished. Yet, there is no consensus between researchers on which optimization technique is much suitable in every aspect, not even for a specific optimizing problem. This leads to develop new algorithms as well as compare their performances when it is needed to.

The present chapter aims to apply seven meta-heuristics algorithms: DA, ALO, GWO, ABC, PSO, WOA, and HPSOGWO to seek out the optimal values of the three SVR parameters; the constant C (box-Constraint), ε the epsilon, and σ the parameter of the kernel function (Kernel-Scale) of the model, while gaussian was chosen to be the kernel function. The developed model was carried out by the ***fitrsvm*** function of MATLAB® R2020b. The Cross-validation "holdout" method was set to 0.3; 70% of the data was used for the training set, and 30% for the test set. The meta-heuristics algorithm fed the SVR initially with a random combination of hyper-parameters within their ranges. During several iterations depends on the algorithm used, the steps began with the division of data and up until the development of the SVR model; it was repeated and the minimum value of AARD obtained was saved as best. Then, each meta-heuristic algorithm generated a new population of hyper-parameters for the SVR algorithm. The same set of steps were repeated to achieve a new best AARD. That was for 100 trials; amongst which the minimum AARD corresponded to the resultant optimal hybrid SVR model. The association of the SVR model with the meta-heuristic algorithms selected (DA, ALO, GWO, ABC, PSO, WOA, and HPSOGWO) can build the following hybrid models (DA-SVR, ALO-SVR, GWO-SVR, ABC-SVR, PSO-SVR, WOA-SVR, and HPSOGWO-SVR) respectively. The flowchart of the proposed SVR hybrid models is summarised in **Figure III. 1.**

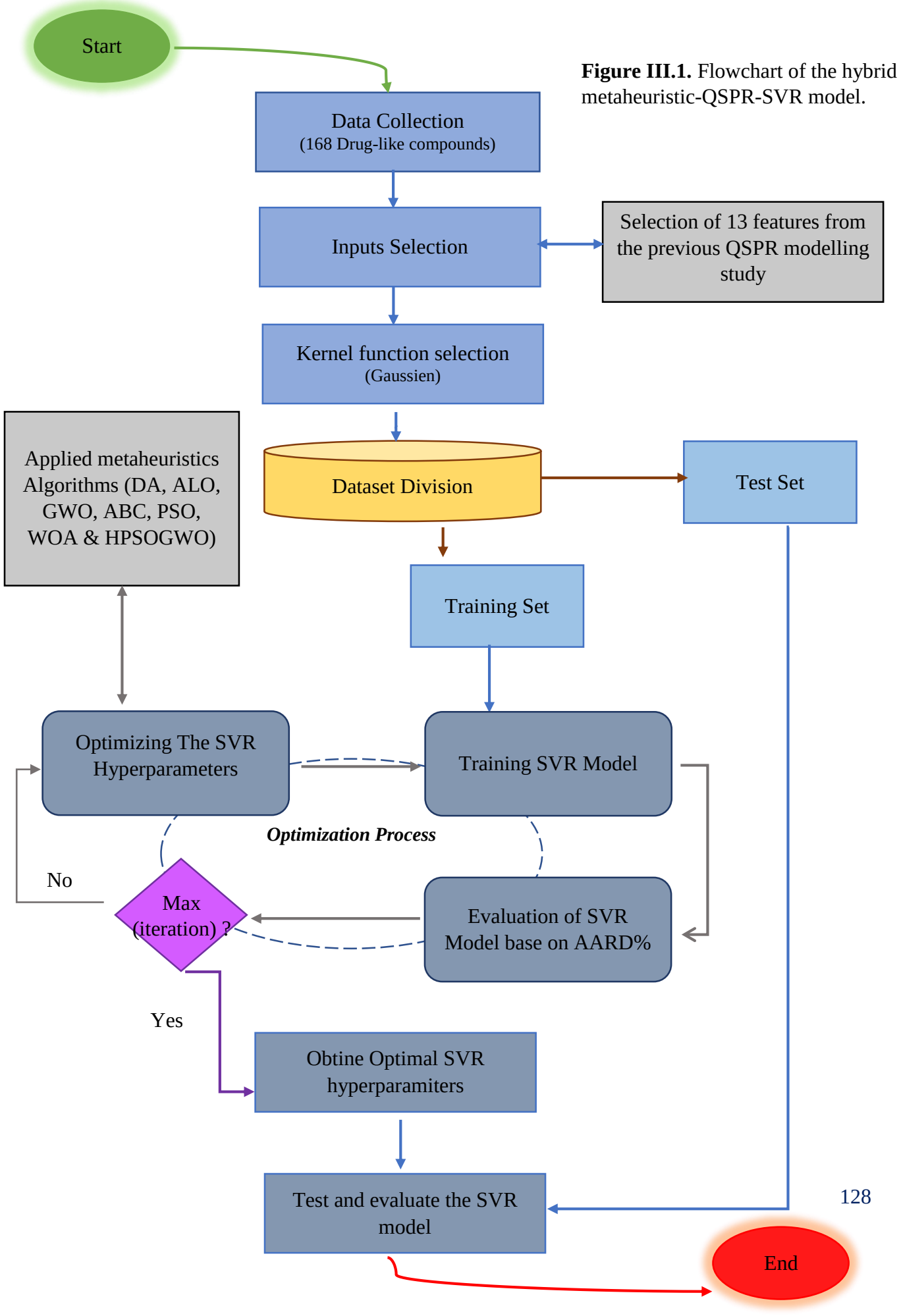


Figure III.1. Flowchart of the hybrid metaheuristic-QSPR-SVR model.

III.2.5 Statistical Criteria for Evaluation of Models Performances

The evaluation of the performance of a regression model, also known as validation, is now known as the most important concept for the development and validation of any model. This process confirms the reliability of the developed model for its possible application on a new set of data, and confidence of prediction can thus be judged along with the possibility of receiving a statistical comparison judgment between several models and various inputs of a model. Hence, a model can lead to the false prediction of response if the developed model is not validated correctly. The prediction accuracy of the seven models was examined by using various statistical criteria and through graphical appraisal, i.e., scatter plot, bar diagrams, and so on. More details are addressed in the R&D section. The statistical performance indicators are given by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(y_i^{obs}-y_i^{pred}\right)^2}{N}} \quad \text{(III.48)}$$

$$R = r = \frac{\sum_{i=1}^{N}\left(y_i^{obs}-\bar{y}^{obs}\right)\cdot\left(y_i^{pred}-\bar{y}^{pred}\right)}{\sqrt{\sum_{i=1}^{N}\left(y_i^{obs}-\bar{y}^{obs}\right)^2}\cdot\sqrt{\sum_{i=1}^{N}\left(y_i^{pred}-\bar{y}^{pred}\right)^2}} \quad \text{(III.49)}$$

$$R^2 = r^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{obs}-y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{obs}-\bar{y}^{obs}\right)^2} \quad \text{(III.50)}$$

$$AARD\% = 100 \cdot \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i^{obs}-y_i^{pred}}{y_i^{obs}}\right| \quad \text{(III.51)}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(y_i^{obs}-y_i^{pred}\right)^2 \quad \text{(III.52)}$$

$$RE\% = 100 \cdot \left|\frac{y_i^{obs}-y_i^{pred}}{y_i^{pred}}\right| \quad \text{(III.53)}$$

$$Q_{F1}^2 = 1 - \frac{\sum_{i=1}^{Nout}\left(y_i^{obs}-y_i^{pred}\right)^2}{\sum_{i=1}^{N_{out}}\left(y_i^{obs}-\bar{y}_{TR}\right)^2} \quad \text{(III.54)}$$

$$Q_{F2}^2 = 1 - \frac{\sum_{i=1}^{Nout}\left(y_i^{obs}-y_i^{pred}\right)^2}{\sum_{i=1}^{N_{out}}\left(y_i^{obs}-\bar{y}_{OUT}\right)^2} \quad \text{(III.55)}$$

$$Q_{F3}^2 = 1 - \frac{\sum_{i=1}^{N_{out}}\left(y_i^{obs}-y_i^{pred}\right)^2/N_{OUT}}{\sum_{i=1}^{N_{TR}}\left(y_i^{obs}-\bar{y}_{TR}\right)^2/N_{TR}} \quad \text{(III.56)}$$

$$AICs = N\ln\left(\frac{SSE}{N}\right) + 2n_p + \frac{2n_p(n_p+1)}{N-(n_p+1)} \quad \text{(III.57)}$$

$$SSE = \sum_{i=0}^{N}(y_i^{obs} - y_i^{pred})^2 \quad \text{(III.58)}$$

$$MAE = \frac{1}{N} X \sum_{i=1}^{N}\left|y_i^{obs} - y_i^{pred}\right| \quad \text{(III.59)}$$

$$Bf = 10^{\left(\frac{\sum_{i=1}^{N}\log\frac{y_i^{obs}}{y_i^{pred}}}{N}\right)} \quad \text{(III.60)}$$

$$Af = 10^{\left(\frac{\sum_{i=1}^{N}\left|\log\frac{y_i^{obs}}{y_i^{pred}}\right|}{N}\right)} \quad \text{(III.61)}$$

Where r is the correlation coefficient between the observed and predicted value of compounds with an interception,R^2 is the coefficient of determination,$RMSE$ is the root-Mean Squared Error, AARD is the average absolute relative deviation, MSE is the mean squared error, RE is the relative error, *AICs* is the Akaike's information criterion, MAE is the mean absolute error , Af is the precision factor and Bf is the bias factor . The error parameters namely AARD, RMSE, MSE, and MAE are usually used for comparison between different model and not to validate their performance while the statistics R, R², Q_{F1}^2, Q_{F2}^2, Q_{F3}^2, A_f, B_f and *AICs* are the statistical criteria used for evaluation of models performance and validation. Moreover, the statistics *AICs* and Q_{F3}^2 were found to the best suitable statistics for the evaluation of similar studies as explained in [409] and [410] , respectively. More detail can be found elsewhere [363,365,366,411].

III.2.6 Applicability Domain

Based on the third OECD principle of the QSAR modeling validation, any model must have a defined applicability domain (AD). The AD is a theoretical region demonstrated by the descriptors of the model used to define the area each time a model can be utilised for screening new chemicals. Viz, chemicals whose structures are "similar" to the ones in the training set are the only molecules that can be predicted reliably. Consequently, it is quite impossible to predict all chemicals using a single model. This study addressed the use of the Williams plot and the Insubria's plot [365,370–372,412].

III.2.6.1 William's plot

William's plot is a plot of standardised residuals vs. the leverage values denote to Hat (hi) diagonal values; it was applied to visualise the respective AD and to discover the outlier's response. those outliers are generally compounds with standardized residuals greater than two ± 3 standard deviation units for y-axes and structurally influential chemicals (hi>h*) for x-axes in the model. Where h* is a threshold leverage value. The leverage (h) value of a compound is given as follows:

$$hi = x_i(X^T X)^{-1} x_i \quad \text{(III.62)}$$

Where x_i is the descriptor-row vector of the query compound, and X is a k x n matrix containing the k descriptor values for each one of the n training compounds.

The warning leverage h* is defined by:

$$h^* = \frac{3(p+1)}{n} \quad \text{(III.63)}$$

Where n is the total number of samples in the training set, and p is the number of descriptors involved in the correlation.

When h < h*, the observed and predicted values have a high probability while molecules with hi>h* hardly influence the quality of fit of the developed model. However, these compounds may not be an outlier because of low residual.

It can be noted that compounds with a high value of leverage and good fitting in the developed model may stabilise the model. In contrast, compounds with bad fitting in the developed model can be outliers. Hence, the standardised residual and the leverage must be used at once for the description of the AD of the expanded model.

III.2.6.2 Insubria Graph

The Insubria graph is another plot used to defined the AD of a model. It represent a diagonal values vs. predicted values plot that are used in the case of chemicals without experimental data to provide a visualisation of interpolated and extrapolated predictions. A zone of higher reliability is always provided for both structures (<h*) and (-ln (y2)) predictions (between the maximum and the minimum value of the observed solubility of the training set).

III.3 Results and Discussion

III.3.1 Comparisons of Performances

In this section, numerous analysis methods such as convergence curve were adopted to compare the performance of the seven hybrid QSPR-SVR models and to select the best-suitable meta-heuristic algorithms so far for tuning the hyper-parameters of the SVR model for a dataset of 168 drug compounds, accounting for 4490 experimental data points regarding solubility in supercritical carbon dioxide, and 13 selected inputs.

Figure III.2 shows the convergence curve of the seven selected optimization approaches. According to the objective function (AARD%) for an amount of 100 iterations, the figure clearly demonstrated that the hybrid HPSOGWO, coloured in blue, led to the lowest error (0.7063) with the smallest needed number of iterations, subsequently the WOA algorithm with an error of 0.7271. The performance of this approach was asymptotic to HPSOGWO. Thereafter, PSO, ABC, ALO, GWO, then DA, with errors of 0.7280, 0.7345, 0.7419, 0.7481, 0.7508, respectively. This conforms to the over-perform of the meta-heuristic HPSOGWO over the rest six algorithms in terms of minimising the cost function (AARD%) and the convergence capacity.

Table III.1 represents the hyper-parameters of the best performances so far from each meta-heuristic's algorithms used for tuning the parameters of the hybrid QSPR-SVR algorithm.

Figure III.3-III.9 represents the scatter plots, i.e., predicted versus observed values of the solubility fraction of drug compounds in SC-CO_2 for the global, train, and test sets for all seven hybrid SVR models of the best iteration. In general, all the plots of all algorithms could give a positive correlation between the calculated values and the experimental values with a few outliers. This confirms the good quality of all models and proves their ability in predicting.

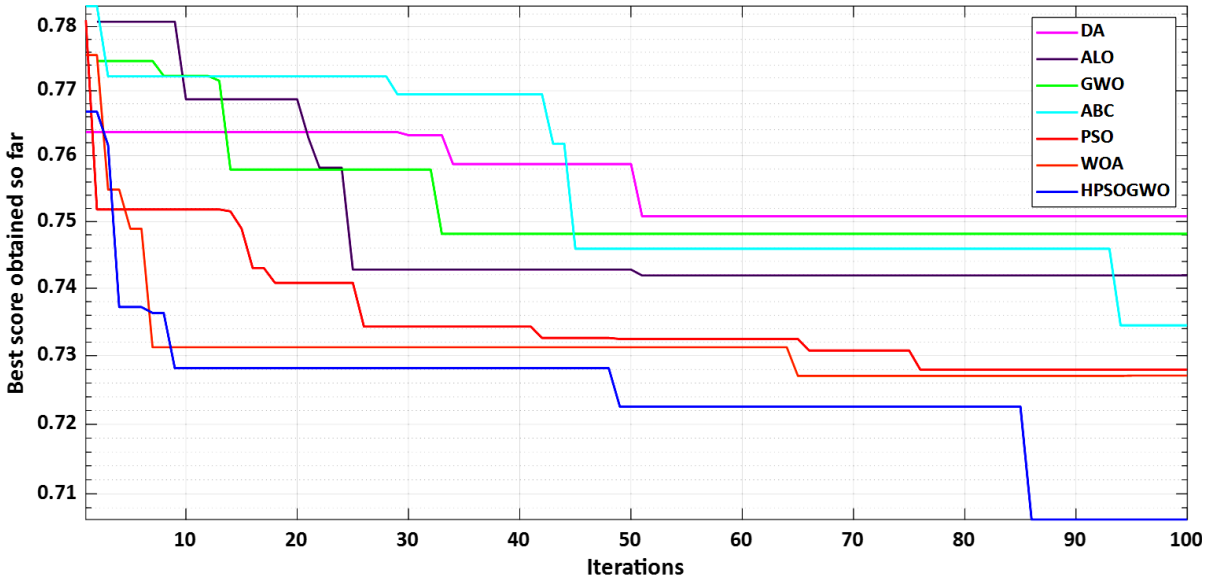
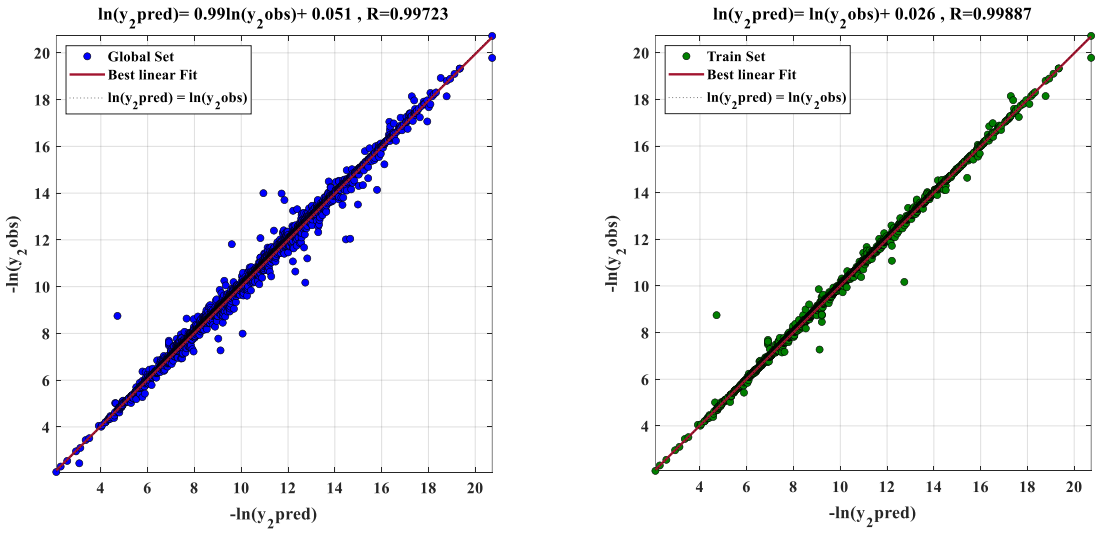


Figure III.2. Convergence curve of the seven algorithms.

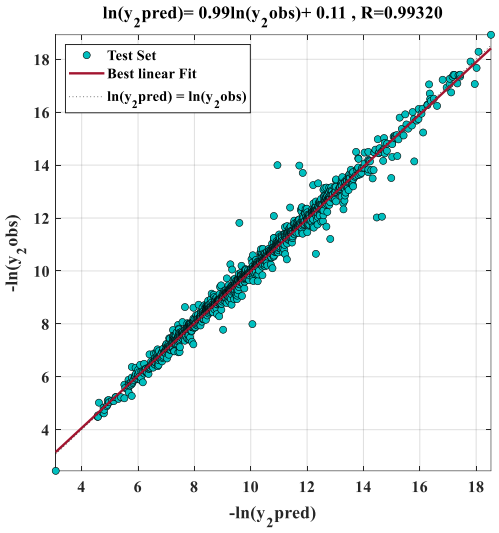
Table III.1 Hyperparameter of the SVR models.

Parameters / Models	*C*	σ	ε
DA-SVR	102.2490	1.7118	0.0076
ALO-SVR	65.0157	1.7484	0.0066
GWO-SVR	60	2	0.0068
ABC-SVR	60	1.6724	0.0067
PSO-SVR	62.4902	1.7045	0.0068
WOA-SVR	52.0924	1.5791	0.0063
HPSOGWO-SVR	62.9392	1.6780	0.0065



(a)

(b)



(c)

Figure III.3. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for DA-SVR model.

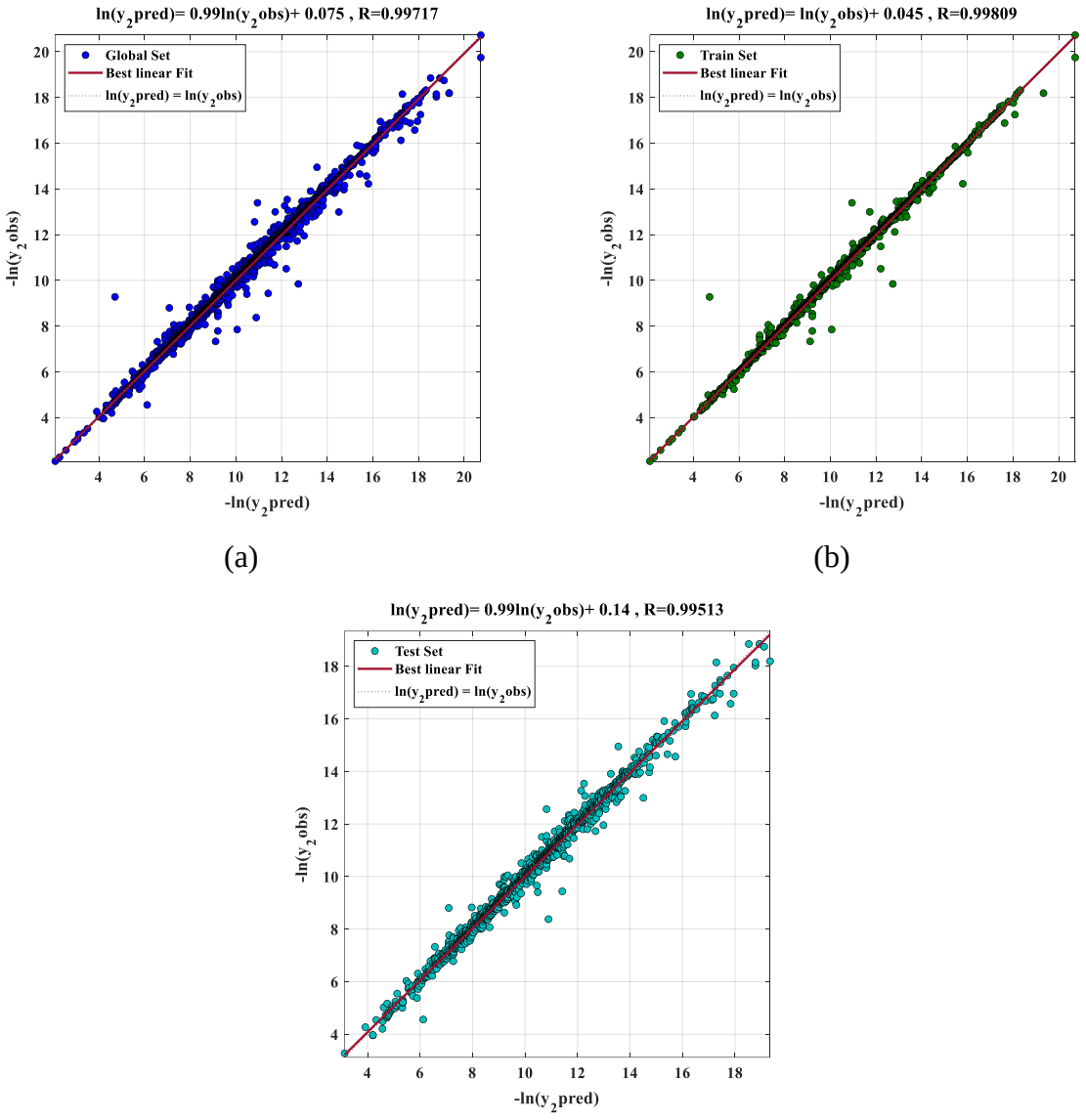
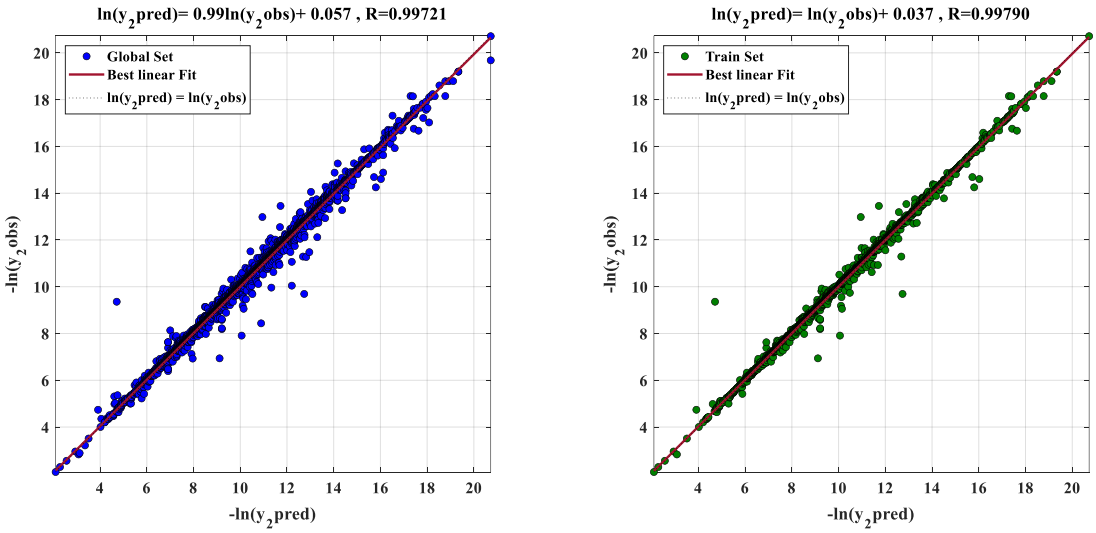
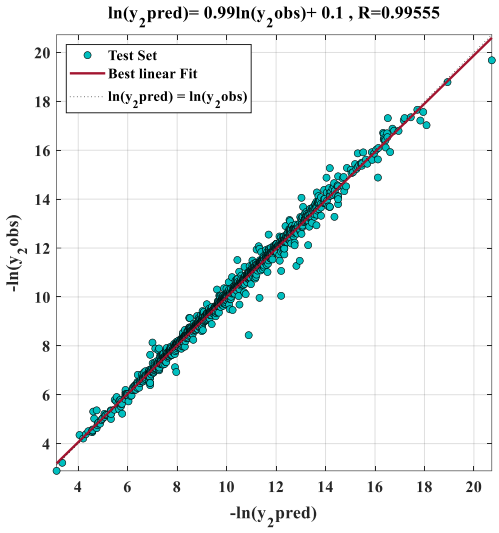


Figure III.4. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for ALO-SVR model.



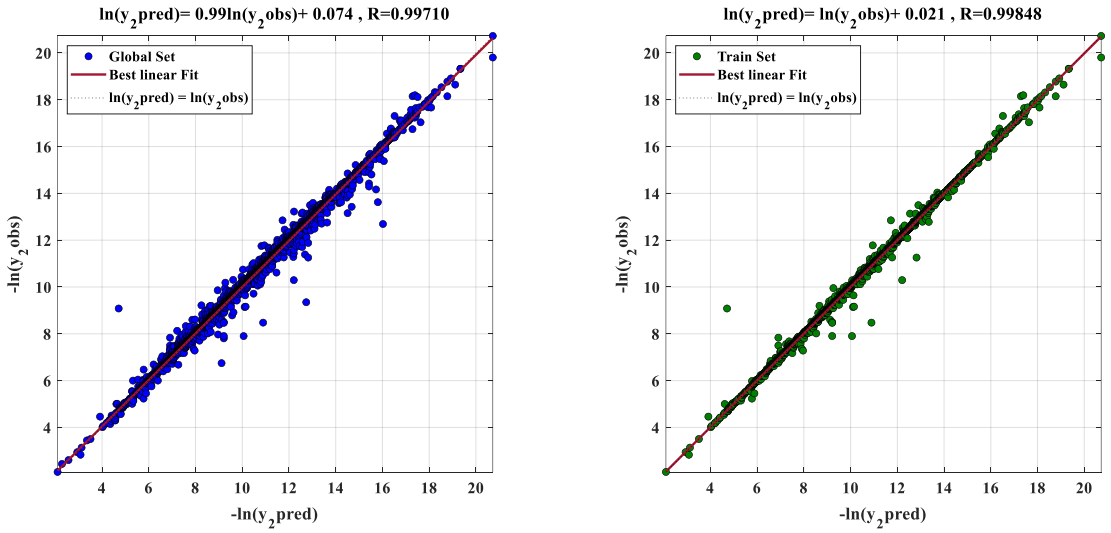
(a)

(b)

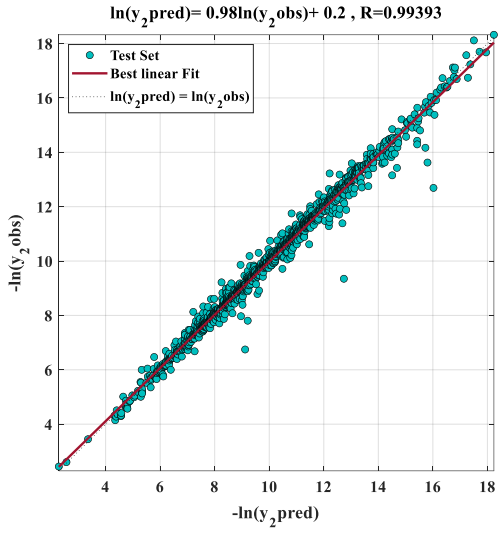


(c)

Figure III.5. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for GWO-SVR model.

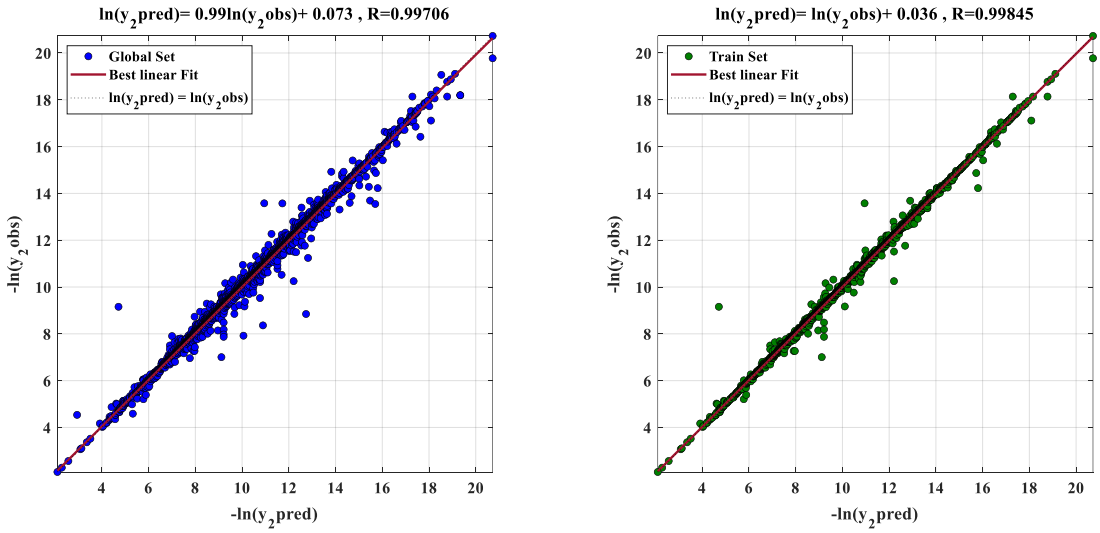


(a) (b)



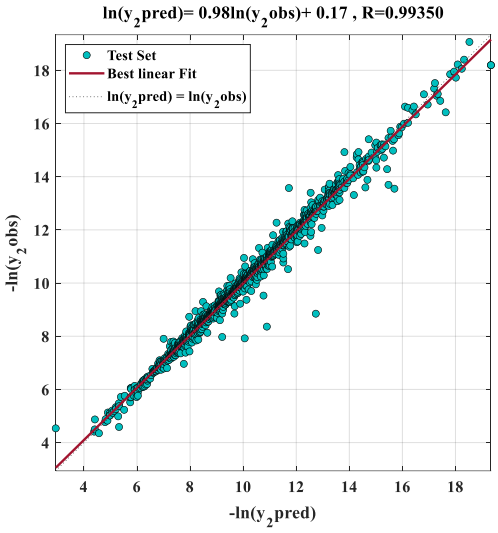
(c)

Figure III.6. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for ABC-SVR model.



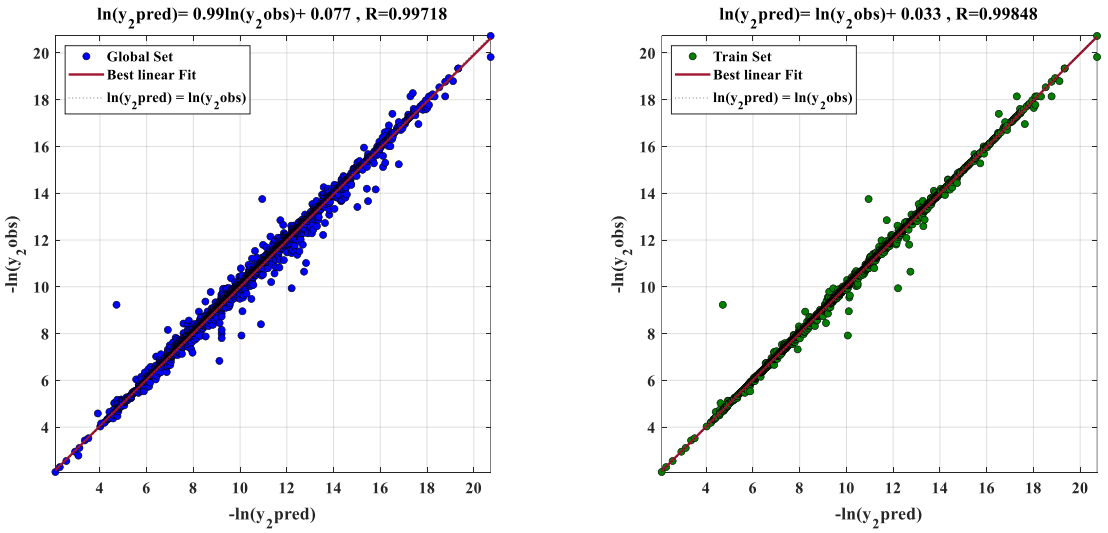
(a)

(b)



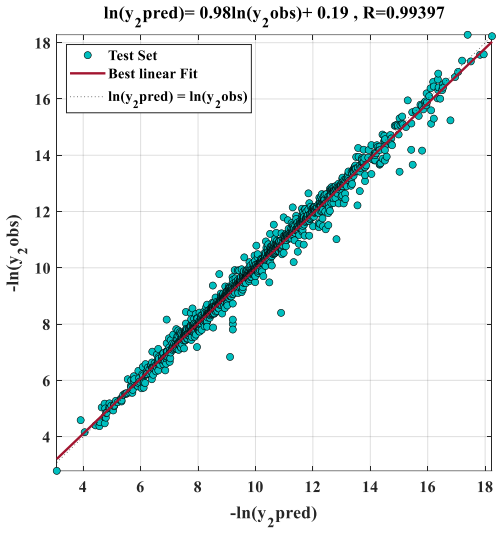
(c)

Figure III.7. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for PSO-SVR model.



(a)

(b)



(c)

Figure III.8. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for WOA-SVR.

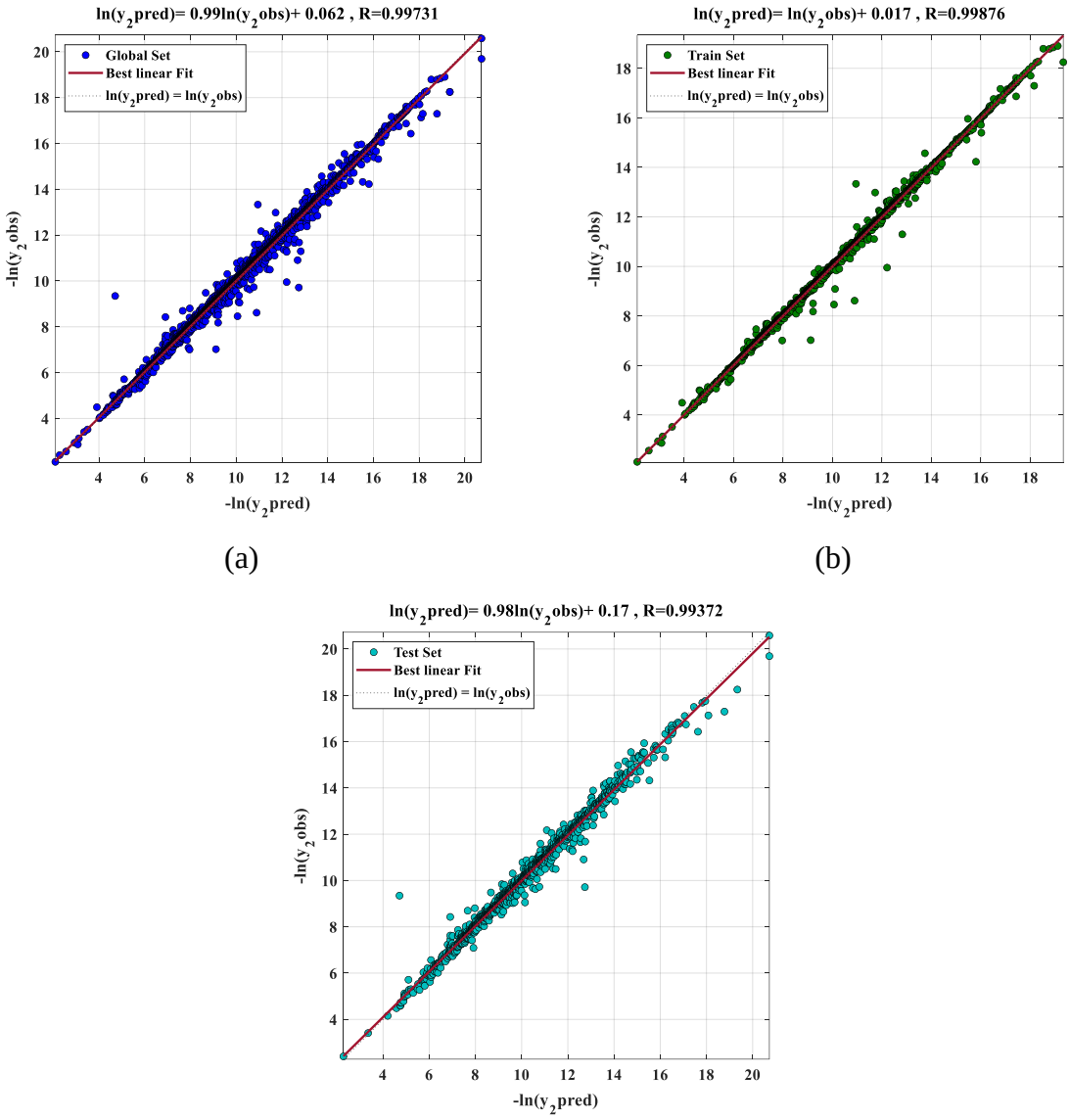


Figure III.9. Scatter plots predicted vs. observed values for the (a) global, (b) train and (c) test sets, for HPSOGWO-SVR model.

Table III.2 illustrates the statistics results about AARD, RMSE, MSE, MAE, r, R^2, Q_{F1}^2, Q_{F2}^2, Q_{F3}^2, and *AIC* for the seven developed models. It can be observed from the Table that the calculated parameters were all valid with high correlation coefficient (r) as well as great robustness and the coefficients of determination (R^2) were all above 1. In addition to the error parameters of all models which were below 0.3.

Table III.2 Statistics parameters of the seven SVR models

Statistics / Models	*DA-SVR*	*ALO-SVR*	*GWO-SVR*	*ABC-SVR*	*PSO-SVR*	*WOA-SVR*	*HPSOGWO -SVR*
AARD%	0.7508	0.7419	0.7481	0.7345	0.7280	0.7271	0.7063
RMSE	0.1909	0.1931	0.1915	0.1957	0.1968	0.1929	0.1883
MSE	0.0365	0.0373	0.0367	0.0383	0.0387	0.0372	0.0355
MAE	0.0730	0.0720	0.0731	0.0714	0.0709	0.0705	0.0693
r	0.9972	0.9972	0.9972	0.9939	0.9971	0.9972	0.9973
R^2	0.9945	0.9943	0.9944	0.9942	0.9941	0.9944	0.9946
Q^2_{F1}	0.9945	0.9943	0.9944	0.9942	0.9941	0.9944	0.9946
Q^2_{F2}	0.9945	0.9943	0.9944	0.9942	0.9941	0.9944	0.9946
Q^2_{F3}	0.9946	0.9942	0.9945	0.9942	0.9941	0.9945	0.9947
A_f	1.0075	1.0074	1.0075	1.0074	1.0073	1.0073	1.0073
B_f	1.0000	1.0002	0.9998	0.9996	0.9999	0.9997	0.9996
AICs	-14313.0819	-14213.7697	-14285.5572	-14100.3559	-14051.7736	-14225.7070	-14434.2487

In summary, the results above were in total agreement with the acceptability criteria of the statistical validation of regression models. Therefore, all the models performed well in model development, assuming an excellent prediction ability of future data. The models were thus stable, robust, and predictive, even if the HPSOGWO-SVR led to the best performances and the best AIC value of -14434.2487.

In order to quantitatively visualise the distribution of relative errors, and to compare the generalisation capacity, the tendency of over-fitting of the different models, **Figure III.10**, and **Figure III.11**, shows a bar distribution of the relative error for both training and test set, respectively.

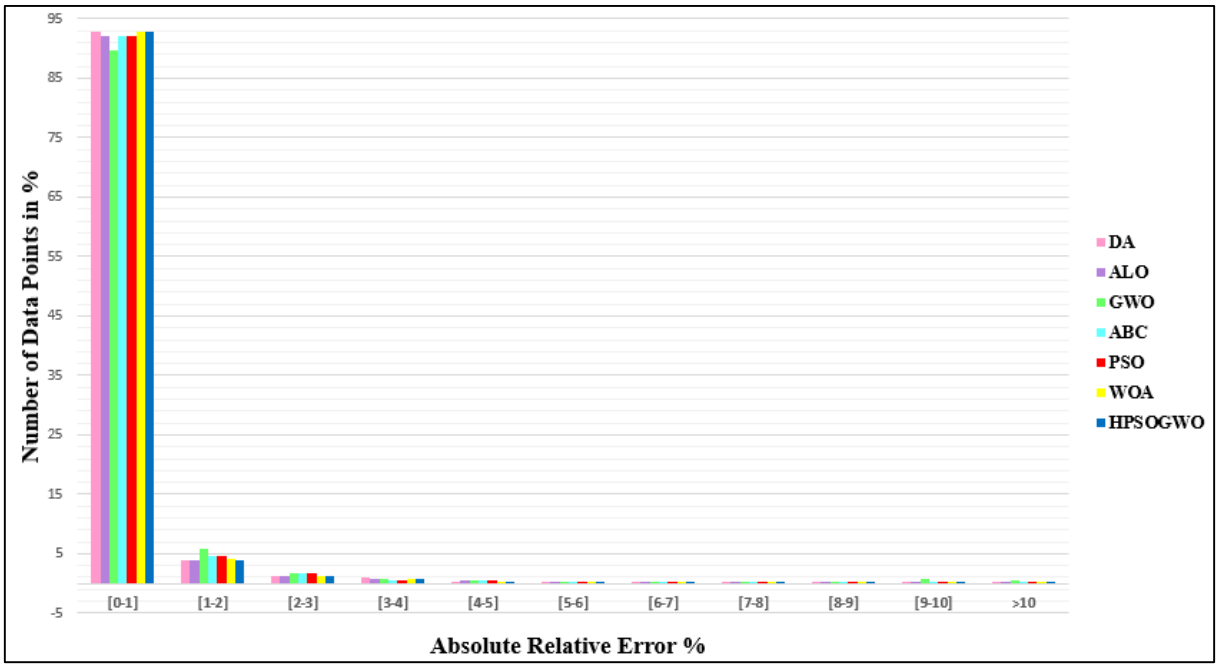


Figure III.10. Bar distribution diagrams of the absolute relative error [%] vs. number of data points in percentage of the training set.

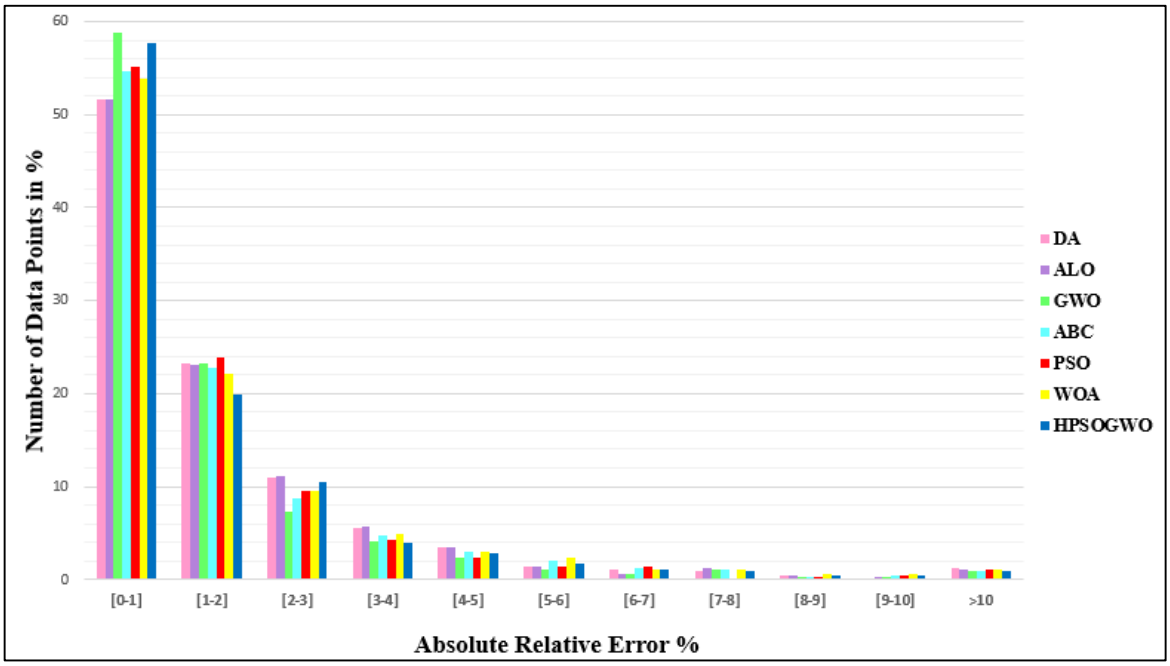


Figure III.11. Bar distribution diagrams of the absolute relative error [%] vs. number of data points in percentage of the test set.

From both Figures, it can observe that all models showed good distribution, as the majority of the data points fall to the range of 0 to less than 1, more than 90% for the training set, and more than 51% for the test set. In other words, all models had similar generalisation capacity and showed no over-fitting.

In conclusion, these results demonstrate that the HPSOGWO algorithm is superior to the other six models in terms of optimising the SVR hyper-parameters. Overall, the results also illustrate that the HPSOGWO-SVR model has a high quality of fast learning and great generalisation capacity to solve nonlinear regression issues. The SVR is an effective and powerful technique for forecasting the solubility of the drug compounds in SC-CO_2.

Figure III.12 shows a plot of standardised residual (Y-axis) vs. leverage values (X-axis) referred to as the Williams plot, for the HPSOGWO-SVR model while the reliable prediction zone of the model based on structural similarity to the training compounds (leverage value) and the predicted value of solubility was defined with the Insubria graph in **Figure III.13**.

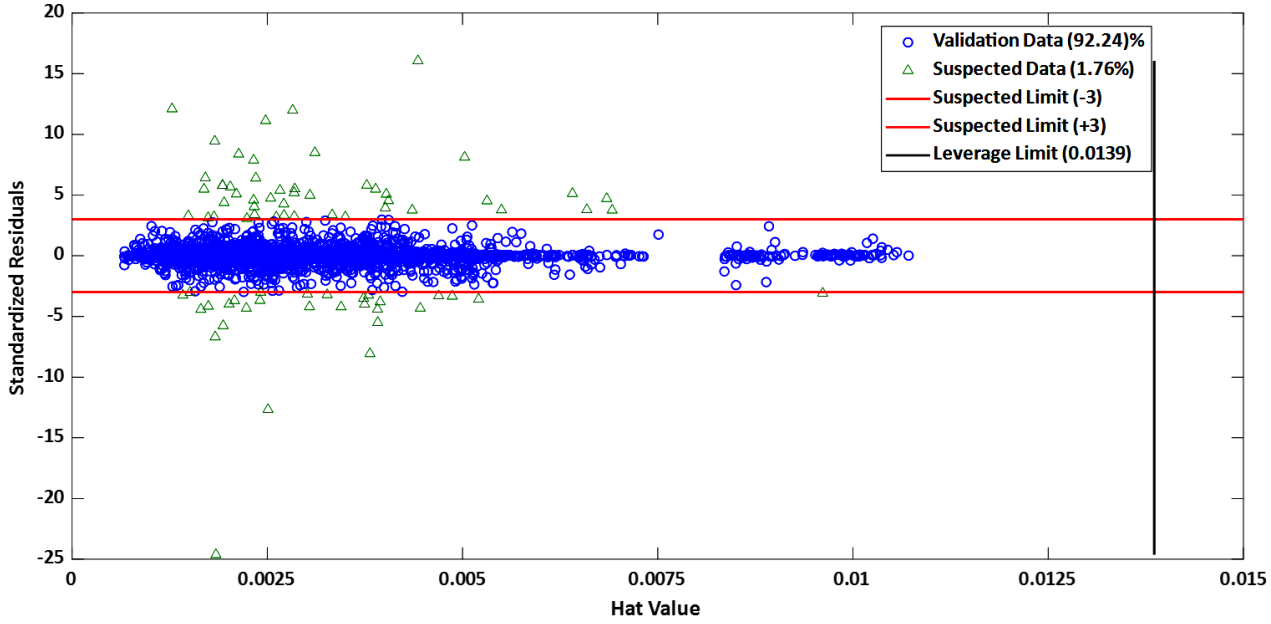


Figure III.12. William's plot for HPSOGWO-SVR model.

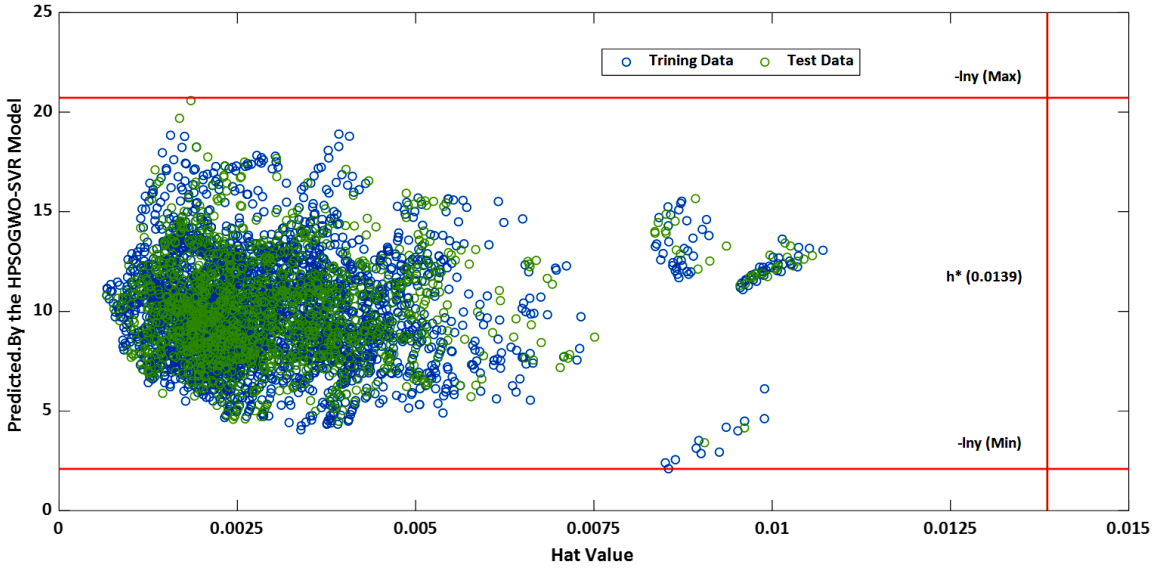


Figure III.13. Insubria graph for the HPSOGWO-SVR model.

Analysing **Figure III.12** indicates that there was no structural influential compound in the whole dataset that had leverage higher than the warning h* (0.0139). 98.24% of the whole dataset was located within the horizontal lines (range of ± 3), and only 1.76% showed a distribution without the suspected limit. Due to its high predictive ability, the proposed model could be used to screen existing databases or virtual chemical structures to identify solubility. In this case, the applicability domain can serve as a valuable tool for filtering out "dissimilar" chemical structures.

In **Figure III.13** a plot of leverages values (X-axis) versus predicted values (Y-axis) for both training and test set referred to as the Insubria graph of the HPSOGWO was considered. The terms "-lny (Max)" and "-lny (Min)" refer to the maximum (20.7233) and the minimum (2.0956) value of observed solubility in the training set, respectively. The predicted results are reliable, if both conditions: hi< h* and -lny (Min) <-lny (Pred) < -lny (Max).

It was found that all compounds from the test set were located within the model's applicability domain. This demonstrates that the model obtained in this work has high applicability to new drugs, and it can be applied it in order to screen and prioritise them for future experiments or for filling the data gap.

As mentioned earlier in this paper, the whole dataset was randomly divided into two main sets. The first set (4330 EDP) for model development and the second set (160 EDP) was kept hidden from the build model HPSOGWO-SVR to check the model's capacity in predicting untested data (**Figure III.14**), and to illustrate a graphical comparison between the observed solubility isotherms and the predicted solubility isotherms calculated by the HPSOGWO-SVR model (**Figure III.15**). Note that the average temperature was taken for each isotherm.

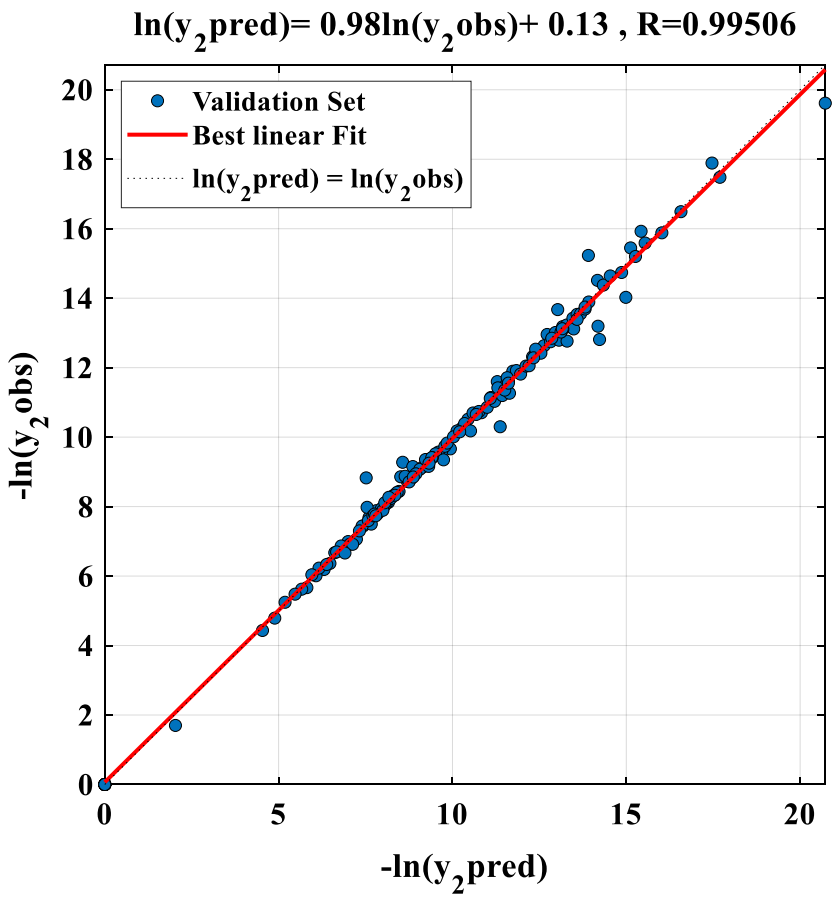


Figure III.14. Scatter plot Predicted vs. Observed for HPSOGWO-SVR model, validation set.

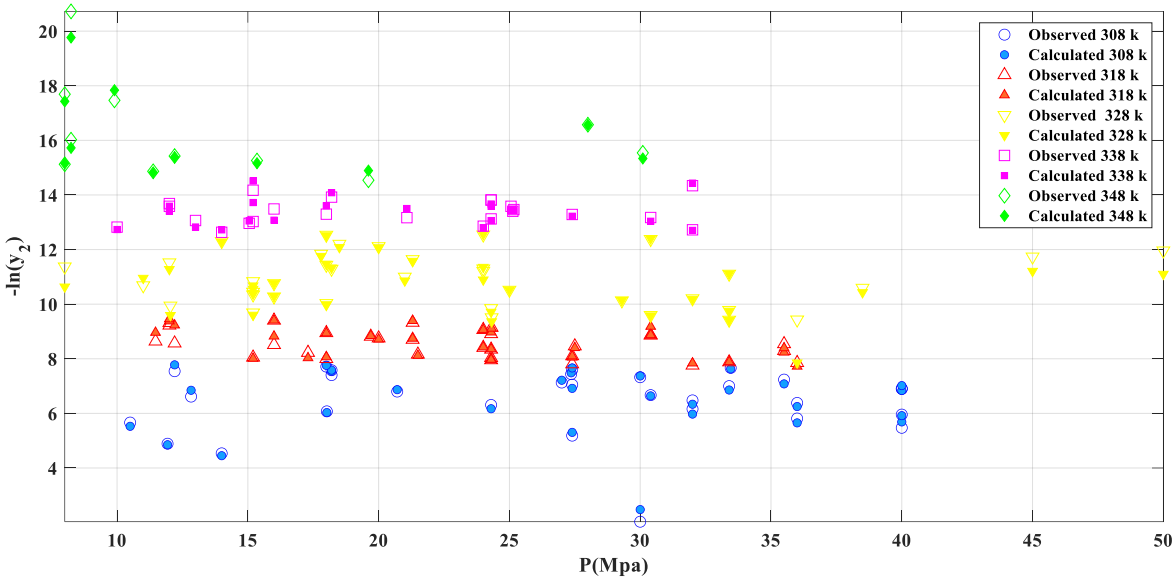


Figure III.15. Comparison diagram of observed solubility isotherms and those calculated by the HPSOGWO-SVR model, for validation set.

Good-of-fit, robustness and predictive power were confirmed by the values of *R* (0. 9951), R^2 (0.9900),Q^2_{F1} (0.9900), Q^2_{F2} (0.9900), Q^2_{F3} (0.9900), relatively low values of errors, *AARD%* (1.6034), *RMSE* (0.3102), *MSE* (0.0962). Moreover, the scatterplot of **Figure III.14** shows a visual correlation between observed and predicted solubility fraction values for the validation set.

In addition, the solubility isotherm's diagram of **Figure III.10** shows that the majority of data points of observed and predicted solubility were identical. It can also be noticed, the superposition of full symbols (solubilities calculated by the HPSOGWO-SVR model) and empty symbols (experiment) for practically all points (77.44% superposed, 18.05% semi-superposed, and 4.51% not-superposed). The isotherms are ordered from the highest superposition between the observed and predicted values to the lowest by 308k, 338k, 318k, 328k, 348k highlighted in blue, pink, red, yellow, and green, respectively.

Those results confirmed the high quality of the model HPSOGWO-SVR. Since the error values were close to the training and test sets and there were no significantly large residual values for the validation set displayed in **Figure III.14**, it can be concluded in the absence of over-fitting of the model. This means that the model predicts correctly not only for the training and test compounds but also for other (external) compounds. In summary, the HPSOGWO-

SVR model has by far the best robustness (>0.9) and high predictive power and can be recommended for predicting future data falling into the defined applicability domain.

III.4 Conclusions

This chapter is the first comparative study between seven meta-heuristic algorithms in terms of fine-tuning hyper-parameters of a hybrid QSPR-SVR model defined for predicting drug solubility in supercritical carbon dioxide. A solubility database of 168 drug compounds was first collected and then correlated with their molecular structure by QSPR technique and two independent intensive state variables (temperature and pressure) of a previous study (represented in **Chapter II**).

The seven meta-heuristic algorithms namely, Dragonfly optimizer (DA), Ant Lion optimizer (ALO), Grey Wolf optimizer (GWO), Artificial Bee Colony optimizer (ABC), Particle Swarm optimizer (PSO), Whale optimizer (WOA), and a hybrid Particle Swarm with Grey Wolf optimizer (HPSOGWO) responsible for the seven models DA-SVR, ALO-SVR, GWO-SVR, ABC-SVR, PSO-SVR, WOA-SVR, and HPSOGWO-SVR were all statistically and graphically approved, while the hybrid HPSOGWO-SVR model over-performed the six other models.

The proposed HPSOGWO-SVR model confirms to have good predictivity ability and robustness, therefore, it can be used to estimate the solubility of drug compounds without experimental data available in the literature. The validity of the model predictions was further guaranteed by the external test on 160 EDP compared to experimental values, considering new compounds which should belong to the applicability domain (William's plot & Insubria graph).

The SVR model presented in this work showed better statistical parameter values and better predictability results. However, due to the stochastic nature of all swarm intelligence algorithms, it is never guaranteed to find an optimal solution for any problem, which always opens the door for new possibilities.

Chapter V

Third Contribution

Chapter IV

Third Contribution : A Hybrid Improved Grey Wolf Support Vector Regression Algorithm for Modeling the Solubility of APIs in Pure Ionic Liquids Using σ-profile Descriptors

IV.1 Introduction

This chapter introduces an advanced artificial intelligence algorithm based on support vector regression (SVR) coupled with an improved version of a well-known nature-inspired method namely grey wolf optimizer (I-GWO) for the prediction of active pharmaceutical ingredients (APIs) solubilities in various ionic liquids (ILs) at various temperature as a function of the σ-moments of cations, anions, and APIs. Those 13 σ-moments were considered as molecular descriptors calculated using the σ- profiles of three categories of descriptors based on a conductor-like screening model for Real Solvents (COSMO-RS) based approach which will be explored for establishing the QSPR model along with the temperature "T(K)" as input parameters.

The chapter will carry on three main sections. The first section 'Procedures' will describe in detail all the passed steps to build the model from data collection to model validations. The second section ''Results & Discussion', will discuss the results obtained of the proposed model and compared it with 11 different models namely, ANN-lm-tan, ANN-lm-log, ANN-br-tan, ANN-br-log, I-GWO-SVR(g), MVO-SVR(g), MVO-SVR(r), MFO-SVR(g), MFO-SVR(r), SCA-SVR(g), SCA-SVR(r). The model will be validated in terms of various statistical factors, graphical analyzes including an external validation test, analysis of variance (ANOVA), residual and outlier detection, as well a sensitivity analysis. The chapter will end with a 'Conclusion' as the third section, which will summarize all the important remarques and results achieved from this study.

IV.2 Procedures

IV.2.1 Overview

Experimental data solubility of active pharmaceutical ingredients (APIs) in pure ionic liquids at different temperature ranges were collected from the literature (see **Table A3**). In total, the dataset for 83 API-ILs systems composed of 14 APIs, 12 cations and 7 anions (32 ILs combinations), was collected as 850 experimental data points. The solubility fraction ranges from 0.0000084 to 0.992 and the temperature ranges from 248.92 K up to 423.73 K. All the details regarding the database were provided in **Table A3**.

IV.2.2 COSMO-RS and σ-profile

Currently, COSMO-RS is one of the most common theoretical methods used for modeling solubility [413]. This technique is an emerging model that attempts to connect the molecular and thermodynamic levels of chemistry [346,414]. In QSPR modeling, σ-moments are uses as molecular descriptors which are obtained based on only molecular structures without the need for experimentally determined data [83]. These descriptors are derived from the σ-profile. They are like moments of a statistical distribution and can be utilized as a method to reduce the high-dimensional information existing in a σ-profile to fewer descriptors that describe that σ-profile. Therefore, these descriptors were reported as potential/valuable descriptors in QSPR models [414]. Numerous advantages of COSMO-RS can be mentioned. The most attractive is that this approach is fast and requires only the chemical constitution and the topology of the molecules to perform the calculation [346,349,414–417].

In this present study, there are three categories of descriptors based on COSMO-RS, namely, σ-profiles of the cations, anions, and APIs. The charge distribution areas of the cations and anions were taken from Paduszynski's [415] paper. The σ-profiles of independent counterions (cation and anion) were both divided into four different regions, respectively S_{1-4} and S_{5-8} for the cations and the anions. Each section is defined by an interval of σ (e/nm²). The sections were divided according to Gonfa's [418] paper (see **Table IV.1**). The COSMOconf (version ; as a part of COSMOtherm suite, COSMOlogic GmbH) software was used to calculate the charge distribution area of the 14 APIs at BP-COSMO-TZVP level (functional: Becke-Perdew [415]; triple-ζ valence polarized basis set [415]; with the standard COSMO cavity[415]); plots presenting σ -profiles of 14 APIs are available in the annex. A set of the

most relevant conformers for each API was obtained with a screening charge density range of -2.5-2.7 e/nm². The average of the probability distribution (or histogram, Peak) $P^{x}(\sigma)$ was considered for calculating σ-moments. **Figure IV.1.** illustrations the σ-profile distribution of the 14 APIs while **Table IV.2** shows the σ-profile descriptors used and the margins limiting of cation, anion, and APIs descriptors as this last were also divided into four regions $S_{\sigma\text{-profile}}(S_{9\text{-}12})$,12 $S_{\sigma\text{-profile}}$ altogether. Those 12 molecular descriptors were applied along with the temperature "T(K)" as input parameters for carrying out the corresponding QSPR model of this work.

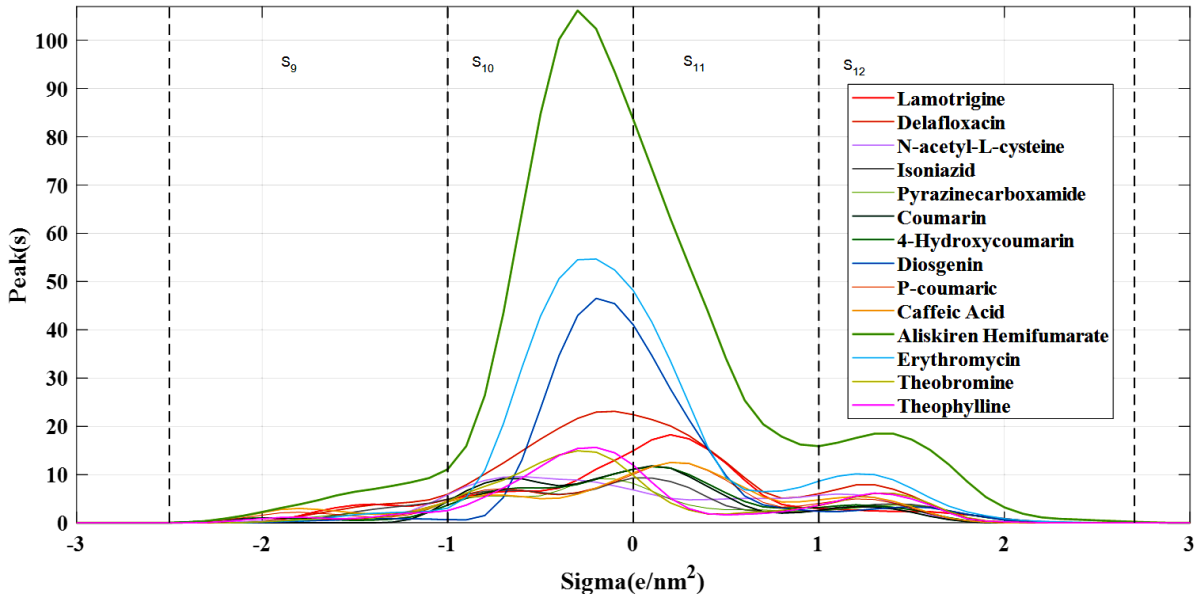


Figure IV.1. σ-profile distribution of the 14 APIs and their margins limiting each section.

IV.2.3 Statistical Data Analysis

A table summarizes the dependent and independent variables, their ranges, unit of measure, and some of their popular statistic metrics is given as **Table IV.2**.

Table IV.1. σ-profile descriptors typical cations and anions of the ILs, the APIs and their margins limiting each section.

	Cation's [418] molecular descriptors margins limit			
	S_1	S_2	S_3	S_4
$S_{\sigma\text{-profile}}$ (s.e/nm²)	[-2.5,-1.8]	[-1.8,-1]	[-1,1]	[1,2]
	Anion's [418] molecular descriptors margins limit			
	S_5	S_6	S_7	S_8
$S_{\sigma\text{-profile}}$ (s.e/nm²)	[-1.1,1]	[1,1.5]	[1.5,2]	[2,3]
	API's molecular descriptors margins limit			
	S_9	S_{10}	S_{11}	S_{12}
$S_{\sigma\text{-profile}}$ (s.e/nm²)	[-2.5,-1]	[-1,0]	[0,1]	[1,2.7]

Table IV.2. Summary of statistics analysis of the dataset.

Variable Category	Factors	Units	Min	Max	Variance	Std	Kurtosis	Skewness
Input	T	K	248.92	423.73	1379.9461	37.1476	1.9953	0.0643
	S_1	s.e/nm²	-0.0005	0.1262	0.0018	0.0425	1.2313	0.1965
	S_2	s.e/nm²	0.9058	2.7157	0.0863	0.2938	16.6792	-3.7118
	S_3	s.e/nm²	13.2625	65.4836	71.9256	8.4809	7.8790	2.0235
	S_4	s.e/nm²	-0.0004	0.6485	0.0097	0.0983	40.5241	6.2868
	S_5	s.e/nm²	-1.43E-08	15.0938	27.6884	5.2620	1.6242	-0.0090
	S_6	s.e/nm²	-0.0007	7.2302	2.9185	1.7084	2.7727	-0.3734
	S_7	s.e/nm²	-0.0029	5.8591	1.0383	1.0190	7.0502	1.8272
	S_8	s.e/nm²	-0.0002	0.6128	0.0022	0.0466	132.0266	10.4076
	S_9	s.e/nm²	0.1914	6.4366	0.6830	0.8265	21.4101	3.3785
	S_{10}	s.e/nm²	6.3735	69.6853	106.3542	10.3128	20.3403	3.9691
	S_{11}	s.e/nm²	2.8743	38.5946	31.5765	5.6193	20.1728	3.8006
	S_{12}	s.e/nm²	1.8810	14.7455	3.3721	1.8363	28.4752	4.7162
Output	x_2	-	8.4E-06	0.992	0.05676	0.2382	3.8133	1.3059

IV.2.4 Dataset Splitting and Output Normalization

The final dataset of 13 inputs, and 850 experimental data points was divided into two main sets. The first set was applied for model development (770 data points) and the second set (80 data points) was kept hidden from the model for external validation test. The first set was

then divided based on the cross-validation technique into a training set (616 data points) to build the model, and a test set (154 data points) to evaluate the predictive ability of the obtained model in a proportion of 80% to 20%.

IV.2.5 Model

Support vector regression (SVR) is one of the most known machine learning algorithms based on statistical learning theory. It was first introduced by Vapnik and Chervonenkis [419] in year 1974. This technique is one of the two categories of support vector machine, SVC for classification and SVR for regression. There is only one difference between those two branches that is the SVR has one kind of sample points [420]. In brief, the SVC approach first uses a nonlinear kernel function to map the input space into a high-dimensional feature space then conducts linear regression in this new space [280]. The approximation function of the SVR model is given by:

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x, x_i) + b \quad \text{(IV.1)}$$

Where x is the input vector and $f(x)$ is the output. x_i is the i^{th} feature vector and α_i is the weight for this feature vector. Support vectors are training points with nonzero weights. The vector and the bias (α & b, respectively) are adjustable model parameters that are optimized through the training process. $K(x, x_i)$ is the select kernel function.

The chosen kernel function and the corresponding kernel parameters together with a few internal parameters of the algorithm is an essential step for better prediction/performance capacity. In this chapter, the four kernel functions (Gaussian, Gaussian radial basis function, polynomial, and linear) were tested. Both the Gaussian radial basis function (RBF) kernel and Gaussian were very effective while RBF has the best results. Its formula is defined as:

$$K(x, x_i) = exp(-r \cdot \|x - x_i\|^2) \quad \text{(IV.2)}$$

r is an adjustable parameter of the kernel controlling the amplitude of the RBF function.

The so-called hyper-parameters (cost function C, the radius of the insensitive tube ε, and the kernel function's parameter σ) were optimized by a natural-inspired algorithm namely I-GWO.

IV.3.6 Model Optimization

The performance ability of a fitted SVR model to handle a problem is mainly related with the best selection of hyper-parameters. Mathematically, there are an infinite number of possible combinations that those parameters may take. Clearly, there is no way to test all combinations. Thus, introducing of optimization methods such as the swarm intelligence (SI) algorithm is necessary. For the present work, an improved grey wolf (I-GWO) has been coupled with an SVR algorithm for better modelling performance.

The I-GWO algorithm was first introduced by Nadimi-Shahraki and co-authors, in 2020 [394] as a solution to overcome few issues related to the grey wolf (GWO) algorithm such as falling into the local optimum instead of finding the optimal solution and the reduction of the diversity of the population. These improvements covered a new search strategy associated with selecting and updating steps. Also, it includes these three stages:

1-Initializing phase:

In the following phase, N wolves are randomly distributed with a range of $[l_i, u_j]$ in the search space.

$$X_{ij} = l_j + rand_j[0,1] \times (u_j - l_j), i\epsilon\ [1, N], j\ \epsilon[1, D] \quad \text{(IV.3)}$$

The position of the ith wolf in the tth iteration is represented as a vector of real values:

$$X_i(t) = \{X_{i1}, X_{i2}, \cdots, X_{iD}\} \quad \text{(IV.4)}$$

D is the dimension number of the problem.

The entire population of wolves is stored in a matrix Pop, which has N rows and D columns. The fitness function, $f(X_i(t))$ calculate the fitness value of $X_i(t)$.

2-Movement phase:

Dimension learning-based hunting (DLH) search strategy was integrated into the I-GWO algorithm. In this new movement strategy, each individual wolf is considered that is learned by its neighbours to be another candidate for the new position of $X_i(t)$.

The following steps describe why the canonical DLH search strategies were adopted and how it generates candidates:

- In the original GWO: using the three leader wolves of the Pop, a new position is generated for each wolf. The GWO algorithm ends up showing slow convergence where the population loses diversity too early and trapped wolves in the local optima. To overcome these defects, the DLH search strategy was applied.
- In the new search strategy: First, the radius $R_i(t)$ is calculated by Equation. (IV.5):

$$R_i(t) = \|X_i(t) - X_{i-GWO}(t+1)\| \quad \text{(IV.5)}$$

$X_i(t)$ is the current position and $X_{i-GWO}(t+1)$ is the candidate position.

Then, $N_i(t)$ the neighbours of $X_i(t)$ is calculated by Equation. (IV.6) respected to the radius $R_i(t)$

$$N_i(t) = \left\{X_i(t) \middle| D_i\left(X_i(t), X_j(t)\right) \leq R_i(t), X_j(t) \in Pop\right\} \quad \text{(IV.6)}$$

D_i is Euclidean distance between $X_i(t)$ and $X_j(t)$.

Once $N_i(t)$ is obtained, multi neighbours learning is accomplished by Equation. (IV.7)

$$X_{i-DLH,d}(t+1) = X_{i,d}(t) + rand \times \left(X_{n,d}(t) - X_{n,d}(t) - X_{r,d}(t)\right) \quad \text{(IV.7)}$$

As explained above, each dimension of the new position of wolf $X_i(t)$ is constructed by Equation. (IV.7). The individual wolf is learned by its different neighbours and a randomly

selected wolf from Pop. Then, besides $X_{i-DLH}(t+1)$, the DLH search strategy generates another candidate for the new position of wolf $X_i(t)$ named $X_{i-DLH}(t+1)$.

3-Selecting and updating phase:

Here, is a comparison between the fitness values of the two candidates $X_{i-GWO}(t+1)$ and $X_{i-DLH}(t+1)$ for selecting the superior candidate using Equation. (IV.8):

$$X_i(t+1) = \begin{cases} X_{i-GWO}(t+1), & if\ f(X_{i-GWO}) < f(X_{i-DLH}) \\ X_{i-DLH}(t+1) & otherwise \end{cases} \quad \text{(IV.8)}$$

Then, if $X_i(t+1) < X_i(t)$, $X_i(t)$ is updated. Else, it remains unchanged in the Pop. This step keeps reaping for all individuals until retching the maximal number of iterations (Maxiter). Noted that each time the counter of iterations (iter) increased by one. More detailed information is available elsewhere [394].

IV.2.7 Performance Evaluation

Numerous statistic metrics, offered in the literature, are used to evaluate and check the model's reliability in predicting (or the ability of a model to fit) the outcome. In the present chapter, correlation coefficient (R), coefficient of determination (R^2) and their adjusted (R_{adj} & R^2_{adj}, respectively), cross-validated correlation coefficient (Q^2), concordance correlation coefficient (Q^2ccc), root-Mean Squared Error ($RMSE$), mean squared error (MSE), mean absolute error (MAE), F-value, t-student test (t-value) and chi-square value (x^2) are introduced [30,366,411,421–423] to determine the precision and performance of the suggested model. Their formula are presented below:

$$R = \frac{\sum_{i=1}^{N}\left(x_i^{obs}-\bar{x}^{obs}\right)\cdot\left(x_i^{pred}-\bar{x}^{pred}\right)}{\sqrt{\sum_{i=1}^{N}\left(x_i^{obs}-\bar{x}^{obs}\right)^2}\cdot\sqrt{\sum_{i=1}^{N}\left(x_i^{pred}-\bar{x}^{pred}\right)^2}} \quad \text{(IV.9)}$$

$$R_{adj} = \sqrt{\left|1-(1-R^2)\left(\frac{N-1}{N-N_p-1}\right)\right|} \quad \text{(IV.10)}$$

$$R^2 = r^2 = 1 - \frac{\sum_{i=1}^{N}\left(x_i^{obs} - x_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(x_i^{obs} - \bar{x}^{obs}\right)^2} \quad \text{(IV.11)}$$

$$R_{adj}^2 = 1 - (1 - R^2)\left(\frac{N-1}{N-N_p-1}\right) \quad \text{(IV.12)}$$

$$Q^2 = 1 - \frac{\sum_{i=1}^{N}\left(x_{i\,(training)}^{obs} - x_{i\,(training)}^{pred}\right)^2}{\sum_{i=1}^{N}\left(x_{i\,(training)}^{obs} - \bar{x}_{(training)}^{obs}\right)^2} \quad \text{(IV.13)}$$

$$Q_{CCC}^2 = \frac{2\cdot\sum_{i=1}^{N_{OUT}}\left(x_i^{obs} - \bar{x}^{obs}\right)\cdot\left(x_i^{pred} - \bar{x}^{pred}\right)}{\sum_{i=1}^{N_{OUT}}\left(x_i^{obs} - \bar{x}^{obs}\right)^2 + \sum_{i=1}^{N_{OUT}}\left(x_i^{pred} - \bar{x}^{pred}\right)^2 + N_{OUT}\cdot\left(\bar{x}^{obs} - \bar{x}^{pred}\right)^2} \quad \text{(IV.14)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(x_i^{obs} - x_i^{pred}\right)^2}{N}} \quad \text{(IV.15)}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(x_i^{obs} - x_i^{pred}\right)^2 \quad \text{(IV.16)}$$

$$MAE = \frac{1}{N} * \sum_{i=1}^{N}\left|x_i^{obs} - x_i^{pred}\right| \quad \text{(IV.17)}$$

$$F - value = \frac{MS_R}{MS_E} \quad \text{(IV.18)}$$

$$t - value = \frac{\bar{x}^{obs} - \bar{x}^{pred}}{s\sqrt{\frac{2}{N}}} \quad \text{(IV.19)}$$

$$x^2 = \sum_{i=1}^{N}\frac{\left(x_i^{obs} - x_i^{pred}\right)^2}{x_i^{pred}} \quad \text{(IV.20)}$$

Akaike's Information Criteria (AICs) and Bayesian Information Criteria (BIC) are commonly uses statistical parameters. A model is determined as the best model that represented the experimental result if that model has the smallest value of AIC (or AICs, BIC) [39,424]:

$$AIC = Nln\left(\frac{SSE}{N}\right) \quad \text{(IV.21)}$$

$$AICs = AICs + 2n_p + \frac{2n_p(n_p+1)}{N-(n_p+1)} \quad \text{(IV.22)}$$

$$SSE = \sum_{i=0}^{N}\left(x_i^{obs} - x_i^{pred}\right)^2 \quad \text{(IV.23)}$$

$$BIC = N \cdot ln(MSE) + n_p \cdot ln(N) \quad \text{(IV.24)}$$

Where, N is is the Number of data points in the data set, x_i^{obs} is observed value while $\bar{x}^{obs}$ is its average, x_i^{pred} is the predicted value while $\bar{x}^{pred}$ is its average, N_p is the number of independent variables in each model, MS_R is the mean square regression, MS_E is the mean square residual, s is the pooled standard deviation, and SSE is the error sum of squares. More details can be found elsewhere [30,39,366,411,421–423,425].

Typically, a single statistic value is meaningless; the principal model studied in this chapter was compared to 11 different topologies.

IV.2.8 Visualization

For more evaluation, various graphical metrics were proposes to visually assess the quality and accuracy of the proposed model. A short representation is given [39,280,308]:

- Cross (scatter) plot: is a two-dimensional plot with experimental data (X-axis) and model predictions (Y-axis). The concentration of data around the line Y=X gives an insight into the model's performance. A high concentration shows that the model fits well the experimental data.
- Error distribution graph: illustrations the errors between the predicted and experimental solubilities $\left(x_i^{obs} - x_i^{pred}\right)$. The concentration around the zero-error line is checked in this graph.
- Error distribution histogram: shows the distribution of the prediction error between the predicted and experimental solubilities $\left(x_i^{obs} - x_i^{pred}\right)$ in form of ranges, percentage, or number of data points.
- Insurbia plot: is a plot of diagonal values against predicted values that are used in the case of chemicals without experimental data to provide a visualisation of interpolated and extrapolated predictions.
- William's plot: is a plot of standardised residuals vs. the leverage values. it was utilized to visualise the respective Applicability Domain and to discover the outlier's response (a comprehensive outlier detection).

IV.2.9 Sensibility Analyses

The relevance factor *r* was adopted to determine the influence of each input variable and its contribution to the output. More details is available in Chapter II.

IV.2.10 Model Implementation

The calculations were implemented in the MATLAB® R2021B environment using the ***fitrsvm*** function. The kernel function RBF gave the best results. The Cross-validation method was set to 0.2. At beginning, the meta-heuristic algorithm initially fed the SVR with a random combination of hyper-parameters within their ranges. During numerous iterations depending on the algorithm used, the steps began with the division of data and up until the development of the SVR model. This operation was repeated many times until the minimum value of RMSE was obtained. Then, the I-GWO algorithm generated a new population of hyper-parameters for the SVR algorithm. The same set of steps was repeated until a best new fitness value was reached for 100 iterations. The processors were repeated serval times to ensure the robustness. **Figure IV.2.** illustrate the support vector regression structure while the flowchart of the proposed I-GWO-SVR(r) is represented in **Figure IV.3**.

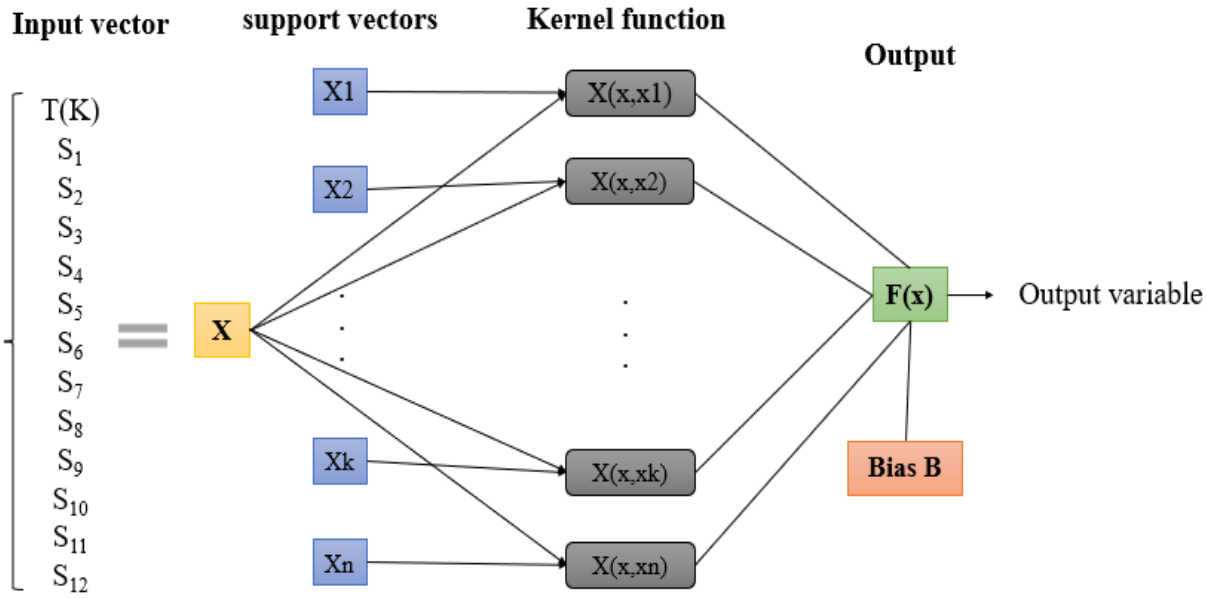


Figure IV.2. Support vector regression structure.

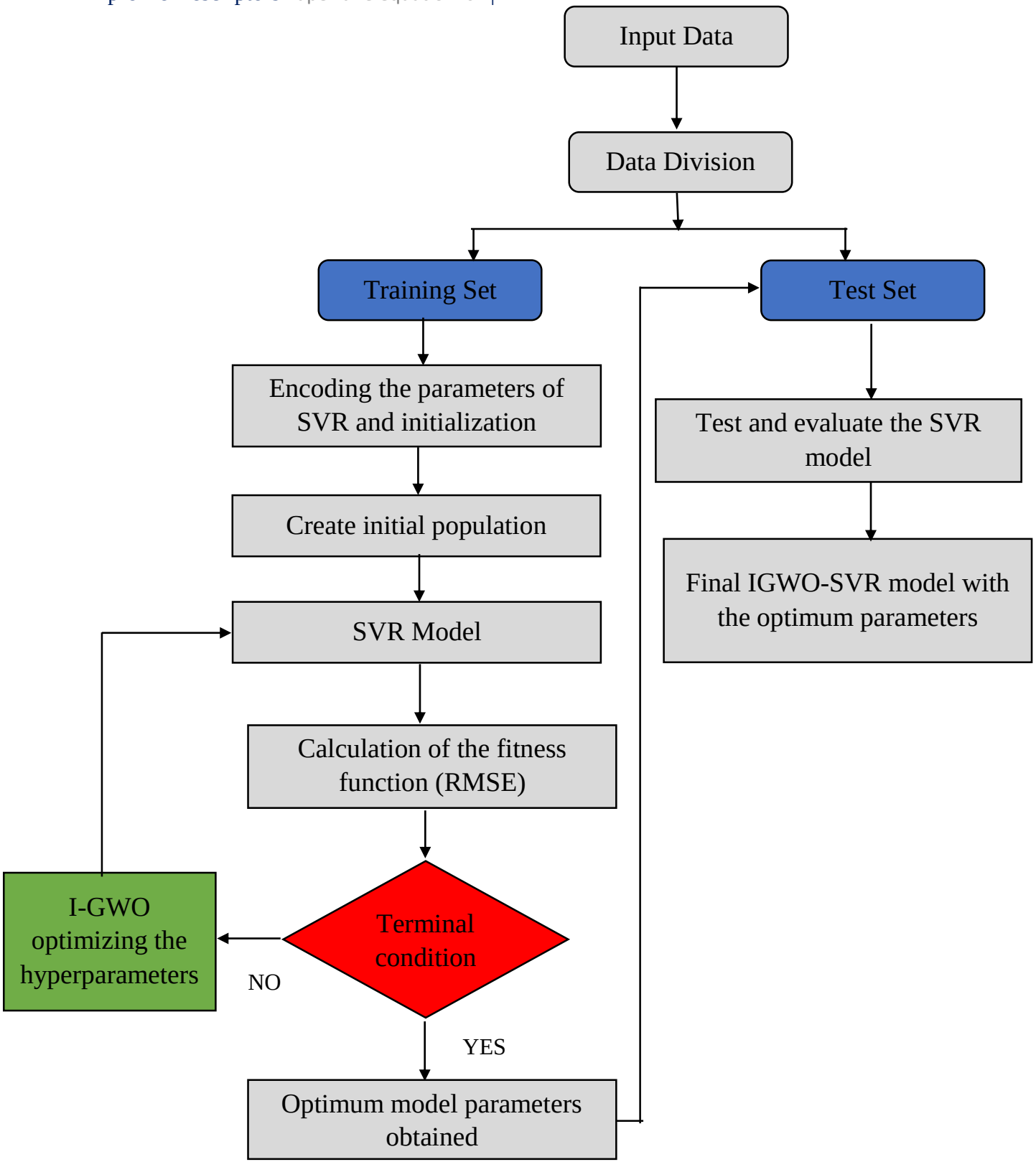


Figure IV.3. The flowchart of the proposed I-GWO-SVR(r) model

IV.3 Results and Discussion

IV.3.1 Performance Comparison

In this chapter, the proposed I-GWO-SVR(r) approach was compared with 11 different models namely four artificial neural network topologies (**Table IV.2.**) and 7 hybrid support vector regression algorithms (**Table IV.3.**) in order to verify and to validate the ability of the proposed model to predict the solubility of API-IL as well to confirm the feasibility and the efficiency of the proposed metaheuristic algorithm (I-GWO) to optimize the SVR hyper-parameters. All 12 models were elaborated using a MATLAB environment and employed using the same set of input descriptors (13) and the same training and testing percentages (80% for training and 20% for test). The RMSE was set as an objective function.

Table IV.2. shows the four (4) selected artificial neural network topologies, the chosen training, and activation functions, and their architectures. Note that, a popular ANN architecture including a three-layer feed forward network along with a very effective combination of transfer functions used in this paper. More details regarding this algorithm can be found in **Chapter II**. The number of hidden neurons was fixed based on the trial rule.

Table IV.2. The artificial neural network models and their selected abbreviation, transfer functions, and obtained architecture.

Model Abbreviation	Training Function	Transfer Function hidden layer	Transfer function output layer	Architecture
ANN-lm-tan	Trainlm	Tansig	Purelin	{13-12-1}
ANN-lm-log	Trainlm	Logsig	Purelin	{13-14-1}
ANN-br-tan	Trainbr	Tansig	Purelin	{13-11-1}
ANN-br-log	Trainbr	Logsig	Purelin	{13-12-1}

Table IV.3. illustrations the eight (8) hybrid support vector regression models together with their selected kernel function and optimized hyperparameters. As mentioned above, the predictive performance of the I-GWO was compared with other common algorithms such as Multi-Verse Optimizer (MVO) [426], Moth-Flame Optimization (MFO) [427], and Sine Cosine Algorithm (SCA) [428] applied to solve this kind of problem. The parameters of these

algorithms were defined as basic. For the population size and the number of iterations, they were assumed to be the same as in I-GWO.

Table IV.3. The hybrid support vector regression models and their selected abbreviation, kernel function, and obtained hyperparameters.

Model Abbreviation	Metaheuristic	Kernel Function	Hyperparameters		
			C	**σ**	**ε**
I-GWO-SVR(g)	I-GWO	Gaussian	155.4449	1.2331	0.0069
I-GWO-SVR(r)	I-GWO	RBF	199.4019	1.3181	0.0073
MVO-SVR(g)	MVO	Gaussian	89.3403	1.5466	0.0060
MVO-SVR(r)	MVO	RBF	200.0000	1.4729	0.0060
MFO-SVR(g)	MFO	Gaussian	173.6828	1.6016	0.0063
MFO-SVR(r)	MFO	RBF	175.7414	1.2007	0.0060
SCA-SVR(g)	SCA	Gaussian	154.9736	1.4375	0.0070
SCA-SVR(r)	SCA	RBF	191.5265	1.8484	0.0061

The computational efficiency and the prediction ability of the model was highlight. **Tables IV.5 to IV.7** summarize a statistical comparison between all the 12 proposed models in terms of relevant statistics, the most commonly used statistics while **Figure IV.4 to IV.15** represents the scatter plots of all the 12 developed models. i.e., the predicted versus observed values for the training, test, and global set and cross-validation set for both ANN-lm-tan and ANN-lm-log models.

The analysis of the statistical metrics results listed in **Table IV.5.** and **Table IV.6.** of the 12 models for their different sets indicated that their values are in agreement with the acceptability criteria with a global correlation coefficient (R & R_{adj}) greater than 0.9980, and an overall coefficient of determination (R^2 & R^2_{adj}) above 0.9959 similar to the robustness (Q^2 & Q^2_{ccc}) greater than 0.9960 as well the three error parameters RMSE, MSE, and MAE showed high accuracies less than 0.0150, 0.0002 and 0.0082, respectively.

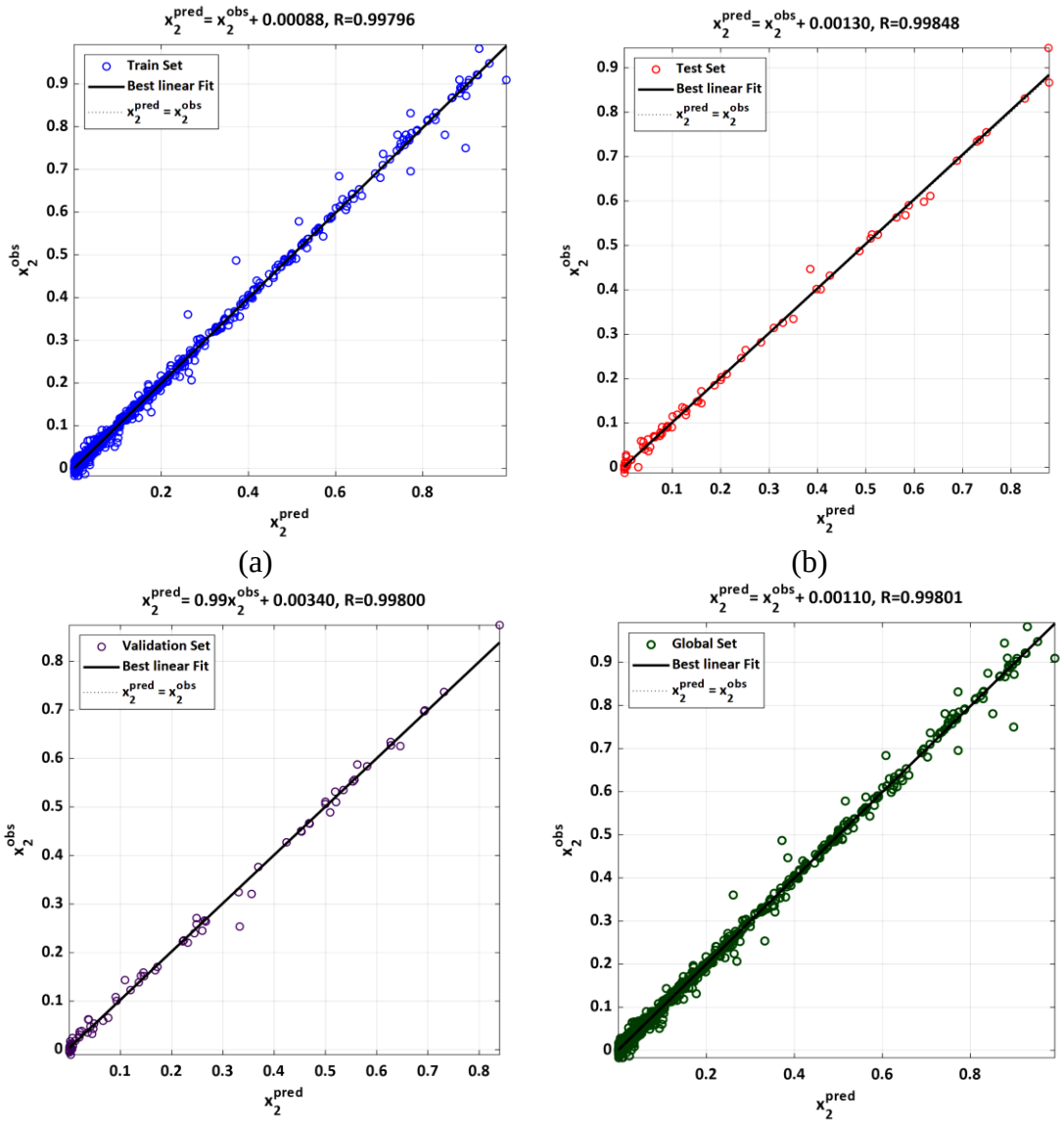


Figure IV.4. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-lm-tan : (a) Train set, (b) Test set, (c) Cross-validation, (d) Global.

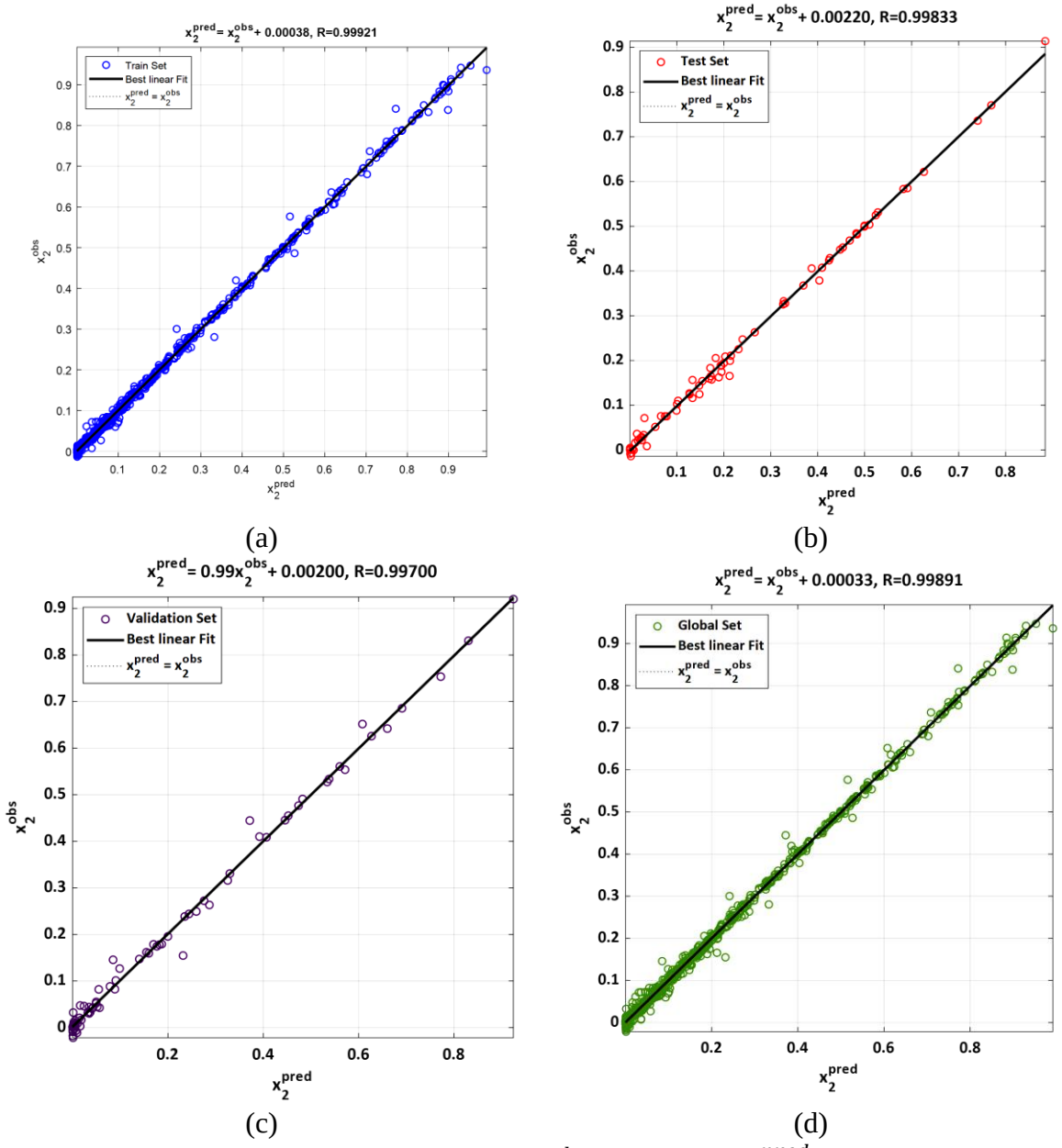


Figure IV.5. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-lm-log : (a) Train set, (b) Test set, (c) Cross-validation, (d) Global.

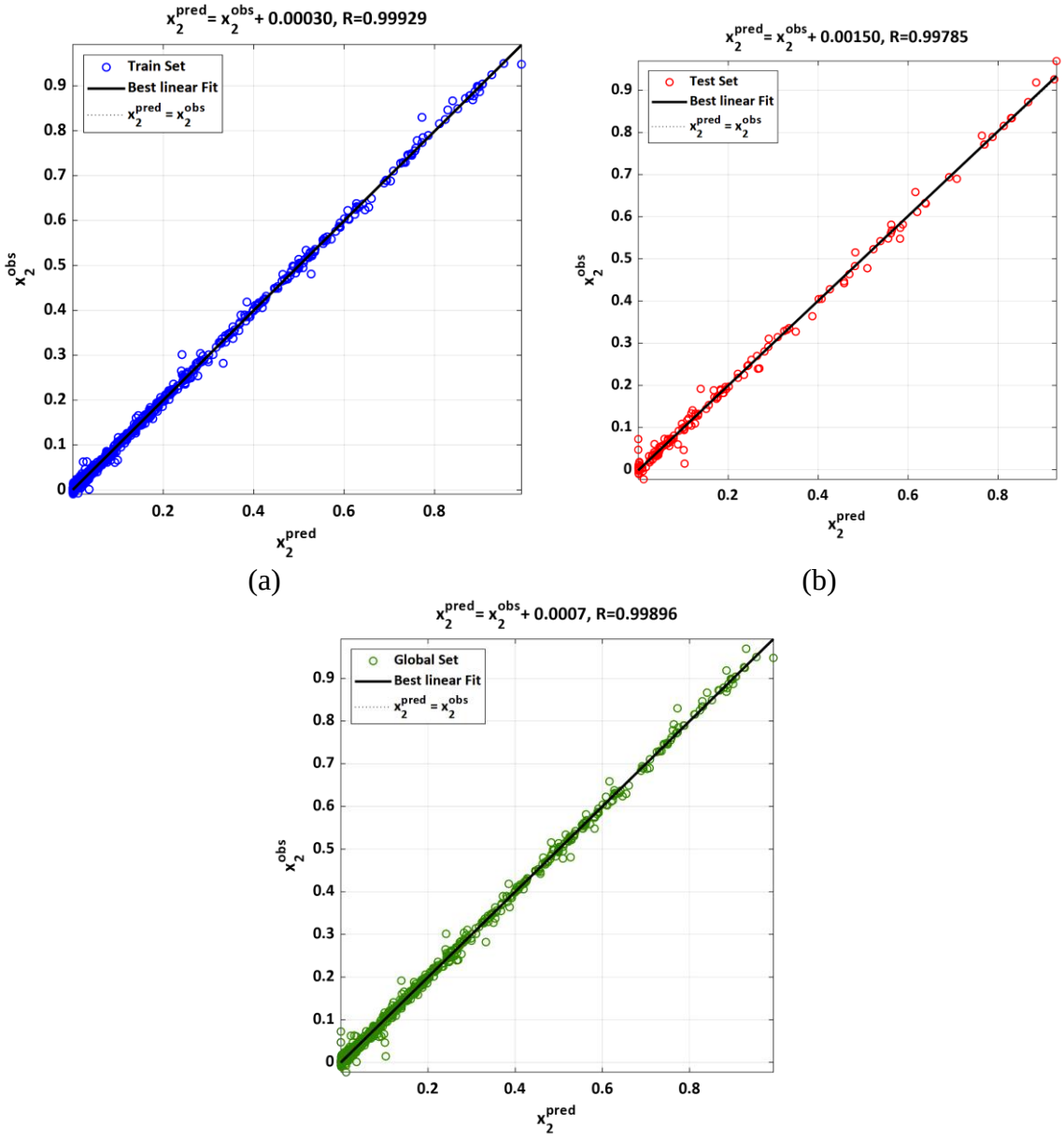


Figure IV.6. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-br-tan : (a) Train set, (b) Test set, (c) Global.

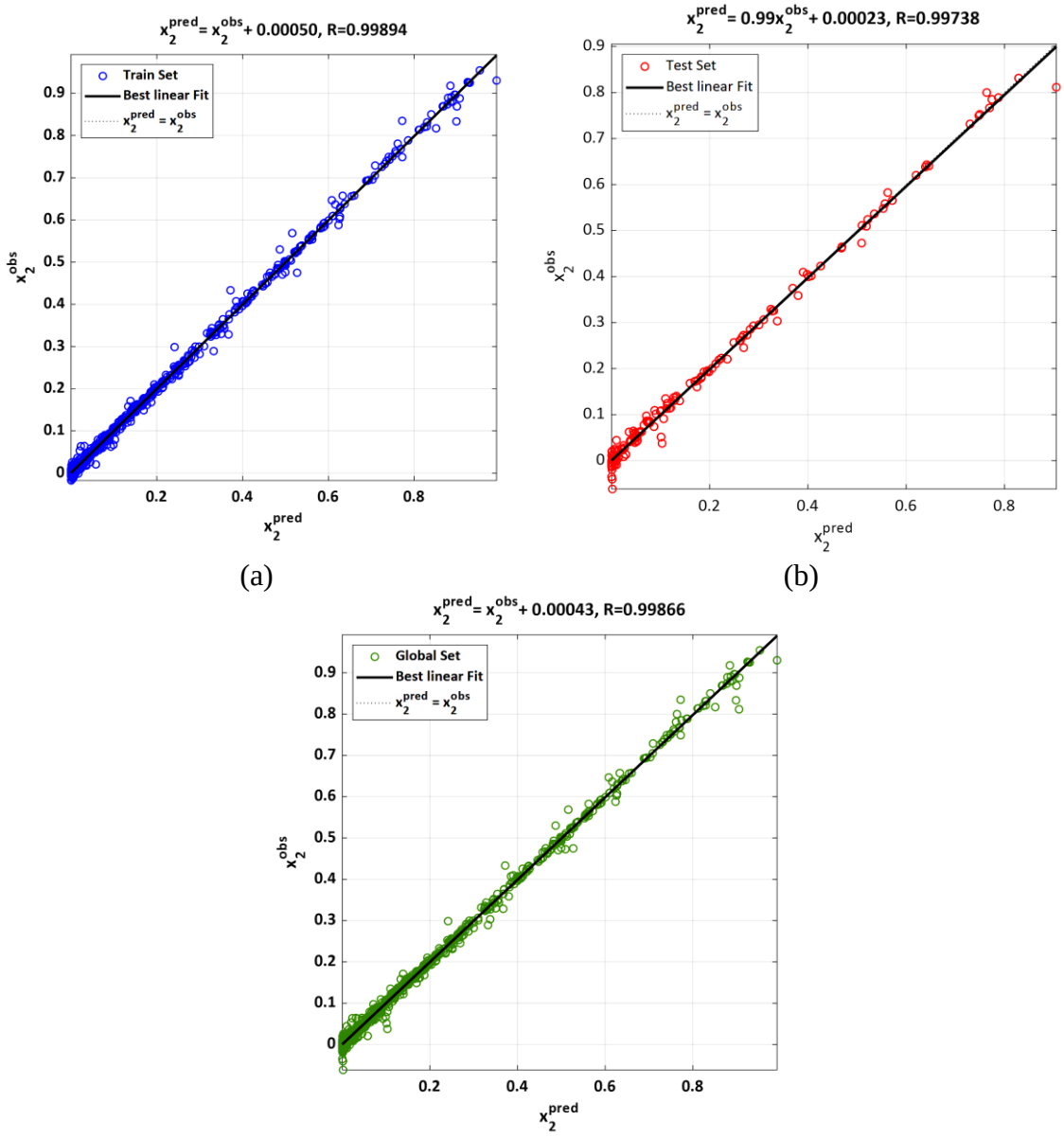


Figure IV.7. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model ANN-br-log : (a) Train set, (b) Test set, (c) Global.

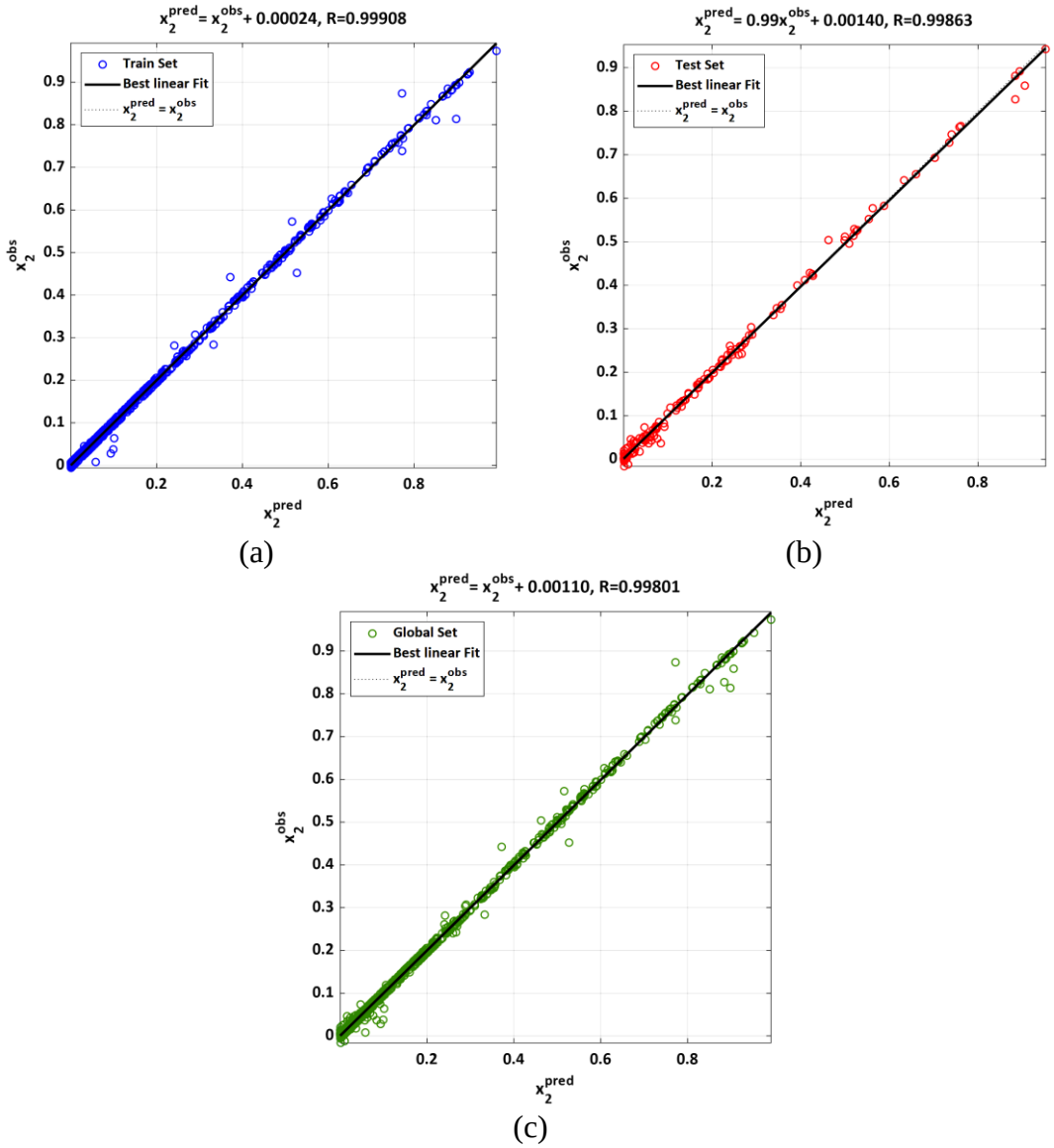


Figure IV.8. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model I-GWO-SVR(g) : (a) Train set, (b) Test set, (c) Global.

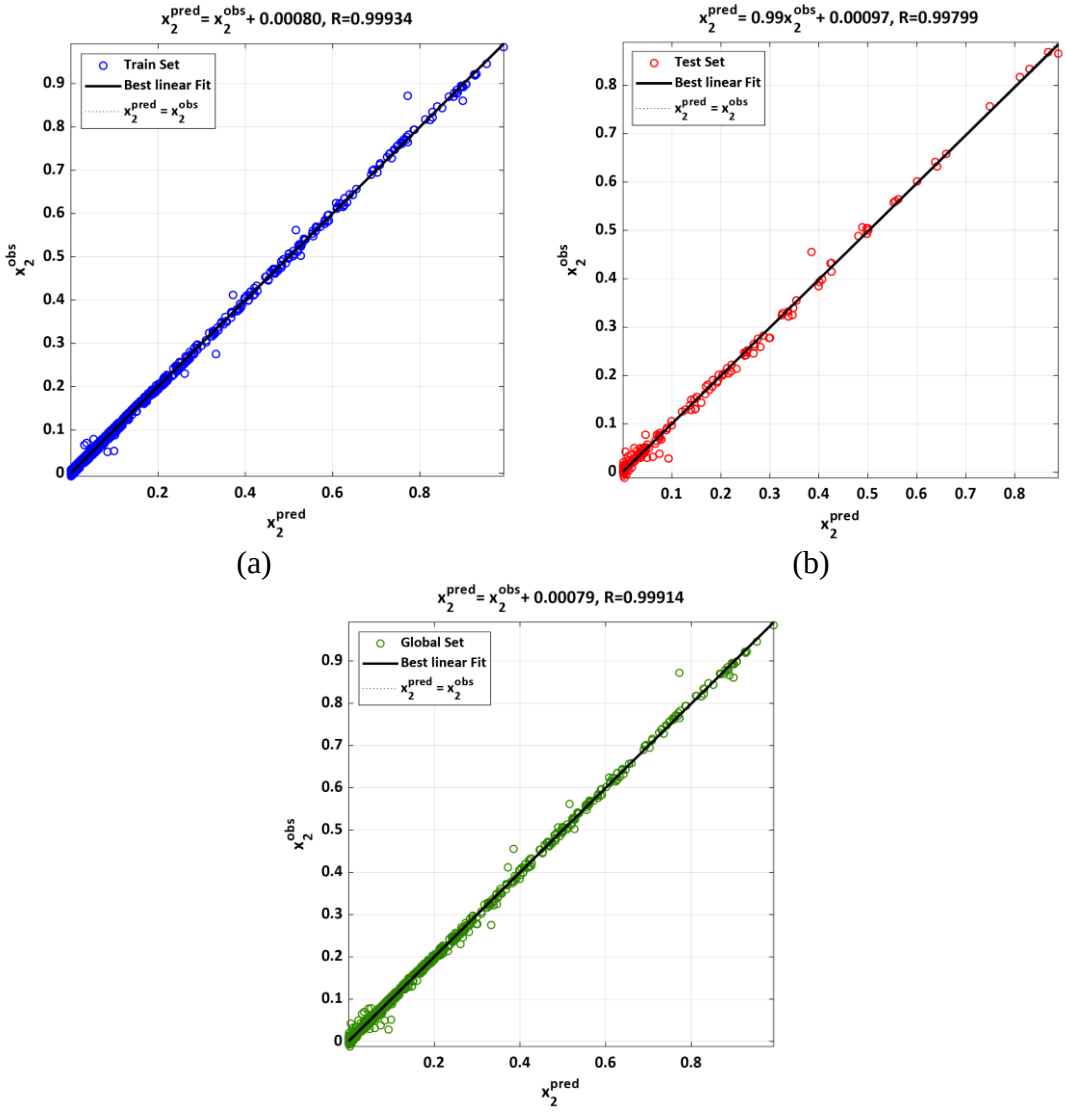


Figure IV.9. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model I-GWO-SVR(r) : (a) Train set, (b) Test set, (c) Global.

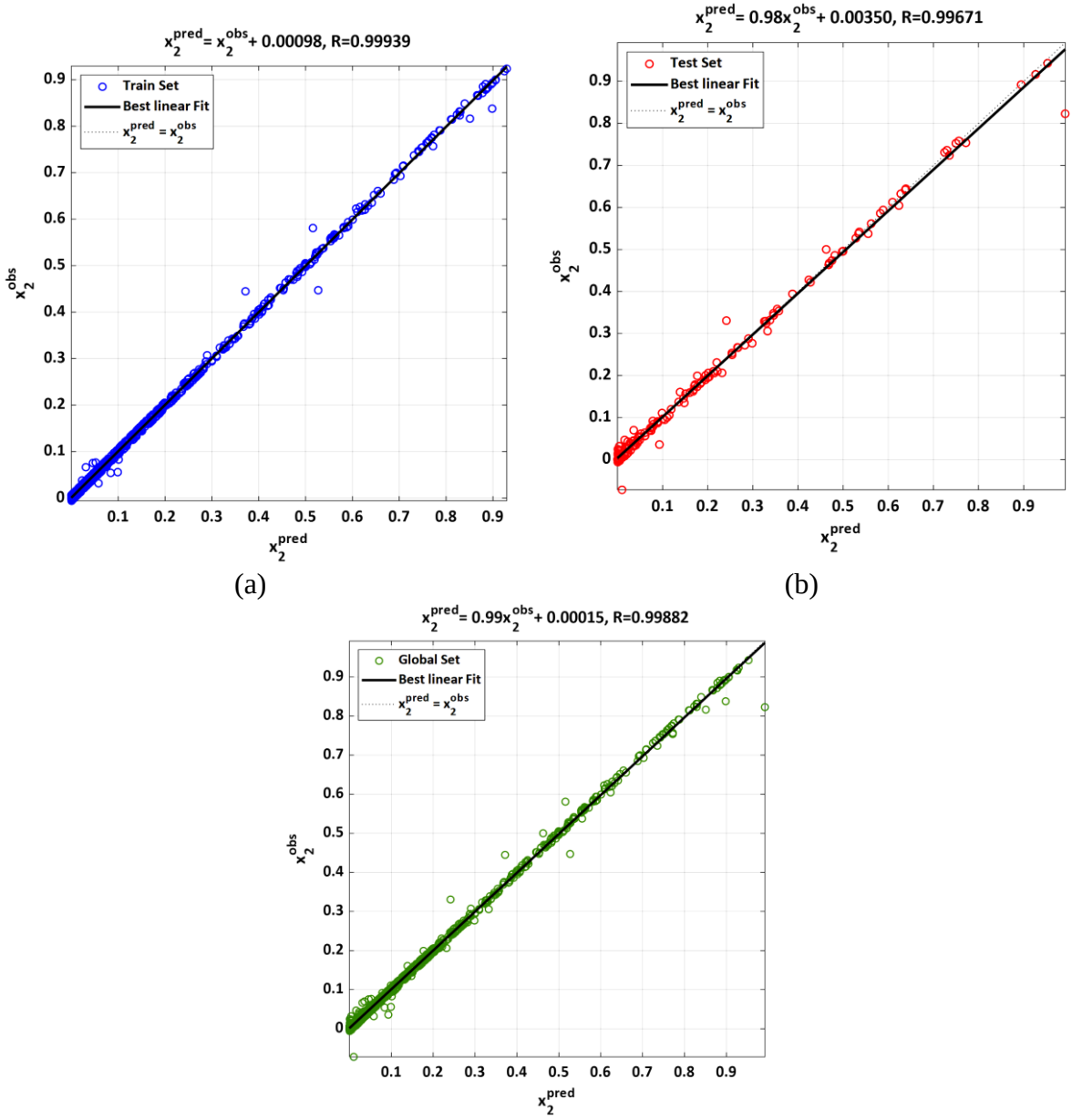


Figure IV.10. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MVO-SVR(g) : (a) Train set, (b) Test set, (c) Global.

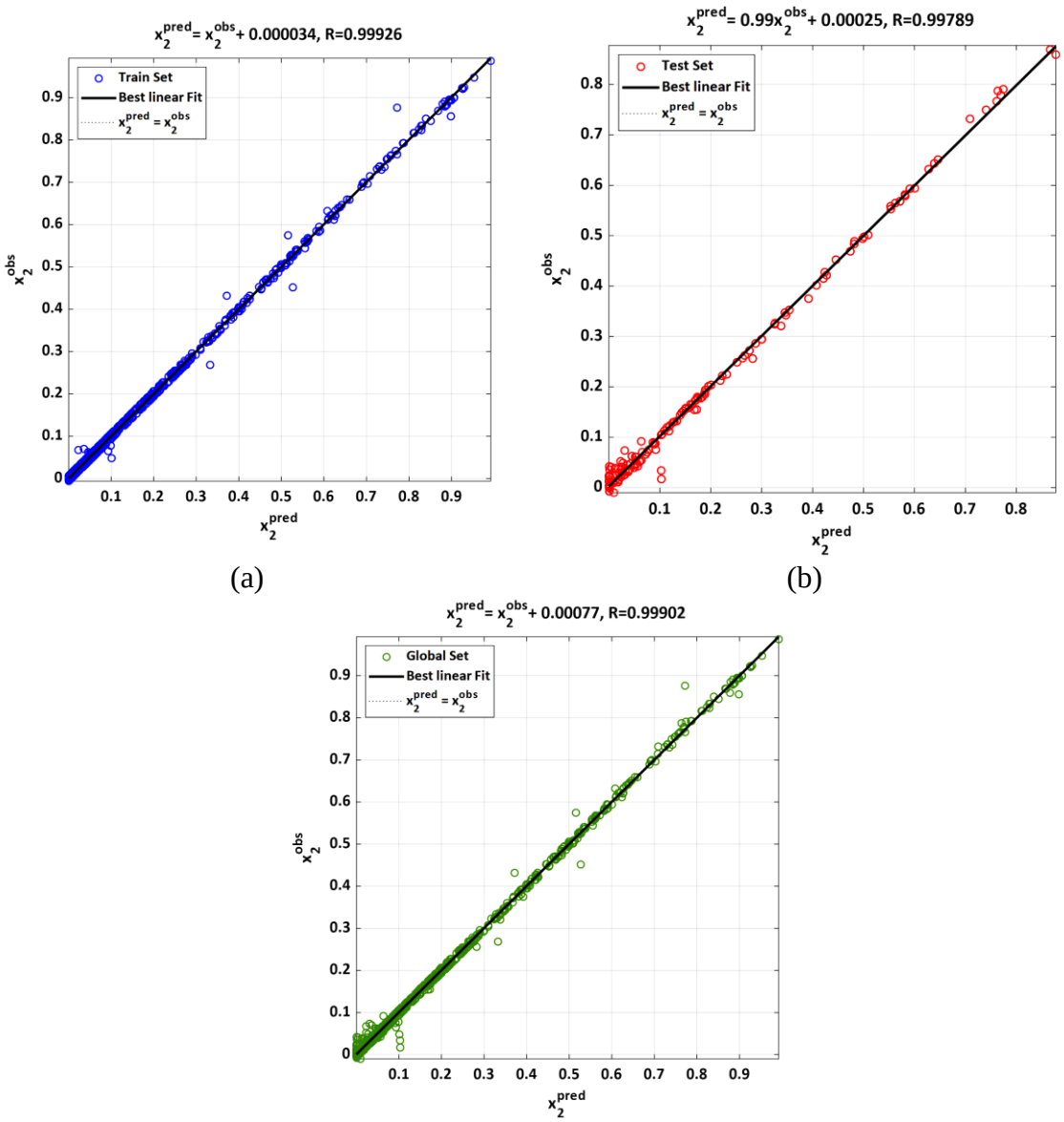


Figure IV.11. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MVO-SVR(r) : (a) Train set, (b) Test set, (c) Global.

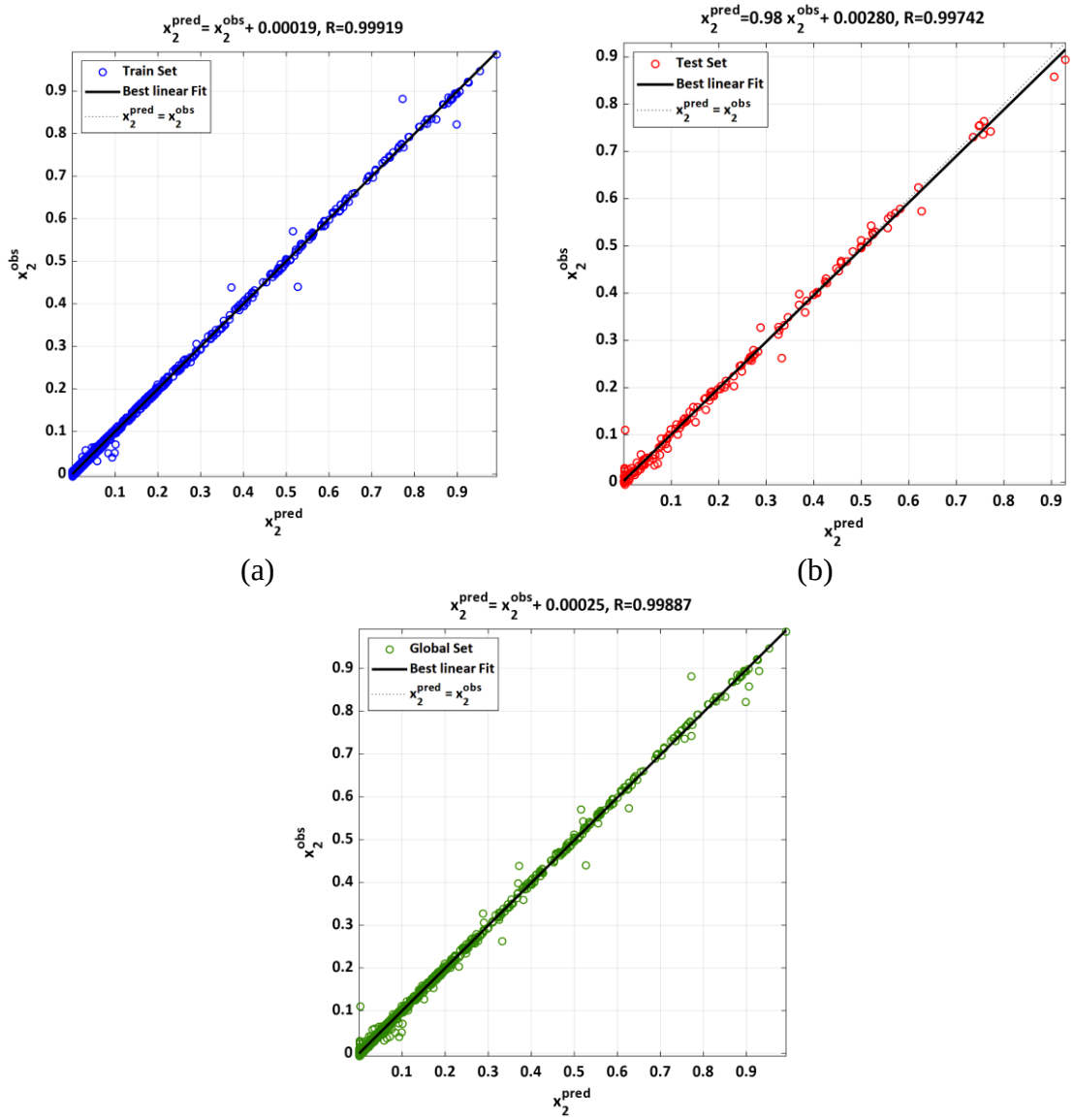


Figure IV.12. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MFO-SVR(g) : (a) Train set, (b) Test set, (c) Global.

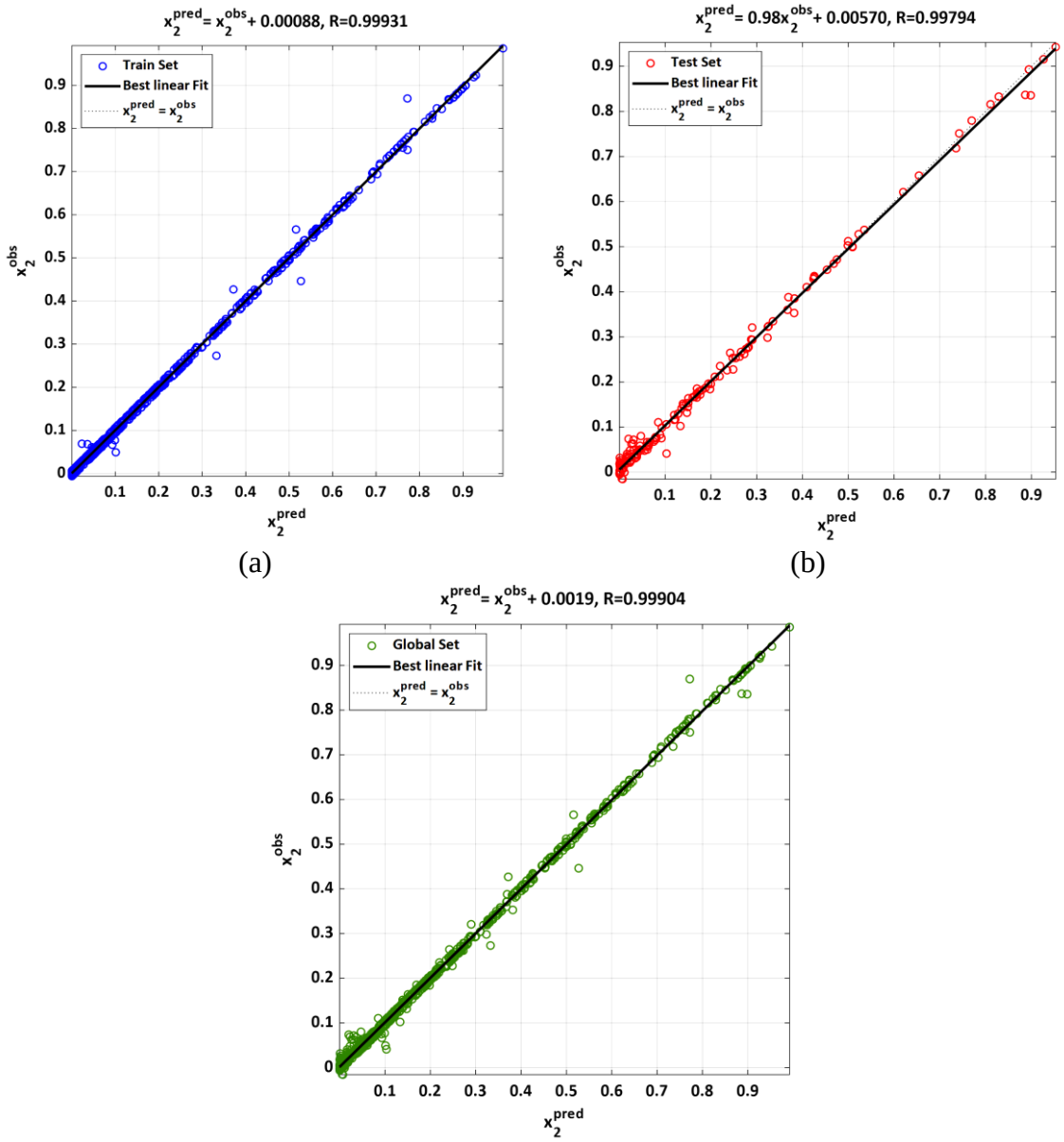


Figure IV.13. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model MFO-SVR(r) : (a) Train set, (b) Test set, (c) Global.

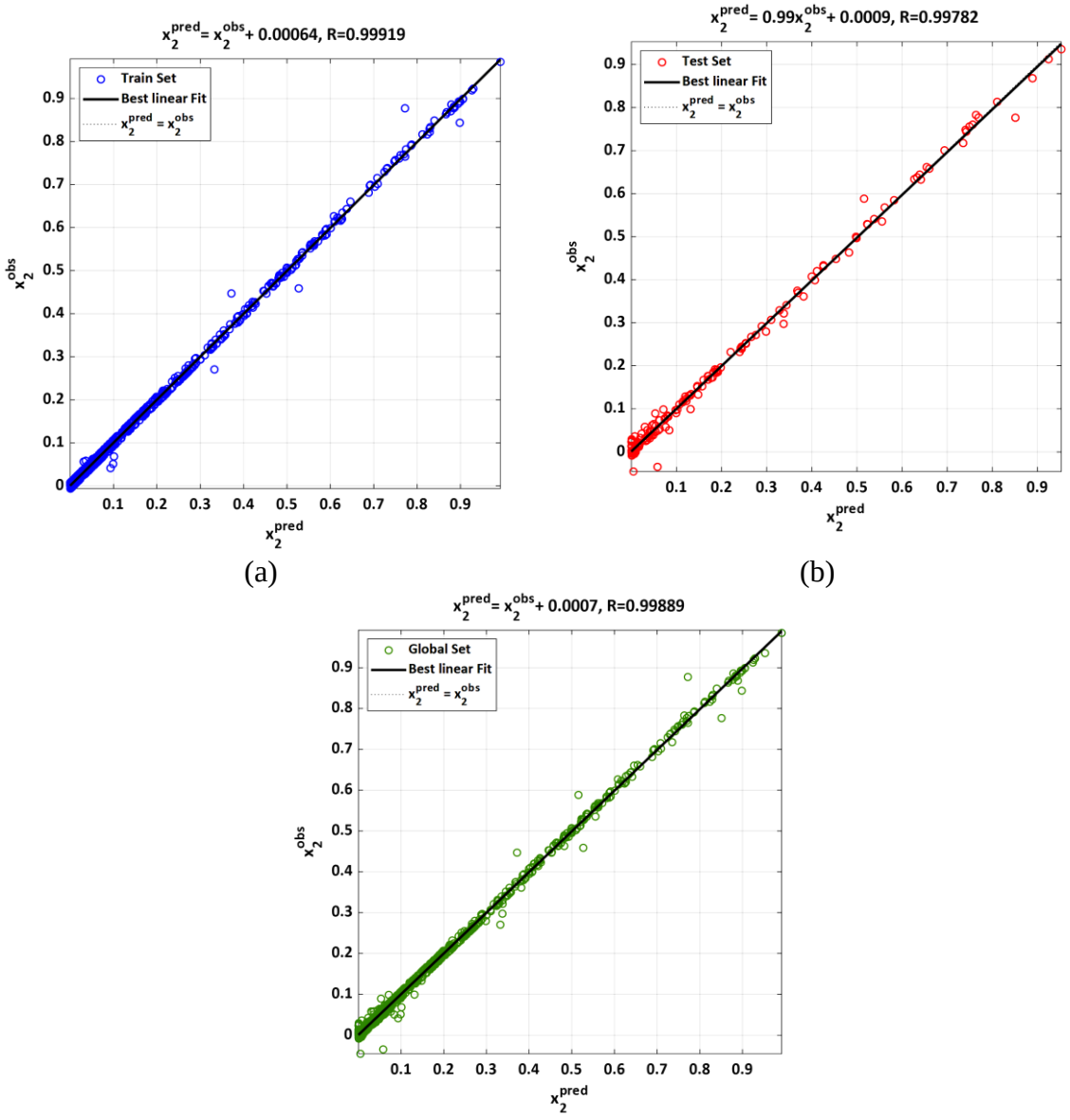


Figure IV.14. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model SCA-SVR(g) : (a) Train set, (b) Test set, (c) Global.

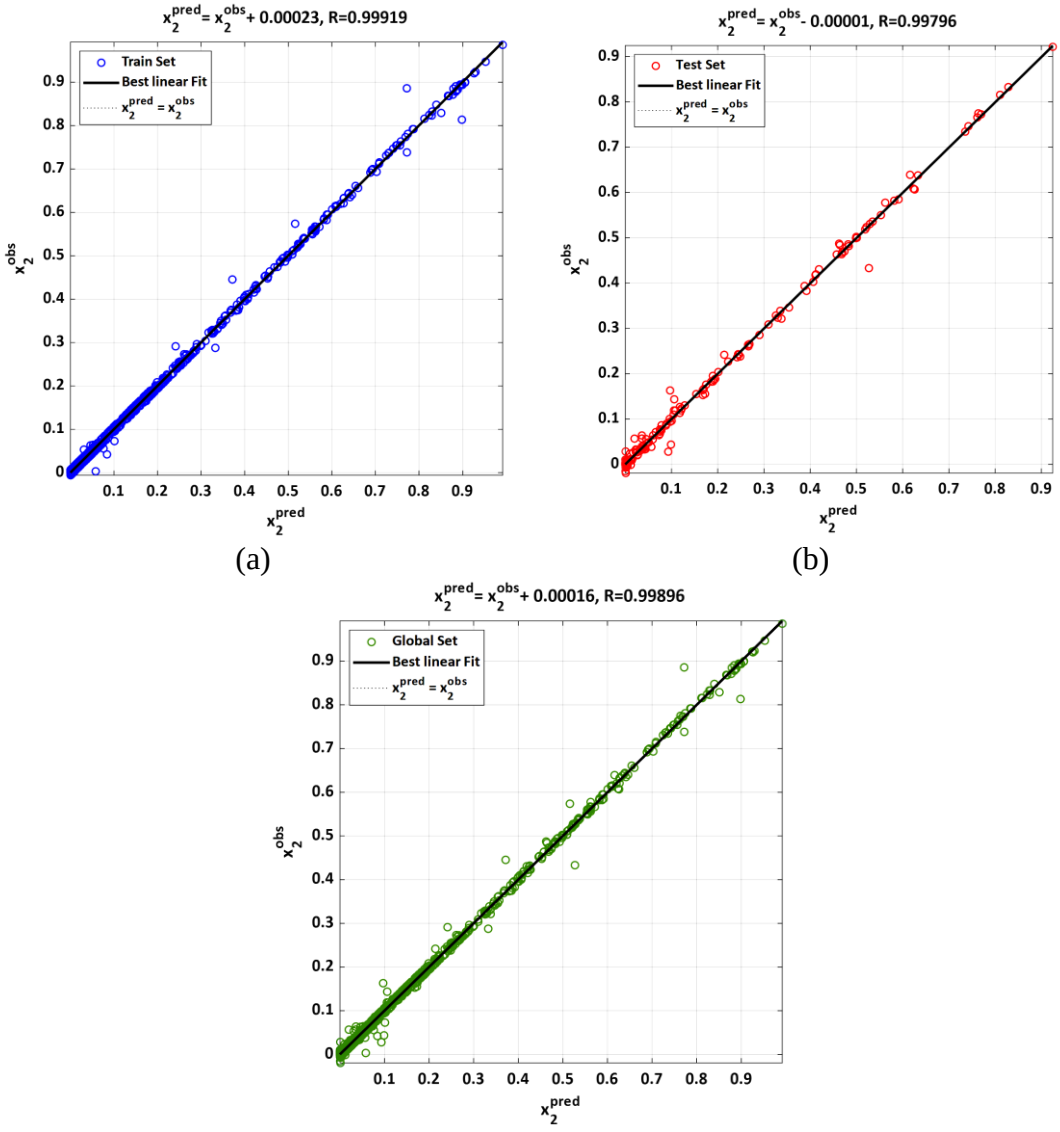


Figure IV.15. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in IL, model SCA-SVR(r) : (a) Train set, (b) Test set, (c) Global.

Table IV.5. Statistical metrics for the global train, test, and inter-validation set for the 4 ANN models.

Model	**ANN-lm-tan**				**ANN-lm-log**				**ANN-br-tan**			**ANN-br-log**		
Set	**Global**	**Train**	**Test**	**Val***	**Global**	**Train**	**Test**	**Val***	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**
R	0.9980	0.9980	0.9985	0.9980	0.9989	0.9992	0.9983	0.9970	0.9990	0.9993	0.9979	0.9987	0.9989	0.9974
R_{adj}	0.9980	0.9979	0.9981	0.9976	0.9989	0.9992	0.9980	0.9964	0.9989	0.9993	0.9976	0.9986	0.9989	0.9971
R^2	0.9960	0.9959	0.9968	0.9959	0.9978	0.9984	0.9966	0.9940	0.9979	0.9986	0.9956	0.9973	0.9979	0.9947
R^2_{adj}	0.9959	0.9958	0.9962	0.9951	0.9978	0.9984	0.9960	0.9929	0.9979	0.9986	0.9952	0.9973	0.9979	0.9942
Q^2	0.9960	0.9959	0.9968	0.9960	0.9978	0.9984	0.9966	0.9940	0.9979	0.9986	0.9956	0.9973	0.9979	0.9948
Q^2_{ccc}	0.9980	0.9980	0.9984	0.9979	0.9989	0.9992	0.9983	0.9970	0.9990	0.9993	0.9978	0.9987	0.9989	0.9974
RMSE	0.0150	0.0151	0.0142	0.0152	0.0111	0.0096	0.0125	0.0184	0.0109	0.0089	0.0166	0.0111	0.0163	0.0123
MSE	2.2462e-04	2.2691e-04	2.0116e-04	2.2971e-04	1.2287e-04	9.1513e-05	1.5712e-04	3.3949e-04	1.1773e-04	7.8373e-05	2.7516e-04	1.5096e-04	1.2260e-04	2.6439e-04
MAE	0.0082	0.0079	0.0088	0.0098	0.0064	0.0056	0.0082	0.0109	0.0065	0.0057	0.0099	0.0074	0.0064	0.0099
F-value	0.0570	0.0014	0.2674	0.2278	0.0026	0.0086	0.1655	0.1038	0.0046	0.0002	0.0153	0.0281	0.0003	0.1033
t- value	0.1619	0.0034	0.2303	0.2471	0.0059	0.0148	0.1832	0.1244	0.0105	0.0004	0.0213	0.0718	0.0005	0.1790
x^2	0.0009	0.0011	0.0008	0.0006	0.0004	0.0003	0.0003	0.0008	0.0003	0.0003	0.0005	0.0005	0.0005	0.0008

*Cross validation set of the activation function “trainlm”.

Table IV.6. Statistical metrics for the global train, and test set for the 8 hybrid SVR models.

Model	**I-GWO-SVR(g)**			**I-GWO-SVR(r)**			**MVO-SVR(g)**			**MVO-SVR(r)**		
Set	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**
R	0.9990	0.9991	0.9986	**0.9991**	0.9993	0.9980	0.9988	0.9994	0.9967	0.9990	0.9993	0.9979
R_{adj}	0.9990	0.9991	0.9985	**0.9991**	0.9993	0.9978	0.9988	0.9994	0.9964	0.9990	0.9992	0.9977
R^2	0.9980	0.9982	0.9972	**0.9983**	0.9987	0.9960	0.9976	0.9988	0.9933	0.9980	0.9985	0.9957
R^2_{adj}	0.9980	0.9982	0.9970	**0.9983**	0.9987	0.9957	0.9976	0.9988	0.9927	0.9980	0.9985	0.9953
Q^2	0.9980	0.9982	0.9972	**0.9983**	0.9987	0.9962	0.9976	0.9988	0.9933	0.9980	0.9985	0.9958
Q^2_{ccc}	0.9990	0.9991	0.9986	**0.9991**	0.9993	0.9980	0.9988	0.9994	0.9966	0.9990	0.9993	0.9979
RMSE	0.0107	0.0121	0.0103	**0.0098**	0.0089	0.0131	0.0116	0.0083	0.0199	0.0105	0.0093	0.0145
MSE	1.1428e-04	1.0595e-04	1.4756e-04	**9.6987e-05**	7.8547e-05	1.7075e-04	1.3423e-04	6.8593e-05	3.9677e-04	1.1125e-04	8.6233e-05	2.1132e-04
MAE	0.0061	0.0056	0.0082	**0.0067**	0.0062	0.0085	0.0060	0.0052	0.0094	0.0063	0.0056	0.0090
F-value	0.0141	0.0046	0.0651	**0.0705**	0.0991	0.0134	0.0335	0.0780	0.0654	0.1010	0.0753	0.1647
t-value	0.0300	0.0078	0.0885	**0.1637**	0.1818	0.0235	0.0704	0.1276	0.0915	0.2204	0.1272	0.2631
x^2	0.0005	0.0004	0.0005	**0.0003**	0.0003	0.0004	0.0006	0.0003	0.0018	0.0003	0.0004	0.0002

Table IV.6. Statistical metrics for the global train, and test set for the 8 hybrid SVR models (continued).

Model	**MFO-SVR(g)**			**MFO-SVR(r)**			**SCA-SVR(g)**			**SCA-SVR(r)**		
Set	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**	**Global**	**Train**	**Test**
R	0.9989	0.9992	0.9974	0.9990	0.9993	0.9979	0.9989	0.9992	0.9978	0.9990	0.9992	0.9980
R_{adj}	0.9988	0.9992	0.9971	0.9990	0.9993	0.9976	0.9994	0.9992	0.9976	0.9989	0.9992	0.9978
R^2	0.9977	0.9984	0.9946	0.9981	0.9986	0.9956	0.9989	0.9984	0.9956	0.9979	0.9984	0.9959
R^2_{adj}	0.9977	0.9984	0.9941	0.9981	0.9986	0.9952	0.9989	0.9984	0.9952	0.9979	0.9984	0.9956
Q^2	0.9977	0.9984	0.9947	0.9981	0.9986	0.9956	0.9978	0.9984	0.9957	0.9979	0.9984	0.9959
Q^2_{ccc}	0.9989	0.9992	0.9973	0.9990	0.9993	0.9977	0.9989	0.9992	0.9978	0.9990	0.9992	0.9980
RMSE	0.0113	0.0098	0.0161	0.0105	0.0089	0.0153	0.0112	0.0094	0.0164	0.0109	0.0096	0.0149
MSE	1.2795e-04	9.5162e-05	2.5911e-04	1.0999e-04	7.9009e-05	2.3394e-04	1.2527e-04	8.9270e-05	2.6925e-04	1.1800e-04	9.2243e-05	2.2104e-04
MAE	0.0064	0.0056	0.0093	0.0065	0.0056	0.0102	0.0067	0.0059	0.0098	0.0060	0.0055	0.0082
F-value	0.0243	0.0244	0.1427	0.1379	0.1194	0.1783	0.0064	0.0301	0.0507	0.0237	0.0442	0.0313
t-value	0.0536	0.0413	0.2396	0.3115	0.2038	0.3027	0.0149	0.0567	0.0710	0.0497	0.0746	0.0416
x^2	0.0005	0.0004	0.0007	0.0004	0.0003	0.0006	0.0004	0.0004	0.0006	0.0005	0.0004	0.0005

In addition, the evaluation of the three ANOVA statistics: f-value, t-value, and chi-square appear to be statistically significant. In fact, all the statistical results were found to be similar for all three stages of model development, namely, training, cross-validation, and testing. Such a result allowed to confirm that the APIs solubility in ILs is indeed a very suitable property to be correlated with σ-profile descriptors and the temperature. The predictive capacity of the models was confirmed by very similar values of the three error parameters obtained for the training and testing set.

A " fortunate " data-sharing question can be assumed but for the data partition, instead of being randomly selected, it was determined based on predefined and well-established MATLAB functions. The "holdout" and the "dividerand" for cross-validation were used for the ANN and SVR algorithms, respectively to ensure that the hybrid SVR optimal model is independent of the validation data. Consequently, the improvement in the predictive accuracy of the model was the result of the swarm optimization of the learning machine hyperparameters and not the result of over-fitting. More details are available elsewhere [429–431]. Noted that the calculation was performed several times to ensure robustness.

Moreover, the analysis of the results of all models revealed that, in almost all cases the hydride SVR models showed better results than the ANN architectures and strongly proved the high exploration ability of I-GWO by exploring the search area extensively and give promising regions of the search area. i.e., better in fine-tuning the hyperparameters of the SVR algorithm compared to the others.

Both values of the Akaike (AIC & AICs) and Bayesian (BIC) Information Criteria represented in **Table IV.7.** proved that the experimental data are well-fitted by the I-GWO-SVR(r) model and confirmed that the best intelligence performance was obtained by the I-GWO-SVR(r) algorithm. As shown in **Figure IV.4.-IV.15** The observed and the predicted solubility data were compared with an excellent agreement for the elements of the global (panel a), training (panel b), test (panel c), and validation (panel d) sets. In fact, for all systems, the predicted values did not show a considerable deviation from the experimental data. Therefore, all 12 models showed significant ability to correlate the experimental data accurately. Also, according to the results listed in **Table IV.5-IV.7**, the intelligent model I-GWO-SVR(r) had the best estimates among all predictive tools for estimating APIs solubility in ILs.

Table IV.7. AIC & BIC values for the global set of the 12 applied models.

Model	SSE	MSE	N	Np	AIC	AICs	BIC
ANN-lm-tan	0.1730	2.25E-04	770	13	-6442.66	-6442.29	-6381.14
ANN-lm-log	0.0946	1.23E-04	770	13	-7620.95	-7620.58	-7484.25
ANN-br-tan	0.0907	1.18E-04	770	13	-7656.36	-7655.99	-7519.15
ANN-br-log	0.1162	1.51E-04	770	13	-7448.00	-7447.63	-7311.76
SVR-IGWO-G	0.0749	1.14E-04	770	13	-7817.33	-7816.96	-7369.39
SVR-IGWO-R	**0.0747**	**9.70E-05**	**770**	**13**	**-7819.58**	**-7819.21**	**-7683.96**
SVR-MVO-G	0.1034	1.34E-04	770	13	-7546.15	-7545.78	-7412.21
SVR-MVO-R	0.0857	1.11E-04	770	13	-7704.05	-7703.68	-7570.58
SVR-MFO-G	0.0985	1.28E-04	770	13	-7586.98	-7586.61	-7450.74
SVR-MFO-R	0.0847	1.10E-04	770	13	-7713.92	-7713.55	-7578.19
SVR-SCA-G	0.0965	1.25E-04	770	13	-7604.23	-7603.86	-7470.68
SVR-SCA-R	0.0909	1.18E-04	770	13	-7654.51	-7654.14	-7519.15

IV.3.2 External Validation Test for the I-GWO-SVR(r) Model

As mentioned in the previous sections, the database was consisted of 13 inputs, and 850 experimental data points were divided into two main sets. The first set formed of 770 data points was used for the model development, while the second set formed of 80 data points was hidden for the external validation test. **Figure IV.16.** illustrates the cross plots of the experimental APIs solubility in various ILs against the predictions of the developed I-GWO-SVR(r) model from the external validation dataset only. Further analysis of **Figure IV.16.** showed that the data points remained close to the Y=X line, analogous to the training and test sets, indicating the high accuracy of the proposed model.

The validation of the I-GWO-SVR(r) model developed was carried out using the external validation metrics R, R^2, Q^2, RMSE, MSE and MAE. The results are summarized in Table. As shown in **Table IV.8.**, the model has a good predictive ability. Noted that the obtained values are similar to the results of the training and test sets confirming the robustness, good-of-fit, accuracy, and relevance of the new model (predictive power).

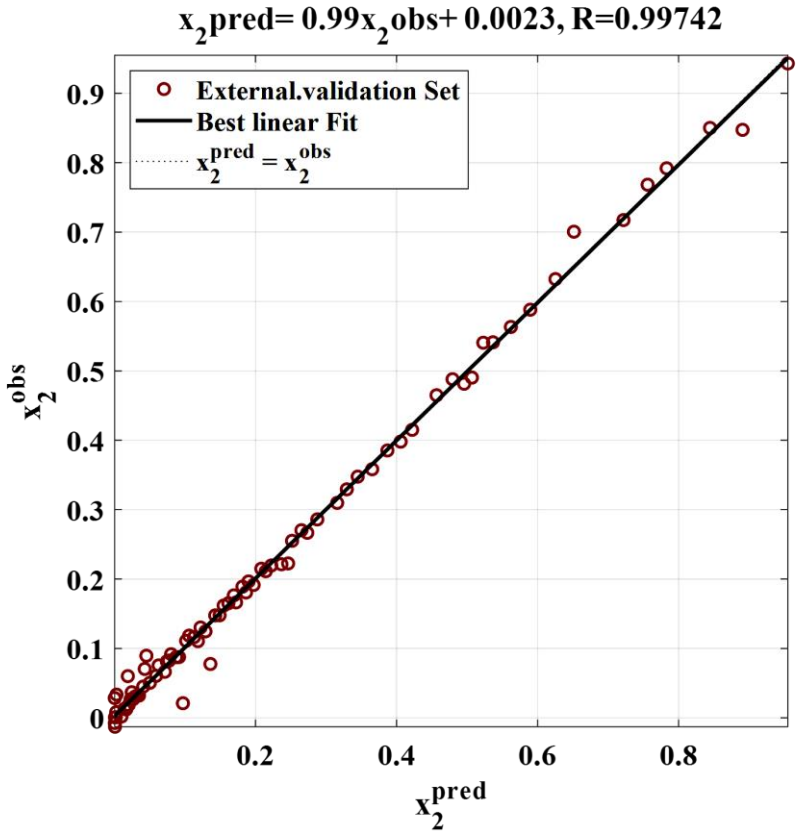


Figure IV.16. Cross plot of the experimental (x_2^{obs}) vs. predict (x_2^{pred}) API solubility in Ils of the external validation set.

Table IV.8. Statistical metrics for the external validation set for the I-GWO-SVR(r) model.

External Validation Metrics	*R*	*R²*	*Q²*	*RMSE*	*MSE*	*MAE*
Value	0.9974	0.9948	0.9948	0.0173	2.9916e-04	0.0105

IV.3.3 Error Distribution for the I-GWO-SVR(r) Model

The errors between the experimental and predicted solubility are presented in **Figure IV.17.** and the histogram of the prediction deviation is illustrated in **Figure IV.18**. The aim of this study was to assess the performance.

As shown, the most deviations are close to the zero-error trend and a minor proportion shows an absolute error greater than 0.02 for all three sets. All errors are in the range of [-0.08,0.08], indicating excellent performance of the established model in estimating the target

property. These plots demonstrated that the obtained I-GWO-SVR(r) model very significantly represents the APIs solubility in ILs at various temperature conditions.

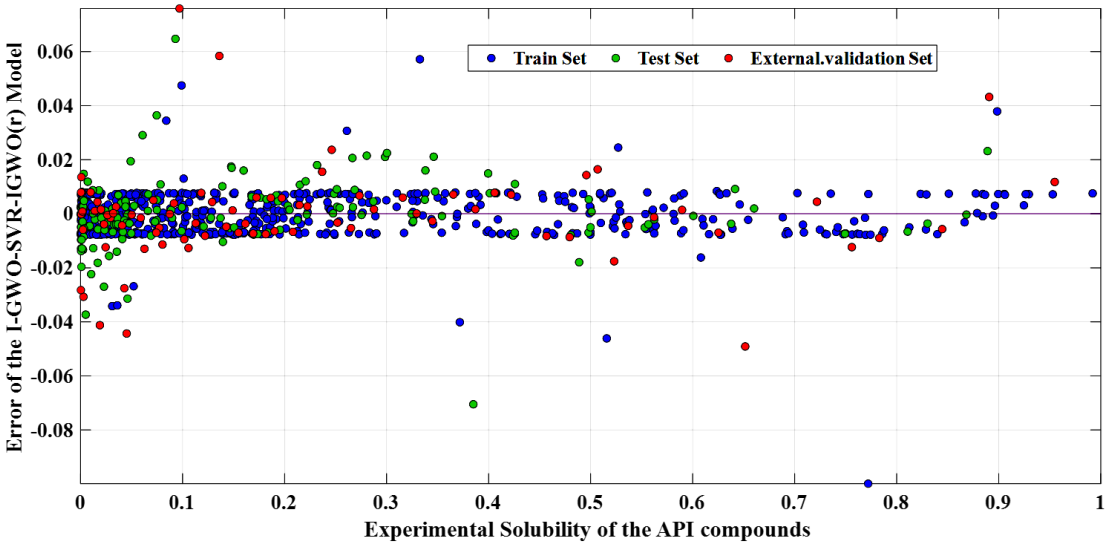


Figure IV.17. Error of the I-GWO-SVR(r) model in predicting the AIPs solubility in ILs.

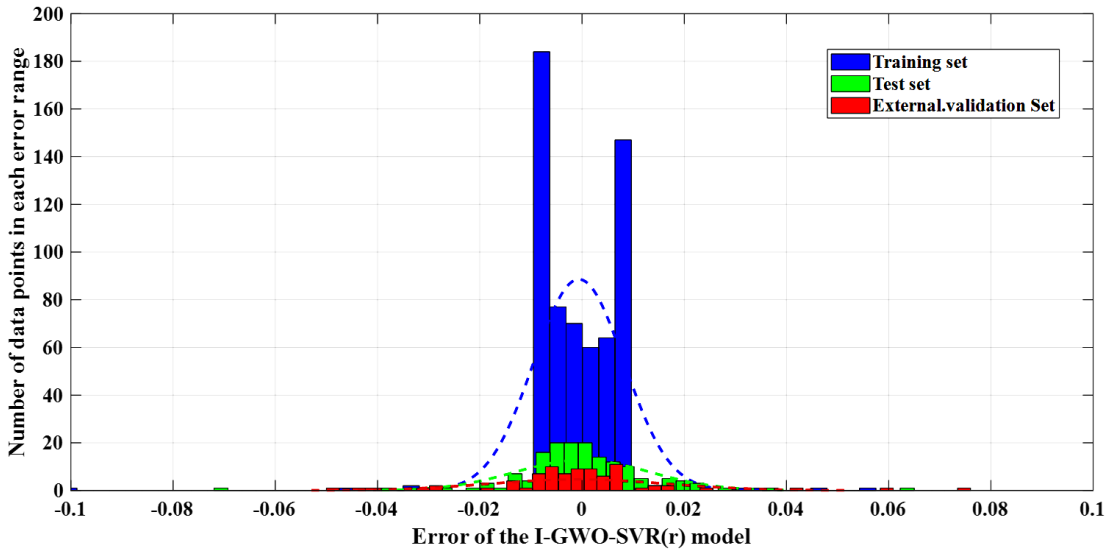


Figure IV.18. Error distribution of the I-GWO-SVR(r) model in predicting the AIPs solubility in ILs.

IV.3.4 Outlier Detection for the I-GWO-SVR(r) Model

All data points that are outstandingly detached from the whole dataset are considered as outliers. In this regard, a comprehensive outlier detection, known as the Leverage method, was performed on the developed model. The mathematical Leverage approach was exploited to detect outliers' data in the collected solubility database, where the residual values and a Hat matrix were utilized. More details can be found in previous chapters.

Figure IV.19. shows William's plot for the I-GWO-SVR(r) model, which is a plot of leverage values against standardized residuals. **Figure IV.20.** depicts the reliable prediction area of the I-GWO-SVR(r) model based on the structural similarity with the training data points and the estimated values of solubility well-known as the Insubria plot.

The analysis of **Figure IV.19.** shows that more than 97.29% of the whole database fall between the two horizontal lines -3 & +3 and less than 2.71%, only 21 data points lie outside the suspected limits. This plot clearly highlighted the applicability scope of the proposed model.

The Insubria plot is a modified version of the well-known Williams plot, which can be used to screen the domain for chemicals without experimental data, as shown in **Figure IV.20**. This figure allows to verify the high statistical performance and the wide structural applicability domain of the I-GWO-SVR(r) model.

Considering both **Figure IV.19.** and **IV.20.**, few data points of the training/test and external validation sets appears to have a leverage greater than the warning h* (0.0682). These data can be identified as outliers. In fact, these points were deliberately not excluded from this study, since it was aimed to cover a wide database, which is possible depending on the availability of data. Furthermore, outliers are probably due to inaccurate measurements.

In addition, it is clearly shown that there are few data points whose leverage value h is greater than the critical leverage value (h*=0.0682) only that have a standard deviation less than ± 3; these are data points are called "good high leverage points", which stabilize the model. Regardless of that, most data points are placed in the standard range of the Leverage method. Thus, the proposed model indicates a high predictive ability.

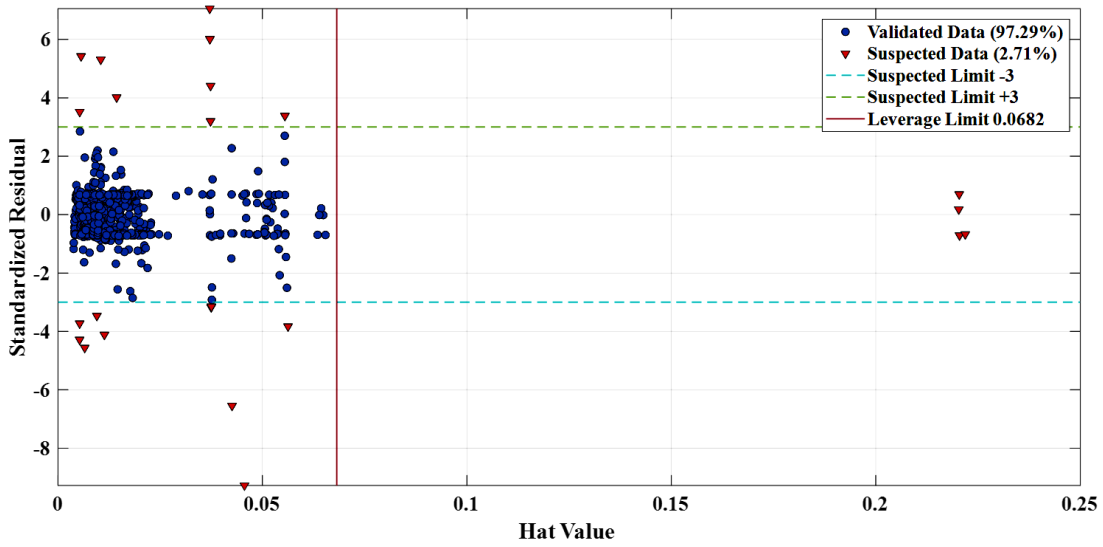


Figure IV.19. William's plot of the I-GWO-SVR(r) model.

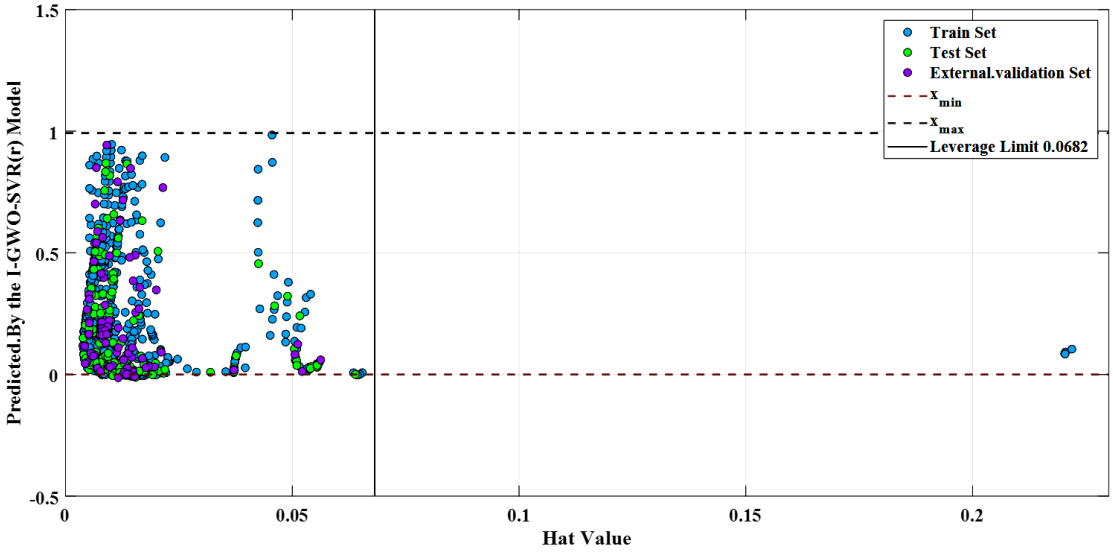


Figure IV.20. Insurbia plot of the I-GWO-SVR(r) model.

IV.3.5 Sensibility Analysis for the I-GWO-SVR(r) Model

To assess the effect of the considered inputs on the studied output property, a sensitivity analysis was performed on the best-established model: I-GWO-SVR(r). For this purpose, the relevance factor (*r*) was calculated to highlight the significance descriptor for the prediction of

APIs solubility in pure ILs. The dependence of the proposed model on the information contained in each descriptor can be represented by the sensitivity of the input variables (the σ-moments and the temperature).

Note that the value of the relevance factor is between -1 and +1. A negative value indicates a decreasing relationship between the input parameter and the target property, while a positive value indicates an increasing relationship. The relevance factor of zero means that no relation can be detected between the input parameter and output target.

Figure IV.21. illustrates the relevance factors of the operational input parameters for the experimental outputs of the developed model. As observed, the solubility showed a straight dependency (positive and direct effects) on the descriptors named T(K), S_1, S_2, S_6 and an inverse dependency (negative and opposite effects) on the rest.

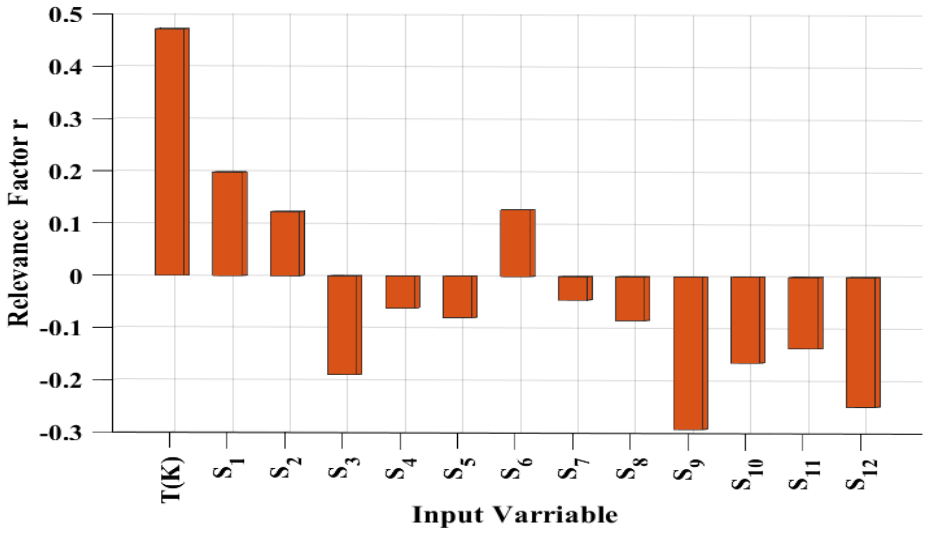


Figure IV.21. Results of the relevance factor of the input parameters of of the I-GWO-SVR(r) model.

The obtained results of the of the relevance factor analysis revealed that the solubility of API-IL is strongly dependent on the temperature, since it is an essential parameter in the measurements of solubility.

Various effects were also observed for the other parameters. The overall effects can be classified as T(K), σ-moment of the APIs (S_9 to S_{12}) then σ-moment of ILs (S_1 to S_4 of the cations and S_5 to S_8 of the anions). The results showed their importance (no zero dependence) and all the selected input parameters in this study were chosen appropriately and with which allowed to conclude the significant effect on the solubility of APIs in various ILs.

IV.3.6 I-GWO-SVR(r) Model Equation

For this purpose, these results confirmed the high performance of the I-GWO-SVR(r) model, since the error values were close to the training and the test sets. Also, there were no significantly large residual values for the validation set, which allowed to conclude that there is no overfitting of the model. In summary, the I-GWO-SVR(r) model showed high predictive power and can be recommended to predict future data falling within the defined applicability domain.

As presented in the annex, for the obtained model, there were 325x13 support vectors in total of 4225. The fitted kernel scale was of 1.3181 and r was equal to 0.0001. The regressed weight vector Alpha 325X1 and bias 1X1 were also provided in the annex. For a given input vector x we have:

$$x_{svr-IGWO-R} = \sum_{i=1}^{325}[\alpha_i \, exp(-r \cdot \|x - x_i\|^2)] + b \qquad \text{(IV.25)}$$

Where r is an adjustable parameter of the kernel controlling the amplitude of the RBF function. More details of all determined parameters for the I-GWO-SVR(r) model are presented in the annex.

IV.4 Conclusions

In this chapter, a new QSPR model namely I-GWO-SVR(r) was established for the solubility prediction of 14 APIs in pure ILs as a function of 13 descriptors, the temperature, and COSMO-RS σ-profile descriptors for both ILs and APIs. Experimental data was carefully selected from the literature and was closely analyzed statistically.

The model ran a database for 83 API-ILs systems composed of 14 APIs, 12 cations, and 7 anions (25 ILs combinations), representing 13 inputs, 1 output, and 850 data points. 770 data points were used for model development (80% for training and 20% for test set) while 80 data points were hidden for external validation test.

The model was then compared to 11 different techniques, namely 4 ANN topologies: ANN-lm-tan, ANN-lm-log, ANN-br-tan, ANN-br-log, and 7 hybridized SVR: I-GWO-SVR(g), MVO-SVR(g), MVO-SVR(r), MFO-SVR(g), MFO-SVR(r), SCA-SVR(g), SCA-SVR(r). Based on various statistical factors, all the models presented a good performance for both training and testing sets, indicating that the developed models can be used for the prediction of the selected property. While the proposed approaches showed significantly better results.

The non-linear I-GWO-SVR(r) prediction model showed high accuracy prediction ability, strong generalization capacity, excellent robustness, fitness, and predictive ability. The performed sensitivity analysis showed a significant relative importance of the inputs parameters where the highest belongs to the operating condition namely, the temperature.

Moreover, the Leverage approach demonstrated that most of the data points are within the acceptable range, which proved the statistical acceptability of the model. It is also worth mentioning that this prioritization provided useful information to reduce the time, the cost, and the number of chemicals tested. The overall results can be used to predict future data falls into the range of the acceptability domain provided the availability of the σ-profile of those future data.

To this end, this study implicated the great potential of combining COSMO-RS theory with the QSPR technique for the prediction and addressed the calculation of the σ-profile of APIs for the first time. We believe this idea will generate so much attention.

General Conclusion

General Conclusion

The main objective of this thesis was to investigate new advanced cheminformatics techniques for predicting the solubility of pharmaceutical compounds in two separate green solvents, including supercritical carbon dioxide, and pure ionic liquids. In this thesis, solubility data were collected from the literature, and inputs were calculated based on molecular information in the form of two types of QSPR descriptors, namely PaDEL descriptors, and σ-moment descriptors for the solubility data in SC-CO_2, and solubility data in ILs, respectively. With the operation conditions, the values were composed to form different databases. Different machine learning algorithms were calculated and validated.

In the first contribution, a quantitative structure-property relationship Artificial neural networks based technique was proposed to predict the solubility of 148 drug (or drug-like) compounds in SC-CO_2 as a function of temperatures, pressure, and 11 PaDEL descriptors, including AATS3v, MATS2e, GATS4c, GATS3v, GATS4e, GATS3s, nBondsM, AVP-0, SHBd, MLogP, and MLFER_S. The ANN-QSPR model was built using a dataset of 3971 experimental data points, representing the solubility (mole fraction) of the drug in SC-CO_2 over large temperatures (K) and pressure (MPa) ranges. According to the results, the best ANN-QSPR model was achieved with the application of the training function trainbr "Bayesian regularization back-propagation", logsig "Log-sigmoid transfer function" activation function for the hidden layer, and an optimized architecture of {13-10-1}. Moreover, the results of the developed model show that all the statistic values fell within the acceptability domain, including average absolute relative deviation (AARD=3.7748 %), root mean square error (RMSE=0.5162), coefficient of correlation (r=0.9761), coefficient of determination (R^2=0.9528), and robustise (Q^2=0.9528). Hence, the built QSAR model was robust, satisfactory, meets all the OECD principles for QSAR validation, and can be used to correlate reliably the solubility of drug-like compounds, for interpolation purposes for drugs of the data set within the model space and for extrapolation purposes for drugs not included in the data set...

In the second contribution, a comparative study between seven meta-heuristic algorithms in terms of fine-tuning hyper-parameters of a hybrid QSPR-SVR model defined for predicting 168 drug solubilities in SC-CO_2, a database of 13 inputs, and 4490 rows (an extended database of the first contribution). The used meta-heuristic algorithms were: Dragonfly

optimizer (DA), Ant Lion optimizer (ALO), Grey Wolf optimizer (GWO), Artificial Bee Colony optimizer (ABC), Particle Swarm optimizer (PSO), Whale optimizer (WOA), and a hybrid Particle Swarm with Grey Wolf optimizer (HPSOGWO) responsible for the seven models DA-SVR, ALO-SVR, GWO-SVR, ABC-SVR, PSO-SVR, WOA-SVR, and HPSOGWO-SVR, respectively. All models were statistically and graphically accepted, while the hybrid HPSOGWO-SVR model performed the best with an average absolute relative deviation (AARD) of 0.706% and an AIC of -14,434,249. The validity of the model predictions was further guaranteed by the external test on 160 EDP compared to experimental values, considering new compounds that should belong to the applicability domain (William's plot & Insubria graph). The overall results proved that the obtained model has good predictive ability and robustness.

The third contribution, presents the application of a QSPR model namely I-GWO-SVR(r) for the prediction of the solubility of APIs in pure ILs as a function of the temperature, and COSMO-RS σ-profile descriptors for both ILs and APIs, counting 83 API-ILs systems composed of 14 APIs, 12 cations, and 7 anions (25 ILs combinations), representing 13 inputs, 1 output, and 850 data points. The results were compared to 11 different techniques, namely 4 ANN topologies: ANN-lm-tan, ANN-lm-log, ANN-br-tan, ANN-br-log, and 7 hybridized SVR: I-GWO-SVR(g), MVO-SVR(g), MVO-SVR(r), MFO-SVR(g), MFO-SVR(r), SCA-SVR(g), SCA-SVR(r). It was found that all the models presented great statistical results, while the proposed approach showed significantly better results, including root mean square error (RMSE=0.0098), mean absolute error (MAE=0.0067), coefficient of correlation (r=0.9991), coefficient of determination (R^2=0.9983), and robustise (Q^2=0.9983). Moreover, the non-linear I-GWO-SVR(r) prediction model showed also high accuracy prediction ability, strong generalization capacity, excellent robustness, fitness, and predictive ability. The overall results can be used to predict future data falls into the range of the acceptability domain provided the availability of the σ-profile of those future data.

To this end, this thesis implicated the great potential of (1) replacing the experimental input parameters with molecular structure information, which makes the new model independent of experimental data (2) PaDEL descriptors along with T(K) and P as inputs parameters for predicting the solubility of drugs in SC-CO_2 (3) combining COSMO-RS theory with the QSPR technique (4) σ-moment as a new molecular descriptor for predicting the solubility of APIs in ILs, where σ-profile of APIs were calculated and used for the first time, and (5) substituting traditional intelligence algorithms with more effective techniques, in order to avoid the defects of traditional

thermodynamic dissolution forecasting methods and limit the use of tedious and expensive experiments. Therefore, it's safe to say that QSPR/QSAR methods remain the current state-of-the-art for solubility prediction.

- **Future works.**

Indeed, this realization project was a very fruitful, beneficial, and enriching experience. It allowed us to exploit the different theoretical concepts of chemoinformatics, including the QSPR-based technique and the COSMO-RS-based approaches. It also allowed us to explore the practical world of machine learning developers using the powerful programming language MATLAB, as well as various software specialized in this field.

The final results obtained in these studies were very satisfactory. However, as a researcher, there is no doubt such as results can be and will be furthermore improved in the future. The proposed perspectives are as follows:

✓ Improving the model accuracy with the application of various other methods, such as those reported in **Chapter I**.

✓ Testing new inputs based on molecular information.

✓ The utilization of co-solvents, mainly ethanol and methanol, along with SC-CO_2 (or ILs) was confirmed to increase the solvating power of solvent (SC-CO_2, or ILs). The co-solvent interacts strongly with the solute and significantly increases mixture density enhancing the solubility and the physical interactions such as dipole-dipole, dipole-induced dipole, and induced dipole-induced dipole interactions...Therefore, it would be interesting to collect such as dataset and study it.

✓ It would be interesting too to collect solubility data of drugs in other green solvents.

✓ The era of innovation has no limitation, hence there will always be a new technique or method that would be worth investigating…

Bibiography

Bibliography

[1] P. Jain, Prediction of aqueous solubility from SCRATCH, The University of Arizona, 2008. https://doi.org/http://hdl.handle.net/10150/193517.

[2] A.B. Varandal, D.D. Magar, R.B. Saudagar, Different approaches toward the enhancement of Drug Solubility: A Review, J. Adv. Pharm. Educ. Res. 3 (2013) 415–426. www.japer.in.

[3] D. Christopher Vimalson, S. Parimalakrishnan, N.S. Jeganathan, S. Anbazhagan, Techniques to enhance solubility of hydrophobic drugs: An overview, Asian J. Pharm. 10 (2016) S67–S75.

[4] V.B. Patel, Solubility Enhancement and Dissolution Method Development of Some Poorly Soluble Drugs, (2018).

[5] B. Dhillon, N.K. Goyal, R. Malviya, P.K. Sharma, Poorly water soluble drugs: Change in solubility for improved dissolution characteristics a review, Glob. J. Pharmacol. 8 (2014) 26–35. https://doi.org/10.5829/idosi.gjp.2014.8.1.81201.

[6] V.R. Vemula, V. Lagishetty, S. Lingala, Solubility enhancement techniques, Int. J. Pharm. Sci. Rev. Res. 5 (2010) 41–51.

[7] K.T. Savjani, A.K. Gajjar, J.K. Savjani, Drug Solubility: Importance and Enhancement Techniques, ISRN Pharm. 2012 (2012) 1–10. https://doi.org/10.5402/2012/195727.

[8] M. Kakran, L. Li, R.H. Müller, Overcoming the challenge of poor drug solubility, Pharm. Eng. 32 (2012) 82–89.

[9] S.M. Abuzar, S.M. Hyun, J.H. Kim, H.J. Park, M.S. Kim, J.S. Park, S.J. Hwang, Enhancing the solubility and bioavailability of poorly water-soluble drugs using supercritical antisolvent (SAS) process, Int. J. Pharm. 538 (2018) 1–13. https://doi.org/10.1016/j.ijpharm.2017.12.041.

[10] Y.A. Abramov, Major source of error in QSPR prediction of intrinsic thermodynamic solubility of drugs: Solid vs nonsolid state contributions?, Mol. Pharm. 12 (2015) 2126–2141. https://doi.org/10.1021/acs.molpharmaceut.5b00119.

[11] R.M. Alzhrani, A.H. Almalki, S. l. Alaqel, S. Alshehri, Novel numerical simulation of drug solubility in supercritical CO2 using machine learning technique: Lenalidomide case study, Arab. J. Chem. 15 (2022) 104180. https://doi.org/10.1016/j.arabjc.2022.104180.

[12] B.W. Cue, J. Zhang, Green process chemistry in the pharmaceutical industry, Green Chem. Lett. Rev. 2 (2009) 193–211. https://doi.org/10.1080/17518250903258150.

[13] S.N. Pedro, C.S.R. Freire, A.J.D. Silvestre, M.G. Freire, The role of ionic liquids in the pharmaceutical field: An overview of relevant applications, Int. J. Mol. Sci. 21 (2020) 1–50. https://doi.org/10.3390/ijms21218298.

[14] A.K. El-Zhry El-Yafi, H. El-Zein, Technical crystallization for application in pharmaceutical material engineering: Review article, Asian J. Pharm. Sci. 10 (2014) 283–291. https://doi.org/10.1016/j.ajps.2015.03.003.

[15] M. Cvjetko Bubalo, S. Vidović, I. Radojčić Redovniković, S. Jokić, Green solvents for green technologies, J. Chem. Technol. Biotechnol. 90 (2015) 1631–1639. https://doi.org/10.1002/jctb.4668.

[16] R.D.R. Julia L. Shamshina, Paula Berton, Hui Wang, Xiaosi Zhou, Gabriela Gurau, Part IV Green techniques and Estrategies in the pharmaceutical Industry, Green Tech. Org. Syntesis Med. Chem. (2018) 539–577. onlinelibrary.wiley.com/doi/10.1002/9781119288152.

[17] N. Adawiyah, M. Moniruzzaman, S. Hawatulaila, M. Goto, Ionic liquids as a potential tool for drug delivery systems, Medchemcomm. 7 (2016) 1881–1897. https://doi.org/10.1039/c6md00358c.

[18] C. Capello, U. Fischer, K. Hungerbühler, What is a green solvent? A comprehensive framework for the environmental assessment of solvents, Green Chem. 9 (2007) 927–93. https://doi.org/10.1039/b617536h.

[19] K. Grodowska, A. Parczewski, Organic solvents in the pharmaceutical industry, Acta Pol. Pharm. - Drug Res. 67 (2010) 3–12.

[20] P.T. Anastas, J.B. Zimmerman, Innovations in Green Chemistry and Green Engineering, Springer, 2012. https://doi.org/10.1007/978-1-4614-5817-3.

[21] B. Kosjek, J. Nti-Gyabaah, K. Telari, L. Dunne, J.C. Moor, Preparative asymmetric

synthesis of 4,4-dimethoxytetrahydro-2H-pyran-3-ol with a ketone reductase and in situ cofactor recycling using glucose dehydrogenase, Org. Process Res. Dev. 12 (2008) 584–588. https://doi.org/10.1021/op700255b.

[22] J. Sylvestre, V. Gena, R.H. Desbrue, E. V Cedex, Re V iews Directed Evolution of Biocatalysts Abstract :, (2006).

[23] P.J. Dunn, A.S. Wells, M.T. Williams, Asymmetric Catalysis on Industrial Scale Green Chemistry and Catalysis Handbook of Green Chemistry – Green Catalysis Microwaves in Organic Synthesis Handbook of Green Chemistry – Green Solvents Supercritical Carbon Dioxide Solvent-free Organic Synthesis, 2010.

[24] Y. Cao, A. Khan, S. Zabihi, A.B. Albadarin, Neural simulation and experimental investigation of Chloroquine solubility in supercritical solvent, J. Mol. Liq. 333 (2021) 115942. https://doi.org/10.1016/j.molliq.2021.115942.

[25] B. The, 1. Introduction 1.1, (n.d.) 1–29.

[26] G. Brunner, Supercritical fluids: Technology and application to food processing, J. Food Eng. 67 (2005) 21–33. https://doi.org/10.1016/j.jfoodeng.2004.05.060.

[27] F.M.V.T. Thakkar, T.G. Soni, M.C. Gohel, T.R. Gandhi, Supercritical fluid technology: A promising approach to enhance the drug solubility, J. Pharm. Sci. Res. 1 (2009) 1–14.

[28] G. Liao, Y. Du, feng zhang, J. E, The Applications of Molecular Dynamics Simulation in Supercritical Carbon Dioxide: A Review, SSRN Electron. J. (2022). https://doi.org/10.2139/ssrn.4097929.

[29] M. Baghbanbashi, N. Hadidi, G. Pazuki, Solubility of pharmaceutical compounds in supercritical carbon dioxide: Application, experimental, and mathematical modeling, Elsevier Inc., 2019. https://doi.org/10.1016/B978-0-12-817388-6.00010-6.

[30] G. Sodeifian, S.A. Sajadian, F. Razmimanesh, N.S. Ardestani, A comprehensive comparison among four different approaches for predicting the solubility of pharmaceutical solid compounds in supercritical carbon dioxide, Korean J. Chem. Eng. 35 (2018) 2097–2116. https://doi.org/10.1007/s11814-018-0125-6.

[31] X.Q. Bian, Q. Zhang, L. Zhang, J. Chen, A grey wolf optimizer-based support vector machine for the solubility of aromatic compounds in supercritical carbon dioxide, Chem. Eng. Res. Des. 123 (2017) 284–294. https://doi.org/10.1016/j.cherd.2017.05.008.

[32] I. Medina, J.L. Bueno, Solubilities of 2-nitroanisole and 3-phenyl-1-propanol in supercritical carbon dioxide, J. Chem. Eng. Data. 45 (2000) 298–300. https://doi.org/10.1021/je9901482.

[33] M. Hojjati, Y. Yamini, M. Khajeh, A. Vatanara, Solubility of some statin drugs in supercritical carbon dioxide and representing the solute solubility data with several density-based correlations, J. Supercrit. Fluids. 41 (2007) 187–194. https://doi.org/10.1016/j.supflu.2006.10.006.

[34] A.A. Amooey, A simple correlation to predict drug solubility in supercritical carbon dioxide, Fluid Phase Equilib. 375 (2014) 332–339. https://doi.org/10.1016/j.fluid.2014.05.025.

[35] A. Baghban, A. Jalali, A.H. Mohammadi, S. Habibzadeh, Efficient modeling of drug solubility in supercritical carbon dioxide, J. Supercrit. Fluids. 133 (2018) 466–478. https://doi.org/10.1016/j.supflu.2017.10.032.

[36] C.G. Pereira, M.A.A. Meireles, Supercritical fluid extraction of bioactive compounds: Fundamentals, applications and economic perspectives, Food Bioprocess Technol. 3 (2010) 340–372. https://doi.org/10.1007/s11947-009-0263-2.

[37] R.K. Kankala, Y.S. Zhang, S. Bin Wang, C.H. Lee, A.Z. Chen, Supercritical Fluid Technology: An Emphasis on Drug Delivery and Related Biomedical Applications, Adv. Healthc. Mater. 6 (2017). https://doi.org/10.1002/adhm.201700433.

[38] B. Damjanović, Ž. Lepojević, V. Živković, A. Tolić, Extraction of fennel (Foeniculum vulgare Mill.) seeds with supercritical CO2: Comparison with hydrodistillation, Food Chem. 92 (2005) 143–149. https://doi.org/10.1016/j.foodchem.2004.07.019.

[39] I. Euldji, C. Si-Moussa, M. Hamadache, O. Benkortbi, QSPR Modelling of the Solubility of Drug and Drug-like Compounds in Supercritical Carbon Dioxide, Mol. Inform. 2200026 (2022) 1–17. https://doi.org/10.1002/minf.202200026.

[40] G. Liao, L. Liu, J. E, F. Zhang, J. Chen, Y. Deng, H. Zhu, Effects of technical progress on performance and application of supercritical carbon dioxide power cycle: A review, Energy Convers. Manag. 199 (2019) 111986. https://doi.org/10.1016/j.enconman.2019.111986.

[41] R. Davarnejad, K.M. Kassim, A. Zainal, S.A. Sata, Supercritical fluid extraction of β-carotene from crude palm oil using CO2, J. Food Eng. 89 (2008) 472–478.

https://doi.org/10.1016/j.jfoodeng.2008.05.032.

[42] C. Perakis, V. Louli, E. Voutsas, K. Magoulas, Supercritical CO2 extraction of dittany oil: Experiments and modelling, J. Supercrit. Fluids. 55 (2010) 573–578. https://doi.org/10.1016/j.supflu.2010.10.020.

[43] T. Wu, Q. Xue, X. Li, Y. Tao, Y. Jin, C. Ling, S. Lu, Extraction of kerogen from oil shale with supercritical carbon dioxide: Molecular dynamics simulations, J. Supercrit. Fluids. 107 (2016) 499–506. https://doi.org/10.1016/j.supflu.2015.07.005.

[44] T. Fang, J. Shi, X. Sun, Y. Shen, Y. Yan, J. Zhang, B. Liu, Supercritical CO2 selective extraction inducing wettability alteration of oil reservoir, J. Supercrit. Fluids. 113 (2016) 10–15. https://doi.org/10.1016/j.supflu.2015.09.016.

[45] C.M.V. Taylor, G.B. Jacobson, Supercritical Fluids, EMagRes. 2007 (2007). https://doi.org/10.1002/9780470034590.emrstm0544.

[46] F. Jutz, J.M. Andanson, A. Baiker, Ionic liquids and dense carbon dioxide: A beneficial biphasic system for catalysis, Chem. Rev. 111 (2011) 322–353. https://doi.org/10.1021/cr100194q.

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

[48] M. Aghaie, S. Zendehboudi, Estimation of CO2 solubility in ionic liquids using connectionist tools based on thermodynamic and structural characteristics, Fuel. 279 (2020) 117984. https://doi.org/10.1016/j.fuel.2020.117984.

[49] A.R. Dias, J. Costa-Rodrigues, M.H. Fernandes, R. Ferraz, C. Prudêncio, The Anticancer Potential of Ionic Liquids, ChemMedChem. 12 (2017) 11–18. https://doi.org/10.1002/cmdc.201600480.

[50] A.J. Greer, J. Jacquemin, C. Hardacre, ILS als LM (nicht entflammbar meist, kein/kaum dampfdruck, leitfähig), 2020.

[51] D. MacFarlane, M. Kar, J. Pringle, Fundamentals of Ionic Liquids Ionic Liquids in Biotransformations and Electrodeposition from Ionic Handbook of Green Chemistry – Green Solvents Electrochemical Aspects of Ionic Liquids , 2nd edition Nanocatalysis in Ionic Liquids, 2017.

[52] K. Paduszyński, Extensive Databases and Group Contribution QSPRs of Ionic Liquid

Properties. 3: Surface Tension, Ind. Eng. Chem. Res. 60 (2021) 5705–5720. https://doi.org/10.1021/acs.iecr.1c00783.

[53] W. Zhuang, K. Hachem, D. Bokov, M. Javed Ansari, A. Taghvaie Nakhjiri, Ionic liquids in pharmaceutical industry: A systematic review on applications and future perspectives, J. Mol. Liq. 349 (2022) 118145. https://doi.org/10.1016/j.molliq.2021.118145.

[54] J. Ranke, K. Mölter, F. Stock, U. Bottin-Weber, J. Poczobutt, J. Hoffmann, B. Ondruschka, J. Filser, B. Jastorff, Biological effects of imidazolium ionic liquids with varying chain lengths in acute Vibrio fischeri and WST-1 cell viability assays, Ecotoxicol. Environ. Saf. 58 (2004) 396–404. https://doi.org/10.1016/S0147-6513(03)00105-2.

[55] A. Marjani, A. Taghvaie Nakhjiri, M. Adimi, H. Fathinejad Jirandehi, S. Shirazian, Modification of polyethersulfone membrane using MWCNT-NH2 nanoparticles and its application in the separation of azeotropic solutions by means of pervaporation, PLoS One. 15 (2020) e0236529. https://doi.org/10.1371/journal.pone.0236529.

[56] A.T. Nakhjiri, M.H. Roudsari, Modeling and Simulation of Natural Convection Heat Transfer Process in Porous and Non-Porous Media, (2016) 199–204.

[57] J. Dupont, R.F. De Souza, P.A.Z. Suarez, Ionic liquid (molten salt) phase organometallic catalysis, Chem. Rev. 102 (2002) 3667–3692. https://doi.org/10.1021/cr010338r.

[58] M. Kosmulski, J. Gustafsson, J.B. Rosenholm, Thermal stability of low temperature ionic liquids revisited, Thermochim. Acta. 412 (2004) 47–53. https://doi.org/10.1016/j.tca.2003.08.022.

[59] X. Cui, S. Zhang, F. Shi, Q. Zhang, X. Ma, L. Lu, Y. Deng, The influence of the acidity of ionic liquids on catalysis, ChemSusChem. 3 (2010) 1043–1047. https://doi.org/10.1002/cssc.201000075.

[60] F. Van Rantwijk, R.A. Sheldon, cr050946x-OK.pdf, Chem. Rev. 107 (2007) 2757–2785.

[61] M. Babanezhad, I. Behroyan, A.T. Nakhjiri, A. Marjani, S. Shirazian, Computational Modeling of Transport in Porous Media Using an Adaptive Network-Based Fuzzy Inference System, ACS Omega. 5 (2020) 30826–30835. https://doi.org/10.1021/acsomega.0c04497.

[62] W. Huang, X. Wu, J. Qi, Q. Zhu, W. Wu, Y. Lu, Z. Chen, Ionic liquids: green and tailor-

made solvents in drug delivery, Drug Discov. Today. 25 (2020) 901–908. https://doi.org/10.1016/j.drudis.2019.09.018.

[63] M. Antonietti, D. Kuang, B. Smarsly, Y. Zhou, Ionic liquids for the convenient synthesis of functional nanoparticles and other inorganic nanostructures, Angew. Chemie - Int. Ed. 43 (2004) 4988–4992. https://doi.org/10.1002/anie.200460091.

[64] M. Ghadiri, A. Hemmati, A.T. Nakhjiri, S. Shirazian, Modelling tyramine extraction from wastewater using a non-dispersive solvent extraction process, Environ. Sci. Pollut. Res. 27 (2020) 39068–39076. https://doi.org/10.1007/s11356-020-09943-2.

[65] J.P. Hallett, T. Welton, Room-temperature ionic liquids: Solvents for synthesis and catalysis. 2, Chem. Rev. 111 (2011) 3508–3576. https://doi.org/10.1021/cr1003248.

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

[67] I.M. Marrucho, L.C. Branco, L.P.N. Rebelo, Ionic liquids in pharmaceutical applications, Annu. Rev. Chem. Biomol. Eng. 5 (2014) 527–546. https://doi.org/10.1146/annurev-chembioeng-060713-040024.

[68] D. Psimadas, P. Georgoulias, V. Valotassiou, G. Loudos, Molecular Nanomedicine Towards Cancer :, J. Pharm. Sci. 101 (2012) 2271–2280. https://doi.org/10.1002/jps.

[69] M.J. Earle, P.B. McCormac, K.R. Seddon, The first high yield green route to a pharmaceutical in a room temperature ionic liquid, Green Chem. 2 (2000) 261–262. https://doi.org/10.1039/b006612p.

[70] V. Kumar, V.S. Parmar, S. V. Malhotra, Enhanced solubility and selective benzoylation of nucleosides in novel ionic liquid, Tetrahedron Lett. 48 (2007) 809–812. https://doi.org/10.1016/j.tetlet.2006.11.155.

[71] H. Mizuuchi, V. Jaitely, S. Murdan, A.T. Florence, Room temperature ionic liquids and their mixtures: Potential pharmaceutical solvents, Eur. J. Pharm. Sci. 33 (2008) 326–331. https://doi.org/10.1016/j.ejps.2008.01.002.

[72] M. Moniruzzaman, N. Kamiya, M. Goto, Ionic liquid based microemulsion with pharmaceutically accepted components: Formulation and potential applications, J. Colloid Interface Sci. 352 (2010) 136–142. https://doi.org/10.1016/j.jcis.2010.08.035.

[73] N. Adawiyah, M. Moniruzzaman, S. Hawatulaila, M. Goto, Ionic liquids as a potential

tool for drug delivery systems, Medchemcomm. 7 (2016) 1881–1897. https://doi.org/10.1039/c6md00358c.

[74] S. Kalepu, V. Nekkanti, Insoluble drug delivery strategies: Review of recent advances and business prospects, Acta Pharm. Sin. B. 5 (2015) 442–453. https://doi.org/10.1016/j.apsb.2015.07.003.

[75] D.A.S. Agostinho, F. Santos, J.M.S.S. Esperança, A.R.C. Duarte, P.M. Reis, New non-toxic biocompatible dianionic ionic liquids that enhance the solubility of oral drugs from BCS class II, J. Ion. Liq. 1 (2021) 100003. https://doi.org/10.1016/j.jil.2021.100003.

[76] A. Balk, U. Holzgrabe, L. Meinel, Pro et contra' ionic liquid drugs - Challenges and opportunities for pharmaceutical translation, Eur. J. Pharm. Biopharm. 94 (2015) 291–304. https://doi.org/10.1016/j.ejpb.2015.05.027.

[77] Y. Chen, G.M. Kontogeorgis, J.M. Woodley, Group Contribution Based Estimation Method for Properties of Ionic Liquids, Ind. Eng. Chem. Res. 58 (2019) 4277–4292. https://doi.org/10.1021/acs.iecr.8b05040.

[78] V. Villazón-León, A. Bonilla-Petriciolet, J.C. Tapia-Picazo, J.G. Segovia-Hernández, M.L. Corazza, A review of group contribution models to calculate thermodynamic properties of ionic liquids for process systems engineering, Chem. Eng. Res. Des. 185 (2022) 458–480. https://doi.org/10.1016/j.cherd.2022.07.033.

[79] J. Alsenz, M. Kansy, High throughput solubility measurement in drug discovery and development, Adv. Drug Deliv. Rev. 59 (2007) 546–567. https://doi.org/10.1016/j.addr.2007.05.007.

[80] W. Stiver, The Journal of Supercritical Fluids Chrastil revisited for supercritical carbon dioxide, J. Supercrit. Fluids. 177 (2021) 105348. https://doi.org/10.1016/j.supflu.2021.105348.

[81] J.L. Shamshina, P.S. Barber, R.D. Rogers, Ionic liquids: Novel applications in drug delivery, Res. J. Pharm. Technol. 6 (2013) 1274–1278.

[82] J.L. Shamshina, R.D. Rogers, Overcoming the problems of solid state drug formulations with ionic liquids, Ther. Deliv. 5 (2014) 489–491. https://doi.org/10.4155/tde.14.28.

[83] G. Járvás, J. Kontos, G. Babics, A. Dallos, A novel method for the surface tension estimation of ionic liquids based on COSMO-RS theory, Fluid Phase Equilib. 468 (2018)

9–17. https://doi.org/10.1016/j.fluid.2018.04.010.

[84] F. Yan, Y. Shi, Y. Wang, Q. Jia, Q. Wang, S. Xia, QSPR models for the properties of ionic liquids at variable temperatures based on norm descriptors, Chem. Eng. Sci. 217 (2020) 115540. https://doi.org/10.1016/j.ces.2020.115540.

[85] D. Ghaslani, Z. Eshaghi Gorji, A. Ebrahimpoor Gorji, S. Riahi, Descriptive and predictive models for Henry's law constant of CO2 in ionic liquids: A QSPR study, Chem. Eng. Res. Des. 120 (2017) 15–25. https://doi.org/10.1016/j.cherd.2016.12.020.

[86] R.T. Kurnlk, S.J. Holla, R.C. Reid, Solubility of Solids in Supercritical Carbon Dioxide and Ethylene, J. Chem. Eng. Data. 26 (1981) 47–51. https://doi.org/10.1021/je00023a016.

[87] K.D. Bartle, A.A. Clifford, G.F. Shilstone, Prediction of solubilities for tar extraction by supercritical carbon dioxide, J. Supercrit. Fluids. 2 (1989) 30–34. https://doi.org/10.1016/0896-8446(89)90007-7.

[88] K.D. Bartle, A.A. Clifford, G.F. Shilstone, Estimation of solubilities in supercritical carbon dioxide: A correlation for the peng-robinson interaction parameters, J. Supercrit. Fluids. 5 (1992) 220–225. https://doi.org/10.1016/0896-8446(92)90011-8.

[89] P.C. Chen, Y.P. Chen, D.S.H. Wong, Correlation of steroid solubilities in supercritical carbon dioxide, Fluid Phase Equilib. 83 (1993) 175–182. https://doi.org/10.1016/0378-3812(93)87020-2.

[90] Y.J. Kwon, G. Ali Mansoori, Solubility modeling of solids in supercritical fluids using the Kirkwood-Buff fluctuation integral with the hard-sphere expansion (RISE) theory, J. Supercrit. Fluids. 6 (1993) 173–180. https://doi.org/10.1016/0896-8446(93)90016-Q.

[91] I. Ashour, H. Hammam, Experimental Measurements and Model Prediction, (1993) 3–8.

[92] H.K. Lee, C.H. Kim, S. Kim, C.S. Choi, Solid Solubilities of Methoxyphenylacetic Acid Isomer Compounds in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 39 (1994) 163–165. https://doi.org/10.1021/je00013a046.

[93] S. Garnier, E. Neau, P. Alessi, A. Cortesi, I. Kikic, Modelling solubility of solids in supercritical fluids using fusion properties, Fluid Phase Equilib. 158–160 (1999) 491–500. https://doi.org/10.1016/s0378-3812(99)00151-x.

[94] C.C. Huang, M. Tang, W.H. Tao, Y.P. Chen, Calculation of the solid solubilities in supercritical carbon dioxide using a modified mixing model, Fluid Phase Equilib. 179 (2001) 67–84. https://doi.org/10.1016/S0378-3812(00)00483-0.

[95] C. Zhong, H. Yang, Representation of the solubility of solids in supercritical fluids using the SAFT equation of state, Ind. Eng. Chem. Res. 41 (2002) 4899–4905. https://doi.org/10.1021/ie020374w.

[96] G. Madras, C. Kulkarni, J. Modak, Modeling the solubilities of fatty acids in supercritical carbon dioxide, Fluid Phase Equilib. 209 (2003) 207–213. https://doi.org/10.1016/S0378-3812(03)00148-1.

[97] J.O. Valderrama, V.H. Alvarez, Temperature independent mixing rules to correlate the solubility of solids in supercritical carbon dioxide, J. Supercrit. Fluids. 32 (2004) 37–46. https://doi.org/10.1016/j.supflu.2003.12.016.

[98] T.Y. Kwak, G. a Mansoori, Rules for Cubic Equations, Chem. Eng. Sci. 41 (1986) 1303–1986.

[99] M.R. Housaindokht, B. Haghighi, M.R. Bozorgmehr, e d i x o i d n o b r a c l a c i t i r c r e p u s n i s d i l o s e m o s f o y t i l i b u l o s e h t g n i t c i d e r p n i e t a t s f o s n o i t a u q e e e r h t g n o m a n o s i r a p m o c A, 24 (2007) 102–105.

[100] A. Danesh, D.H. Xu, D.H. Tehrani, A.C. Todd, Improving predictions of equation of state by modifying its parameters for super critical components of hydrocarbon reservoir fluids, Fluid Phase Equilib. 112 (1995) 45–61. https://doi.org/10.1016/0378-3812(95)02782-A.

[101] K.A.M. Gasem, W. Gao, Z. Pan, R.L. Robinson, A modified temperature dependence for the Peng-Robinson equation of state, Fluid Phase Equilib. 181 (2001) 113–125. https://doi.org/10.1016/S0378-3812(01)00488-5.

[102] P. Coimbra, C.M.M. Duarte, H.C. De Sousa, Cubic equation-of-state correlation of the solubility of some anti-inflammatory drugs in supercritical carbon dioxide, Fluid Phase Equilib. 239 (2006) 188–199. https://doi.org/10.1016/j.fluid.2005.11.028.

[103] H.A. Martinez-Correa, D.C.A. Gomes, S.L. Kanehisa, F.A. Cabral, Measurements and thermodynamic modeling of the solubility of squalene in supercritical carbon dioxide, J. Food Eng. 96 (2010) 43–50. https://doi.org/10.1016/j.jfoodeng.2009.06.041.

[104] R. Ruivo, A. Paiva, P.C. Simões, Dynamic Studies on a Scf Countercurrent Extraction Process, (n.d.).

[105] M. Yazdizadeh, A. Eslamimanesh, F. Esmaeilzadeh, Thermodynamic modeling of solubilities of various solid compounds in supercritical carbon dioxide: Effects of equations of state and mixing rules, J. Supercrit. Fluids. 55 (2011) 861–875. https://doi.org/10.1016/j.supflu.2010.10.019.

[106] M. Yazdizadeh, A. Eslamimanesh, F. Esmaeilzadeh, Applications of cubic equations of state for determination of the solubilities of industrial solid compounds in supercritical carbon dioxide: A comparative study, Chem. Eng. Sci. 71 (2012) 283–299. https://doi.org/10.1016/j.ces.2011.10.055.

[107] M. Khamda, M.H. Hosseini, M. Rezaee, Measurement and correlation solubility of cefixime trihydrate and oxymetholone in supercritical carbon dioxide (CO2), J. Supercrit. Fluids. 73 (2013) 130–137. https://doi.org/10.1016/j.supflu.2012.09.006.

[108] A.A. El Hadj, C. Si-Moussa, S. Hanini, M. Laidi, Primena PC-SAFT i kubne jednačine stanja za koerelisanje rastvorljivosti nekih farmaceutskih i statinskih aktivnih supstanci u supekritičnom CO2, Chem. Ind. Chem. Eng. Q. 19 (2013) 449–460. https://doi.org/10.2298/CICEQ120407005E.

[109] H. Rostamian, M.N. Lotfollahi, A new simple equation of state for calculating solubility of solids in supercritical carbon dioxide, Period. Polytech. Chem. Eng. 59 (2015) 174–185. https://doi.org/10.3311/PPch.7714.

[110] M.A. Khansary, A. Hallaji Sani, H. Shahbeig, A. Arif, H. Dalirsefat, M.A. Salehi, A Novel equation of state: Determination and validation for dyes and drugs solubility calculations in supercritical carbon dioxide, Period. Polytech. Chem. Eng. 58 (2014) 179–194. https://doi.org/10.3311/PPch.7436.

[111] M. Ardjmand, M. Mirzajanzadeh, F. Zabihi, Measurement and correlation of solid drugs solubility in supercritical systems, Chinese J. Chem. Eng. 22 (2014) 549–558. https://doi.org/10.1016/S1004-9541(14)60073-2.

[112] R. Hartono, G.A. Mansoori, A. Suwono, Prediction of molar volumes, vapor pressures and supercritical solubilities of alkanes by equations of state, Chem. Eng. Commun. 173 (1999) 23–42. https://doi.org/10.1080/00986449908912774.

[113] H. Rostamian, M.N. Lotfollahi, Modified Redlich-Kwong and Peng-Robinson Equations

of State for solubility calculation of solid compounds in supercritical carbon dioxide, Indian J. Sci. Technol. 9 (2016). https://doi.org/10.17485/ijst/2016/v9i16/52344.

[114] J.A.P. Coelho, R.M. Filipe, G.P. Naydenova, D.S. Yankov, R.P. Stateva, Semi-empirical models and a cubic equation of state for correlation of solids solubility in scCO2: Dyes and calix[4]arenes as illustrative examples, Fluid Phase Equilib. 426 (2016) 37–46. https://doi.org/10.1016/j.fluid.2016.01.026.

[115] J. Sang, Y. Wang, H. Wang, R. Abro, J. Jin, Simplified calculation of solubility of solid solutes in supercritical carbon dioxide by EoS models, Elsevier Ltd, 2017. https://doi.org/10.1016/j.fluid.2016.10.012.

[116] H. Bagheri, G. Ali Mansoori, H. Hashemipour, A novel approach to predict drugs solubility in supercritical solvents for RESS process using various cubic EoS-mixing rule, J. Mol. Liq. 261 (2018) 174–188. https://doi.org/10.1016/j.molliq.2018.03.081.

[117] G. Sodeifian, F. Razmimanesh, S.A. Sajadian, H. Soltani Panah, Solubility measurement of an antihistamine drug (Loratadine) in supercritical carbon dioxide: Assessment of qCPA and PCP-SAFT equations of state, Fluid Phase Equilib. 472 (2018) 147–159. https://doi.org/10.1016/j.fluid.2018.05.018.

[118] L. Antonio Estévez, F.J. Colpas, E.A. Müller, A simple thermodynamic model for the solubility of thermolabile solids in supercritical fluids, Chem. Eng. Sci. 232 (2021). https://doi.org/10.1016/j.ces.2020.116268.

[119] Z. Cai, H. Liang, W. Chen, S. Lin, C. Hsieh, Fluid Phase Equilibria First-principles prediction of solid solute solubility in supercritical carbon dioxide using PR þ COSMOSAC EOS, Fluid Phase Equilib. 522 (2020) 112755. https://doi.org/10.1016/j.fluid.2020.112755.

[120] I. Ushiki, R. Fujimitsu, S. Takishima, Predicting the solubilities of metal acetylacetonates in supercritical CO2: Thermodynamic approach using PC-SAFT, J. Supercrit. Fluids. 164 (2020) 104909. https://doi.org/10.1016/j.supflu.2020.104909.

[121] R.S. Alwi, C. Garlapati, New correlations for the solubility of anticancer drugs in supercritical carbon dioxide, Chem. Pap. 76 (2022) 1385–1399. https://doi.org/10.1007/s11696-021-01943-x.

[122] M.M. Azim, I. Ushiki, A. Miyajima, S. Takishima, Estimating the solubility of salsalate in supercritical CO2 via PC-SAFT modeling using its experimental solubility data in

organic solvents, J. Supercrit. Fluids. 189 (2022) 105725. https://doi.org/10.1016/j.supflu.2022.105725.

[123] G. Sodeifian, C.M. Hsieh, R. Derakhsheshpour, Y.M. Chen, F. Razmimanesh, Measurement and modeling of metoclopramide hydrochloride (anti-emetic drug) solubility in supercritical carbon dioxide, Arab. J. Chem. 15 (2022) 103876. https://doi.org/10.1016/j.arabjc.2022.103876.

[124] E.T. de Souza, P.B. Staudt, R. de P. Soares, Prediction of solid solubility in supercritical carbon dioxide using a pairwise surface contact equation of state — COSMO-SAC-Phi, J. Supercrit. Fluids. 191 (2022) 105765. https://doi.org/10.1016/j.supflu.2022.105765.

[125] H.W. Wang, C.M. Hsieh, Prediction of solid solute solubility in supercritical carbon dioxide from PSRK EOS with only input of molecular structure, J. Supercrit. Fluids. 180 (2022) 105446. https://doi.org/10.1016/j.supflu.2021.105446.

[126] O. Faraz, M. Poustchi, E. Nazari Denyani, P. Movahedi, F. Rajabi Kouchi, R. Shahriari, Thermodynamic modeling of pharmaceuticals solubility in pure, mixed and supercritical solvents, J. Mol. Liq. 353 (2022) 118809. https://doi.org/10.1016/j.molliq.2022.118809.

[127] B. Honarvar, S.A. Sajadian, A. Rojas, M.J. Galotto, A. Jouyban, Solubility and thermodynamic modeling of sildenafil citrate in supercritical carbon dioxide, Fluid Phase Equilib. 566 (2023). https://doi.org/10.1016/j.fluid.2022.113677.

[128] M. Majrashi, A. Salah Al-Shati, M. Grishina, S.M. Sarkar, M.T. Nguyen-Le, S. Shirazian, Experimental measurement and thermodynamic modeling of Chlorothiazide solubility in supercritical carbon dioxide, Case Stud. Therm. Eng. 41 (2023) 102621. https://doi.org/10.1016/j.csite.2022.102621.

[129] G. Sodeifian, H. Bagheri, M. Arbab Nooshabadi, F. Razmimanesh, A. Roshanghias, Experimental solubility of fexofenadine hydrochloride (antihistamine) drug in SC-CO2: Evaluation of cubic equations of state, J. Supercrit. Fluids. 200 (2023) 106000. https://doi.org/10.1016/j.supflu.2023.106000.

[130] N. Saadati Ardestani, S.A. Sajadian, A. Rojas, R.S. Alwi, L.A. Estévez, Solubility of famotidine in supercritical carbon dioxide: Experimental measurement and thermodynamic modeling, J. Supercrit. Fluids. 201 (2023). https://doi.org/10.1016/j.supflu.2023.106031.

[131] G. Sodeifian, C.M. Hsieh, A. Tabibzadeh, H.C. Wang, M. Arbab Nooshabadi, Solubility

of palbociclib in supercritical carbon dioxide from experimental measurement and Peng–Robinson equation of state, Sci. Rep. 13 (2023) 1–12. https://doi.org/10.1038/s41598-023-29228-1.

[132] H. Nateghi, G. Sodeifian, F. Razmimanesh, J. Mohebbi Najm Abad, A machine learning approach for thermodynamic modeling of the statically measured solubility of nilotinib hydrochloride monohydrate (anti-cancer drug) in supercritical CO2, Sci. Rep. 13 (2023) 1–17. https://doi.org/10.1038/s41598-023-40231-4.

[133] G. Sodeifian, M. Arbab Nooshabadi, F. Razmimanesh, A. Tabibzadeh, Solubility of buprenorphine hydrochloride in supercritical carbon dioxide: Study on experimental measuring and thermodynamic modeling, Arab. J. Chem. 16 (2023) 105196. https://doi.org/10.1016/j.arabjc.2023.105196.

[134] K. Venkatesan, S.M. Alshahrani, A.M. Alsubaiyel, M.H. Abduljabbar, M.E. Alosaimi, M. Grishina, S. Shirazian, Experimental - Theoretical approach for determination of Metformin solubility in supercritical carbon dioxide: Thermodynamic modeling, Case Stud. Therm. Eng. 41 (2023) 102649. https://doi.org/10.1016/j.csite.2022.102649.

[135] S. Edrisi, H. Bakhshi, Cubic plus association (CPA/qCPA) and classical thermodynamic modeling of pharmaceuticals solubility in supercritical carbon dioxide, J. Therm. Anal. Calorim. (2023). https://doi.org/10.1007/s10973-023-12211-3.

[136] G. Sodeifian, L. Nasri, F. Razmimanesh, M. Arbab Nooshabadi, Solubility of ibrutinib in supercritical carbon dioxide (Sc-CO2): Data correlation and thermodynamic analysis, J. Chem. Thermodyn. 182 (2023) 107050. https://doi.org/10.1016/j.jct.2023.107050.

[137] D.H. Zlger, C.A. Eckert, Correlation and Prediction of Solid-Supercritical Fluid Phase Equilibria, Ind. Eng. Chem. Process Des. Dev. 22 (1983) 582–588. https://doi.org/10.1021/i200023a005.

[138] A. Vetere, A short-cut method to predict the solubilities of solids in supercritical carbon dioxide, Fluid Phase Equilib. 148 (1998) 83–93. https://doi.org/10.1016/S0378-3812(98)00180-0.

[139] A. Jouyban, H.K. Chan, N.R. Foster, Mathematical representation of solute solubility in supercritical carbon dioxide using empirical expressions, J. Supercrit. Fluids. 24 (2001) 19–35. https://doi.org/10.1016/S0896-8446(02)00015-3.

[140] C.S. Su, Y.P. Chen, Measurement and correlation for the solid solubility of non-steroidal

anti-inflammatory drugs (NSAIDs) in supercritical carbon dioxide, J. Supercrit. Fluids. 43 (2008) 438–446. https://doi.org/10.1016/j.supflu.2007.08.006.

[141] D.L. Sparks, R. Hernandez, L.A. Estévez, Evaluation of density-based models for the solubility of solids in supercritical carbon dioxide and formulation of a new model, Chem. Eng. Sci. 63 (2008) 4292–4301. https://doi.org/10.1016/j.ces.2008.05.031.

[142] C. Garlapati, G. Madras, New empirical expressions to correlate solubilities of solids in supercritical carbon dioxide, Thermochim. Acta. 500 (2010) 123–127. https://doi.org/10.1016/j.tca.2009.12.004.

[143] X.Y. Gong, X.J. Cao, Measurement and correlation of solubility of artemisinin in supercritical carbon dioxide, Fluid Phase Equilib. 284 (2009) 26–30. https://doi.org/10.1016/j.fluid.2009.05.018.

[144] S. Jafari Nejad, H. Abolghasemi, M.A. Moosavian, M.G. Maragheh, Prediction of solute solubility in supercritical carbon dioxide: A novel semi-empirical model, Chem. Eng. Res. Des. 88 (2010) 893–898. https://doi.org/10.1016/j.cherd.2009.12.006.

[145] A. Tabernero, E.M.M. del Valle, M.Á. Galán, A comparison between semiempirical equations to predict the solubility of pharmaceutical compounds in supercritical carbon dioxide, J. Supercrit. Fluids. 52 (2010) 161–174. https://doi.org/10.1016/j.supflu.2010.01.009.

[146] N. De Zordi, I. Kikic, M. Moneghini, D. Solinas, Solubility of pharmaceutical compounds in supercritical carbon dioxide, J. Supercrit. Fluids. 66 (2012) 16–22. https://doi.org/10.1016/j.supflu.2011.09.018.

[147] Y. Yamini, M. Moradi, Measurement and correlation of antifungal drugs solubility in pure supercritical CO2 using semiempirical models, J. Chem. Thermodyn. 43 (2011) 1091–1096. https://doi.org/10.1016/j.jct.2011.02.020.

[148] L. Nasri, Z. Bensetiti, S. Bensaad, Correlation of the solubility of some organic aromatic pollutants in supercritical carbon dioxide based on the UNIQUAC equation, Energy Procedia. 18 (2012) 1261–1270. https://doi.org/10.1016/j.egypro.2012.05.142.

[149] R. Chim, S. Marceneiro, M.B.C. De Matos, M.E.M. Braga, A.M.A. Dias, H.C. De Sousa, Solubility of Poorly Soluble Drugs in Supercritical Carbon Dioxide : Experimental Measurement and Density-Based Correlations, (2013) 1–6.

[150] A. Tabernero, S.A.B. Vieira De Melo, R. Mammucari, E.M. Martín Del Valle, N.R. Foster, Modelling solubility of solid active principle ingredients in sc-CO2 with and without cosolvents: A comparative assessment of semiempirical models based on Chrastil's equation and its modifications, J. Supercrit. Fluids. 93 (2014) 91–102. https://doi.org/10.1016/j.supflu.2013.11.017.

[151] H. Asiabi, Y. Yamini, M. Tayyebi, M. Moradi, A. Vatanara, K. Keshmiri, Measurement and correlation of the solubility of two steroid drugs in supercritical carbon dioxide using semi empirical models, J. Supercrit. Fluids. 78 (2013) 28–33. https://doi.org/10.1016/j.supflu.2013.03.018.

[152] R. Haghbakhsh, H. Hayer, M. Saidi, S. Keshtkari, F. Esmaeilzadeh, Density estimation of pure carbon dioxide at supercritical region and estimation solubility of solid compounds in supercritical carbon dioxide: Correlation approach based on sensitivity analysis, Fluid Phase Equilib. 342 (2013) 31–41. https://doi.org/10.1016/j.fluid.2012.12.029.

[153] A. Tabernero, S.A.B. Vieira de Melo, R. Mammucari, A.M. Martin del Valle, N.R. Foster, Modeling Solubility of Solids in Supercritical Carbon Dioxide Using q-Chrastil Equation, III Iberoam. Conf. Supercrit. Fluids Cart. Indias. (2013) 1–8.

[154] A.Z. Hezave, M. Lashkarbolooki, A new simple correlation for calculating solubility of drugs in supercritical carbon dioxide, J. Theor. Comput. Chem. 12 (2013). https://doi.org/10.1142/S0219633613500624.

[155] M. Asgarpour Khansary, F. Amiri, A. Hosseini, A. Hallaji Sani, H. Shahbeig, Representing solute solubility in supercritical carbon dioxide: A novel empirical model, Chem. Eng. Res. Des. 93 (2015) 355–365. https://doi.org/10.1016/j.cherd.2014.05.004.

[156] S.B. Hozhabr, S.H. Mazloumi, J. Sargolzaei, Correlation of solute solubility in supercritical carbon dioxide using a new empirical equation, Chem. Eng. Res. Des. 92 (2014) 2734–2739. https://doi.org/10.1016/j.cherd.2014.01.026.

[157] L. Nasri, S. Bensaad, Z. Bensetiti, Modeling the solubility of dihydroxybenzoic acid and methylbenzoic acid isomers in supercritical carbon dioxide, Int. J. Thermodyn. 17 (2014) 81–85. https://doi.org/10.5541/ijot.513.

[158] X.Q. Bian, Q. Zhang, Z.M. Du, J. Chen, J.N. Jaubert, A five-parameter empirical model for correlating the solubility of solid compounds in supercritical carbon dioxide, Fluid

Phase Equilib. 411 (2016) 74–80. https://doi.org/10.1016/j.fluid.2015.12.017.

[159] H. Asiabi, Y. Yamini, F. Latifeh, A. Vatanara, Solubilities of four macrolide antibiotics in supercritical carbon dioxide and their correlations using semi-empirical models, J. Supercrit. Fluids. 104 (2015) 62–69. https://doi.org/10.1016/j.supflu.2015.05.018.

[160] M. Banchero, L. Manna, Solubility of fenamate drugs in supercritical carbon dioxide by using a semi-flow apparatus with a continuous solvent-washing step in the depressurization line, J. Supercrit. Fluids. 107 (2016) 400–407. https://doi.org/10.1016/j.supflu.2015.10.008.

[161] T.A. Reddy, C. Srividya R, Garlapati, A new empirical model to correlate solubility of pharmaceutical compounds in supercritical carbon dioxide, J. Appl. Sci. Eng. Methodol. 2 (2018) 220–223.

[162] W. Wang, X. Gui, Z. Yun, New models for correlating and predicting the solubility of some compounds in supercritical CO2, Fluid Phase Equilib. 430 (2016) 135–142. https://doi.org/10.1016/j.fluid.2016.09.028.

[163] C. Si-moussa, A. Belghait, L. Khaouane, S. Hanini, Comptes Rendus Chimie Novel density-based model for the correlation of solid drugs solubility in supercritical carbon dioxide, Comptes Rendus - Chim. (2016). https://doi.org/10.1016/j.crci.2016.09.009.

[164] T.A. Reddy, R. Srividya, C. Garlapati, A new empirical model to correlate solubility of pharmaceutical compounds in supercritical carbon dioxide, J. Appl. Sci. Eng. Methodol. 4 (2018) 575–590.

[165] A. Belghait, C. Si-Moussa, M. Laidi, S. Hanini, Corrélation semi-empirique de solubilité des solutés solides dans le dioxyde de carbone supercritique: étude comparative et proposition d'un nouveau modèle basé sur la densité, Comptes Rendus Chim. 21 (2018) 494–513. https://doi.org/10.1016/j.crci.2018.02.006.

[166] T.A. Reddy, C. Garlapati, Dimensionless Empirical Model to Correlate Pharmaceutical Compound Solubility in Supercritical Carbon Dioxide, Chem. Eng. Technol. 42 (2019) 2621–2630. https://doi.org/10.1002/ceat.201900283.

[167] Z. Vatani, S. Ramezanian Bajgiran, G. Amini, S. Tayyebi, Solubility modeling of supercritical fluid extraction in a wide range compounds: comparison between fuzzy-genetic and new empirical models, Energy Sources, Part A Recover. Util. Environ. Eff. 42 (2020) 365–374. https://doi.org/10.1080/15567036.2019.1587083.

[168] M. Pishnamazi, S. Zabihi, S. Jamshidian, F. Borousan, A.Z. Hezave, S. Shirazian, Thermodynamic modelling and experimental validation of pharmaceutical solubility in supercritical solvent, J. Mol. Liq. 319 (2020) 114120. https://doi.org/10.1016/j.molliq.2020.114120.

[169] M. Bazaei, B. Honarvar, N. Esfandiari, S.A. Sajadian, Z.A. Aboosadi, Measurement and thermodynamic modeling of solubility of Erlotinib hydrochloride, as an anti-cancer drug, in supercritical carbon dioxide, Fluid Phase Equilib. 573 (2023) 113877. https://doi.org/10.1016/j.fluid.2023.113877.

[170] E. Ansari, B. Honarvar, S.A. Sajadian, Z.A. Aboosadi, M. Azizi, Experimental solubility of aripiprazole in supercritical carbon dioxide and modeling, Sci. Rep. 13 (2023) 1–12. https://doi.org/10.1038/s41598-023-40537-3.

[171] G. Sodeifian, C. Garlapati, M. Arbab Nooshabadi, F. Razmimanesh, A. Tabibzadeh, Solubility measurement and modeling of hydroxychloroquine sulfate (antimalarial medication) in supercritical carbon dioxide, Sci. Rep. 13 (2023) 1–16. https://doi.org/10.1038/s41598-023-34900-7.

[172] S.M. Alshahrani, A.M. Alsubaiyel, M.H. Abduljabbar, M.A.S. Abourehab, Measurement of metoprolol solubility in supercritical carbon dioxide; Experimental and modeling study, Case Stud. Therm. Eng. 42 (2023) 102764. https://doi.org/10.1016/j.csite.2023.102764.

[173] J.S. Cheng, M. Tang, Y.P. Chen, Correlation of solid solubility for biological compounds in supercritical carbon dioxide: Comparative study using solution model and other approaches, Fluid Phase Equilib. 194–197 (2002) 483–491. https://doi.org/10.1016/S0378-3812(01)00657-4.

[174] C.S. Su, Y.P. Chen, Correlation for the solubilities of pharmaceutical compounds in supercritical carbon dioxide, Fluid Phase Equilib. 254 (2007) 167–173. https://doi.org/10.1016/j.fluid.2007.03.004.

[175] C.S. Su, Correlation and estimation of solubilities of fatty acids in supercritical carbon dioxide using solution model approach, J. Supercrit. Fluids. 81 (2013) 79–85. https://doi.org/10.1016/j.supflu.2013.05.001.

[176] C.Y. Huang, L.S. Lee, C.S. Su, Correlation of solid solubilities of pharmaceutical compounds in supercritical carbon dioxide with solution model approach, J. Taiwan Inst.

Chem. Eng. 44 (2013) 349–358. https://doi.org/10.1016/j.jtice.2012.12.004.

[177] V. Mani Rathnam, N. Lamba, G. Madras, Evaluation of new density based model to correlate the solubilities of ricinoleic acid, methyl ricinoleate and methyl 10-undecenoate in supercritical carbon dioxide, J. Supercrit. Fluids. 130 (2017) 357–363. https://doi.org/10.1016/j.supflu.2017.07.017.

[178] V.R. Vasconcellos, F.A. Cabral, Vasconcellos_JAOC_solubilidad, 78 (2001) 827–829.

[179] S.A.B. Vieira De Melo, R.L.F. Vieira De Melo, G.M.N. Costa, T.L.M. Alves, Solubility of l-Dopa in supercritical carbon dioxide: Prediction using a cubic equation of state, J. Supercrit. Fluids. 34 (2005) 231–236. https://doi.org/10.1016/j.supflu.2004.11.019.

[180] B. Tomberli, S. Goldman, C.G. Gray, M.D.A. Saldaña, F. Temelli, Using solute structure to predict solubility of organic molecules in supercritical carbon dioxide, J. Supercrit. Fluids. 37 (2006) 333–341. https://doi.org/10.1016/j.supflu.2005.10.008.

[181] A. Tabernero, E.M. Martín Del Valle, M.A. Galán, An empirical analysis of the solubility of pharmaceuticals in supercritical carbon dioxide using sublimation enthalpies, Ind. Eng. Chem. Res. 52 (2013) 18447–18457. https://doi.org/10.1021/ie403031k.

[182] M. Hosseini Anvari, G. Pazuki, A study on the predictive capability of the SAFT-VR equation of state for solubility of solids in supercritical CO2, J. Supercrit. Fluids. 90 (2014) 73–83. https://doi.org/10.1016/j.supflu.2014.03.005.

[183] S.A. Jahromi, A. Roosta, Estimation of critical point, vapor pressure and heat of sublimation of pharmaceuticals and their solubility in supercritical carbon dioxide, Fluid Phase Equilib. 488 (2019) 1–8. https://doi.org/10.1016/j.fluid.2019.01.020.

[184] Z.Z. Cai, C.M. Hsieh, Prediction of solid solute solubility in supercritical carbon dioxide with and without organic cosolvents from PSRK EOS, J. Supercrit. Fluids. 158 (2020) 104735. https://doi.org/10.1016/j.supflu.2019.104735.

[185] C. Morales-Díaz, A.L. Cabrera, J.C. de la Fuente, A. Mejía, Modelling of solubility of vitamin K3 derivatives in supercritical carbon dioxide using cubic and SAFT equations of state, J. Supercrit. Fluids. 167 (2021). https://doi.org/10.1016/j.supflu.2020.105040.

[186] H. Bezaze, Optimization of The Binary Interaction Coefficient Kij for Calculating the Solubility of Drug Compounds in Co2 Using a Genetic Algorithm Chemical compounds,

8 (2022) 1–17.

[187] M. Loreno, R.A. Reis, S. Mattedi, M.L.L. Paredes, Statistical analysis of parameterization strategies for GC-sPC-SAFT equation of state to predict carbon dioxide solubility in imidazolium-based ionic liquids, Fluid Phase Equilib. 574 (2023). https://doi.org/10.1016/j.fluid.2023.113899.

[188] M. Loreno, R.A. Reis, S. Mattedi, M.L.L. Paredes, Predicting the solubility of carbon dioxide or methane in imidazolium-based ionic liquids with GC-sPC-SAFT equation of state, Fluid Phase Equilib. 479 (2019) 85–98. https://doi.org/10.1016/j.fluid.2018.09.010.

[189] A.K. Mutlag, COMPARISON STUDY FOR THE SOLUBILITY PREDICTION OF PHENANTHRENE IN SUPERCRITICAL CO 2 ENTRAINED WITH n- PENTANE WHEN USING EQUATION OF STATE AND ARTIFICIAL NEURAL NETWORKS, 16 (2011) 81–90.

[190] B. Mehdizadeh, K. Movagharnejad, A comparative study between LS-SVM method and semi empirical equations for modeling the solubility of different solutes in supercritical carbon dioxide, Chem. Eng. Res. Des. 89 (2011) 2420–2427. https://doi.org/10.1016/j.cherd.2011.03.017.

[191] B. Mehdizadeh, K. Movagharnejad, A comparison between neural network method and semi empirical equations to predict the solubility of different compounds in supercritical carbon dioxide, Fluid Phase Equilib. 303 (2011) 40–44. https://doi.org/10.1016/j.fluid.2010.12.010.

[192] F. Gharagheizi, A. Eslamimanesh, A.H. Mohammadi, D. Richon, Artificial neural network modeling of solubilities of 21 commonly used industrial solid compounds in supercritical carbon dioxide, Ind. Eng. Chem. Res. 50 (2011) 221–226. https://doi.org/10.1021/ie101545g.

[193] Ö. Güçlü Üstündağ, U. Güçlü, S. Erşan, Application of artificial neural networks for the modeling of \nlipid solubility in supercritical carbon dioxide, 10th Int. Symp. Supercrit. Fluids. (2012). http://issf2012.com/handouts/documents/569_003.pdf.

[194] I.A.A. Hamid, A.N. Mustapa, N. Ismail, Z. Abdullah, Solubility prediction of mangosteen peel oil in Supercritical Carbon Dioxide using Neural Network, BEIAC 2013 - 2013 IEEE Bus. Eng. Ind. Appl. Colloq. (2013) 91–96.

https://doi.org/10.1109/BEIAC.2013.6560269.

[195] B. Vaferi, M. Karimi, M. Azizi, H. Esmaeili, Comparison between the artificial neural network, SAFT and PRSV approach in obtaining the solubility of solid aromatic compounds in supercritical carbon dioxide, J. Supercrit. Fluids. 77 (2013) 44–51. https://doi.org/10.1016/j.supflu.2013.02.027.

[196] M.R. Dadkhah, A. Tatar, A. Mohebbi, A. Barati-Harooni, A. Najafi-Marghmaleki, M.M. Ghiasi, A.H. Mohammadi, F. Pourfayaz, Prediction of solubility of solid compounds in supercritical CO2 using a connectionist smart technique, J. Supercrit. Fluids. 120 (2017) 181–190. https://doi.org/10.1016/j.supflu.2016.06.006.

[197] M. Soleimani Lashkenari, A. KhazaiePoul, Application of KNN and Semi-Empirical Models for Prediction of Polycyclic Aromatic Hydrocarbons Solubility in Supercritical Carbon Dioxide, Polycycl. Aromat. Compd. 37 (2017) 415–425. https://doi.org/10.1080/10406638.2015.1129976.

[198] A. Abdallah el hadj, M. Laidi, C. Si-Moussa, S. Hanini, Novel approach for estimating solubility of solid drugs in supercritical carbon dioxide and critical properties using direct and inverse artificial neural network (ANN), Neural Comput. Appl. 28 (2017) 87–99. https://doi.org/10.1007/s00521-015-2038-1.

[199] A. Baghban, A. Jalali, A.H. Mohammadi, S. Habibzadeh, Efficient modeling of drug solubility in supercritical carbon dioxide, J. Supercrit. Fluids. 133 (2018) 466–478. https://doi.org/10.1016/j.supflu.2017.10.032.

[200] A. Aminian, Estimating the solubility of different solutes in supercritical CO2 covering a wide range of operating conditions by using neural network models, J. Supercrit. Fluids. 125 (2017) 79–87. https://doi.org/10.1016/j.supflu.2017.02.007.

[201] A. Baghban, J. Sasanipour, Z. Zhang, A new chemical structure-based model to estimate solid compound solubility in supercritical CO2, J. CO2 Util. 26 (2018) 262–270. https://doi.org/10.1016/j.jcou.2018.05.009.

[202] M. Abdi-Khanghah, A. Bemani, Z. Naserzadeh, Z. Zhang, Prediction of solubility of N-alkanes in supercritical CO2 using RBF-ANN and MLP-ANN, J. CO2 Util. 25 (2018) 108–119. https://doi.org/10.1016/j.jcou.2018.03.008.

[203] M. Li, S. Lian, F. Wang, Y. Zhou, B. Chen, L. Guan, Y. Wu, Neural network modeling based double-population chaotic accelerated particle swarm optimization and diffusion

theory for solubility prediction, Chem. Eng. Res. Des. 155 (2020) 98–107. https://doi.org/10.1016/j.cherd.2020.01.003.

[204] T. Wang, C.H. Su, Medium Gaussian SVM, Wide Neural Network and stepwise linear method in estimation of Lornoxicam pharmaceutical solubility in supercritical solvent, J. Mol. Liq. 349 (2022) 118120. https://doi.org/10.1016/j.molliq.2021.118120.

[205] H. Zhu, L. Zhu, Z. Sun, A. Khan, Machine learning based simulation of an anti-cancer drug (busulfan) solubility in supercritical carbon dioxide: ANFIS model and experimental validation, J. Mol. Liq. 338 (2021) 116731. https://doi.org/10.1016/j.molliq.2021.116731.

[206] M. Najmi, M.A. Ayari, H. Sadeghsalehi, B. Vaferi, A. Khandakar, M.E.H. Chowdhury, T. Rahman, Z.H. Jawhar, Estimating the Dissolution of Anticancer Drugs in Supercritical Carbon Dioxide with a Stacked Machine Learning Model, Pharmaceutics. 14 (2022). https://doi.org/10.3390/pharmaceutics14081632.

[207] F. An, B.T. Sayed, R.M.R. Parra, M.H. Hamad, R. Sivaraman, Z. Zanjani Foumani, A.A. Rushchitc, E. El-Maghawry, R.M. Alzhrani, S. Alshehri, K. M. AboRas, Machine learning model for prediction of drug solubility in supercritical solvent: Modeling and experimental validation, J. Mol. Liq. 363 (2022) 119901. https://doi.org/10.1016/j.molliq.2022.119901.

[208] Y. Liu, D. Ahmed Smait, A. Yaseen Naser, F. M. A. Altalbawy, H. Bahri, A. Abdul Kadhim Ruhaima, T. Zayad Fathallah, S.K. Hadrawi, R.E. Alsaddon, A. Alshetaili, A.M. Alsubaiyel, Optimization and validation of drug solubility by development of advanced artificial intelligence models, J. Mol. Liq. 372 (2023) 121113. https://doi.org/10.1016/j.molliq.2022.121113.

[209] H. Chinh Nguyen, F. Alamray, M. Kamal, T. Diana, A. Mohamed, M. Algarni, C.H. Su, Computational prediction of drug solubility in supercritical carbon dioxide: Thermodynamic and artificial intelligence modeling, J. Mol. Liq. 354 (2022) 118888. https://doi.org/10.1016/j.molliq.2022.118888.

[210] E. V. Kostyrin, V. V. Ponkratov, A. Salah Al-Shati, Development of machine learning model and analysis study of drug solubility in supercritical solvent for green technology development, Arab. J. Chem. 15 (2022) 104346. https://doi.org/10.1016/j.arabjc.2022.104346.

[211] Z. Zhao, P. Liu, Y. Li, S. Zhang, L. Guo, S. Ghazali, A.S. El-Shafay, Multi support vector models to estimate solubility of Busulfan drug in supercritical carbon dioxide, J. Mol. Liq. 350 (2022) 118573. https://doi.org/10.1016/j.molliq.2022.118573.

[212] A. Byabani-Givo, M. Khajeh, M. Bohlooli, B. Keshtegar, M. Ghaffari-Moghaddam, M.A. Lakzahi Moghaddam, Prediction of solubility of some dihydropyridine derivative drugs in supercritical fluid carbon dioxide by RBFNN, Netw. Model. Anal. Heal. Informatics Bioinforma. 11 (2022) 1–12. https://doi.org/10.1007/s13721-022-00380-4.

[213] T. Rezaei, V. Nazarpour, N. Shahini, S. Bahmani, A. Shahkar, M. Abdihaji, S. Ahmadi, F.T. Shahdost, A universal methodology for reliable predicting the non-steroidal anti-inflammatory drug solubility in supercritical carbon dioxide, Sci. Rep. 12 (2022) 1–12. https://doi.org/10.1038/s41598-022-04942-4.

[214] Z. Tianhao, H. Sh. Majdi, B. Dmitry Olegovich, W.K. Abdelbasset, L. Thangavelu, C.H. Su, H. Chinh Nguyen, M. Alashwal, S. Ghazali, Prediction of busulfan solubility in supercritical CO2 using tree-based and neural network-based methods, J. Mol. Liq. 351 (2022) 118630. https://doi.org/10.1016/j.molliq.2022.118630.

[215] W. Liu, R. Zhao, X. Su, A. Mohamed, T. Diana, Development and validation of machine learning models for prediction of nanomedicine solubility in supercritical solvent for advanced pharmaceutical manufacturing, J. Mol. Liq. 358 (2022) 119208. https://doi.org/10.1016/j.molliq.2022.119208.

[216] W.K. Abdelbasset, S.M. Elkholi, K. Ahmed Ismail, T. A.A.M. Alalwani, K. Hachem, A. Mohamed, T. Agustiono Kurniawan, A. Andreevna Rushchtc, Modeling and computational study on prediction of pharmaceutical solubility in supercritical CO2 for manufacture of nanomedicine for enhanced bioavailability, J. Mol. Liq. 359 (2022) 119306. https://doi.org/10.1016/j.molliq.2022.119306.

[217] Y. Li, A.A. Alameri, Z.A. Farhan, H. l. AI_Sadi, M.E. Alosaimi, A. Ghaleb AbdalSalam, D. Jumaah Jasim, S.K. Hadrawi, M. Mohammed Al-Taee, A.H. Lafta, H.A. Othman, S. Mousa Alzahrani, A.A. Moniem, T. Alqadi, Theoretical modeling study on preparation of nanosized drugs using supercritical-based processing: Determination of solubility of Chlorothiazide in supercritical carbon dioxide, J. Mol. Liq. 370 (2023) 120984. https://doi.org/10.1016/j.molliq.2022.120984.

[218] N.I. Namazi, S. Alshehri, R. Bafail, B. Huwaimel, A.M. Alsubaiyel, A.H. Alamri, A.D.

Alatawi, H. Kotb, M.S. Sarjadi, M.L. Rahman, M.A.S. Abourehab, Solubility enhancement of decitabine as anticancer drug via green chemistry solvent: Novel computational prediction and optimization, Arab. J. Chem. 15 (2022) 104259. https://doi.org/10.1016/j.arabjc.2022.104259.

[219] B. Huwaimel, T.N. Alharby, Development of computational intelligence models for assessment of drug nanonization using green chemistry technique: Improvement of drug solubility, Case Stud. Therm. Eng. 45 (2023) 103005. https://doi.org/10.1016/j.csite.2023.103005.

[220] M.Y. Begum, Advanced modeling based on machine learning for evaluation of drug nanoparticle preparation via green technology: Theoretical assessment of solubility variations, Case Stud. Therm. Eng. 45 (2023) 103029. https://doi.org/10.1016/j.csite.2023.103029.

[221] B. Huwaimel, A.S. Abouzied, Development of green technology based on supercritical solvent for production of nanomedicine: Solubility prediction using computational methods, J. Mol. Liq. 376 (2023) 121471. https://doi.org/10.1016/j.molliq.2023.121471.

[222] A.A. Waskita, S. Yushady, C.H. Bissa, I.A. Satya, R.S. Alwi, Development of Novel Machine Learning to Optimize the Solubility of Azathioprine as Anticancer Drug in Supercritical Carbon Dioxide, 9 (2023) 49–57. https://doi.org/10.26555/jiteki.v9i1.25608.

[223] U. Hani, Z. Ali Bu sinnah, A.J. Obaidullah, B. Huwaimel, M. Alanazi, T. Nafea Alharby, A.A. Lahiq, A. Ali Alshehri, Computational intelligence modeling of nanomedicine preparation using advanced processing: Solubility of fludrocortisone acetate in supercritical carbon dioxide, Case Stud. Therm. Eng. 45 (2023) 102968. https://doi.org/10.1016/j.csite.2023.102968.

[224] A.J. Obaidullah, Advanced AI modeling and optimization for determination of pharmaceutical solubility in supercritical processing for production of nanosized drug particles, Case Stud. Therm. Eng. 49 (2023) 103199. https://doi.org/10.1016/j.csite.2023.103199.

[225] M. Zhang, W.A. Mahdi, Development of SVM-based machine learning model for estimating lornoxicam solubility in supercritical solvent, Case Stud. Therm. Eng. 49 (2023) 103268. https://doi.org/10.1016/j.csite.2023.103268.

[226] H. Xiao Li, U. Abdul-Reda Hussein, I. Waleed, S. Hassan Zain Al-Abdeen, F.M.A. Altalbawy, Z. Hussein Adhab, A. Faisal, M.Y. Alshahrani, H. Kamil Zaidan, M. Suliman, X. Ben Hu, An advanced computational method for studying drug nanonization using green supercritical-based processing for improvement of pharmaceutical bioavailability in aqueous media, J. Mol. Liq. 381 (2023) 121805. https://doi.org/10.1016/j.molliq.2023.121805.

[227] C. Hao, Y. Li, L. Zhang, M. Sheng, Y. Cui, A.S. Al-Shati, M. Khalid, Computational study and experimental validation on the solubility of drugs in supercritical solvent for assessment of nanomedicine production via green technology for enhanced drug bioavailability, J. Mol. Liq. 382 (2023) 121835. https://doi.org/10.1016/j.molliq.2023.121835.

[228] M. Ghazwani, M. Yasmin Begum, A.M. Naglah, H.M. Alkahtani, A.A. Almehizia, Development of advanced model for understanding the behavior of drug solubility in green solvents: Machine learning modeling for small-molecule API solubility prediction, J. Mol. Liq. 386 (2023) 122446. https://doi.org/10.1016/j.molliq.2023.122446.

[229] M. Ghazwani, M. Yasmin Begum, Machine learning aided drug development: Assessing improvement of drug efficiency by correlation of solubility in supercritical solvent for nanomedicine preparation, J. Mol. Liq. 387 (2023) 122511. https://doi.org/10.1016/j.molliq.2023.122511.

[230] P.M. Agrawal, D.C. Sorescu, B.M. Rice, D.L. Thompson, A model for predicting the solubility of 1,3,5-trinitro-1,3,5-s-triazine (RDX) in supercritical CO2: Isothermal-isobaric Monte Carlo simulations, Fluid Phase Equilib. 155 (1999) 177–191. https://doi.org/10.1016/S0378-3812(98)00471-3.

[231] J. García, F. Rodríguez, J.A. Revenga, Modelling solubility of solids in supercritical fluids using response surface methodology, J. Chem. Technol. Biotechnol. 75 (2000) 245–251. https://doi.org/10.1002/(SICI)1097-4660(200003)75:3<245::AID-JCTB200>3.0.CO;2-C.

[232] A. Jouybana, S. Soltanib, K.-A. Zeynalic, Ar ch ive, 6 (2007) 243–250.

[233] Y. Shimoyama, Y. Iwai, Development of activity coefficient model based on COSMO method for prediction of solubilities of solid solutes in supercritical carbon dioxide, J. Supercrit. Fluids. 50 (2009) 210–217. https://doi.org/10.1016/j.supflu.2009.06.004.

[234] L.H. Wang, S.T. Lin, A predictive method for the solubility of drug in supercritical carbon dioxide, J. Supercrit. Fluids. 85 (2014) 81–88. https://doi.org/10.1016/j.supflu.2013.10.019.

[235] J. huan Li, Z. Huang, J. lin Wei, L. Xu, A new optimization method for parameter determination in modeling solid solubility in supercritical CO2, Fluid Phase Equilib. 344 (2013) 117–124. https://doi.org/10.1016/j.fluid.2013.01.028.

[236] J. Sakabe, H. Uchida, Y. Shimoyama, Modeling of drug solubility in supercritical carbon dioxide using equation of state based on hole theory with molecular surface charge density, Chem. Eng. Res. Des. 92 (2014) 2970–2976. https://doi.org/10.1016/j.cherd.2014.08.003.

[237] H. Li, D. Jia, R. Liu, B. Shen, Correlating and estimating the solubilities of solid organic compounds in supercritical CO2 based on the modified expanded liquid model and the group contribution method, Fluid Phase Equilib. 385 (2015) 10–24. https://doi.org/10.1016/j.fluid.2014.10.036.

[238] H. Li, D. Jia, Q. Zhu, B. Shen, Determination, correlation and prediction of the solubilities of niflumic acid, clofenamic acid and tolfenamic acid in supercritical CO2, Fluid Phase Equilib. 392 (2015) 95–103. https://doi.org/10.1016/j.fluid.2015.02.007.

[239] X.Q. Bian, J. Li, J. Chen, M.J. Li, Z.M. Du, A combined model for the solubility of different compounds in supercritical carbon dioxide, Chem. Eng. Res. Des. 104 (2015) 416–428. https://doi.org/10.1016/j.cherd.2015.08.028.

[240] J. Sakabe, H. Uchida, Y. Shimoyama, Mixed solid phase model using equation of state based on hole-theory for solubility prediction of pharmaceutical compound in supercritical CO2, J. Supercrit. Fluids. 100 (2015) 26–33. https://doi.org/10.1016/j.supflu.2015.02.002.

[241] A.G. Reveco-Chilla, L.M. Valenzuela, J.M. del Valle, E.J. Maginn, Use of molecular dynamics simulations to estimate the solubility of menadione in supercritical CO2 using Chrastil's model, Fluid Phase Equilib. 433 (2017) 112–118. https://doi.org/10.1016/j.fluid.2016.10.017.

[242] G. Sodeifian, C. Garlapati, F. Razmimanesh, F. Sodeifian, The solubility of Sulfabenzamide (an antibacterial drug) in supercritical carbon dioxide : Evaluation of a new thermodynamic model, J. Mol. Liq. 335 (2021) 116446.

https://doi.org/10.1016/j.molliq.2021.116446.

[243] M.M. Azim, I. Ushiki, A. Miyajima, S. Takishima, Modeling the solubility of non-steroidal anti-inflammatory drugs (ibuprofen and ketoprofen) in supercritical CO2 using PC-SAFT, J. Supercrit. Fluids. 186 (2022) 105626. https://doi.org/10.1016/j.supflu.2022.105626.

[244] A. Abdallah el hadj, M. Laidi, S. Hanini, AI-PCSAFT approach: New high predictive method for estimating PC-SAFT pure component properties and phase equilibria parameters, Fluid Phase Equilib. 555 (2022). https://doi.org/10.1016/j.fluid.2021.113297.

[245] M. Abadian, G. Sodeifian, F. Razmimanesh, S. Zarei Mahmoudabadi, Experimental measurement and thermodynamic modeling of solubility of Riluzole drug (neuroprotective agent) in supercritical carbon dioxide, Fluid Phase Equilib. 567 (2023) 113711. https://doi.org/10.1016/j.fluid.2022.113711.

[246] R.S. Alwi, A. Rojas, N. Esfandiari, S.A. Sajadian, N.S. Ardestani, A. Jouyban, Experimental study and thermodynamic modeling of clonazepam solubility in supercritical carbon dioxide, Fluid Phase Equilib. 574 (2023) 113880. https://doi.org/10.1016/j.fluid.2023.113880.

[247] T. Khayamian, M. Esteki, Prediction of solubility for polycyclic aromatic hydrocarbons in supercritical carbon dioxide using wavelet neural networks in quantitative structure property relationship, J. Supercrit. Fluids. 32 (2004) 73–78. https://doi.org/10.1016/j.supflu.2004.02.003.

[248] B. Hemmateenejad, M. Shamsipur, R. Miri, M. Elyasi, F. Foroghinia, H. Sharghi, Linear and nonlinear quantitative structure-property relationship models for solubility of some anthraquinone, anthrone and xanthone derivatives in supercritical carbon dioxide, Anal. Chim. Acta. 610 (2008) 25–34. https://doi.org/10.1016/j.aca.2008.01.011.

[249] L.M. Valenzuela, A.G. Reveco-Chilla, J.M. Del Valle, Modeling solubility in supercritical carbon dioxide using quantitative structure-property relationships, J. Supercrit. Fluids. 94 (2014) 113–122. https://doi.org/10.1016/j.supflu.2014.06.022.

[250] P. Gurikov, I. Lebedev, A. Kolnoochenko, N. Menshutina, Prediction of the Solubility in Supercritical Carbon Dioxide: a Hybrid Thermodynamic/QSPR Approach, Elsevier Masson SAS, 2016. https://doi.org/10.1016/B978-0-444-63428-3.50269-1.

[251] M. Moussaoui, M. Laidi, S. Hanini, M. Hentabli, Artificial Neural Network and Support Vector Regression Applied in Quantitative Structure-property Relationship Modelling of Solubility of Solid Solutes in Supercritical CO2, Kem. u Ind. 69 (2020) 611–630. https://doi.org/10.15255/kui.2020.004.

[252] Y.S. Kim, W.Y. Choi, J.H. Jang, K.P. Yoo, C.S. Lee, Solubility measurement and prediction of carbon dioxide in ionic liquids, Fluid Phase Equilib. 228–229 (2005) 439–445. https://doi.org/10.1016/j.fluid.2004.09.006.

[253] M.C. Kroon, E.K. Karakatsani, I.G. Economou, G.J. Witkamp, C.J. Peters, Modeling of the carbon dioxide solubility in imidazolium-based ionic liquids with the tPC-PSAFT equation of state, J. Phys. Chem. B. 110 (2006) 9262–9269. https://doi.org/10.1021/jp060300o.

[254] J.S. Torrecilla, J. Palomar, J. García, E. Rojo, F. Rodríguez, Modelling of carbon dioxide solubility in ionic liquids at sub and supercritical conditions by neural networks and mathematical regressions, Chemom. Intell. Lab. Syst. 93 (2008) 149–159. https://doi.org/10.1016/j.chemolab.2008.05.004.

[255] J.E. Kim, J.S. Lim, J.W. Kang, Measurement and correlation of solubility of carbon dioxide in 1-alkyl-3-methylimidazolium hexafluorophosphate ionic liquids, Fluid Phase Equilib. 306 (2011) 251–255. https://doi.org/10.1016/j.fluid.2011.04.017.

[256] S. Mortazavi-manesh, M. Satyro, R.A. Marriott, Fluid Phase Equilibria A semi-empirical Henry ' s law expression for carbon dioxide dissolution in ionic liquids, Fluid Phase Equilib. 307 (2011) 208–215. https://doi.org/10.1016/j.fluid.2011.05.006.

[257] S.S. Ashrafmansouri, S. Raeissi, Modeling gas solubility in ionic liquids with the SAFT-γ group contribution method, J. Supercrit. Fluids. 63 (2012) 81–91. https://doi.org/10.1016/j.supflu.2011.12.014.

[258] A. Lymperiadis, C.S. Adjiman, A. Galindo, G. Jackson, A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ), J. Chem. Phys. 127 (2007). https://doi.org/10.1063/1.2813894.

[259] A. Eslamimanesh, F. Gharagheizi, A.H. Mohammadi, D. Richon, Artificial Neural Network modeling of solubility of supercritical carbon dioxide in 24 commonly used ionic liquids, Chem. Eng. Sci. 66 (2011) 3039–3044. https://doi.org/10.1016/j.ces.2011.03.016.

[260] Z. Lei, J. Han, B. Zhang, Q. Li, J. Zhu, B. Chen, Solubility of CO 2 in binary mixtures of room-temperature ionic liquids at high pressures, J. Chem. Eng. Data. 57 (2012) 2153–2159. https://doi.org/10.1021/je300016q.

[261] M. Gonzalez-Miquel, M. Talreja, A.L. Ethier, K. Flack, J.R. Switzer, E.J. Biddinger, P. Pollet, J. Palomar, F. Rodriguez, C.A. Eckert, C.L. Liotta, COSMO-RS Studies: Structure-property relationships for CO2 capture by reversible ionic liquids, Ind. Eng. Chem. Res. 51 (2012) 16066–16073. https://doi.org/10.1021/ie302449c.

[262] Y. Chen, F. Mutelet, J.N. Jaubert, Modeling the solubility of carbon dioxide in imidazolium-based ionic liquids with the PC-SAFT equation of state, J. Phys. Chem. B. 116 (2012) 14375–14388. https://doi.org/10.1021/jp309944t.

[263] S. Mortazavi-Manesh, M. Satyro, R.A. Marriott, Modelling carbon dioxide solubility in ionic liquids, Can. J. Chem. Eng. 91 (2013) 783–789. https://doi.org/10.1002/cjce.21687.

[264] R. MacÍas-Salinas, J.A. Chávez-Velasco, M.A. Aquino-Olivos, J.L. Mendoza De La Cruz, J.C. Sánchez-Ochoa, Accurate modeling of CO2 solubility in ionic liquids using a cubic EoS⊥, Ind. Eng. Chem. Res. 52 (2013) 7593–7601. https://doi.org/10.1021/ie400106g.

[265] R. Haghbakhsh, H. Soleymani, S. Raeissi, A simple correlation to predict high pressure solubility of carbon dioxide in 27 commonly used ionic liquids, J. Supercrit. Fluids. 77 (2013) 158–166. https://doi.org/10.1016/j.supflu.2013.02.022.

[266] M.A. Sedghamiz, A. Rasoolzadeh, M.R. Rahimpour, The ability of artificial neural network in prediction of the acid gases solubility in different ionic liquids, J. CO2 Util. 9 (2015) 39–47. https://doi.org/10.1016/j.jcou.2014.12.003.

[267] A. Baghban, M.A. Ahmadi, B.H. Shahraki, Prediction carbon dioxide solubility in presence of various ionic liquids using computational intelligence approaches, Elsevier B.V., 2015. https://doi.org/10.1016/j.supflu.2015.01.002.

[268] A. Tatar, S. Naseri, M. Bahadori, A.Z. Hezave, T. Kashiwao, A. Bahadori, H. Darvish, Prediction of carbon dioxide solubility in ionic liquids using MLP and radial basis function (RBF) neural networks, J. Taiwan Inst. Chem. Eng. 60 (2016) 151–164. https://doi.org/10.1016/j.jtice.2015.11.002.

[269] H.F. Hizaddin, M.K. Hadj-Kali, I.M. Alnashef, F.S. Mjalli, M.A. Hashim, Prediction of

CO2 solubility in ionic liquids using the PSRK model, J. Supercrit. Fluids. 100 (2015) 184–193. https://doi.org/10.1016/j.supflu.2015.02.009.

[270] E. Ali, M.K. Hadj-Kali, I. Alnashef, Modeling of CO2 Solubility in Selected Imidazolium-Based Ionic Liquids, Chem. Eng. Commun. 204 (2017) 205–215. https://doi.org/10.1080/00986445.2016.1254086.

[271] I. Mehraein, S. Riahi, The QSPR models to predict the solubility of CO2 in ionic liquids based on least-squares support vector machines and genetic algorithm-multi linear regression, J. Mol. Liq. 225 (2017) 521–530. https://doi.org/10.1016/j.molliq.2016.10.133.

[272] A. Tatar, A. Barati, A. Yarahmadi, A. Najafi, M. Lee, A. Bahadori, Prediction of carbon dioxide solubility in aqueous mixture of methyldiethanolamine and N-methylpyrrolidone using intelligent models, Int. J. Greenh. Gas Control. 47 (2016) 122–136. https://doi.org/10.1016/j.ijggc.2016.01.048.

[273] M. Zoubeik, M. Mohamedali, A. Henni, Experimental solubility and thermodynamic modeling of CO2 in four new imidazolium and pyridinium-based ionic liquids, Fluid Phase Equilib. 419 (2016) 67–74. https://doi.org/10.1016/j.fluid.2016.03.009.

[274] V. Venkatraman, B.K. Alsberg, Predicting CO2 capture of ionic liquids using machine learning, J. CO2 Util. 21 (2017) 162–168. https://doi.org/10.1016/j.jcou.2017.06.012.

[275] H. Soltani Panah, Modeling H2S and CO2 solubility in ionic liquids using the CPA equation of state through a new approach, Fluid Phase Equilib. 437 (2017) 155–165. https://doi.org/10.1016/j.fluid.2017.01.023.

[276] H. Al-fnaish, L. Lue, Modelling the solubility of H2S and CO2 in ionic liquids using PC-SAFT equation of state, Fluid Phase Equilib. 450 (2017) 30–41. https://doi.org/10.1016/j.fluid.2017.07.008.

[277] T. Deng, F. hai Liu, G. zhu Jia, Prediction carbon dioxide solubility in ionic liquids based on deep learning, Mol. Phys. 118 (2020) 1–8. https://doi.org/10.1080/00268976.2019.1652367.

[278] Y. Zhang, X. Xu, Solubility predictions through LSBoost for supercritical carbon dioxide in ionic liquids, New J. Chem. 44 (2020) 20544–20567. https://doi.org/10.1039/d0nj03868g.

[279] H. Moosanezhad-Kermani, F. Rezaei, A. Hemmati-Sarapardeh, S.S. Band, A. Mosavi, Modeling of carbon dioxide solubility in ionic liquids based on group method of data handling, Eng. Appl. Comput. Fluid Mech. 15 (2021) 23–42. https://doi.org/10.1080/19942060.2020.1842250.

[280] Z. Song, H. Shi, X. Zhang, T. Zhou, Prediction of CO2 solubility in ionic liquids using machine learning methods, Chem. Eng. Sci. 223 (2020) 115752. https://doi.org/10.1016/j.ces.2020.115752.

[281] H. Ouaer, A.H. Hosseini, M.N. Amar, M.E.A. Ben Seghier, M.A. Ghriga, N. Nabipour, P.Ø. Andersen, A. Mosavi, S. Shamshirband, Rigorous connectionist models to predict carbon dioxide solubility in various ionic liquids, Appl. Sci. 10 (2020). https://doi.org/10.3390/app10010304.

[282] P. Valeh-e-Sheyda, M. Faridi Masouleh, P. Zarei-Kia, Prediction of CO2 solubility in pyridinium-based ionic liquids implementing new descriptor-based chemoinformatics models, Fluid Phase Equilib. 546 (2021) 113136. https://doi.org/10.1016/j.fluid.2021.113136.

[283] J. Wang, Z. Song, L. Chen, T. Xu, L. Deng, Z. Qi, Prediction of CO2 solubility in deep eutectic solvents using random forest model based on COSMO-RS-derived descriptors, Green Chem. Eng. 2 (2021) 431–440. https://doi.org/10.1016/j.gce.2021.08.002.

[284] R. Zhang, X. He, T. Liu, C. Li, M. Xiao, H. Ling, X. Hu, X. Zhang, F. Tang, H. Luo, Thermodynamic studies for improving the prediction of CO2 equilibrium solubility in aqueous 2-dimethylamino-2-methyl-1-propanol, Sep. Purif. Technol. 295 (2022) 121292. https://doi.org/10.1016/j.seppur.2022.121292.

[285] H. Feng, P. Zhang, W. Qin, W. Wang, H. Wang, Estimation of solubility of acid gases in ionic liquids using different machine learning methods, J. Mol. Liq. 349 (2022) 118413. https://doi.org/10.1016/j.molliq.2021.118413.

[286] Y. Tian, X. Wang, Y. Liu, W. Hu, Prediction of CO2 and N2 solubility in ionic liquids using a combination of ionic fragments contribution and machine learning methods, J. Mol. Liq. 383 (2023) 122066. https://doi.org/10.1016/j.molliq.2023.122066.

[287] S. Heidari, A. Aleosfoor, M. Mohammadi, F. Esmaeilzadeh, Solubility prediction of CO2 in ionic liquids under subcritical and supercritical carbon dioxide conditions by the adaptive neuro-fuzzy inference system, Case Stud. Chem. Environ. Eng. 7 (2023)

100317. https://doi.org/10.1016/j.cscee.2023.100317.

[288] S.M. Ali, A.J. Hadi, S.Z. Ahmad, Prediction the Experimental Data of CO2 and H2S Solubility in Ionic Liquids Using PR-EoS with Three Different Mixing Rules, Diyala J. Eng. Sci. 8716 (2023) 97–112. https://doi.org/10.24237/djes.2023.16209.

[289] A.H. Jalili, M. Rahmati-Rostami, C. Ghotbi, M. Hosseini-Jenab, A.N. Ahmadi, Solubility of H2S in ionic liquids [bmim][PF6], [bmim][BF4], and [bmim][Tf2N], J. Chem. Eng. Data. 54 (2009) 1844–1849. https://doi.org/10.1021/je8009495.

[290] A.H. Jalili, M. Rahmati-rostami, C. Ghotbi, M. Hosseini-jenab, A.N. Ahmadi, Solubility of H 2 S in Ionic Liquids [bmim][PF 6], [bmim][BF 4], and [bmim][Tf 2 N], J. Chem. Eng. Data. 55 (2010) 5839–5845.

[291] A. Shafiei, M.A. Ahmadi, S.H. Zaheri, A. Baghban, A. Amirfakhrian, R. Soleimani, Estimating hydrogen sulfide solubility in ionic liquids using a machine learning approach, J. Supercrit. Fluids. 95 (2014) 525–534. https://doi.org/10.1016/j.supflu.2014.08.011.

[292] Y. Hiraga, Y. Sato, R.L. Smith, Development of a simple method for predicting CO2 enhancement of H2 gas solubility in ionic liquids, J. Supercrit. Fluids. 96 (2015) 162–170. https://doi.org/10.1016/j.supflu.2014.09.010.

[293] M.A. Ahmadi, R. Haghbakhsh, R. Soleimani, M.B. Bajestani, Estimation of H2S solubility in ionic liquids using a rigorous method, J. Supercrit. Fluids. 92 (2014) 60–69. https://doi.org/10.1016/j.supflu.2014.05.003.

[294] C.A. Faúndez, E.N. Fierro, J.O. Valderrama, Solubility of hydrogen sulfide in ionic liquids for gas removal processes using artificial neural networks, J. Environ. Chem. Eng. 4 (2016) 211–218. https://doi.org/10.1016/j.jece.2015.11.008.

[295] Y. Zhao, J. Gao, Y. Huang, R.M. Afzal, X. Zhang, S. Zhang, Predicting H2S solubility in ionic liquids by the quantitative structure-property relationship method using: S σ-profile molecular descriptors, RSC Adv. 6 (2016) 70405–70413. https://doi.org/10.1039/c6ra15429h.

[296] H.R. Amedi, A. Baghban, M.A. Ahmadi, Evolving machine learning models to predict hydrogen sulfide solubility in the presence of various ionic liquids, J. Mol. Liq. 216 (2016) 411–422. https://doi.org/10.1016/j.molliq.2016.01.060.

[297] A. Kamgar, F. Esmaeilzadeh, Prediction of H2S solubility in [hmim][Pf6], [hmim][Bf4] and [hmim][Tf2N] using UNIQUAC, NRTL and COSMO-RS, J. Mol. Liq. 220 (2016) 631–634. https://doi.org/10.1016/j.molliq.2016.04.110.

[298] Y. Zhao, H. Gao, X. Zhang, Y. Huang, D. Bao, S. Zhang, Hydrogen Sulfide Solubility in Ionic Liquids (ILs): An Extensive Database and a New ELM Model Mainly Established by Imidazolium-Based ILs, J. Chem. Eng. Data. 61 (2016) 3970–3978. https://doi.org/10.1021/acs.jced.6b00449.

[299] A. Shojaeian, Thermodynamic modeling of solubility of hydrogen sulfide in ionic liquids using Peng Robinson-Two State equation of state, J. Mol. Liq. 229 (2017) 591–598. https://doi.org/10.1016/j.molliq.2016.12.001.

[300] X. Kang, J. Qian, J. Deng, U. Latif, Y. Zhao, Novel molecular descriptors for prediction of H2S solubility in ionic liquids, J. Mol. Liq. 265 (2018) 756–764. https://doi.org/10.1016/j.molliq.2018.06.113.

[301] M. Mesbah, E. Soroush, M. Momeni, S. Shahsavari, M. Mofidi, S. Soltanali, Rigorous correlations for predicting the solubility of H2S in methylimidazolium-based ionic liquids, Can. J. Chem. Eng. 98 (2020) 441–452. https://doi.org/10.1002/cjce.23604.

[302] M. Nait Amar, M.A. Ghriga, H. Ouaer, On the evaluation of solubility of hydrogen sulfide in ionic liquids using advanced committee machine intelligent systems, J. Taiwan Inst. Chem. Eng. 118 (2021) 159–168. https://doi.org/10.1016/j.jtice.2021.01.007.

[303] S.P. Mousavi, S. Atashrouz, R. Nakhaei-Kohani, F. Hadavimoghaddam, A. Shawabkeh, A. Hemmati-Sarapardeh, A. Mohaddespour, Modeling of H2S solubility in ionic liquids using deep learning: A chemical structure-based approach, J. Mol. Liq. 351 (2022) 118418. https://doi.org/10.1016/j.molliq.2021.118418.

[304] J. Abdi, M. Hadipoor, S.H. Esmaeili-Faraj, B. Vaferi, A modeling approach for estimating hydrogen sulfide solubility in fifteen different imidazole-based ionic liquids, Sci. Rep. 12 (2022) 1–18. https://doi.org/10.1038/s41598-022-08304-y.

[305] S.P. Mousavi, R. Nakhaei-Kohani, S. Atashrouz, F. Hadavimoghaddam, A. Abedi, A. Hemmati-Sarapardeh, A. Mohaddespour, Modeling of H2S solubility in ionic liquids: comparison of white-box machine learning, deep learning and ensemble learning approaches, Sci. Rep. 13 (2023) 1–23. https://doi.org/10.1038/s41598-023-34193-w.

[306] A.R. Bahmani, F. Sabzi, M. Bahmani, Prediction of solubility of sulfur dioxide in ionic

liquids using artificial neural network, J. Mol. Liq. 211 (2015) 395–400. https://doi.org/10.1016/j.molliq.2015.07.009.

[307] H. Mokarizadeh, S. Atashrouz, H. Mirshekar, A. Hemmati-Sarapardeh, A. Mohaddes Pour, Comparison of LSSVM model results with artificial neural network model for determination of the solubility of SO2 in ionic liquids, J. Mol. Liq. 304 (2020) 112771. https://doi.org/10.1016/j.molliq.2020.112771.

[308] M.R. Mohammadi, F. Hadavimoghaddam, S. Atashrouz, A. Abedi, A. Hemmati-Sarapardeh, A. Mohaddespour, Toward predicting SO2 solubility in ionic liquids utilizing soft computing approaches and equations of state, J. Taiwan Inst. Chem. Eng. 133 (2022) 104220. https://doi.org/10.1016/j.jtice.2022.104220.

[309] D. Camper, P. Scovazzo, C. Koval, R. Noble, Gas solubilities in room-temperature ionic liquids, Ind. Eng. Chem. Res. 43 (2004) 3049–3054. https://doi.org/10.1021/ie034097k.

[310] T. Wang, C. Peng, H. Liu, Y. Hu, Description of the pVT behavior of ionic liquids and the solubility of gases in ionic liquids using an equation of state, Fluid Phase Equilib. 250 (2006) 150–157. https://doi.org/10.1016/j.fluid.2006.10.010.

[311] Y.S. Kim, J.H. Jang, B.D. Lim, J.W. Kang, C.S. Lee, Solubility of mixed gases containing carbon dioxide in ionic liquids: Measurements and predictions, Fluid Phase Equilib. 256 (2007) 70–74. https://doi.org/10.1016/j.fluid.2006.11.019.

[312] T.K. Carlisle, J.E. Bara, C.J. Gabriel, R.D. Noble, D.L. Gin, Interpretation of CO2 solubility and selectivity in nitrile-functionalized room-temperature ionic liquids using a group contribution approach, Ind. Eng. Chem. Res. 47 (2008) 7005–7012. https://doi.org/10.1021/ie8001217.

[313] Á. Martín, D. Méndez, M.D. Bermejo, Application of a group contribution equation of state for the thermodynamic modeling of binary systems (gas + ionic liquids) with bis[(trifluoromethyl)sulfonyl]imide anion, J. Chem. Thermodyn. 42 (2010) 524–529. https://doi.org/10.1016/j.jct.2009.11.008.

[314] N. Ab Manan, C. Hardacre, J. Jacquemin, D.W. Rooney, T.G.A. Youngs, Evaluation of gas solubility prediction in ionic liquids using COSMOthermX, J. Chem. Eng. Data. 54 (2009) 2005–2022. https://doi.org/10.1021/je800857x.

[315] A. Sharma, C. Julcour, A.A. Kelkar, R.M. Deshpande, H. Delmas, Mass transfer and solubility of CO and H2 in ionic liquid. Case of [Bmim][PF6] with gas-inducing stirrer

reactor, Ind. Eng. Chem. Res. 48 (2009) 4075–4082. https://doi.org/10.1021/ie801584p.

[316] A. Yokozeki, M.B. Shiflett, Gas solubilities in ionic liquids using a generic van der Waals equation of state, J. Supercrit. Fluids. 55 (2010) 846–851. https://doi.org/10.1016/j.supflu.2010.09.015.

[317] J. Blath, M. Christ, N. Deubler, T. Hirth, T. Schiestel, Gas solubilities in room temperature ionic liquids - Correlation between RTiL-molar mass and Henry's law constant, Chem. Eng. J. 172 (2011) 167–176. https://doi.org/10.1016/j.cej.2011.05.084.

[318] D. Camper, C. Becker, C. Koval, R. Noble, Low pressure hydrocarbon solubility in room temperature ionic liquids containing imidazolium rings interpreted using regular solution theory, Ind. Eng. Chem. Res. 44 (2005) 1928–1933. https://doi.org/10.1021/ie049312r.

[319] S.S. Moganty, R.E. Baltus, Regular solution theory for low pressure carbon dioxide solubility in room temperature ionic liquids: Ionic liquid solubility parameter from activation energy of viscosity, Ind. Eng. Chem. Res. 49 (2010) 5846–5853. https://doi.org/10.1021/ie901837k.

[320] A.A. Oliferenko, P. V. Oliferenko, K.R. Seddon, J.S. Torrecilla, Prediction of gas solubilities in ionic liquids, Phys. Chem. Chem. Phys. 13 (2011) 17262–17272. https://doi.org/10.1039/c1cp20336c.

[321] R. Shahriari, M.R. Dehghani, B. Behzadi, A modified polar PHSC model for thermodynamic modeling of gas solubility in ionic liquids, Fluid Phase Equilib. 313 (2012) 60–72. https://doi.org/10.1016/j.fluid.2011.09.029.

[322] J.M.M.V. Sousa, J.F.O. Granjo, A.J. Queimada, A.G.M. Ferreira, N.M.C. Oliveira, I.M.A. Fonseca, Solubilities of hydrofluorocarbons in ionic liquids: Experimental and modelling study, J. Chem. Thermodyn. 73 (2014) 36–43. https://doi.org/10.1016/j.jct.2013.07.013.

[323] R. Singh, E. Marin-Rimoldi, E.J. Maginn, A Monte Carlo simulation study to predict the solubility of carbon dioxide, hydrogen, and their mixture in the ionic liquids 1-alkyl-3-methylimidazolium bis(trifluoromethanesulfonyl)amide ([Cnmim+][Tf2N-], n = 4, 6), Ind. Eng. Chem. Res. 54 (2015) 4385–4395. https://doi.org/10.1021/ie503086z.

[324] A. Kamgar, S. Mohsenpour, F. Esmaeilzadeh, Solubility prediction of CO2, CH4, H2, CO and N2 in Choline Chloride/Urea as a eutectic solvent using NRTL and COSMO-RS models, J. Mol. Liq. 247 (2017) 70–74. https://doi.org/10.1016/j.molliq.2017.09.101.

[325] M. Nait Amar, M.A. Ghriga, M.E.A. Ben Seghier, H. Ouaer, Predicting solubility of nitrous oxide in ionic liquids using machine learning techniques and gene expression programming, J. Taiwan Inst. Chem. Eng. 128 (2021) 156–168. https://doi.org/10.1016/j.jtice.2021.08.042.

[326] S. Asensio-Delgado, F. Pardo, G. Zarca, A. Urtiaga, Machine learning for predicting the solubility of high-GWP fluorinated refrigerants in ionic liquids, J. Mol. Liq. 367 (2022) 120472. https://doi.org/10.1016/j.molliq.2022.120472.

[327] Q. Lv, T. Zhou, R. Zheng, R. Nakhaei-Kohani, M. Riazi, A. Hemmati-Sarapardeh, J. Li, W. Wang, Application of group method of data handling and gene expression programming for predicting solubility of CO2-N2 gas mixture in brine, Fuel. 332 (2023) 126025. https://doi.org/10.1016/j.fuel.2022.126025.

[328] R. Nakhaei-Kohani, S. Atashrouz, M. Pourmahdi, F. Hadavimoghaddam, K. Jabbour, A. Hemmati-Sarapardeh, A. Mohaddespour, Hydrogen solubility in ionic liquids: Application of a structure-based deep learning approach and equations of state, Int. J. Hydrogen Energy. 48 (2023) 31234–31253. https://doi.org/10.1016/j.ijhydene.2023.04.018.

[329] A.R. Katritzky, M. Kuanar, I.B. Stoyanova-Slavova, S.H. Slavov, D.A. Dobchev, M. Karelson, W.E. Acree, Quantitative structure-property relationship studies on Ostwald solubility and partition coefficients of organic solutes in ionic liquids, J. Chem. Eng. Data. 53 (2008) 1085–1092. https://doi.org/10.1021/je700607b.

[330] S.F. Pan, G.X. Hu, Y. Lü, J.W. Zou, Q. Sen Yu, QSPR model analysis on the solubility of organic compounds in ionic liquids, Wuli Huaxue Xuebao/ Acta Phys. - Chim. Sin. 26 (2010) 2494–2502. https://doi.org/10.3866/pku.whxb20100902.

[331] S.S. Alsaleem, W.M. Zahid, I.M. Alnashef, M.K. Hadj-Kali, Solubility of Halogenated Hydrocarbons in Hydrophobic Ionic Liquids: Experimental Study and COSMO-RS Prediction, J. Chem. Eng. Data. 60 (2015) 2926–2936. https://doi.org/10.1021/acs.jced.5b00310.

[332] M.B. Shiflett, A. Yokozeki, Phase behavior of carbon dioxide in ionic liquids: [emim][acetate], [emim][trifluoroacetate], and [emim][acetate] + [emim][trifluoroacetate] mixtures, J. Chem. Eng. Data. 54 (2009) 108–114. https://doi.org/10.1021/je800701j.

[333] E. Can, A. Jalal, I.G. Zirhlioglu, A. Uzun, R. Yildirim, Predicting water solubility in ionic liquids using machine learning towards design of hydro-philic/phobic ionic liquids, J. Mol. Liq. 332 (2021) 115848. https://doi.org/10.1016/j.molliq.2021.115848.

[334] A.P. Carneiro, C. Held, O. Rodríguez, G. Sadowski, E.A. Macedo, Solubility of sugars and sugar alcohols in ionic liquids: Measurement and PC-SAFT modeling, J. Phys. Chem. B. 117 (2013) 9980–9995. https://doi.org/10.1021/jp404864c.

[335] K. Paduszyński, M. Okuniewski, U. Domańska, "Sweet-in-Green" systems based on sugars and ionic liquids: New solubility data and thermodynamic analysis, Ind. Eng. Chem. Res. 52 (2013) 18482–18491. https://doi.org/10.1021/ie4033186.

[336] M. Mohan, V. V. Goud, T. Banerjee, Solubility of glucose, xylose, fructose and galactose in ionic liquids: Experimental and theoretical studies using a continuum solvation model, Fluid Phase Equilib. 395 (2015) 33–43. https://doi.org/10.1016/j.fluid.2015.03.020.

[337] A. Bakhtyari, A. Rasoolzadeh, B. Vaferi, A. Khandakar, Application of machine learning techniques to the modeling of solubility of sugar alcohols in ionic liquids, Sci. Rep. 13 (2023) 1–20. https://doi.org/10.1038/s41598-023-39441-7.

[338] J. Kahlen, K. Masuch, K. Leonhard, Modelling cellulose solubilities in ionic liquids using COSMO-RS, Green Chem. 12 (2010) 2172–2181. https://doi.org/10.1039/c0gc00200c.

[339] N.L. Mai, C.K. Kim, B. Park, H.J. Park, S.H. Lee, Y.M. Koo, Prediction of cellulose dissolution in ionic liquids using molecular descriptors based QSAR model, J. Mol. Liq. 215 (2016) 541–548. https://doi.org/10.1016/j.molliq.2016.01.040.

[340] Y. Chu, X. Zhang, M. Hillestad, X. He, Computational prediction of cellulose solubilities in ionic liquids based on COSMO-RS, Fluid Phase Equilib. 475 (2018) 25–36. https://doi.org/10.1016/j.fluid.2018.07.026.

[341] A. Shojaeian, H. Fatoorehchi, Modeling solubility of refrigerants in ionic liquids using Peng Robinson-Two State equation of state, Fluid Phase Equilib. 486 (2019) 80–90. https://doi.org/10.1016/j.fluid.2019.01.003.

[342] Y. Chu, X. He, MoDoop: An Automated Computational Approach for COSMO-RS Prediction of Biopolymer Solubilities in Ionic Liquids, ACS Omega. 4 (2019) 2337–2343. https://doi.org/10.1021/acsomega.8b03255.

[343] M. Lotfi, M. Moniruzzaman, M. Sivapragasam, S. Kandasamy, M.I. Abdul Mutalib, N.B. Alitheen, M. Goto, Solubility of acyclovir in nontoxic and biodegradable ionic liquids: COSMO-RS prediction and experimental verification, J. Mol. Liq. 243 (2017) 124–131. https://doi.org/10.1016/j.molliq.2017.08.020.

[344] Z. Qiao, Z. Wang, C. Zhang, S. Yuan, Y. Zhu, J. Wang, PVAm–PIP/PS composite membrane with high performance for CO_2/N_2 separation, AIChE J. 59 (2012) 215–228. https://doi.org/10.1002/aic.

[345] S.Z. Mahmoudabadi, G. Pazuki, A predictive PC-SAFT EOS based on COSMO for pharmaceutical compounds, Sci. Rep. 11 (2021) 1–18. https://doi.org/10.1038/s41598-021-85942-8.

[346] A. Klamt, Conductor-like screening model for real solvents: A new approach to the quantitative calculation of solvation phenomena, J. Phys. Chem. 99 (1995) 2224–2235. https://doi.org/10.1021/j100007a062.

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

[348] A. Klamt, V. Jonas, T. Bürger, J.C.W. Lohrenz, Refinement and parametrization of COSMO-RS, J. Phys. Chem. A. 102 (1998) 5074–5085. https://doi.org/10.1021/jp980017s.

[349] A. Klamt, F. Eckert, COSMO-RS: a novel and efficient method for the a priori prediction of thermophysical data of liquids, Fluid Phase Equilib. 172 (2000) 43–72. https://doi.org/10.1016/S0378-3812(00)00357-5.

[350] F. Eckert, A. Klamt, Fast Solvent Screening via Quantum Chemistry: COSMO-RS Approach, AIChE J. 48 (2002) 369–385. https://doi.org/10.1002/aic.690480220.

[351] M. Diedenhofen, A. Klamt, COSMO-RS as a tool for property prediction of IL mixtures-A review, Fluid Phase Equilib. 294 (2010) 31–38. https://doi.org/10.1016/j.fluid.2010.02.002.

[352] K.W. Ko, J.H. Her, Toilet house, Space. 482 (2008) 48–53.

[353] D.O. Abranches, Y. Zhang, E.J. Maginn, Y.J. Colón, Sigma profiles in deep learning:

towards a universal molecular descriptor, Chem. Commun. (2022) 5630–5633. https://doi.org/10.1039/d2cc01549h.

[354] C. Si-Moussa, A. Belghait, L. Khaouane, S. Hanini, A. Halilali, Novel density-based model for the correlation of solid drugs solubility in supercritical carbon dioxide, Comptes Rendus Chim. 20 (2017) 559–572. https://doi.org/10.1016/j.crci.2016.09.009.

[355] A. Allouche, Software News and Updates Gabedit — A Graphical User Interface for Computational Chemistry Softwares, J. Comput. Chem. 32 (2012) 174–182. https://doi.org/10.1002/jcc.

[356] M. Yasen, N. Al-Madi, N. Obeid, Optimizing Neural Networks using Dragonfly Algorithm for Medical Prediction, 2018 8th Int. Conf. Comput. Sci. Inf. Technol. CSIT 2018. (2018) 71–76. https://doi.org/10.1109/CSIT.2018.8486178.

[357] H. Harandizadeh, D. Jahed Armaghani, M. Khari, A new development of ANFIS–GMDH optimized by PSO to predict pile bearing capacity based on experimental datasets, Eng. Comput. 37 (2021) 685–700. https://doi.org/10.1007/s00366-019-00849-3.

[358] M. Caudill, Neural networks primer, Mill. Free. Inc. (1993).

[359] R. Matsumura, K. Harada, Y. Domae, W. Wan, Learning based industrial bin-picking trained with approximate physics simulator, Adv. Intell. Syst. Comput. 867 (2019) 786–798. https://doi.org/10.1007/978-3-030-01370-7_61.

[360] D.J. Armaghani, E.T. Mohamad, M.S. Narayanasamy, N. Narita, S. Yagiz, Development of hybrid intelligent models for predicting TBM penetration rate in hard rock condition, Tunn. Undergr. Sp. Technol. 63 (2017) 29–43. https://doi.org/10.1016/j.tust.2016.12.009.

[361] D. González-Arjona, G. López-Pérez, A. Gustavo González, Non-linear QSAR modeling by using multilayer perceptron feedforward neural networks trained by back-propagation, Talanta. 56 (2002) 79–90. https://doi.org/10.1016/S0039-9140(01)00537-9.

[362] E.T. Mohamad, R.S. Faradonbeh, D.J. Armaghani, M. Monjezi, M.Z.A. Majid, An optimized ANN model based on genetic algorithm for predicting ripping production, Neural Comput. Appl. 28 (2017) 393–406. https://doi.org/10.1007/s00521-016-2359-8.

[363] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, C.P. Varghese, R.K. Agrawal, Validation of QSAR Models - Strategies and Importance, Int. J. Drug Des. Disocovery. 2 (2011) 511–519.

[364] K. Roy, R.N. Das, P. Ambure, R.B. Aher, Be aware of error measures. Further studies on validation of predictive QSAR models, Elsevier B.V., 2016. https://doi.org/10.1016/j.chemolab.2016.01.008.

[365] K. Roy, I. Mitra, On Various Metrics Used for Validation of Predictive QSAR Models with Applications in Virtual Screening and Focused Library Design, Comb. Chem. High Throughput Screen. 14 (2011) 450–474. https://doi.org/10.2174/138620711795767893.

[366] R. Soleimani, A.H. Saeedi Dehaghani, N.A. Shoushtari, P. Yaghoubi, A. Bahadori, Toward an intelligent approach for predicting surface tension of binary mixtures containing ionic liquids, Korean J. Chem. Eng. 35 (2018) 1556–1569. https://doi.org/10.1007/s11814-017-0326-4.

[367] O. Falyouna, O. Eljamal, I. Maamoun, A. Tahara, Y. Sugihara, Magnetic zeolite synthesis for efficient removal of cesium in a lab-scale continuous treatment system, J. Colloid Interface Sci. 571 (2020) 66–79. https://doi.org/10.1016/j.jcis.2020.03.028.

[368] F.K. Nievergelt, S. von F. T. Roth, J. Guélat, B. Almasi, P.K. Nievergelt, Bayesian data analysis in ecology using linear models with R,BUGS, and STAN, Acad. Press. (2015) 175–196.

[369] E.M. de Haas, T. Eikelboom, T. Bouwman, Internal and external validation of the long-term QSARs for neutral organics to fish from ECOSAR tm, SAR QSAR Environ. Res. 22 (2011) 545–559. https://doi.org/10.1080/1062936X.2011.569949.

[370] M. Hamadache, S. Hanini, O. Benkortbi, A. Amrane, L. Khaouane, C.S. Moussa, Artificial neural network-based equation to predict the toxicity of herbicides on rats, Chemom. Intell. Lab. Syst. 154 (2016) 7–15. https://doi.org/10.1016/j.chemolab.2016.03.007.

[371] X. Zhao, Y. Pan, J. Jiang, S. Xu, J. Jiang, L. Ding, Thermal Hazard of Ionic Liquids: Modeling Thermal Decomposition Temperatures of Imidazolium Ionic Liquids via QSPR Method, Ind. Eng. Chem. Res. 56 (2017) 4185–4195. https://doi.org/10.1021/acs.iecr.6b04762.

[372] M. Mansourian, L. Saghaie, A. Fassihi, A. Madadkar-Sobhani, K. Mahnam, Linear and

nonlinear QSAR modeling of 1,3,8-substituted-9-deazaxanthines as potential selective A2BAR antagonists, Med. Chem. Res. 22 (2013) 4549–4567. https://doi.org/10.1007/s00044-012-0453-8.

[373] M. Gevrey, I. Dimopoulos, S. Lek, Review and comparison of methods to study the contribution of variables in artificial neural network models, Ecol. Modell. 160 (2003) 249–264. https://doi.org/10.1016/S0304-3800(02)00257-0.

[374] A. Baghban, A. Jalali, M. Shafiee, M.H. Ahmadi, K. wing Chau, Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids, Eng. Appl. Comput. Fluid Mech. 13 (2019) 26–39. https://doi.org/10.1080/19942060.2018.1542345.

[375] J. Leszczynski, Advances in QSAR Modeling, 2017. https://doi.org/10.1007/978-3-319-56850-8.

[376] X. Hu, Support Vector Machine and Its Application to Regression and Classification, Thesis. (2017). https://bearworks.missouristate.edu/theses/3177.

[377] V. Jakkula, Tutorial on Support Vector Machine (SVM), Sch. EECS, Washingt. State Univ. (2011) 1–13. http://www.ccs.neu.edu/course/cs5100f11/resources/jakkula.pdf.

[378] M. Nait Amar, N. Zeraibi, Application of hybrid support vector regression artificial bee colony for prediction of MMP in CO2-EOR process, Petroleum. 6 (2020) 415–422. https://doi.org/10.1016/j.petlm.2018.08.001.

[379] N. Neisani Samani, S.M. Miforughy, H. Safari, O. Mohammadzadeh, M.H. Panahbar, S. Zendehboudi, Solubility of hydrocarbon and non-hydrocarbon gases in aqueous electrolyte solutions: A reliable computational strategy, Fuel. 241 (2019) 1026–1035. https://doi.org/10.1016/j.fuel.2018.11.150.

[380] C. Xu, M. Nait Amar, M.A. Ghriga, H. Ouaer, X. Zhang, M. Hasanipanah, Evolving support vector regression using Grey Wolf optimization; forecasting the geomechanical properties of rock, Eng. Comput. (2020). https://doi.org/10.1007/s00366-020-01131-7.

[381] E. Pourbasheer, S. Riahi, M.R. Ganjali, P. Norouzi, Prediction of solubility of fullerene C60 in various organic solvents by genetic algorithm-multiple linear regression, Fullerenes Nanotub. Carbon Nanostructures. 19 (2011) 585–598. https://doi.org/10.1080/1536383X.2010.504952.

[382] V.N. Vapnik, The nature of statistical learning, Theory. 2 (1995).

[383] W. De Cheng, C.Z. Cai, Accurate model to predict the solubility of fullerene C60 in organic solvents by using support vector regression, Fullerenes Nanotub. Carbon Nanostructures. 25 (2017) 58–64. https://doi.org/10.1080/1536383X.2016.1252336.

[384] A. Baghban, A.H. Mohammadi, M.S. Taleghani, Rigorous modeling of CO2 equilibrium absorption in ionic liquids, Int. J. Greenh. Gas Control. 58 (2017) 19–41. https://doi.org/10.1016/j.ijggc.2016.12.009.

[385] N.H. Farhat, Photonit neural networks and learning mathines the role of electron-trapping materials, IEEE Expert. Syst. Their Appl. 7 (1992) 63–72. https://doi.org/10.1109/64.163674.

[386] M. Amroune, T. Bouktir, I. Musirin, Power System Voltage Stability Assessment Using a Hybrid Approach Combining Dragonfly Optimization Algorithm and Support Vector Regression, Arab. J. Sci. Eng. 43 (2018) 3023–3036. https://doi.org/10.1007/s13369-017-3046-5.

[387] S. Mirjalili, Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems, Neural Comput. Appl. 27 (2016) 1053–1073. https://doi.org/10.1007/s00521-015-1920-1.

[388] M.A. Salam, H.M. Zawbaa, E. Emary, K.K.A. Ghany, B. Parv, A hybrid dragonfly algorithm with extreme learning machine for prediction, Proc. 2016 Int. Symp. Innov. Intell. Syst. Appl. INISTA 2016. (2016). https://doi.org/10.1109/INISTA.2016.7571839.

[389] C.W. Reynolds, Flocks, herds, and schools: A distributed behavioral model, Proc. 14th Annu. Conf. Comput. Graph. Interact. Tech. SIGGRAPH 1987. 21 (1987) 25–34. https://doi.org/10.1145/37401.37406.

[390] S. Mirjalili, The ant lion optimizer, Adv. Eng. Softw. 83 (2015) 80–98. https://doi.org/10.1016/j.advengsoft.2015.01.010.

[391] S. Saha, V. Mukherjee, A novel quasi-oppositional chaotic antlion optimizer for global optimization, Appl. Intell. 48 (2018) 2628–2660. https://doi.org/10.1007/s10489-017-1097-7.

[392] E. Gupta, A. Saxena, Performance Evaluation of Antlion Optimizer Based Regulator in Automatic Generation Control of Interconnected Power System, J. Eng. (United

Kingdom). 2016 (2016). https://doi.org/10.1155/2016/4570617.

[393] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey Wolf Optimizer, Adv. Eng. Softw. 69 (2014) 46–61. https://doi.org/10.1016/j.advengsoft.2013.12.007.

[394] M.H. Nadimi-Shahraki, S. Taghian, S. Mirjalili, An improved grey wolf optimizer for solving engineering problems, Expert Syst. Appl. 166 (2021) 113917. https://doi.org/10.1016/j.eswa.2020.113917.

[395] E. Emary, W. Yamany, A.E. Hassanien, V. Snasel, Multi-Objective Gray-Wolf Optimization for Attribute Reduction, Procedia Comput. Sci. 65 (2015) 623–632. https://doi.org/10.1016/j.procs.2015.09.006.

[396] A. Kumar, S. Pant, M. Ram, System Reliability Optimization Using Gray Wolf Optimizer Algorithm, Qual. Reliab. Eng. Int. 33 (2017) 1327–1335. https://doi.org/10.1002/qre.2107.

[397] D. Karaboga, B. Basturk, On the performance of artificial bee colony (ABC) algorithm, Appl. Soft Comput. J. 8 (2008) 687–697. https://doi.org/10.1016/j.asoc.2007.05.007.

[398] D. Karaboga, An idea based on honey bee swarm for numerical optimization, J. Mater. Process. Technol. 1 (2005) 1–8. http://dx.doi.org/10.1016/j.cirp.2016.06.001%0Ahttp://dx.doi.org/10.1016/j.powtec.2016.12.055%0Ahttps://doi.org/10.1016/j.ijfatigue.2019.02.006%0Ahttps://doi.org/10.1016/j.matlet.2019.04.024%0Ahttps://doi.org/10.1016/j.matlet.2019.127252%0Ahttp://dx.doi.o.

[399] P.W. Tsai, J.S. Pan, B.Y. Liao, S.C. Chu, Enhanced artificial bee colony optimization, Int. J. Innov. Comput. Inf. Control. 5 (2009) 5081–5092. https://doi.org/10.1002/elps.200900194.

[400] M.O. Okwu, L.K. Tartibu, Particle Swarm Optimisation, Stud. Comput. Intell. 927 (2021) 5–13. https://doi.org/10.1007/978-3-030-61111-8_2.

[401] R. Eberhart, J. Kennedy, New optimizer using particle swarm theory, Proc. Int. Symp. Micro Mach. Hum. Sci. (1995) 39–43. https://doi.org/10.1109/mhs.1995.494215.

[402] H. Garg, A hybrid PSO-GA algorithm for constrained optimization problems, Appl. Math. Comput. 274 (2016) 292–305. https://doi.org/10.1016/j.amc.2015.11.001.

[403] F.A. Şenel, F. Gökçe, A.S. Yüksel, T. Yiğit, A novel hybrid PSO–GWO algorithm for

optimization problems, Eng. Comput. 35 (2019) 1359–1373. https://doi.org/10.1007/s00366-018-0668-5.

[404] S. Mirjalili, A. Lewis, The Whale Optimization Algorithm, Adv. Eng. Softw. 95 (2016) 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008.

[405] A.M. Hemeida, S. Alkhalaf, A. Mady, E.A. Mahmoud, M.E. Hussein, A.M. Baha Eldin, Implementation of nature-inspired optimization algorithms in some data mining tasks, Ain Shams Eng. J. 11 (2020) 309–318. https://doi.org/10.1016/j.asej.2019.10.003.

[406] J. Nasiri, F.M. Khiyabani, A whale optimization algorithm (WOA) approach for clustering, Cogent Math. Stat. 5 (2018) 1483565. https://doi.org/10.1080/25742558.2018.1483565.

[407] G. Kaur, S. Arora, Chaotic whale optimization algorithm, J. Comput. Des. Eng. 5 (2018) 275–284. https://doi.org/10.1016/j.jcde.2017.12.006.

[408] N. Singh, S.B. Singh, Hybrid Algorithm of Particle Swarm Optimization and Grey Wolf Optimizer for Improving Convergence Performance, J. Appl. Math. 2017 (2017). https://doi.org/10.1155/2017/2030489.

[409] D. Kuonen, Book Review: Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis, 2004. https://doi.org/10.1177/096228020401300512.

[410] V. Consonni, D. Ballabio, R. Todeschini, Comments on the definition of the Q2 parameter for QSAR validation, J. Chem. Inf. Model. 49 (2009) 1669–1678. https://doi.org/10.1021/ci900115y.

[411] R. Todeschini, D. Ballabio, F. Grisoni, Beware of Unreliable Q2! A Comparative Study of Regression Metrics for Predictivity Assessment of QSAR Models, J. Chem. Inf. Model. 56 (2016) 1905–1913. https://doi.org/10.1021/acs.jcim.6b00277.

[412] N. Bouarra, S. Kherouf, A. Bouakkadia, D. Messadi, QSPR Application on Modeling of Boiling Point of Polycyclic Aromatic Hydrocarbons, Res. J. Pharm. Biol. Chem. Sci. 8 (2017) 19–28.

[413] M.B. Shiflett, E.J. Maginn, The solubility of gases in ionic liquids, AIChE J. 63 (2017) 4722–4737. https://doi.org/10.1002/aic.15957.

[414] A. Klamt, COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and

Drug Design, Elsevier, 2016.

[415] K. Paduszyński, An overview of the performance of the COSMO-RS approach in predicting the activity coefficients of molecular solutes in ionic liquids and derived properties at infinite dilution, Phys. Chem. Chem. Phys. 19 (2017) 11835–11850. https://doi.org/10.1039/c7cp00226b.

[416] E. Mullins, R. Oldland, Y.A. Liu, S. Wang, S.I. Sandler, C.C. Chen, M. Zwolak, K.C. Seavey, Sigma-profile database for using COSMO-based thermodynamic methods, Ind. Eng. Chem. Res. 45 (2006) 4389–4415. https://doi.org/10.1021/ie060370h.

[417] Book Reviews, Auto / Biogr. 14 (2006) 381–400. https://www.proquest.com/scholarly-journals/book-reviews/docview/217169337/se-2?accountid=8630%0Ahttps://birmingham-primo.hosted.exlibrisgroup.com/openurl/44BIR/44BIR_Services?genre=article&issn=09675507&title=Book+Reviews&volume=14&issue=4&date=2006&atit.

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

[419] V. Vladimir, A. Chervonenkis, Theory of pattern recognition, (1974).

[420] Q. Quan, Z. Hao, H. Xifeng, L. Jingchun, Research on water temperature prediction based on improved support vector regression, Neural Comput. Appl. 34 (2022) 8501–8510. https://doi.org/10.1007/s00521-020-04836-4.

[421] J.G. Lugo-Armenta, L.R. Pino-Fan, Inferential Reasoning Levels for t-Student Statistical, Bolema - Math. Educ. Bull. 35 (2021) 1776–1802. https://doi.org/10.1590/1980-4415V35N71A25.

[422] O. Harel, The estimation of R2 and adjusted R2 in incomplete data sets using multiple imputation, J. Appl. Stat. 36 (2009) 1109–1118. https://doi.org/10.1080/02664760802553000.

[423] R. Veerasamy, H. Rajak, A. Jain, S. Sivadasan, Validation of QSAR Models - Strategies and Importance Validation of QSAR Models - Strategies and Importance, (2011).

[424] F. Abascal, R. Zardoya, D. Posada, ProtTest: Selection of best-fit models of protein

evolution, Bioinformatics. 21 (2005) 2104–2105. https://doi.org/10.1093/bioinformatics/bti263.

[425] I. Euldji, A. Belghait, A. Amrane, A new hybrid quantitative structure property relationships-support vector regression (QSPR-SVR) approach for predicting the solubility of drug compounds in supercritical carbon dioxide, (2023) 1–17. https://doi.org/10.1002/aic.18115.

[426] S. Mirjalili, S.M. Mirjalili, A. Hatamlou, Multi-Verse Optimizer: a nature-inspired algorithm for global optimization, Neural Comput. Appl. 27 (2016) 495–513. https://doi.org/10.1007/s00521-015-1870-7.

[427] S. Mirjalili, Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm, Knowledge-Based Syst. 89 (2015) 228–249. https://doi.org/10.1016/j.knosys.2015.07.006.

[428] S. Mirjalili, SCA: A Sine Cosine Algorithm for solving optimization problems, Knowledge-Based Syst. 96 (2016) 120–133. https://doi.org/10.1016/j.knosys.2015.12.022.

[429] C. Si-Moussa, S. Hanini, R. Derriche, M. Bouhedda, A. Bouzidi, Prediction of high-pressure vapor liquid equilibrium of six binary systems, carbon dioxide with six esters, using an artificial neural network model, Brazilian J. Chem. Eng. 25 (2008) 183–199. https://doi.org/10.1590/S0104-66322008000100019.

[430] S.W. Lin, K.C. Ying, S.C. Chen, Z.J. Lee, Particle swarm optimization for parameter determination and feature selection of support vector machines, Expert Syst. Appl. 35 (2008) 1817–1824. https://doi.org/10.1016/j.eswa.2007.08.088.

[431] C.L. Huang, C.J. Wang, A GA-based feature selection and parameters optimizationfor support vector machines, Expert Syst. Appl. 31 (2006) 231–240. https://doi.org/10.1016/j.eswa.2005.09.024.

[432] Y.M. Chen, Y.P. Chen, Measurements for the solid solubilities of antipyrine, 4-aminoantipyrine and 4-dimethylaminoantipyrine in supercritical carbon dioxide, Fluid Phase Equilib. 282 (2009) 82–87. https://doi.org/10.1016/j.fluid.2009.04.019.

[433] J.S. Jin, X. Fan, H. Zhang, Y.W. Wang, Z.T. Zhang, Solubility of 4-aminosalicylic acid in supercritical carbon dioxide and subcritical 1,1,1,2-tetrafluoroethane, J. Chem. Eng. Data. 59 (2014) 2095–2100. https://doi.org/10.1021/je5002736.

[434] D. Suleiman, L. Antonio Estévez, J.C. Pulido, J.E. García, C. Mojica, Solubility of anti-inflammatory, anti-cancer, and anti-HIV drugs in supercritical carbon dioxide, J. Chem. Eng. Data. 50 (2005) 1234–1241. https://doi.org/10.1021/je049551l.

[435] N. Lamba, R.C. Narayan, J. Modak, G. Madras, Solubilities of 10-undecenoic acid and geraniol in supercritical carbon dioxide, J. Supercrit. Fluids. 107 (2016) 384–391. https://doi.org/10.1016/j.supflu.2015.09.026.

[436] A.B.S. Rosa, S. Marceneiro, M.E.M. Braga, A.M.A. Dias, H.C. De Sousa, Solubility of all-trans retinoic acid in supercritical carbon dioxide, J. Supercrit. Fluids. 98 (2015) 70–78. https://doi.org/10.1016/j.supflu.2014.12.027.

[437] G. Sodeifian, S.A. Sajadian, F. Razmimanesh, Solubility of an antiarrhythmic drug (amiodarone hydrochloride) in supercritical carbon dioxide: Experimental and modeling, Fluid Phase Equilib. 450 (2017) 149–159. https://doi.org/10.1016/j.fluid.2017.07.015.

[438] G. Sodei, C. Garlapati, F. Razmimanesh, F. Sodei, Solubility of Amlodipine Besylate (Calcium Channel Blocker Drug) in Supercritical Carbon Dioxide : Measurement and Correlations, (2021). https://doi.org/10.1021/acs.jced.0c00913.

[439] M. Hojjati, A. Vatanara, Y. Yamini, M. Moradi, A.R. Najafabadi, Supercritical CO2 and highly selective aromatase inhibitors: Experimental solubility and empirical data correlation, J. Supercrit. Fluids. 50 (2009) 203–209. https://doi.org/10.1016/j.supflu.2009.06.015.

[440] G. Sodeifian, S.A. Sajadian, N.S. Ardestani, Determination of solubility of Aprepitant (an antiemetic drug for chemotherapy) in supercritical carbon dioxide: Empirical and thermodynamic models, J. Supercrit. Fluids. 128 (2017) 102–111. https://doi.org/10.1016/j.supflu.2017.05.019.

[441] Z. Huang, W.D. Lu, S. Kawi, Y.C. Chiew, Solubility of aspirin in supercritical carbon dioxide with and without Acetone, J. Chem. Eng. Data. 49 (2004) 1323–1327. https://doi.org/10.1021/je0499465.

[442] A.R. Karami, Y. Yamini, A.R. Ghiasvand, H. Sharghi, M. Shamsipur, Solubilities of some 9-anthrone derivatives in supercritical carbon dioxide, J. Chem. Eng. Data. 46 (2001) 1371–1374. https://doi.org/10.1021/je0101045.

[443] G. Sodeifian, F. Razmimanesh, N. Saadati Ardestani, S.A. Sajadian, Experimental data and thermodynamic modeling of solubility of Azathioprine, as an immunosuppressive

and anti-cancer drug, in supercritical carbon dioxide, 2020. https://doi.org/10.1016/j.molliq.2019.112179.

[444] M. Sahihi, H.S. Ghaziaskar, M. Hajebrahimi, Solubility of maleic acid in supercritical carbon dioxide, J. Chem. Eng. Data. 55 (2010) 2596–2599. https://doi.org/10.1021/je9008393.

[445] A. Garmroodi, J. Hassan, Y. Yamini, Solubilities of the drugs benzocaine, metronidazole benzoate, and naproxen in supercritical carbon dioxide, J. Chem. Eng. Data. 49 (2004) 709–712. https://doi.org/10.1021/je020218w.

[446] M. Asghari-Khiavi, Y. Yamini, Solubility of the drugs bisacodyl, methimazole, methylparaben, and iodoquinol in supercritical carbon dioxide, J. Chem. Eng. Data. 48 (2003) 61–65. https://doi.org/10.1021/je020080h.

[447] A. Vatanara, A.R. Najafabadi, M. Khajeh, Y. Yamini, Solubility of some inhaled glucocorticoids in supercritical carbon dioxide, J. Supercrit. Fluids. 33 (2005) 21–25. https://doi.org/10.1016/j.supflu.2004.05.002.

[448] M. Johannsen, G. Brunner, Solubilities of the xanthines caffeine, theophylline and theobromine in supercritical carbon dioxide, Fluid Phase Equilib. 95 (1994) 215–226. https://doi.org/10.1016/0378-3812(94)80070-7.

[449] Y. Yamini, M. Hojjati, P. Kalantarian, M. Moradi, A. Esrafili, A. Vatanara, Solubility of capecitabine and docetaxel in supercritical carbon dioxide: Data and the best correlation, Thermochim. Acta. 549 (2012) 95–101. https://doi.org/10.1016/j.tca.2012.09.010.

[450] K. Ongkasin, M. Sauceau, Y. Masmoudi, J. Fages, E. Badens, Solubility of cefuroxime axetil in supercritical CO2: Measurement and modeling, J. Supercrit. Fluids. 152 (2019) 104498. https://doi.org/10.1016/j.supflu.2019.03.010.

[451] C.C. Tsai, H.M. Lin, M.J. Lee, Solubility of niflumic acid and celecoxib in supercritical carbon dioxide, J. Supercrit. Fluids. 95 (2014) 17–23. https://doi.org/10.1016/j.supflu.2014.07.026.

[452] A.Z. Hezave, A. Mowla, F. Esmaeilzadeh, Cetirizine solubility in supercritical CO2 at different pressures and temperatures, J. Supercrit. Fluids. 58 (2011) 198–203. https://doi.org/10.1016/j.supflu.2011.05.017.

[453] B. Li, W. Guo, E.D. Ramsey, Solubility measurements of chloramphenicol in

supercritical fluid co2 using static solubility apparatus interfaced with online supercritical fluid chromatography, J. Chem. Eng. Data. 65 (2020) 153–159. https://doi.org/10.1021/acs.jced.9b00831.

[454] C.A. Lee, M. Tang, S.L. Ho, Y.P. Chen, Solubilities of chlormezanone, metaxalone and methocarbamol in supercritical carbon dioxide, J. Supercrit. Fluids. 85 (2014) 11–16. https://doi.org/10.1016/j.supflu.2013.10.006.

[455] M. Lashkarbolooki, A.Z. Hezave, Y. Rahnama, H. Rajaei, F. Esmaeilzadeh, Solubility of chlorpheniramine maleate in supercritical carbon dioxide, J. Supercrit. Fluids. 84 (2013) 29–35. https://doi.org/10.1016/j.supflu.2013.09.007.

[456] L. Manna, M. Banchero, Solubility of Tolbutamide and Chlorpropamide in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 63 (2018) 1745–1751. https://doi.org/10.1021/acs.jced.8b00050.

[457] J.T. Paula, I.M.O. Sousa, M.A. Foglio, F.A. Cabral, Solubility of protocatechuic acid, sinapic acid and chrysin in supercritical carbon dioxide, J. Supercrit. Fluids. 112 (2016) 89–94. https://doi.org/10.1016/j.supflu.2016.02.014.

[458] C. Zhang, D. Jia, Y. Sun, L. Zhang, Measurement and Correlation of the Solubility of Cinildipine in Supercritical CO 2, IOP Conf. Ser. Earth Environ. Sci. 208 (2018). https://doi.org/10.1088/1755-1315/208/1/012041.

[459] C.A. Lee, M. Tang, Y.P. Chen, Measurement and correlation for the solubilities of cinnarizine, pentoxifylline, and piracetam in supercritical carbon dioxide, Fluid Phase Equilib. 367 (2014) 182–187. https://doi.org/10.1016/j.fluid.2014.01.041.

[460] K. Shi, L. Feng, L. He, H. Li, Solubility Determination and Correlation of Gatifloxacin, Enrofloxacin, and Ciprofloxacin in Supercritical CO2, J. Chem. Eng. Data. 62 (2017) 4235–4243. https://doi.org/10.1021/acs.jced.7b00601.

[461] G. Sodeifian, C. Garlapati, F. Razmimanesh, Measurement and modeling of clemastine fumarate (antihistamine drug) solubility in supercritical carbon dioxide, Sci. Rep. (2021) 1–16. https://doi.org/10.1038/s41598-021-03596-y.

[462] C.I. Park, M.S. Shin, H. Kim, Solubility of climbazole and triclocarban in supercritical carbon dioxide: Measurement and correlation, J. Chem. Thermodyn. 41 (2009) 30–34. https://doi.org/10.1016/j.jct.2008.08.009.

[463] M.H. Hosseini, N. Alizadeh, A.R. Khanchi, Solubility analysis of clozapine and lamotrigine in supercritical carbon dioxide using static system, J. Supercrit. Fluids. 52 (2010) 30–35. https://doi.org/10.1016/j.supflu.2009.11.006.

[464] J.R. Dean, M. Kane, S. Khundker, C. Dowle, R.L. Tranter, P. Jones, Estimation and determination of steroid solubility in supercritical carbon dioxide, Analyst. 120 (1995) 2153–2157. https://doi.org/10.1039/an9952002153.

[465] G. Sodeifian, N.S. Ardestani, S.A. Sajadian, H.S. Panah, Experimental measurements and thermodynamic modeling of Coumarin-7 solid solubility in supercritical carbon dioxide: Production of nanoparticles via RESS method, Fluid Phase Equilib. (2018). https://doi.org/10.1016/j.fluid.2018.11.006.

[466] M. Lashkarbolooki, A.Z. Hezave, Y. Rahnama, R. Ozlati, H. Rajaei, F. Esmaeilzadeh, Solubility of cyproheptadine in supercritical carbon dioxide; Experimental and modeling approaches, J. Supercrit. Fluids. 84 (2013) 13–19. https://doi.org/10.1016/j.supflu.2013.09.004.

[467] M. Asghari-Khiavi, Y. Yamini, M.A. Farajzadeh, Solubilities of two steroid drugs and their mixtures in supercritical carbon dioxide, J. Supercrit. Fluids. 30 (2004) 111–117. https://doi.org/10.1016/j.supflu.2003.07.002.

[468] G. Sodeifian, R. Surya, F. Razmimanesh, M. Abadian, Solubility of Dasatinib monohydrate (anticancer drug) in supercritical CO 2 : Experimental and thermodynamic modeling, J. Mol. Liq. 346 (2022) 117899. https://doi.org/10.1016/j.molliq.2021.117899.

[469] C. Tang, Y.X. Guan, S.J. Yao, Z.Q. Zhu, Solubility of dexamethasone in supercritical carbon dioxide with and without a cosolvent, J. Chem. Eng. Data. 59 (2014) 3359–3364. https://doi.org/10.1021/je500345n.

[470] R.B. Chim, M.B.C. De Matos, M.E.M. Braga, A.M.A. Dias, H.C. De Sousa, Solubility of dexamethasone in supercritical carbon dioxide, J. Chem. Eng. Data. 57 (2012) 3756–3760. https://doi.org/10.1021/je301065f.

[471] A. Zeinolabedini Hezave, F. Esmaeilzadeh, Solubility measurement of diclofenac acid in the supercritical CO 2, J. Chem. Eng. Data. 57 (2012) 1659–1664. https://doi.org/10.1021/je200012x.

[472] and H.C. de S. Patrı´cia Coimbra, Daniel Fernandes, Maria H. Gil, Solubility of

Diflunisal in supercritical carbon dioxide, J. Chem. Eng. Data. (2008) 1990–1995.

[473] Y. Yamini, P. Kalantarian, M. Hojjati, A. Esrafily, M. Moradi, A. Vatanara, I. Harrian, Solubilities of flutamide, dutasteride, and finasteride as antiandrogenic agents, in supercritical carbon dioxide: Measurement and correlation, J. Chem. Eng. Data. 55 (2010) 1056–1059. https://doi.org/10.1021/je900520a.

[474] G. Sodeifian, R. Detakhsheshpour, S.A. Sajadian, Experimental study and thermodynamic modeling of Esomeprazole (proton-pump inhibitor drug for stomach acid reduction) solubility in supercritical carbon dioxide, J. Supercrit. Fluids. 154 (2019) 104606. https://doi.org/10.1016/j.supflu.2019.104606.

[475] X. Zha, S. Han, W. Wang, Z. Jiao, Experimental measurement and correlation of solubility of ethosuximide in supercritical carbon dioxide, J. Chem. Thermodyn. 131 (2019) 104–110. https://doi.org/10.1016/j.jct.2018.10.032.

[476] A. Zeinolabedini Hezave, H. Rajaei, M. Lashkarbolooki, F. Esmaeilzadeh, Analyzing the solubility of fluoxetine hydrochloride in supercritical carbon dioxide, J. Supercrit. Fluids. 73 (2013) 57–62. https://doi.org/10.1016/j.supflu.2012.11.005.

[477] A.R.C. Duarte, P. Coimbra, H.C. De Sousa, C.M.M. Duarte, Solubility of flurbiprofen in supercritical carbon dioxide, J. Chem. Eng. Data. 49 (2004) 449–452. https://doi.org/10.1021/je034099b.

[478] S.A. Shojaee, A.Z. Hezave, S. Aftab, M. Lashkarbolooki, F. Esmaeilzadeh, Solubility of gabapentin in supercritical carbon dioxide, J. Supercrit. Fluids. 78 (2013) 1–6. https://doi.org/10.1016/j.supflu.2013.02.003.

[479] G. Sodeifian, S.M. Hazaveie, F. Sodeifian, Determination of Galantamine solubility (an anti-alzheimer drug) in supercritical carbon dioxide (CO2): Experimental correlation and thermodynamic modeling, J. Mol. Liq. 330 (2021). https://doi.org/10.1016/j.molliq.2021.115695.

[480] S.T. Xiang, B.Q. Chen, R.K. Kankala, S. Bin Wang, A.Z. Chen, Solubility measurement and RESOLV-assisted nanonization of gambogic acid in supercritical carbon dioxide for cancer therapy, J. Supercrit. Fluids. 150 (2019) 147–155. https://doi.org/10.1016/j.supflu.2019.04.008.

[481] Y.M. Chen, P.C. Lin, M. Tang, Y.P. Chen, Solid solubility of antilipemic agents and micronization of gemfibrozil in supercritical carbon dioxide, J. Supercrit. Fluids. 52

(2010) 175–182. https://doi.org/10.1016/j.supflu.2009.12.012.

[482] S.A. Sodeifian, Gholamhoussein, Razmimanesh, Fariba, Sajadian, Solubility measurement of a chemotherapeutic agent (Imatinib mesylate) in supercritical carbon dioxide: Assessment of new empirical model, J. Supercrit. Fluids. 146 (2019) 89–99. https://doi.org/10.1016/j.supflu.2019.01.006.

[483] E. Jara-Morante, D. Suleiman, L.A. Estévez, Solubilities of imipramine HCl in supercritical carbon dioxide, Ind. Eng. Chem. Res. 42 (2003) 1821–1823. https://doi.org/10.1021/ie0109105.

[484] B.C. Wang, C.S. Su, Solid solubility measurement of ipriflavone in supercritical carbon dioxide and microparticle production through the rapid expansion of supercritical solutions process, J. CO2 Util. 37 (2020) 285–294. https://doi.org/10.1016/j.jcou.2019.12.012.

[485] R. Heryanto, E.C. Abdullah, M. Hasan, Solubility of isoniazid in supercritical carbon dioxide, J. Chem. Eng. Data. 55 (2010) 2306–2309. https://doi.org/10.1021/je900622h.

[486] G. Sodeifian, S.A. Sajadian, F. Razmimanesh, S.M. Hazaveie, Solubility of Ketoconazole (antifungal drug) in SC-CO2 for binary and ternary systems: measurements and empirical correlations, Sci. Rep. 11 (2021) 1–13. https://doi.org/10.1038/s41598-021-87243-6.

[487] M.A. Sabegh, H. Rajaei, F. Esmaeilzadeh, M. Lashkarbolooki, Solubility of ketoprofen in supercritical carbon dioxide, J. Supercrit. Fluids. 72 (2012) 191–197. https://doi.org/10.1016/j.supflu.2012.08.008.

[488] G. Sodeifian, N. Saadati Ardestani, S.A. Sajadian, H.S. Panah, Measurement, correlation and thermodynamic modeling of the solubility of Ketotifen fumarate (KTF) in supercritical carbon dioxide: Evaluation of PCP-SAFT equation of state, Fluid Phase Equilib. 458 (2018) 102–114. https://doi.org/10.1016/j.fluid.2017.11.016.

[489] Y. Wang, T. Meng, D. Jia, Y. Sun, N. Li, Solubility of Nifedipine and Lacidipine in Supercritical CO2: Measurement and Correlation, J. Solution Chem. 46 (2017) 70–88. https://doi.org/10.1007/s10953-016-0550-2.

[490] G. Sodeifian, S.A. Sajadian, R. Derakhsheshpour, Experimental measurement and thermodynamic modeling of Lansoprazole solubility in supercritical carbon dioxide: Application of SAFT-VR EoS, Fluid Phase Equilib. 507 (2020) 112422.

https://doi.org/10.1016/j.fluid.2019.112422.

[491] Y. Yamini, M. Tayyebi, M. Moradi, A. Vatanara, Solubility of megestrol acetate and levonorgestrel in supercritical carbon dioxide, Thermochim. Acta. 569 (2013) 48–54. https://doi.org/10.1016/j.tca.2013.07.018.

[492] G. Sodeifian, L. Nasri, F. Razmimanesh, M. Abadian, Measuring and modeling the solubility of an antihypertensive drug (losartan potassium, Cozaar) in supercritical carbon dioxide, J. Mol. Liq. 331 (2021). https://doi.org/10.1016/j.molliq.2021.115745.

[493] K.W. Cheng, M. Tang, Y.P. Chen, Solubilities of benzoin, propyl 4-hydroxybenzoate and mandelic acid in supercritical carbon dioxide, Fluid Phase Equilib. 201 (2002) 79–96. https://doi.org/10.1016/S0378-3812(02)00070-5.

[494] A. Zeinolabedini Hezave, M.H. Khademi, F. Esmaeilzadeh, Measurement and modeling of mefenamic acid solubility in supercritical carbon dioxide, Fluid Phase Equilib. 313 (2012) 140–147. https://doi.org/10.1016/j.fluid.2011.09.031.

[495] H.C. De Sousa, M.S. Costa, P. Coimbra, A.A. Matias, C.M.M. Duarte, Experimental determination and correlation of meloxicam sodium salt solubility in supercritical carbon dioxide, J. Supercrit. Fluids. 63 (2012) 40–45. https://doi.org/10.1016/j.supflu.2011.12.004.

[496] Y. Yamini, J. Arab, M. Asghari-khiavi, Solubilities of phenazopyridine, propranolol, and methimazole in supercritical carbon dioxide, J. Pharm. Biomed. Anal. 32 (2003) 181–187. https://doi.org/10.1016/S0731-7085(03)00016-5.

[497] G. Sodeifian, N.S. Ardestani, F. Razmimanesh, S.A. Sajadian, Experimental and thermodynamic analyses of supercritical CO2-Solubility of minoxidil as an antihypertensive drug, Fluid Phase Equilib. (2020) 112745. https://doi.org/10.1016/j.fluid.2020.112745.

[498] P. Kotnik, M. Škerget, Ž. Knez, Solubility of nicotinic acid and nicotinamide in carbon dioxide at T = (313.15 to 373.15) K and p = (5 to 30) MPa: Experimental data and correlation, J. Chem. Eng. Data. 56 (2011) 338–343. https://doi.org/10.1021/je100697a.

[499] I. Medina, J.L. Bueno, Solubilities of zopiclone and nimodipine in supercritical carbon dioxide, J. Chem. Eng. Data. 46 (2001) 1211–1214. https://doi.org/10.1021/je010025t.

[500] S. Zhan, L. Cui, Q. Zhao, L. Chen, J. Wang, S. Chen, S. Ding, Measurement and

correlation of solubility of nitrendipine in supercritical carbon dioxide with and without ethanol cosolvent, J. Solution Chem. 44 (2015) 1–15. https://doi.org/10.1007/s10953-014-0262-4.

[501] R. Chim, S. Marceneiro, M.E.M. Braga, A.M.A. Dias, H.C. de Sousa, Solubility of norfloxacin and ofloxacin in supercritical carbon dioxide, Fluid Phase Equilib. 331 (2012) 6–11. https://doi.org/10.1016/j.fluid.2012.06.023.

[502] Z. Tang, J. Jin, Z. Zhang, H. Liu, 24 Ind. Eng. Chem. Res. 2012, 51, 5515−5526 ref (16).pdf, (2012).

[503] G. Sodeifian, S.M. Hazaveie, S.A. Sajadian, F. Razmimanesh, Experimental investigation and modeling of the solubility of oxcarbazepine (an anticonvulsant agent)in supercritical carbon dioxide, Fluid Phase Equilib. 493 (2019) 160–173. https://doi.org/10.1016/j.fluid.2019.04.013.

[504] Z. Jiao, J. Cheng, M. Guo, S. Han, Measurement and Correlation of Paeonol Solubility in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 64 (2019) 4424–4429. https://doi.org/10.1021/acs.jced.9b00483.

[505] M. Ko, V. Shah, P.R. Bienkowski, H.D. Cochran, Solubility of the antibiotic penicillin V in supercritical CO2, J. Supercrit. Fluids. 4 (1991) 32–39. https://doi.org/10.1016/0896-8446(91)90028-5.

[506] H. Rajaei, A.Z. Hezave, M. Lashkarbolooki, F. Esmaeilzadeh, Representing experimental solubility of phenylephrine hydrochloride in supercritical carbon dioxide and modeling solute solubility using semi-empirical correlations, J. Supercrit. Fluids. 75 (2013) 181–186. https://doi.org/10.1016/j.supflu.2012.11.014.

[507] G. Sodeifian, R. Surya, F. Razmimanesh, Fluid Phase Equilibria Solubility of Pholcodine (antitussive drug) in supercritical carbon dioxide : Experimental data and thermodynamic modeling, Fluid Phase Equilib. 556 (2022) 113396. https://doi.org/10.1016/j.fluid.2022.113396.

[508] R.D. Weinstein, J.J. Gribbin, K.R. Muske, Solubility and salting behavior of several β-adrenergic blocking agents in liquid and supercritical carbon dioxide, J. Chem. Eng. Data. 50 (2005) 226–229. https://doi.org/10.1021/je0497202.

[509] S.A. Shojaee, H. Rajaei, A.Z. Hezave, M. Lashkarbolooki, F. Esmaeilzadeh, Experimental measurement and correlation for solubility of piroxicam (a non-steroidal

anti-inflammatory drugs (NSAIDs)) in supercritical carbon dioxide, J. Supercrit. Fluids. 80 (2013) 38–43. https://doi.org/10.1016/j.supflu.2013.03.015.

[510] C.T. Chen, C.A. Lee, M. Tang, Y.P. Chen, Experimental investigation for the solubility and micronization of pyridin-4-amine in supercritical carbon dioxide, J. CO2 Util. 18 (2017) 173–180. https://doi.org/10.1016/j.jcou.2017.01.020.

[511] J. García-González, M.J. Molina, F. Rodríguez, F. Mirada, Solubilities of phenol and pyrocatechol in supercritical carbon dioxide, J. Chem. Eng. Data. 46 (2001) 918–921. https://doi.org/10.1021/je0003795.

[512] G. Sodeifian, R.S. Alwi, F. Razmimanesh, K. Tamura, Solubility of Quetiapine hemifumarate (antipsychotic drug) in supercritical carbon dioxide: Experimental, modeling and Hansen solubility parameter application, Fluid Phase Equilib. 537 (2021) 113003. https://doi.org/10.1016/j.fluid.2021.113003.

[513] B. Li, W. Guo, E.D. Ramsey, Determining the Solubility of Nifedipine and Quinine in Supercritical Fluid Carbon Dioxide Using Continuously Stirred Static Solubility Apparatus Interfaced with Online Supercritical Fluid Chromatography, J. Chem. Eng. Data. 62 (2017) 1530–1537. https://doi.org/10.1021/acs.jced.7b00012.

[514] G. Sodeifian, S.M. Hazaveie, S.A. Sajadian, N. Saadati Ardestani, Determination of the Solubility of the Repaglinide Drug in Supercritical Carbon Dioxide: Experimental Data and Thermodynamic Modeling, J. Chem. Eng. Data. 64 (2019) 5338–5348. https://doi.org/10.1021/acs.jced.9b00550.

[515] E. Reverchon, G. Donsi, D. Gorgoglione, Salicylic acid solubilization in supercritical CO2 and its micronization by RESS, J. Supercrit. Fluids. 6 (1993) 241–248. https://doi.org/10.1016/0896-8446(93)90034-U.

[516] G. Sodeifian, S.A. Sajadian, Experimental measurement of solubilities of sertraline hydrochloride in supercritical carbon dioxide with/without menthol: Data correlation with a pressure of about 60, J. Supercrit. Fluids. 149 (2019) 79–87. https://doi.org/10.1016/j.supflu.2019.03.020.

[517] G. Sodeifian, N. Saadati Ardestani, S.A. Sajadian, M.R. Golmohammadi, A. Fazlali, Prediction of Solubility of Sodium Valproate in Supercritical Carbon Dioxide: Experimental Study and Thermodynamic Modeling, ACS Appl. Mater. Interfaces. (2020). https://doi.org/10.1021/acs.jced.9b01069.

[518] G. Sodeifian, F. Razmimanesh, S.A. Sajadian, S.M. Hazaveie, Experimental data and thermodynamic modeling of solubility of Sorafenib tosylate, as an anti-cancer drug, in supercritical carbon dioxide: Evaluation of Wong-Sandler mixing rule, J. Chem. Thermodyn. 142 (2020) 105998. https://doi.org/10.1016/j.jct.2019.105998.

[519] A.Z. Hezave, S. Shahnazar, H. Rajaei, M. Lashkarbolooki, F. Esmaeilzadeh, Solubility of spironolactone in supercritical carbon dioxide: Experimental and modeling approaches, Fluid Phase Equilib. 355 (2013) 130–134. https://doi.org/10.1016/j.fluid.2013.07.003.

[520] A.Z. Hezave, S. Aftab, F. Esmaeilzadeh, Solubility of sulindac in the supercritical carbon dioxide: Experimental and modeling approach, J. Supercrit. Fluids. 68 (2012) 39–44. https://doi.org/10.1016/j.supflu.2012.04.006.

[521] G. Sodeifian, F. Razmimanesh, S.A. Sajadian, Prediction of solubility of sunitinib malate (an anti-cancer drug) in supercritical carbon dioxide (SC–CO2): Experimental correlations and thermodynamic modeling, J. Mol. Liq. 297 (2020) 111740. https://doi.org/10.1016/j.molliq.2019.111740.

[522] S.M. Hazaveie, G. Sodeifian, S.A. Sajadian, Measurement and thermodynamic modeling of solubility of Tamsulosin drug (anti cancer and anti-prostatic tumor activity) in supercritical carbon dioxide, J. Supercrit. Fluids. 163 (2020) 104875. https://doi.org/10.1016/j.supflu.2020.104875.

[523] G. Sodeifian, L. Nasri, F. Razmimanesh, CO 2 utilization for determining solubility of teriflunomide (immunomodulatory agent) in supercritical carbon dioxide : Experimental investigation and thermodynamic modeling, J. CO2 Util. 58 (2022) 101931. https://doi.org/10.1016/j.jcou.2022.101931.

[524] Y.Y. Sun, S.F. Li, C. Quan, Measurement of the solubility of ferulic acid and tetramethylpyrazine in supercritical carbon dioxide, Gao Xiao Hua Xue Gong Cheng Xue Bao/Journal Chem. Eng. Chinese Univ. 19 (2005) 839–842.

[525] G. Sodeifian, C. Garlapati, S.M. Hazaveie, F. Sodeifian, Solubility of 2,4,7-Triamino-6-phenylpteridine (Triamterene, Diuretic Drug) in Supercritical Carbon Dioxide: Experimental Data and Modeling, J. Chem. Eng. Data. 65 (2020) 4406–4416. https://doi.org/10.1021/acs.jced.0c00268.

[526] S. Han, W. Wang, Z. Jiao, X. Wei, Solubility of Vitamin e Acetate in Supercritical

Carbon Dioxide: Measurement and Correlation, J. Chem. Eng. Data. 62 (2017) 3854–3860. https://doi.org/10.1021/acs.jced.7b00550.

[527] J.M. Ciou, B.C. Wang, C.S. Su, J.J. Liu, M.T. Sheu, Measurement of solid solubility of warfarin in supercritical carbon dioxide and recrystallization study using supercritical antisolvent process, Adv. Powder Technol. 29 (2018) 479–487. https://doi.org/10.1016/j.apt.2017.12.005.

[528] N. Saadati Ardestani, M. Amani, L. Moharrery, Determination of Anthraquinone Violet 3RN solubility in supercritical carbon dioxide with/without co-solvent: Experimental data and modeling (empirical and thermodynamic models), Chem. Eng. Res. Des. 159 (2020) 529–542. https://doi.org/10.1016/j.cherd.2020.04.026.

[529] S. Ismadji, Y.H. Ju, F.E. Soetaredjo, A. Ayucitra, Solubility of azadirachtin in supercritical carbon dioxide at several temperatures, J. Chem. Eng. Data. 56 (2011) 4396–4399. https://doi.org/10.1021/je200239p.

[530] M. Pishnamazi, S. Zabihi, S. Jamshidian, H.Z. Hezaveh, A.Z. Hezave, S. Shirazian, Measuring solubility of a chemotherapy-anti cancer drug (busulfan) in supercritical carbon dioxide, J. Mol. Liq. 317 (2020) 113954. https://doi.org/10.1016/j.molliq.2020.113954.

[531] S.W. Wang, S.Y. Chang, C.M. Hsieh, Measurement and modeling of solubility of gliclazide (hypoglycemic drug) and captopril (antihypertension drug) in supercritical carbon dioxide, J. Supercrit. Fluids. 174 (2021) 105244. https://doi.org/10.1016/j.supflu.2021.105244.

[532] M. Pishnamazi, S. Hosseini, S. Zabihi, F. Borousan, A. Zeinolabedini Hezave, A. Marjani, S. Shirazian, Chloroquine (antimalaria medication with anti SARS-CoV activity) solubility in supercritical carbon dioxide, J. Mol. Liq. 322 (2021) 114539. https://doi.org/10.1016/j.molliq.2020.114539.

[533] S. Zhan, S. Li, Q. Zhao, W. Wang, J. Wang, Measurement and Correlation of Curcumin Solubility in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 62 (2017) 1257–1263. https://doi.org/10.1021/acs.jced.6b00798.

[534] M. Pishnamazi, S. Zabihi, S. Jamshidian, F. Borousan, A.Z. Hezave, A. Marjani, S. Shirazian, Experimental and thermodynamic modeling decitabine anti cancer drug solubility in supercritical carbon dioxide, Sci. Rep. 11 (2021) 1–8.

https://doi.org/10.1038/s41598-020-80399-7.

[535] M. Sauceau, J.J. Letourneau, B. Freiss, D. Richon, J. Fages, Solubility of eflucimibe in supercritical carbon dioxide with or without a co-solvent, J. Supercrit. Fluids. 31 (2004) 133–140. https://doi.org/10.1016/j.supflu.2003.11.004.

[536] S. Zabihi, Y. Rahnama, A. Sharafi, F. Borousan, A. Zeinolabedini Hezave, S. Shirazian, Experimental Solubility Measurements of Fenoprofen in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 65 (2020) 1425–1434. https://doi.org/10.1021/acs.jced.9b00861.

[537] Y. Khayyat, S.M. Kashkouli, F. Esmaeilzadeh, Solubility of fluvoxamine maleate in supercritical carbon dioxide, Fluid Phase Equilib. 399 (2015) 98–104. https://doi.org/10.1016/j.fluid.2015.04.030.

[538] S. Zhan, H. Miao, Y. Zhao, J. Wang, Z. Li, Experimental Determination and Association Model for the Solubility of Laminarin in Supercritical Carbon Dioxide, J. Chem. Eng. Data. 65 (2020) 1814–1823. https://doi.org/10.1021/acs.jced.9b01084.

[539] S. Zabihi, S.H. Esmaeili-Faraj, F. Borousan, A.Z. Hezave, S. Shirazian, Loxoprofen Solubility in Supercritical Carbon Dioxide: Experimental and Modeling Approaches, J. Chem. Eng. Data. 65 (2020) 4613–4620. https://doi.org/10.1021/acs.jced.0c00470.

[540] S. Zabihi, S. Jamshidian, F. Borousan, A. Zeinolabedini Hezave, M. Pishnamazi, A. Marjani, S. Shirazian, Measuring salsalate solubility in supercritical carbon dioxide: Experimental and thermodynamic modelling, J. Chem. Thermodyn. 152 (2021) 106271. https://doi.org/10.1016/j.jct.2020.106271.

[541] M. Pishnamazi, S. Zabihi, P. Sarafzadeh, F. Borousan, A. Marjani, R. Pelalak, S. Shirazian, Using static method to measure tolmetin solubility at different pressures and temperatures in supercritical carbon dioxide, Sci. Rep. 10 (2020) 1–7. https://doi.org/10.1038/s41598-020-76330-9.

[542] S.N. Mirheydari, M. Barzegar-Jalali, H. Shekaari, F. Martinez, A. Jouyban, Experimental determination and correlation of lamotrigine solubility in aqueous mixtures of 1-octyl-3-methylimidazolium bromide ionic liquid at various temperatures, J. Chem. Thermodyn. 135 (2019) 75–85. https://doi.org/10.1016/j.jct.2019.03.024.

[543] M.K. Anwer, M. Mohammad, N.Y. Khalil, F. Imam, M.J. Ansari, M.F. Aldawsari, F. Shakeel, M. Iqbal, Solubility, thermodynamics and molecular interaction studies of delafloxacin in environmental friendly ionic liquids, J. Mol. Liq. 305 (2020).

https://doi.org/10.1016/j.molliq.2020.112854.

[544] A.D. dos Santos, A.R.C. Morais, C. Melo, R. Bogel-Łukasik, E. Bogel-Łukasik, Solubility of pharmaceutical compounds in ionic liquids, Fluid Phase Equilib. 356 (2013) 18–29. https://doi.org/10.1016/j.fluid.2013.07.020.

[545] R.A.M. Faria, T.F.M. Vieira, C.I. Melo, E. Bogel-ŁUkasik, Solubility Data as a Response for a Challenge for Formulation Chemists: Imidazolium-Based Ionic Liquids and Antitubercular Antibiotic Medicines, J. Chem. Eng. Data. 61 (2016) 3116–3126. https://doi.org/10.1021/acs.jced.6b00201.

[546] C.I. Melo, R. Bogel-Łukasik, M. Nunes da Ponte, E. Bogel-Łukasik, Ammonium ionic liquids as green solvents for drugs, Fluid Phase Equilib. 338 (2013) 209–216. https://doi.org/10.1016/j.fluid.2012.11.029.

[547] C. Lourenço, C.I. Melo, R. Bogel-Łukasik, E. Bogel-Łukasik, Solubility advantage of pyrazine-2-carboxamide: Application of alternative solvents on the way to the future pharmaceutical development, J. Chem. Eng. Data. 57 (2012) 1525–1533. https://doi.org/10.1021/je300044x.

[548] L. Ge, L. Guo, K. Yang, K. Tao, J. Su, Y. Long, Solubility of diosgenin in several imidazolium-based ionic liquids, J. Chem. Eng. Data. 60 (2015) 11–15. https://doi.org/10.1021/je5004324.

[549] E.I. Alevizou, E.C. Voutsas, Solubilities of p-coumaric and caffeic acid in ionic liquids and organic solvents, J. Chem. Thermodyn. 62 (2013) 69–78. https://doi.org/10.1016/j.jct.2013.02.013.

[550] M.K. Anwer, M. Muqtader, M. Iqbal, R. Ali, B.K. Almutairy, A. Alshetaili, S.M. Alshahrani, M.F. Aldawsari, A. Alalaiwe, F. Shakeel, Estimating the Solubility, Solution Thermodynamics, and Molecular Interaction of Aliskiren Hemifumarate in Alkylimidazolium Based Ionic Liquids, Molecules. 24 (2019) 1–13. https://doi.org/10.3390/molecules24152807.

[551] M.S. Manic, V. Najdanovic-Visak, Solubility of erythromycin in ionic liquids, J. Chem. Thermodyn. 44 (2012) 102–106. https://doi.org/10.1016/j.jct.2011.08.004.

[552] X. Sun, H. Yang, Z. Wang, F. Li, H. An, W. Xue, Y. Wang, Measurement and Correlation of the Solubility of Theobromine and Theophylline in Seven Imidazolium-Based Ionic Liquids at Various Temperatures, J. Chem. Eng. Data. 65 (2020) 848–855.

https://doi.org/10.1021/acs.jced.9b01043.

Annex

Annex

Contents

1- Parameters of the proposed QSPR-ANN and I-GWO-SVR(r) models.
2- σ-profile of Sigma profile of the 14th pharmaceutical compounds.
3- Databases applied to build the three proposed models.

- **Parameters of the proposed ANN-QSPR model:**

$\mathbf{W_1}$= {0,0223 -0,2368 4,2050 -2,3887 11,0089 -1,4085 -4,4649 4,9215 0,4690 1,7592 7,1743 -12,6234 -4,0089
0,0761 0,0338 -15,9941 -7,1859 12,0244 0,2264 -5,8862 -14,5798 7,0177 0,7574 -12,9555 6,2791 4,1439
0,0730 0,0351 2,5789 -1,8767 -0,3624 2,9787 0,0435 0,4149 -2,5839 -0,4087 2,2881 2,2612 0,9320
0,0271 0,0421 -1,2803 1,8892 -1,7804 -2,5431 2,4371 1,5401 -0,9707 1,7957 -0,9664 0,8885 0,5618
-0,0663 -0,0547 2,5200 -0,0255 0,8692 2,4443 -1,0204 5,5229 1,1954 2,5149 0,9654 -0,2007 -3,2408
0,6704 -0,0939 54,5880 -26,3102 -33,4228 16,6952 49,0114 -24,1445 23,2279 39,3540 -12,5717 -37,8790 -25,4777
0,0159 0,0113 0,9359 0,6486 2,2561 2,6896 -3,3922 3,0546 2,4833 0,2144 -0,1507 -0,7673 -2,7289
-0,1117 -0,1060 -1,5869 3,6293 2,5457 -0,0285 -0,5570 1,9271 -5,4381 -1,0534 -1,0281 6,6338 4,4134
2,8173 -5,9599 -0,5039 0,3791 0,4843 0,0793 -1,2663 -0,3432 0,5633 0,4487 -0,2449 -0,0807 -0,5413
-0,0590 -0,1531 -2,8480 5,1008 2,3252 -8,3377 -0,7938 -2,5180 -1,0227 13,4857 3,1392 6,3347 -1,8259}

$\mathbf{b_1}$=
[-3,0646
9,8384
0,9886
0,6537
-0,1061
-3,1646
-0,2959
0,0467
-4,9608
-5,4594]

$\mathbf{w_2}$= [-0,9881 -0,8044 -3,3372 -4,4161 3,5117 -0,4881 -4,1447 1,8525 0,4593 -1,6098]

$\mathbf{b_2}$= [5,4747]

- **Parameters of the proposed I-GWO-SVR(r) model:**

SupportVectors	{-1,4772	-0,9117	0,3650	0,3635	-0,1472	-1,5787	-2,7602	4,9220	-0,2174	1,5415	-0,2710	0,9887	-0,5411
	-0,9447	-0,9117	0,3650	0,3635	-0,1472	-1,5787	-2,7602	4,9220	-0,2174	1,5415	-0,2710	0,9887	-0,5411
	-1,3441	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-1,3441	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-0,8116	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-1,3441	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-0,8116	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-0,8116	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	1,9376	0,6442	1,1837	0,9790
	-1,1537	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,6604	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,1617	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-1,3398	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-1,4796	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-1,1239	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-1,7397	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,9783	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,7482	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,6487	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,4301	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,2706	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	0,4479	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,3207	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
	-0,3446	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785

0,6056	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,0981	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
-0,0765	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,1756	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,2456	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,3268	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,7030	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,1210	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,4783	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,4200	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,5124	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,6471	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,8140	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,6785	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,7749	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,7831	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,8614	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,8106	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
0,7967	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,6418	-0,1475	-0,2665	-0,1785
-0,6409	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,6678	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,9189	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,2198	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,1857	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,7450	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,2533	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,4692	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,2113	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188

0,1080	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-1,5051	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-1,2575	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-1,4170	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-1,4378	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,8090	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,5523	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,7970	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,3255	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,8297	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
-0,5249	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,8723	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,4224	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
0,8436	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,9519	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,5818	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
-0,4399	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,3508	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
-0,5949	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,5126	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,3704	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,0201	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,5821	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,1050	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,0422	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,7123	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-0,1764	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,4668	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188

0,8092	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,7892	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,8518	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,6412	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,0560	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,8673	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,4330	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,3212	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,7844	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,5696	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,5659	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,5228	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,8582	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	0,4981	-0,4017	-0,2613	-0,1188
1,7961	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,5427	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,7924	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
0,9911	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,7650	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,0116	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,0542	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
1,2184	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,4981	-0,4017	-0,2613	-0,1188
-2,1433	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-2,6561	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,9532	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,4452	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,0464	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,1015	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,4774	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135

-0,9405	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,5517	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,5494	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,5195	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,6047	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,1156	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,2328	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,1439	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,5864	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
-0,3508	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,2504	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,2834	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,0965	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,8291	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,1386	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,2019	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,1287	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,7557	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
-1,8480	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,4926	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,3729	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,3555	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,7290	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,4690	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,3990	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,6692	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,3750	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,4711	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135

0,3984	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,9319	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,2544	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
0,7717	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,3793	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
-0,1346	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
0,9192	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,4179	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,5206	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,1756	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,1849	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,3156	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,7642	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,2722	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
0,3212	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,9216	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
1,5360	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,6641	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,4120	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,4359	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-0,0829	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,4019	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,1298	-0,8947	-2,4833	2,5086	-0,1466	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
1,7333	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,0017	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,9107	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
1,4013	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
0,9367	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135

1,2868	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,1513	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,9979	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,6907	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,2802	-0,9113	-1,2298	1,8411	-0,1470	-1,3496	-0,6448	1,9962	1,1459	-0,3005	-0,3566	-0,5165	-0,4135
1,6236	2,0770	0,6504	-1,0567	6,8051	1,2133	0,3366	-0,6147	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,0688	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,2262	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,4913	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,5198	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,8300	-0,7939	0,2526	0,8270	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,5973	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,7578	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,8201	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,2805	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,7666	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,3409	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,6114	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,4367	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,6817	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,7267	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
1,6119	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
0,9280	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,3005	-0,3566	-0,5165	-0,4135
-1,9481	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-2,2658	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,8717	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,6255	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,2740	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717

-1,4995	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,9798	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,5980	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,9287	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,7948	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,5818	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,8933	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,5794	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,1178	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,9370	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-1,0267	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,7653	-0,9115	0,3638	-0,0961	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,6412	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,7030	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,7552	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,4663	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,5912	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,4341	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,3835	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,2640	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,2986	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-1,8440	-0,2154	0,0073	-0,7717
-0,7269	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-0,2088	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-1,7482	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-0,9647	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
0,4328	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-1,1348	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-0,6338	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238

1,3659	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
0,1812	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
0,3002	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
1,6750	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
0,6892	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
0,6468	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
1,2666	-0,9118	0,4686	-1,0275	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
1,7709	1,0480	0,2097	-0,5567	-0,1462	-0,3099	-0,0644	0,6040	-0,2135	-0,4573	-0,3461	0,1586	-0,6238
-0,4109	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	-1,1853	1,4570	1,4700	-0,3051
-0,2831	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	-1,1853	1,4570	1,4700	-0,3051
0,0417	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
-0,5254	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
0,0603	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	-1,1853	1,4570	1,4700	-0,3051
0,4064	-0,9118	0,4686	-1,0275	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
0,2201	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	-1,1853	1,4570	1,4700	-0,3051
-0,8821	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
-0,1766	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
0,0763	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
0,1828	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
-0,6665	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
0,2866	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
-0,6186	-0,9117	0,3650	0,3635	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-1,1853	1,4570	1,4700	-0,3051
-0,8329	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-1,2136	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-0,8329	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-0,8329	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-1,2136	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-0,8329	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,3957	-0,4030	0,4347	-0,3932

-0,9607	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-1,2136	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-0,9607	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-0,8329	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,3957	-0,4030	0,4347	-0,3932
-1,2136	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-0,9607	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,2136	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,2136	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-0,9607	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-0,8329	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,2136	-0,9117	0,3650	0,3635	-0,1472	-0,3099	-0,0644	0,6040	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,2136	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-0,8329	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,2136	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	1,2790	-0,4640	0,5562	-0,3203
-1,0778	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	6,5567	6,3122	6,3730	7,2401
-0,8116	-0,9117	0,3650	0,3635	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	6,5567	6,3122	6,3730	7,2401
-1,3441	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	6,5567	6,3122	6,3730	7,2401
-0,9447	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	6,5567	6,3122	6,3730	7,2401
-1,2322	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,7115	-0,9049	-4,0900	-0,3798	-0,1471	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,3121	-0,9049	-4,0900	-0,3798	-0,1471	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,3537	-0,9049	-4,0900	-0,3798	-0,1471	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
0,1468	-0,9049	-4,0900	-0,3798	-0,1471	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,9314	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,6928	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,6439	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,6039	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,1630	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714

-0,0129	1,0480	0,2097	-0,5567	-0,1462	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,5241	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,3813	-0,9225	-1,3803	-0,7226	-0,1519	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,3350	-0,9225	-1,3803	-0,7226	-0,1519	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
0,1548	-0,9225	-1,3803	-0,7226	-0,1519	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,6103	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,7264	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
0,2001	-0,7939	0,2526	0,8270	-0,1472	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,6172	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,7717	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,3494	-0,9115	-5,4513	4,9983	-0,1472	-1,5787	-2,7598	3,7792	11,2609	-0,2273	2,6154	2,2800	2,1714
-0,5613	-0,9115	-5,4513	4,9983	-0,1472	-1,5787	-2,7598	3,7792	11,2609	-0,2273	2,6154	2,2800	2,1714
0,2267	-0,9115	-5,4513	4,9983	-0,1472	-1,5787	-2,7598	3,7792	11,2609	-0,2273	2,6154	2,2800	2,1714
-0,7557	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,4335	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-0,0768	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
0,2480	-0,9115	-4,5947	3,3401	-0,1440	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714
-1,2123	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-1,2123	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-1,2123	-0,9115	0,3638	-0,0961	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-0,1473	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-0,4136	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
0,1189	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-0,4136	-0,9115	0,3638	-0,0961	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
0,1189	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-0,4136	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-1,2123	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	-0,4184	0,0900	-0,6747	0,3837
0,1189	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	-0,4184	0,0900	-0,6747	0,3837

-0,6798	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	-0,4184	0,0900	-0,6747	0,3837
-0,1473	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	-0,4184	0,0900	-0,6747	0,3837
-0,4136	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	-0,4184	0,0900	-0,6747	0,3837
0,1189	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	-0,4184	0,0900	-0,6747	0,3837
-1,2123	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	0,0062	0,0272	-0,6465	0,3428
0,1189	1,0480	0,2097	-0,5567	-0,1462	0,2113	-1,7255	-0,6379	-0,2135	0,0062	0,0272	-0,6465	0,3428
-0,4136	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
-0,4136	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	0,0062	0,0272	-0,6465	0,3428
-1,2123	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
-0,6798	-0,9115	0,3638	-0,0961	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
-0,4136	-0,9115	0,3638	-0,0961	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
0,1189	1,0480	0,2097	-0,5567	-0,1462	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
0,1189	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	0,0062	0,0272	-0,6465	0,3428
0,1189	-0,9115	0,3638	-0,0961	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
-0,1473	-0,9117	0,3650	0,3635	-0,1472	-1,2426	1,5301	-0,6406	-0,2135	0,0062	0,0272	-0,6465	0,3428
-1,2123	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	0,0062	0,0272	-0,6465	0,3428
-0,6798	1,0480	0,2097	-0,5567	-0,1462	-1,3496	-0,6448	1,9962	1,1459	0,0062	0,0272	-0,6465	0,3428
-0,6798	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	0,0062	0,0272	-0,6465	0,3428
-0,4136	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	0,0062	0,0272	-0,6465	0,3428
0,1189	-0,9115	0,3638	-0,0961	-0,1472	-1,3496	-0,6448	1,9962	1,1459	0,0062	0,0272	-0,6465	0,3428
0,1189	-0,9115	0,3638	-0,0961	-0,1472	0,2113	-1,7255	-0,6379	-0,2135	-0,4184	0,0900	-0,6747	0,3837
-0,9634	-0,9225	-1,3803	-0,7226	-0,1519	1,2133	0,3366	-0,6147	-0,2135	-0,2273	2,6154	2,2800	2,1714}

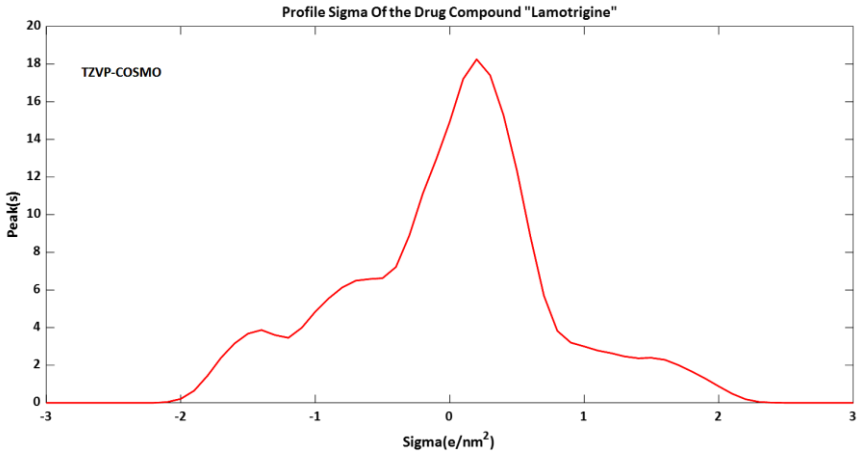
Bias= 0,2645

KernelScale= 1,3181

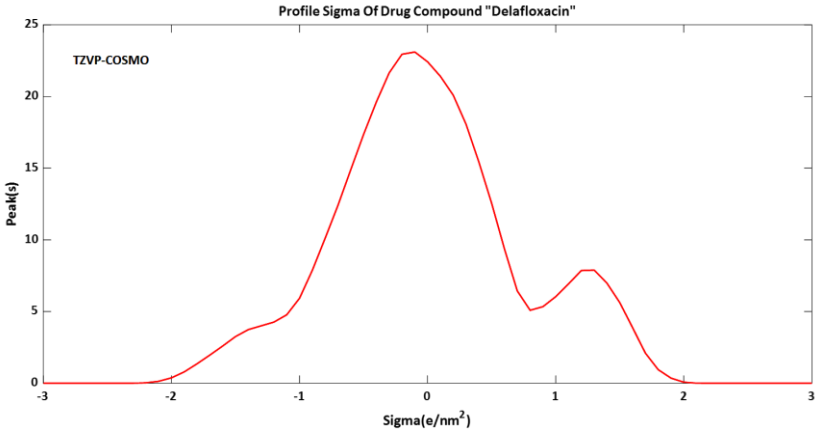
Alpha =	[-0,1388	-0,1391	-0,0870	-0,0374	-0,0622	-0,1253	-0,0636	-0,1066	7,3497	-19,1087	6,5976	-1,9921
0,5154	4,4008	-0,0342	-2,0659	4,5710	-23,1101	4,4372	2,5477	199,4019	33,1859	4,2472	-199,4019	0,0000
-12,8659	-6,4774	-25,3195	-118,7533	-199,4019	-1,4943	199,4019	56,5835	0,0000	-85,1242	10,5950	-127,8890	14,2036
0,0000	199,4019	5,7970	83,3240	-0,5076	0,0039	-0,1165	0,0986	0,2659	0,0000	0,4166	-0,0010	0,0000
-0,2305	-0,1244	-0,0482	-0,0682	-0,2655	-1,6546	-0,0816	0,2713	-0,1400	-1,4238	-0,0242	0,2230	-0,0637
0,3102	0,6203	0,0000	0,0000	4,9451	0,0000	2,1226	-0,0201	-0,4380	0,0000	-7,1817	-0,3266	-0,2031
-2,6016	0,6580	0,1553	0,2421	0,0919	1,9645	0,8425	6,3268	0,0628	-1,5203	0,1206	-12,7476	-1,6928
7,5337	10,6255	1,4613	-1,9408	0,3049	-1,8098	1,8158	-4,0447	-13,4604	14,3227	0,6767	-0,2469	-1,3559
0,1533	-2,1919	1,5733	-0,1489	-0,8275	0,0000	10,2025	1,9032	0,0000	-0,6480	1,4871	0,0000	-0,1271
0,5473	-3,7994	0,0000	-38,1385	-0,1872	-0,3298	-0,1969	-0,1584	3,2908	-0,7252	0,0377	-0,0454	1,1483
-0,0743	41,0944	0,0000	0,8746	6,4282	0,0000	-0,3816	0,1817	0,0000	-5,9388	0,0192	-21,2561	-5,4543
-1,4610	-0,1086	1,9079	0,0314	0,0000	0,0000	41,1081	-5,0888	-0,0650	-32,2992	-0,1255	0,5609	2,1471
-0,5164	-0,5896	-0,0795	14,9213	0,0000	-88,0608	-135,6593	17,8163	-0,3414	1,3202	0,1310	1,4469	199,4019
0,4023	0,8862	0,0753	-67,1591	-7,5238	-0,1307	0,0000	199,4019	-146,3350	0,0000	5,6884	-0,8849	0,0000
-4,1300	0,0000	-4,7207	4,7134	4,8868	-4,0344	-0,3918	-0,5732	15,0780	-27,3508	0,3440	-0,2664	0,6476
-0,4359	61,6073	0,6410	2,1027	-63,5226	0,3486	-0,7764	2,6175	-0,4663	-0,2231	-5,0439	-1,6359	-8,9259
2,0942	3,3643	1,2060	18,9786	7,6439	0,1917	-1,7461	-0,1435	0,0000	2,0337	0,1163	-0,1963	-4,1522
0,6079	-1,7979	1,9730	-0,9244	2,4483	-2,2972	-0,4853	-21,9225	37,7308	-0,5656	-54,2206	-37,3806	-0,2977
21,7706	29,3063	199,4019	-199,4019	-114,7090	-199,4019	169,7149	171,8938	0,3527	-0,0625	-0,1314	-0,0313	-0,1506
0,0050	0,4199	-0,0748	0,1639	0,1862	-0,1787	-0,1590	-0,0439	-0,0470	-0,0817	-0,1178	-0,2745	-0,1813
-0,0011	-0,1074	-0,1148	-0,0648	-0,0609	-0,2055	18,6608	-0,2120	0,0016	0,0013	-0,2067	-175,0106	-0,2125
-199,4019	-199,4019	0,0513	-0,1806	-199,4019	-0,1747	0,0389	-0,1813	-0,1510	0,0142	-0,1617	166,9908	199,4019
-0,1615	0,0353	-0,1462	199,4019	199,4019	-10,6994	-0,2488	-0,2088	-0,3555	-0,1497	0,6569	0,3038	-0,7779
0,0258	-0,4430	0,0357	-0,2746	-0,2141	0,0813	-0,2971	0,0824	-0,6692	-0,1990	-0,1844	0,6110	0,1747
-0,0994	0,0310	0,0304	-0,7105	-0,0561	-0,1267	0,0850	-0,2904	0,2642	0,0083	0,0279	0,4232	-0,2468

0,0164]

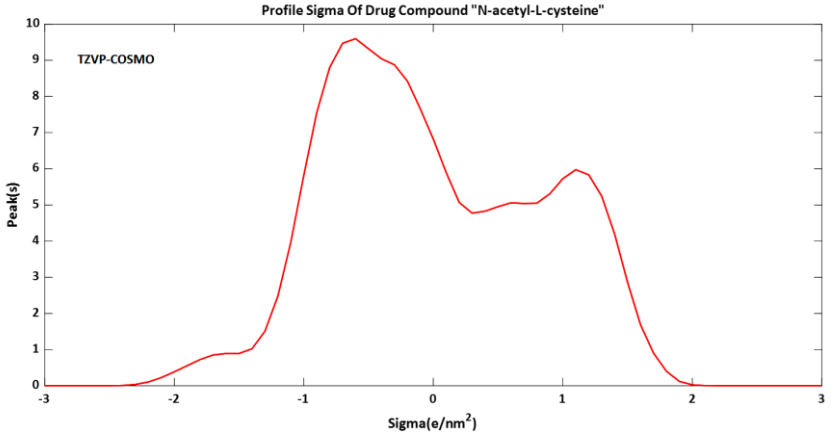
(a)


Profile Sigma Of the Drug Compound "Lamotrigine"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

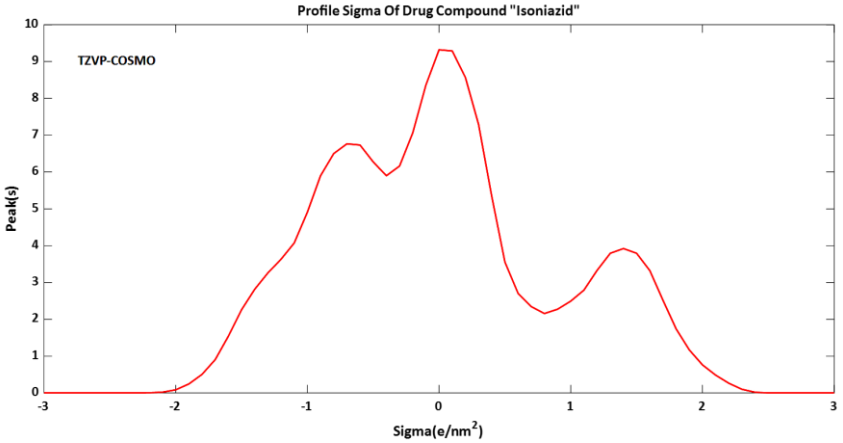
(b)


Profile Sigma Of Drug Compound "Delafloxacin"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

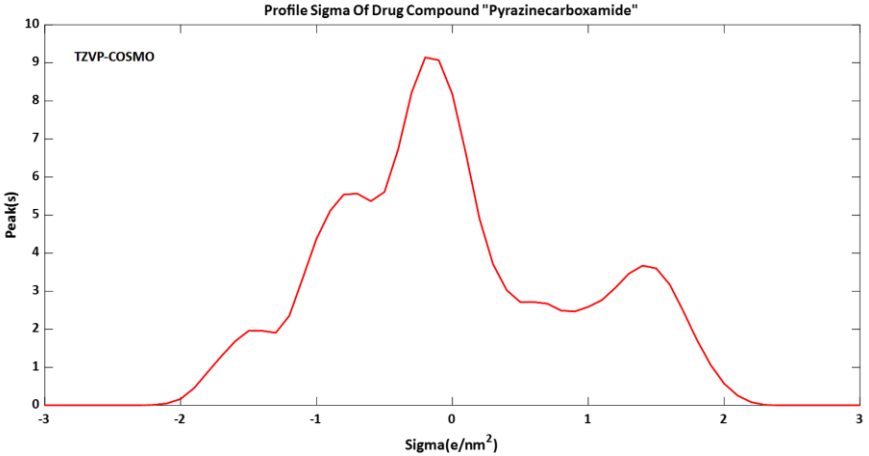
(d)


Profile Sigma Of Drug Compound "N-acetyl-L-cysteine"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

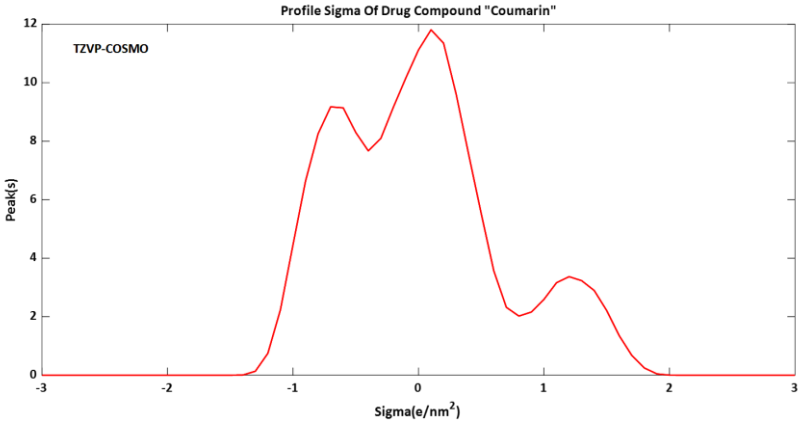
(c)


Profile Sigma Of Drug Compound "Isoniazid"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

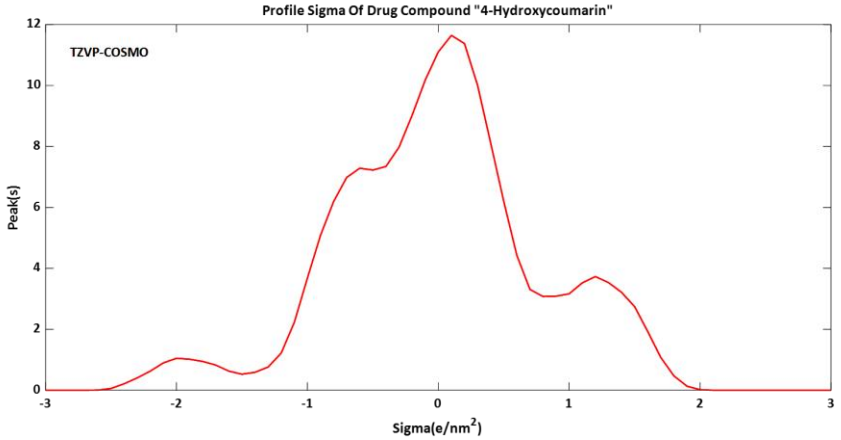
(e)


Profile Sigma Of Drug Compound "Pyrazinecarboxamide"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

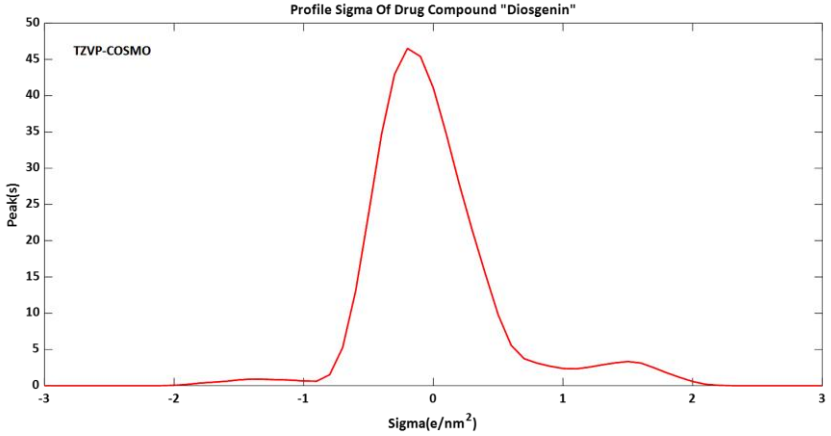
(f)


Profile Sigma Of Drug Compound "Coumarin"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

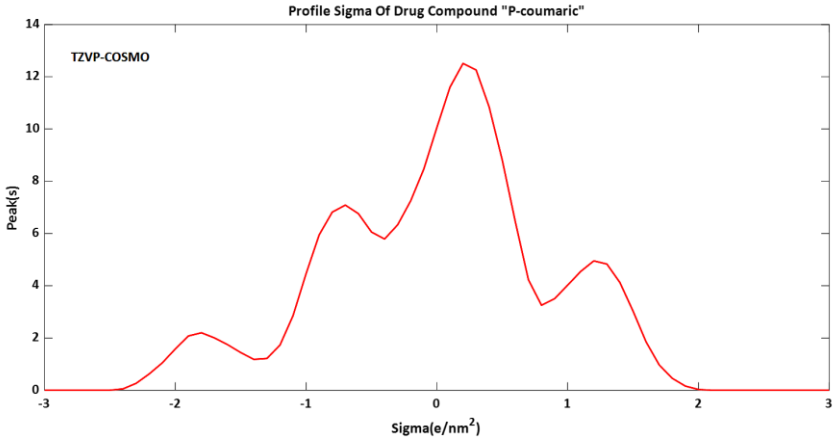
(g)


Profile Sigma Of Drug Compound "4-Hydroxycoumarin"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

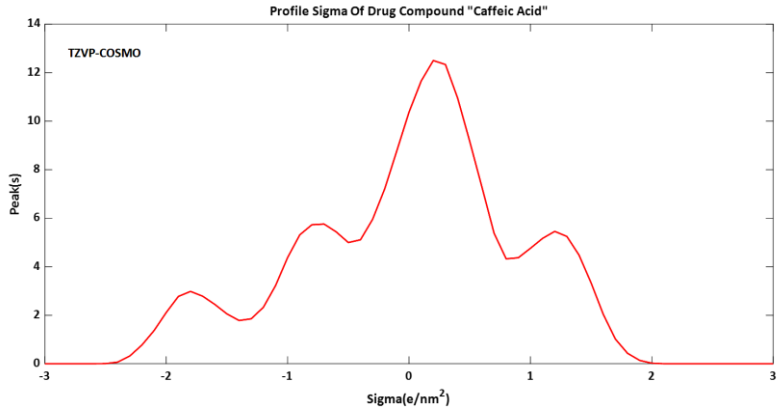
(h)


Profile Sigma Of Drug Compound "Diosgenin"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

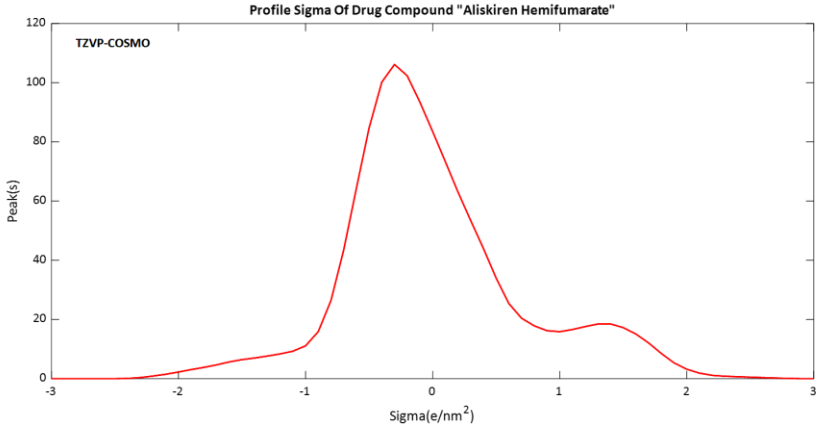
(i)


Profile Sigma Of Drug Compound "P-coumaric"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

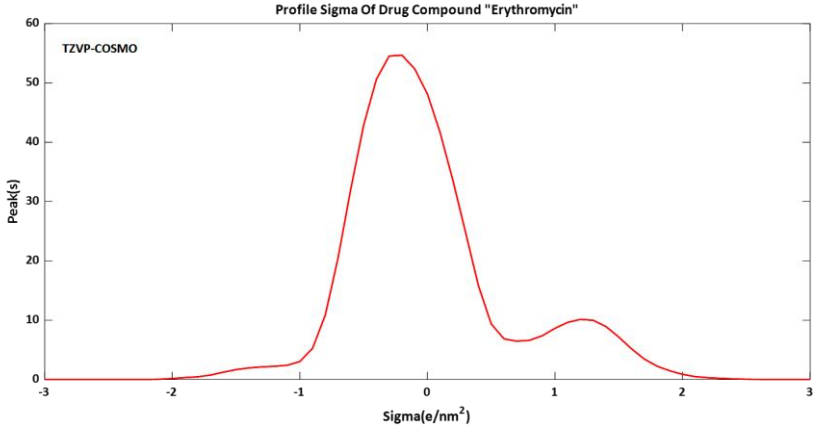
(j)


Profile Sigma Of Drug Compound "Caffeic Acid"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

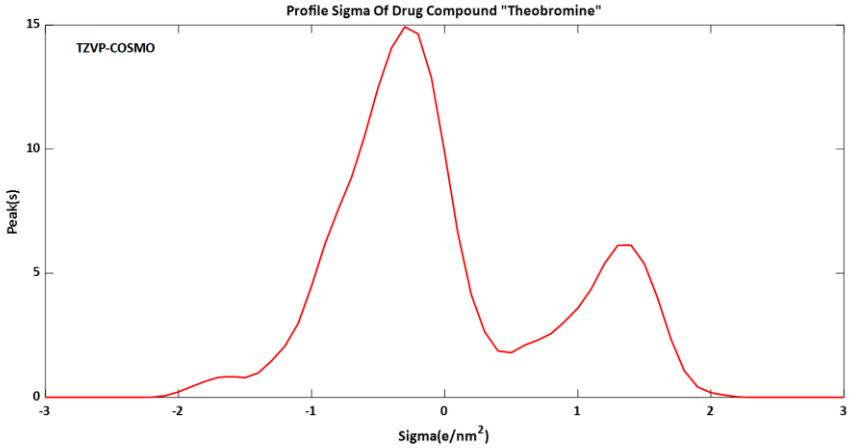
(k)


Profile Sigma Of Drug Compound "Aliskiren Hemifumarate"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

(l)


Profile Sigma Of Drug Compound "Erythromycin"
TZVP-COSMO
Peak(s)
Sigma(e/nm^2)

(m)



(n)

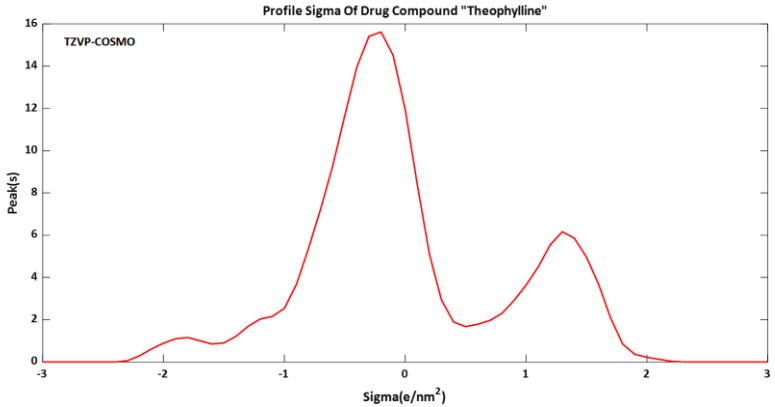


Figure. A Sigma profile distribution of the 14th APIs: (a) Lamotrigine, (b) Delafloxacin, (c) N-acetyl-L-cysteine, (d) Isoniazid, (e) Pyrazine carboxamide, (f) Coumarin, (g) 4-Hydroxycoumarin, (h) Diosgenin, (i) p-coumaric, (j) Caffeic Acid, (k) Aliskiren Hemifumarate, (l) Erythromycin, (m) Theobromine (n) Theophylline.

Table. A1. Summary of Drugs and temperature, pressure and solubility ranges, the 11th QSPR desriptors and their refrences contained in the database used in **chapter II**

No.	Nom	T(K)	P(Mpa)	AATS3v	MATS2e	GATS4c	GATS3v	GATS4e	GATS3s	Nbondsm	AVP-0	Shbd	Mlogp	MLFER_S	Solubility Range (Mole Fraction)	Ref.
1	4-Aminoantipyrine	308.2-328.2	10.03-22.04	208.4654	0.1136	1.418	0.9713	1.4709	0.7436	8	0.5844	0.5395	2.23	1.124	0.000005-0.00036	[432]
2	4-Aminosalicylic Acid	308-328	11.0-21	207.699	0.1805	0.6322	1.2279	0.8292	0.9395	7	0.5101	2.0166	1.79	1.284	0.0000035-0.0000359	[433]
3	5-Fluorouracil	313-328	8.0-20.0	184.356	0.0791	0.9297	1.0804	1.5854	1.1499	3	0.4746	1.4681	1.35	0.495	0.0000013-0.00000525	[434]
4	10-Undecenoic Acid	308-333	10.0-18.0	114.8072	0.2571	0.6702	0.9895	0.7065	0.7811	2	0.6382	0.7291	2.45	0.548	0.0007-0.0174	[435]
5	All-Trans Retinoic Acid	308.2-328.2	10.2-31.2	182.234	0.2772	0.5449	1.3259	0.384	0.6936	6	0.6564	0.71	3.44	0.983	0.00000019-0.0000184	[436]
6	Amiodarone Hydrochloride	313.2-342.2	12.0-30.0	201.279	-0.0381	1.1632	0.7574	1.213	0.8687	17	0.7308	0	3.55	2.451	0.0000251-0.00101	[437]
7	Amlodipine Besylate	308-338	12.0-27.0	208.449	0.1654	1.1063	0.9235	1.2394	0.7368	12	0.5850	1.4976	3	4.24	0.000004146-0.0000022996	[438]
8	Anastrozole	308-348	12.2-35.5	211.0515	0.1176	0.8892	1.394	0.999	0.7343	13	0.6059	0	2.78	2.518	0.0000036-0.00038	[439]
9	Aprepitant	308.15-338.15	12.0-33	206.247	0.2919	0.8433	1.1277	0.9294	0.5845	14	0.5178	1.2461	2.45	1.684	0.0000045-0.0000767	[440]
10	Aspirin	308.15-328.15	12.0-25	208.2891	0.2894	0.826	0.9011	0.9965	1.0858	8	0.537	0.7462	2.01	1.341	0.000063-0.000347	[441]
11	Atorvastatine	308-348	12.16-35.46	191.5763	0.0813	0.8363	0.9608	0.7306	0.9913	27	0.564	1.596	4.21	3.845	0.000001-0.00145	[33]
12	Atropine	308-348	12.2-35.5	165.4072	0.0454	0.762	1.0509	0.9386	0.7612	7	0.5925	0.4974	2.89	1.698	0.00006-0.00167	[442]
13	Azathioprine	313-338	12.0-27.0	262.0937	0.0496	0.9303	1.1366	0.8745	0.575	16	0.5425	0.5242	1.35	2.294	0.0000027-0.0000183	[443]
14	Azelaic Acid	313.15-333.15	10.0-30.0	118.5903	0.2997	0.6753	1.0151	0.7861	0.8502	2	0.5892	1.4901	2.01	0.889	0.00000042-0.0000101	[444]
15	Benzocaine	308-348	12.2-35.5	177.319	0.1807	1.3478	0.9647	1.6374	0.8585	7	0.5758	0.5452	2.12	1.418	0.00027-0.0121	[445]
16	Bisacodyl	308-348	12.2-35.5	215.9504	0.2447	0.7089	0.9187	0.6682	1.2092	20	0.5587	0	3.33	2.353	0.000009-0.000583	[446]
17	Budesonide	338-358	21.3-38.5	191.9097	0.0289	0.8834	1.0587	0.9264	0.7307	4	0.6023	0.464	3.55	2.443	0.00000593-0.0000291	[447]
18	Caffeine	313-353	19.9-34.9	232.0621	0.1562	1.104	1.2438	1.5655	0.9426	7	0.5844	0	1.68	0.996	0.000283-0.00113	[448]
19	Capecitabine	308-348	15.2-35.4	159.97	0.1123	1.2248	0.9445	1.0837	0.8026	4	0.5605	1.6171	2.01	1.522	0.0000027-0.000159	[449]
20	Carbamazepine	308-348	12.2-35.5	234.2513	0.4052	1.0654	0.8578	0.9715	0.8385	14	0.5392	0.5977	2.78	1.449	0.000003-0.000094	[442]
21	Cefixime Trihydrate	308-328	18.3-33.5	229.7306	0.1078	0.9696	0.9675	0.8859	0.9275	12	0.5578	2.3614	1.68	4.226	0.00000016-0.000000302	[107]
22	Cefuroxime Axetil	308-328	8.0-30.0	214.9035	0.122	1.117	0.9133	0.9549	1.0724	12	0.5616	1.1698	2.01	3.944	0.00000022-0.0000112	[450]
23	Celecoxib	323.3-343.2	15-30	228.6141	0.4423	0.7527	1.092	0.8058	0.6159	21	0.528	0.5598	2.34	2.439	0.0000152-0.000152	[451]
24	Cetirizine	308.15-338.15	16-40	182.0835	-0.0552	0.6153	1.0493	0.4804	0.7265	13	0.5931	0.7278	3.11	2.218	0.0000105-0.00492	[452]
25	Chloramphenicol	308-338	10.0-26.0	196.2615	0.0171	1.1274	0.992	1.0764	0.8024	8	0.5717	1.3194	1.68	2.107	0.0000041-0.0000099	[453]
26	Chlormezanone	308.2-328.2	11.91-23.99	207.1075	0.1159	1.2947	1.0358	1.3552	1.1984	9	0.6011	0	2.01	2.799	0.000005-0.000545	[454]

27	Chlorpheniramine Maleate	308-338	20-40	188.5535	-0.1532	0.7767	1.1291	0.7621	1.1446	12	0.6388	0	2.89	1.494	0.0000154-0.000426	[455]
28	Chlorpropamide	313.15-353.15	10.0-30.0	181.153	0.3666	1.139	0.9322	1.2696	0.9861	9	0.5996	0.9782	1.79	2.033	0.00000229-0.0000722	[456]
29	Chrysin	313-333	20-40	233.324	-0.0366	0.5951	0.9708	0.6932	1.0668	14	0.5173	1.175	2.67	1.975	0.000000018-0.000000113	[457]
30	Cinildipine	313-333	12.0-36.0	214.7053	0.0962	1.0382	1.0221	1.0327	0.6622	18	0.5612	0.367	3.44	2.796	0.0000012-0.0000399	[458]
31	Cinnarizine	308.15-328.15	13.97-24	182.4832	-0.2551	0.9535	1.052	1.0177	0.9352	19	0.5829	0	4.1	2.112	0.0000029-0.000205	[459]
32	Ciprofloxacin	313-333	17-36	199.1704	-0.004	0.666	0.7763	0.7512	0.6746	9	0.5453	0.9428	2.56	2.315	0.0000000265-0.000000189	[460]
33	Clemastine Fumarate	308-338	12.0-27.0	178.4961	0.1171	0.8759	1.0630	0.8025	0.6802	9	0.5988	1.7740	3.44	2.289	0.0000161-0.0000941	[461]
34	Climbazole	313.2-333.2	10.55-39.89	192.2391	0.0452	1.4104	1.325	1.5748	1.1387	12	0.6231	0	2.56	1.846	0.00062-0.00488	[462]
35	Clobetasole Propionate	308-348	12.2-35.5	205.2235	0.0599	1.0832	1.0154	1.2446	0.9404	5	0.6211	0.0631	3.44	2.316	0.0000001-0.0000035	[151]
36	Clofenamic Acid	313-333	12.0-36.0	215.9826	0.1819	0.8208	1.0687	0.8842	0.6788	13	0.5651	1.0603	2.45	1.913	0.00000154-0.0000354	[238]
37	Clotrimazole	308-348	12.2-35.5	236.5251	-0.0031	0.954	0.8858	1.0874	0.818	23	0.5737	0	3.55	2.081	0.0000002-0.000107	[147]
38	Clozapine	318-348	12.16-35.46	204.9515	-0.1001	1.0422	0.9839	0.9216	0.9495	13	0.6019	0.306	2.89	2.681	0.0000012-0.0000419	[463]
39	Codeine	308-348	12.2-35.5	212.2692	-0.1339	0.9069	1.0213	0.9318	0.7427	7	0.5886	0.3164	3	2.156	0.00001-0.000797	[442]
40	Cortisone Acetate	308.15-373.15	8.24-22.65	195.8706	0.0818	0.8991	1.1542	0.8232	0.8992	5	0.6018	0.0819	3.33	2.52	0.00000011-0.00000414	[464]
41	Coumarin-7	308-338	9.0-33.0	208.9521	0.1860	0.8908	0.8196	0.8943	0.7036	18	0.5697	0.4406	3.11	2.9	0.0000415-0.0001009	[465]
42	Cyproheptadine	308-338	16-40	212.7801	-0.2504	0.4639	0.9685	0.5417	0.8336	14	0.5931	0	3.66	1.631	0.0000335-0.00309	[466]
43	Cyproterone Acetate	308-348	12.2-35.5	209.6532	0.13	1.0215	1.0128	0.9251	0.8609	5	0.6325	0	3.55	2.127	0.000013-0.000261	[467]
44	Dasatinib Monohydrate	308-338	12.0-27.0	197.2930	0.0390	0.6595	0.9740	0.6940	0	9	0.6052	1.3254	2.56	4.486	0.000000454-0.000009078	[468]
45	Dexamethasone	308.15-328.15	10-35.07	203.3863	-0.0081	0.9111	1.1082	1.1472	0.8336	4	0.5933	0.6436	3.22	2.182	0.00000119-0.00000281	[469,470]
46	Diazepam	308-348	12-35.07	238.9165	0.0302	1.2376	0.9297	1.1518	0.9749	14	0.5881	0	2.78	2.109	0.000039-0.000582	[442]
47	Diclofenac Acid	308.15-338.15	12.0-40.0	224.1334	0.1507	1.0423	1.0117	0.8957	0.9816	13	0.5985	0.9706	2.45	2.1	0.0000234-0.00198	[471]
48	Diflunisal	308.2-328.2	9.1-24.6	238.7639	0.0631	0.7155	1.0607	0.8642	0.9856	14	0.5012	1.3249	2.34	1.394	0.000000544-0.00000807	[472]
49	Docetaxel	308-348	15.2-35.4	205.6952	0.121	1.2119	1.1162	1.1127	0.9009	18	0.5842	0.2923	4.54	3.951	0.000037-0.000702	[107]
50	Dutasteride	308-348	12.2-35.5	195.6823	0.4128	0.6627	1.1882	0.5938	0.4119	9	0.5567	0.6906	3.33	2.185	0.0000001-0.000161	[473]
51	Enrofloxacin	313-333	12.0-36.0	196.8204	0.0006	0.623	0.8051	0.6901	0.6612	9	0.567	0.6212	2.78	2.502	0.000000022-0.0000056	[461]
52	Esomeprazole	308.2-338.2	12.0-27.0	232.2006	-0.079	1.3285	1.0686	1.3089	1.1303	17	0.6042	0.4044	2.56	3.406	0.000011-0.00091	[474]
53	Ethosuximide	313.15-328.15	9.37-14.8	161.7181	0.2617	0.8615	1.3069	1.0784	1.0598	2	0.623	0.548	1.9	0.695	0.00345-0.00993	[475]
54	Exemestane	308-348	12.2-35.5	197.4053	0.0384	0.8557	1.1435	0.8794	0.7824	5	0.6146	0	3.44	1.586	0.0000126-0.00188	[439]

55	Finasteride	308-348	12.2-35.5	175.3167	0.1983	0.8732	1.2519	0.7794	0.7831	3	0.6518	0.3783	3.55	1.942	0.000336-0.00332	[473]
56	Fluoxetine Hydrochloride	308.15-338.15	16-40	182.7174	0.3698	0.9534	1.0135	1.1184	0.4801	12	0.5559	0.3369	2.78	1.236	0.0000265-0.000812	[476]
57	Flurbiprofen	303-323	8.9-24.5	210.3439	0.1535	0.9121	1.019	1.1564	0.8865	14	0.5516	0.6615	2.78	1.388	0.0000167-0.000197	[477]
58	Flutamide	308-348	12.2-35.5	197.4672	0.4344	1.2309	1.1323	1.1207	0.6545	8	0.5392	0.529	1.9	1.863	0.0000046-0.000509	[473]
59	Fluvastatin	308-348	12.16-35.46	192.8681	0.0598	0.6283	0.9066	0.609	1.0013	19	0.5664	1.4356	3.44	2.486	0.000005-0.000601	[33]
60	Gabapentin	308-338	16-40	141.7961	0.2009	0.8215	1.2575	0.5609	0.6744	1	0.6152	1.0533	2.12	0.829	0.0000897-0.00736	[478]
61	Galantamine	308-338	12.0-27.0	206.4365	-0.1472	0.8419	1.0653	0.8497	0.9104	7	0.5954	0.4141	2.89	2.151	0.000006-0.00266	[479]
62	Gambogic Acid	308.15-328.15	10.0-30.0	214.8228	0.0998	1.0269	1.0774	1.1301	0.8507	14	0.6086	0.6876	4.76	2.935	0.00000163-0.0000226	[480]
63	Gatifloxacin	313-333	12.0-36.0	203.8424	-0.0339	0.8266	0.8102	0.7465	0.7	9	0.5662	0.8703	2.67	2.313	0.000000106-0.0000016	[461]
64	Gemfibrozil	308.2-328.2	10.01-22.02	173.1881	0.1698	1.2968	1.3523	1.7007	0.8193	7	0.6453	0.6041	2.78	0.992	0.00000294-0.00419	[481]
65	Ibuprofen	308-318	8.0-13.0	169.2564	0.2781	0.8134	1.1929	0.9692	0.7951	7	0.6351	0.6548	2.67	0.832	0.0000115-0.00326	[434]
66	Imatinib Mesylate	308.2-338.2	12.0-27.0	189.5516	0.2392	0.7186	1.1346	0.5797	0.9800	15	0.5728	1.6131	3.44	6.192	0.0000153-0.0002235	[482]
67	Imipramine Hcl	313-323	30-50	188.3197	-0.1823	1.0557	1.047	1.2625	1.0154	12	0.6213	0	3.33	1.254	0.0000051-0.0000099	[483]
68	Ipriflavone	308.2-328.2	10.0-30	209.4584	-0.0323	0.875	0.8758	0.9364	1.034	14	0.5713	0	3.11	1.921	0.0000054-0.00017	[484]
69	Isoniazid	308-313	13-18.5	188.1176	-0.0365	1.2133	1.0729	1.7202	0.87	7	0.5242	1.3234	1.68	1.447	0.00000053-0.00000605	[485]
70	Ketoconazole	308-348	12.2-35.5	196.7504	-0.059	1.1118	0.9759	1.0464	1.1941	18	0.5969	0	3.22	3.481	0.0000005-0.000175	[147,486]
71	Ketoprofen	308.15-338.15	16-40	208.2778	0.19322	1.0397	1.0351	1.42	0.8396	14	0.5545	0.6399	2.89	1.768	0.0000221-0.000712	[487]
72	Ketotifen Fumarate (Ktf)	308.2-338.2	12.0-30.0	218.951	-0.0975	0.5679	0.9644	0.6782	1.009	13	0.6172	0	3.22	1.902	0.0000211-0.00108	[488]
73	Lacidipine	313-333	12.0-36.0	201.0299	0.2281	1.0778	0.9691	0.9463	0.9087	12	0.6183	0.3455	3.55	2.323	0.0000018-0.0000572	[489]
74	Lamotrigine	318-348	12.16-35.46	261.4388	0.1705	0.841	0.7891	1.1977	0.7752	13	0.5935	1.2079	1.68	2.856	0.0000003-0.0000059	[463]
75	Lansoprazole	308.2-338.2	12.0-36.0	214.9262	0.2991	1.0793	0.9955	0.7079	0.8571	17	0.5436	0.4243	2.23	3.139	0.000012-0.000736	[490]
76	Letrozole	308-348	12.2-35.5	232.1377	0.0055	0.7349	1.1631	1.0866	0.5422	19	0.5266	0	2.78	2.455	0.000001-0.000083	[439]
77	Levonorgestrel	308-338	12.2-35.5	180.7864	0.0366	0.9999	1.0594	0.9834	0.7736	3	0.6166	0.562	3.55	1.391	0.0000004-0.0000029	[491]
78	Loratadine	308.15-338.15	12.0-27.0	208.6064	0.2117	1.201	0.9043	1.2514	0.6893	14	0.6021	0	3.33	1.68	0.0000045-0.0013	[117]
79	Losartan Potassium	308-338	12.0-27.0	206.5251	0.1536	0.9866	0.9161	0.8590	0.5119	24	0.5657	0.4617	3	3.14	0.0000203-0.0001875	[492]
80	Lovastatine	308-348	12.16-35.46	160.7995	0.1706	0.9191	1.0663	1.048	0.8986	4	0.6266	0.3972	3.55	2.21	0.000016-0.000114	[33]
81	Mandelic Acid	308.15-328.15	10.1-23.06	183.1658	0.1404	1.105	1.1004	1.1027	0.6947	7	0.5242	1.3896	2.01	1.055	0.000027-0.0029	[493]
82	Medroxyprogesterone Acetate	308-348	12.2-35.5	191.7921	0.1513	0.9654	1.119	0.8588	0.8466	4	0.6417	0	3.66	1.783	0.000016-0.009	[467]

83	Mefenamic Acid	308.15-338.15	16-40	213.7418	0.263	0.9038	1.1078	0.9024	0.6543	13	0.5776	1.0147	2.78	1.77	0.0000831-0.00594	[494]
84	Megestrol Acetate	308-338	12.2-35.5	199.9641	0.1541	0.9915	1.1091	0.8883	0.8852	5	0.6343	0	3.66	1.898	0.0000029-0.0000871	[490]
85	Meloxicam Sodium Salt	303-323	14.9-25.5	252.8421	0.229	1.4026	1.1312	1.4267	0.933	15	0.5583	0.4982	2.01	2.973	0.00000441-0.0000128	[495]
86	Metaxalone	308.2-328.2	11.95-24	172.6197	0.2069	0.8954	1.2285	0.9023	0.9941	7	0.5905	0.5441	2.34	1.292	0.00000058-0.000148	[454]
87	Methimazole	308-348	12.2-35.5	183.7796	0.0817	0.6811	1.2079	1.1281	1.0643	2	0.6895	0.4563	1.57	1.609	0.0000054-0.00019	[496]
88	Methocarbamol	308.2-328.2	11.99-22.05	162.5395	-0.0777	1.2056	1.0348	1.2422	1.2814	7	0.5739	1.1066	2.23	2.348	0.000028-0.000489	[454]
89	Methylparaben	308-348	12.2-35.5	206.4829	0.0594	1.1821	1.1733	1.6055	0.939	7	0.5521	0.6907	2.01	1.282	0.000108-0.00121	[446]
90	Metronidazole Benzoate	308-348	12.2-35.5	196.6955	0.1749	1.2717	0.9401	1.264	0.6216	13	0.5352	0	2.12	1.87	0.00007-0.00455	[445]
91	Minoxidil	308-338	12.0-27.0	159.2087	0.1781	0.8148	1.0014	0.9025	0.4671	4	0.5653	1.870	1.79	1.855	0.0000027-0.0000339	[497]
92	Nabumetone	308.2-328.2	10.0-22.0	200.5738	-0.0591	0.7723	1.0898	0.7202	1.445	12	0.5996	0	2.89	1.527	0.0000393-0.00268	[140]
93	Naproxen	308-348	12.2-35.5	216.5803	0.1069	0.8528	1.0591	1.012	0.8845	12	0.5767	0.6465	2.67	1.398	0.00001-0.0002	[445]
94	Nicotinic Acid	313.15-373.15	8.0-30.2	195.7393	0.187	0.8557	1.1989	1.3406	0.8178	7	0.5124	0.8023	1.79	1.107	0.00000031-0.0000104	[498]
95	Nifedipine	313-333	12.0-36.0	244.3145	0.1565	0.9267	1.0163	1.0799	0.6539	11	0.5654	0.4034	2.45	2.085	0.0000023-0.0000705	[488]
96	Niflumic Acid	313-333	12.0-36.0	222.924	0.4391	0.7966	1.131	1.0907	0.6016	13	0.4989	1.2355	2.12	2.075	0.00000163-0.000021	[451]
97	Nimodipine	313-333	10.0-36.0	208.5713	0.1133	1.0184	0.9741	0.9504	0.704	11	0.5844	0.3696	2.78	2.231	0.0000006-0.0000423	[499]
98	Nitrendipine	308-328	8.0-20.0	227.3893	0.1785	1.0325	0.944	1.0861	0.6581	11	0.5708	0.3883	2.56	2.085	0.000001-0.0000632	[500]
99	Norfloxacin	313.15-323.15	9.94-30.33	197.5117	-0.0077	0.6611	0.8374	0.7473	0.6801	9	0.5566	0.9682	2.45	2.204	0.00000122-0.0000244	[501]
100	Ofloxacin	318.15-328.15	10.11-30.33	205.1238	-0.0528	0.7616	0.851	0.7801	0.667	9	0.5484	0.9501	2.45	2.352	0.00000039-0.00000134	[501]
101	O-Nitrobenzoic Acid	308-328	12.0-36.0	222.9693	0.2922	1.0443	1.0917	1.2823	0.7149	8	0.4816	0.7762	1.68	1.124	0.00000152-0.0000356	[502]
102	Oxcarbazepine	313-333	12.0-27.0	235.6584	0.2519	0.9417	0.8755	0.988	0.9602	14	0.5351	0.606	2.67	1.82	0.00000011-0.0000268	[503]
103	Oxymetholone	308-328	12.1-30.5	182.579	0.0218	0.7743	1.2362	0.7991	0.7801	2	0.6394	0.6042	3.44	1.525	0.000000016-0.000000149	[107]
104	Paclitaxel	310.2-329	14.09-34.69	212.3053	0.0987	1.2584	1.0523	1.1885	0.9297	25	0.5682	0.1712	4.98	5.288	0.00000206-0.00000632	[434]
105	Paeonol	313.2-333.2	8.6-14.3	215.1309	-0.0969	0.7256	1.0566	0.7984	1.3742	7	0.5829	0.6174	2.12	1.101	0.00388-0.0132	[504][504]
106	Penicillin G	313.15-333.15	10.0-35.0	199.3197	0.2017	1.1924	1.2814	1.0114	0.9273	9	0.5943	0.9122	2.45	2.761	0.0000042-0.0000633	[142]
107	Penicillin V	314.85-334.85	8.0-28.05	198.3949	0.1103	1.2455	1.2529	1.0672	0.9342	9	0.5865	0.9655	2.34	2.844	0.0000545-0.000576	[505]
108	Pentoxifylline	308.15-328.15	11.98-24	186.6443	0.1796	1.1029	1.0849	1.2261	1.0568	8	0.5959	0	2.12	1.467	0.00003-0.00137	[459]
109	Phenazopyridine	308-348	12.2-35.5	191.539	0.2091	0.9277	1.0141	1.3192	0.989	13	0.5336	1.1849	2.12	2.298	0.0000044-0.000202	[496]
110	Phenylephrine Hydrochloride	308.15-338.15	16-40	162.8124	-0.1167	1.0426	1.1078	1.0227	1.0123	6	0.5823	1.5194	2.12	1.23	0.000101-0.00289	[506]

111	Pholcodine	308-338	12.0-27.0	191.6056	-0.1771	0.7784	1.0073	0.8155	0.8385	7	0.5879	0.2687	3.33	2.739	0.000206-0.000593	[507]
112	Pindolol	298-318	8.0-27.5	156.2058	-0.1023	1.0287	1.003	1.0614	1.1802	10	0.6006	1.1534	2.56	1.474	0.000027-0.000224	[508]
113	Piracetam	308.15-328.15	12.0-24.0	137.1858	0.1394	1.173	1.1344	1.227	1.1751	2	0.5669	0.6283	1.68	1.69	0.0000075-0.0000373	[459]
114	Piroxicam	308.15-338.15	16-40	239.6693	0.204	1.3062	0.9891	1.4006	0.9136	16	0.5346	0.9707	2.23	3.344	0.0000117-0.000512	[509]
115	Prednisolone	308.15-373.15	8.24-22.65	195.318	-0.0136	0.797	1.1443	0.8253	0.7194	4	0.5939	0.7067	3.22	2.228	0.000000001-0.000000728	[465]
116	Prednisone	308.15-373.15	8.24-22.65	201.6644	0.0017	0.8339	1.1362	0.848	0.7567	5	0.5894	0.6249	3.22	2.416	0.000000001-0.000000893	[465]
117	Propranolol	308-348	12.2-35.46	165.2954	-0.0852	1.123	0.9722	1.1398	1.1646	11	0.6034	0.7016	2.89	1.353	0.0000358-0.0024	[495]
118	Protocatechuic Acid	313-333	20-40	210.4681	0.117	1.2791	1.2192	1.4321	0.848	7	0.4983	2.1877	1.79	1.63	0.000000152-0.000000723	[457]
119	Pyridin-4-Amine	308.2-328.2	10.0-24.0	160.3554	-0.1414	0.6104	1.2343	0.5937	1.2442	6	0.5477	0.5606	1.79	1.184	0.0000209-0.000207	[510]
120	Pyrocatechol	308.15-338.15	12.0-40.0	173.2791	-0.021	1.5527	1.1495	1.6125	0.8387	6	0.5254	1.4441	1.9	1.273	0.00066-0.00394	[511]
121	Quetiapine Hemifumarate	308-338	12.0-27.0	193.8001	-0.1956	0.7809	0.8693	0.7208	0.7673	13	0.5999	0.6394	3.11	2.593	0.000003-0.00000903	[512]
122	Quinine	323.15-343.15	12.5-27.5	187.7316	-0.1334	1.0275	1.042	0.9547	0.7903	12	0.5857	0.1959	3.22	1.916	0.000012-0.0000504	[513]
123	Repaglinide	308-338	12.0-27.0	171.4509	0.1245	0.7644	0.9959	0.7364	0.7723	14	0.6052	0.8312	3.77	2.676	0.0000029-0.0000952	[514]
124	Rosuvastatin	308-348	12.16-35.46	190.1097	0.1763	0.8297	1.0194	0.8149	1.1091	17	0.576	1.4526	2.67	3.247	0.000003-0.000244	[33]
125	Salicylamide	308.2-328.2	10.0-22.0	196.9071	0.1322	0.8218	1.0996	1.0555	1.0942	7	0.5242	1.32	1.9	1.403	0.00000845-0.00021	[140]
126	Salicylic Acid	313-333	10.0-35.0	202.4417	0.1948	0.8128	1.1015	1.0632	0.883	7	0.5112	1.4675	1.9	0.888	0.0000855-0.000659	[515]
127	Sertraline Hydrochloride	308-338	12.0-30.0	209.0308	-0.0491	0.9301	0.9282	0.7247	0.8609	12	0.6439	0.2174	3	1.515	0.0000053-0.000093	[516]
128	Simvastatine	308-348	12.16-35.46	164.5879	0.1765	0.978	1.1563	1.1741	0.8803	4	0.6365	0.3912	3.66	2.17	0.000002-0.000535	[33]
129	Sinapic Acid	313-333	20-20	224.6376	-0.0717	0.8698	1.0916	0.8907	0.8611	8	0.558	1.3686	2.12	1.816	0.000000007-0.000000704	[457]
130	Sodium Valproate	308.15-338.15	12.0-27.0	123.5955	0.2823	1.2126	0.9490	1.3502	0.6349	1	0.6314	0	2.12	0.222	0.000005-0.000371	[517]
131	Sorafenib Tosylate	308-338	12.0-27.0	224.3504	0.3286	1.079	0.9818	1.0898	0.727	20	0.5379	1.6217	2.56	3.495	0.00000068-0.0000126	[518]
132	Spironolactone	308-338	16-40	192.5641	0.1583	0.7561	1.1411	0.7154	0.7552	4	0.6416	0	3.55	2.707	0.0000625-0.00501	[519]
133	Sulfabenzamide	308-338	12.0-27.0	210.4415	0.34804	1.1464	1.1258	1.3454	1.0015	15	0.5321	1.0364	2.23	2.873	0.000001-0.00004406	[242]
134	Sulindac	308.15-338.15	16-40	237.552	0.116	0.6604	1.019	0.6298	1.0402	16	0.5733	0.6175	3.11	3.076	0.0000365-0.00869	[520]
135	Sunitinib Malate	308-338	12.0-27.0	194.3161	0.0706	1.0411	0.9237	0.8549	0.9051	14	0.5958	1.1794	3.11	3.02	0.000005-0.0000856	[521]
136	Tamsulosin	308-338	12.0-27.0	180.7768	0.05181	0.8715	0.9336	0.8792	0.9643	14	0.5956	0.7457	2.78	2.769	0.0000018-0.0001013	[522]
137	Teriflunomide	308-338	12.0-27.0	225.4198	0.4297	0.6198	1.1428	0.9912	0.7796	6	0.5129	1.2336	2.01	2.533	0.0000804-0.0005428	[523]
138	Tetramethylpyrazine	318-338	10.0-30.0	198.2682	0.2432	1.6602	1.0242	1.9067	1.9052	6	0.6894	0	2.12	0.725	0.0022-0.131	[524]

139	Theobromine	313-353	19.3-34.5	225.0545	0.1779	0.9163	1.2139	1.102	1.0652	7	0.5565	0.647	1.57	1.013	0.0000088-0.0000471	[448]
140	Theophylline	313-353	19.9-34.9	226.2795	0.1779	1.1128	1.1859	1.7682	0.9489	7	0.5565	0.5649	1.57	1.027	0.0000103-0.0000332	[448]
141	Thymidine	308-328	10.0-30.0	146.6929	0.3309	0.8811	1.1562	1.372	1.0501	2	0.5667	1.147	1.57	0.347	0.0000012-0.000008	[434]
142	Tolbutamide	313.15-353.15	10.0-30.0	167.1545	0.4476	1.0151	1.0224	1.0788	0.9488	9	0.5981	0.9488	2.12	1.876	0.0000166-0.000405	[456]
143	Tolfenamic Acid	313-333	12.0-36.0	225.3529	0.2001	0.9252	1.0254	1.0827	0.6526	13	0.585	1.0134	2.56	1.889	0.00000161-0.0000181	[238]
144	Triamterene	308-338	12.0-27.0	227.9452	0.4132	0.8020	0.8863	0.7956	0.8199	18	0.5214	1.8105	2.01	3.463	0.000003-0.000289	[525]
145	Triclocarban	313.2-333.2	10.93-38.96	230.821	0.1677	1.1382	1.0555	1.0755	1.0864	13	0.6237	0.9365	2.23	2.187	0.00009-0.00087	[462]
146	Vitamin E Acetate	308.15-328.15	8.41-14.19	159.6518	0.2274	1.1469	1.136	1.1251	0.992	7	0.698	0	4.54	1.003	0.000139-0.000726	[526]
147	Warfarin	308.2-328.2	10.0-18.0	218.4587	0.1384	1.0262	0.9961	1.0387	1.045	15	0.5501	0.3817	3.11	2.927	0.000000148-0.00000539	[527]
148	Zopiclone	313-333	10.0-25.0	205.2551	0.0835	1.0419	1.0251	1.1723	0.7797	14	0.5679	0	2.23	2.794	0.0000015-0.0000219	[499]

Table. A2. Summary of Drugs and temperature, pressure and solubility ranges, the 11th QSPR desriptors and their refrences contained in the database used in **chapter III**

No.	Nom	T(K)	P(Mpa)	AATS3v	MATS2e	GATS4c	GATS3v	GATS4e	GATS3s	Nbondsm	AVP-0	Shbd	Mlogp	MLFER_S	Solubility Range (Mole Fraction)	Ref.
1	4-Aminoantipyrine	308.2-328.2	10.03-22.04	208.4654	0.1136	1.418	0.9713	1.4709	0.7436	8	0.5844	0.5395	2.23	1.124	0.000005-0.00036	[432]
2	4-Aminosalicylic Acid	308-328	11.0-21	207.699	0.1805	0.6322	1.2279	0.8292	0.9395	7	0.5101	2.0166	1.79	1.284	0.0000035-0.0000359	[433]
3	5-Fluorouracil	313-328	8.0-20.0	184.356	0.0791	0.9297	1.0804	1.5854	1.1499	3	0.4746	1.4681	1.35	0.495	0.0000013-0.00000525	[434]
4	10-Undecenoic Acid	308-333	10.0-18.0	114.8072	0.2571	0.6702	0.9895	0.7065	0.7811	2	0.6382	0.7291	2.45	0.548	0.0007-0.0174	[435]
5	All-Trans Retinoic Acid	308.2-328.2	10.2-31.2	182.234	0.2772	0.5449	1.3259	0.384	0.6936	6	0.6564	0.71	3.44	0.983	0.00000019-0.0000184	[436]
6	Amiodarone Hydrochloride	313.2-342.2	12.0-30.0	201.279	-0.0381	1.1632	0.7574	1.213	0.8687	17	0.7308	0	3.55	2.451	0.0000251-0.00101	[437]
7	Amlodipine Besylate	308-338	12.0-27.0	208.449	0.1654	1.1063	0.9235	1.2394	0.7368	12	0.5850	1.4976	3	4.24	0.000004146-0.0000022996	[438]
8	Anastrozole	308-348	12.2-35.5	211.0515	0.1176	0.8892	1.394	0.999	0.7343	13	0.6059	0	2.78	2.518	0.0000036-0.00038	[439]
9	Anthraquinone	308-338	10.0-34.0	250.8532	-0.0011	1.3986	0.8715	1.6037	0.8158	1	0.5272	0.0000	2.78	2.023	0.0000047-0.0000546	[528]
10	Aprepitant	308.15-338.15	12.0-33	206.247	0.2919	0.8433	1.1277	0.9294	0.5845	14	0.5178	1.2461	2.45	1.684	0.0000045-0.0000767	[440]
11	Aspirin	308.15-328.15	12.0-25	208.2891	0.2894	0.826	0.9011	0.9965	1.0858	8	0.537	0.7462	2.01	1.341	0.000063-0.000347	[441]
12	Atorvastatine	308-348	12.16-35.46	191.5763	0.0813	0.8363	0.9608	0.7306	0.9913	27	0.564	1.596	4.21	3.845	0.000001-0.00145	[33]
13	Atropine	308-348	12.2-35.5	165.4072	0.0454	0.762	1.0509	0.9386	0.7612	7	0.5925	0.4974	2.89	1.698	0.00006-0.00167	[442]
14	Azadiracthin	308.15-333.15	10.0-26.0	359.0470	0.0534	0.9164	0.9654	1.0049	0.0000	6	0.5766	0.0000	3.55	3.851	0.0000012-0.00000149	[529]
15	Azathioprine	313-338	12.0-27.0	262.0937	0.0496	0.9303	1.1366	0.8745	0.575	16	0.5425	0.5242	1.35	2.294	0.0000027-0.0000183	[443]
16	Azelaic Acid	313.15-333.15	10.0-30.0	118.5903	0.2997	0.6753	1.0151	0.7861	0.8502	2	0.5892	1.4901	2.01	0.889	0.00000042-0.0000101	[444]
17	Azithromycin	308-318	12.2-35.5	169.1742	0.0200	1.2349	1.2005	1.2611	1.1100	1	0.6574	0.2499	4.1	3.2500	0.000069-0.000273	[159]
18	Benzocaine	308-348	12.2-35.5	177.319	0.1807	1.3478	0.9647	1.6374	0.8585	7	0.5758	0.5452	2.12	1.418	0.00027-0.0121	[445]
19	Bisacodyl	308-348	12.2-35.5	215.9504	0.2447	0.7089	0.9187	0.6682	1.2092	20	0.5587	0	3.33	2.353	0.000009-0.000583	[446]
20	Budesonide	338-358	21.3-38.5	191.9097	0.0289	0.8834	1.0587	0.9264	0.7307	4	0.6023	0.464	3.55	2.443	0.00000593-0.0000291	[447]
21	Busulfan	308-338	12.0-40.0	135.6561	0.3507	0.9470	0.9139	1.1212	1.1982	4	0.6073	0.0000	1.24	0.967	0.0000327-0.0000865	[530]
22	Caffeine	313-353	19.9-34.9	232.0621	0.1562	1.104	1.2438	1.5655	0.9426	7	0.5844	0	1.68	0.996	0.000283-0.00113	[448]
23	Capecitabine	308-348	15.2-35.4	159.97	0.1123	1.2248	0.9445	1.0837	0.8026	4	0.5605	1.6171	2.01	1.522	0.0000027-0.000159	[449]
24	Captopril	308.2-328.2	10.0-18.5	156.3398	0.1936	1.3444	0.9858	1.5887	0.8311	2	0.6454	0.9591	1.90	1.950	0.00000359-0.00000932	[531]
25	Carbamazepine	308-348	12.2-35.5	234.2513	0.4052	1.0654	0.8578	0.9715	0.8385	14	0.5392	0.5977	2.78	1.449	0.000003-0.000094	[442]
26	Cefixime Trihydrate	308-328	18.3-33.5	229.7306	0.1078	0.9696	0.9675	0.8859	0.9275	12	0.5578	2.3614	1.68	4.226	0.00000016-0.000000302	[107]

27	Cefuroxime Axetil	308-328	8.0-30.0	214.9035	0.122	1.117	0.9133	0.9549	1.0724	12	0.5616	1.1698	2.01	3.944	0.00000022-0.0000112	[450]
28	Celecoxib	323.3-343.2	15-30	228.6141	0.4423	0.7527	1.092	0.8058	0.6159	21	0.528	0.5598	2.34	2.439	0.0000152-0.000152	[451]
29	Cetirizine	308.15-338.15	16-40	182.0835	-0.0552	0.6153	1.0493	0.4804	0.7265	13	0.5931	0.7278	3.11	2.218	0.0000105-0.00492	[452]
30	Chloramphenicol	308-338	10.0-26.0	196.2615	0.0171	1.1274	0.992	1.0764	0.8024	8	0.5717	1.3194	1.68	2.107	0.0000041-0.0000099	[453]
31	Chlormezanone	308.2-328.2	11.91-23.99	207.1075	0.1159	1.2947	1.0358	1.3552	1.1984	9	0.6011	0	2.01	2.799	0.000005-0.000545	[454]
32	Chloroquine	318-338	12.0-40.0	167.0068	-0.0458	0.8628	0.8887	0.7115	0.9818	11	0.6604	0.1840	3.00	1.721	0.0000164-0.000892	[532]
33	Chlorpheniramine Maleate	308-338	20-40	188.5535	-0.1532	0.7767	1.1291	0.7621	1.1446	12	0.6388	0	2.89	1.494	0.0000154-0.000426	[455]
34	Chlorpropamide	313.15-353.15	10.0-30.0	181.153	0.3666	1.139	0.9322	1.2696	0.9861	9	0.5996	0.9782	1.79	2.033	0.00000229-0.0000722	[456]
35	Chrysin	313-333	20-40	233.324	-0.0366	0.5951	0.9708	0.6932	1.0668	14	0.5173	1.175	2.67	1.975	0.000000018-0.000000113	[457]
36	Cinildipine	313-333	12.0-36.0	214.7053	0.0962	1.0382	1.0221	1.0327	0.6622	18	0.5612	0.367	3.44	2.796	0.0000012-0.0000399	[458]
37	Cinnarizine	308.15-328.15	13.97-24	182.4832	-0.2551	0.9535	1.052	1.0177	0.9352	19	0.5829	0	4.1	2.112	0.0000029-0.000205	[459]
38	Ciprofloxacin	313-333	17-36	199.1704	-0.004	0.666	0.7763	0.7512	0.6746	9	0.5453	0.9428	2.56	2.315	0.0000000265-0.000000189	[460]
39	Clarithromycin	308-318	12.2-35.5	174.3024	0.0110	1.2742	1.1572	1.3217	1.0718	2	0.6519	0.4250	4.1	3.4700	0.000073-0.000528	[159]
40	Clemastine Fumarate	308-338	12.0-27.0	178.4961	0.1171	0.8759	1.0630	0.8025	0.6802	9	0.5988	1.7740	3.44	2.289	0.0000161-0.0000941	[461]
41	Climbazole	313.2-333.2	10.55-39.89	192.2391	0.0452	1.4104	1.325	1.5748	1.1387	12	0.6231	0	2.56	1.846	0.00062-0.00488	[462]
42	Clindamycin	308-318	12.2-35.5	162.8046	-0.0556	1.2445	1.1051	1.1181	0.8933	1	0.6662	1.1306	2.45	2.3640	0.000056-0.000898	[159]
43	Clobetasole Propionate	308-348	12.2-35.5	205.2235	0.0599	1.0832	1.0154	1.2446	0.9404	5	0.6211	0.0631	3.44	2.316	0.0000001-0.0000035	[151]
44	Clofenamic Acid	313-333	12.0-36.0	215.9826	0.1819	0.8208	1.0687	0.8842	0.6788	13	0.5651	1.0603	2.45	1.913	0.00000154-0.0000354	[238]
45	Clotrimazole	308-348	12.2-35.5	236.5251	-0.0031	0.954	0.8858	1.0874	0.818	23	0.5737	0	3.55	2.081	0.0000002-0.000107	[147]
46	Clozapine	318-348	12.16-35.46	204.9515	-0.1001	1.0422	0.9839	0.9216	0.9495	13	0.6019	0.306	2.89	2.681	0.0000012-0.0000419	[463]
47	Codeine	308-348	12.2-35.5	212.2692	-0.1339	0.9069	1.0213	0.9318	0.7427	7	0.5886	0.3164	3	2.156	0.00001-0.000797	[442]
48	Cortisone Acetate	308.15-373.15	8.24-22.65	195.8706	0.0818	0.8991	1.1542	0.8232	0.8992	5	0.6018	0.0819	3.33	2.52	0.00000011-0.00000414	[464]
49	Coumarin-7	308-338	9.0-33.0	208.9521	0.1860	0.8908	0.8196	0.8943	0.7036	18	0.5697	0.4406	3.11	2.9	0.0000415-0.0001009	[465]
50	Curcumin	308.15-328.15	8.0-20.0	214.8248	-0.1104	0.9894	1.1062	0.9658	1.1668	16	0.5559	1.1660	3.11	3.239	0.0000000182-0.000000037	[533]
51	Cyproheptadine	308-338	16-40	212.7801	-0.2504	0.4639	0.9685	0.5417	0.8336	14	0.5931	0	3.66	1.631	0.0000335-0.00309	[466]
52	Cyproterone Acetate	308-348	12.2-35.5	209.6532	0.13	1.0215	1.0128	0.9251	0.8609	5	0.6325	0	3.55	2.127	0.000013-0.000261	[467]
53	Dasatinib Monohydrate	308-338	12.0-27.0	197.2930	0.0390	0.6595	0.9740	0.6940	0	9	0.6052	1.3254	2.56	4.486	0.000000454-0.000009078	[468]
54	Decitabine	308-338	12.0-40.0	167.7847	0.0254	0.7633	1.0454	0.9079	0.8803	3	0.5220	1.8191	1.46	1.982	0.0000284-0.00107	[534]

55	Dexamethasone	308.15-328.15	10-35.07	203.3863	-0.0081	0.9111	1.1082	1.1472	0.8336	4	0.5933	0.6436	3.22	2.182	0.00000119-0.00000281	[469,470]
56	Diazepam	308-348	12-35.07	238.9165	0.0302	1.2376	0.9297	1.1518	0.9749	14	0.5881	0	2.78	2.109	0.000039-0.000582	[442]
57	Diclofenac Acid	308.15-338.15	12.0-40.0	224.1334	0.1507	1.0423	1.0117	0.8957	0.9816	13	0.5985	0.9706	2.45	2.1	0.0000234-0.00198	[471]
58	Diflunisal	308.2-328.2	9.1-24.6	238.7639	0.0631	0.7155	1.0607	0.8642	0.9856	14	0.5012	1.3249	2.34	1.394	0.000000544-0.00000807	[472]
59	Docetaxel	308-348	15.2-35.4	205.6952	0.121	1.2119	1.1162	1.1127	0.9009	18	0.5842	0.2923	4.54	3.951	0.000037-0.000702	[107]
60	Dutasteride	308-348	12.2-35.5	195.6823	0.4128	0.6627	1.1882	0.5938	0.4119	9	0.5567	0.6906	3.33	2.185	0.0000001-0.000161	[473]
61	Eflucimibe	308.15-318.15	9.33-30.15	164.1925	0.1269	1.3504	1.0054	1.2968	0.8924	13	0.6636	0.6746	4.21	2.3120	0.000000014-0.000001567	[535]
62	Enrofloxacin	313-333	12.0-36.0	196.8204	0.0006	0.623	0.8051	0.6901	0.6612	9	0.567	0.6212	2.78	2.502	0.000000022-0.0000056	[461]
63	Erythromycin	308-318	12.2-35.5	172.0505	0.0241	1.2691	1.1696	1.3206	1.1107	2	0.6458	0.3604	3.99	3.541	0.000054-0.000312	[159]
64	Esomeprazole	308.2-338.2	12.0-27.0	232.2006	-0.079	1.3285	1.0686	1.3089	1.1303	17	0.6042	0.4044	2.56	3.406	0.000011-0.00091	[474]
65	Ethosuximide	313.15-328.15	9.37-14.8	161.7181	0.2617	0.8615	1.3069	1.0784	1.0598	2	0.623	0.548	1.9	0.695	0.00345-0.00993	[475]
66	Exemestane	308-348	12.2-35.5	197.4053	0.0384	0.8557	1.1435	0.8794	0.7824	5	0.6146	0	3.44	1.586	0.0000126-0.00188	[439]
67	Fenoprofen	308-338	12.0-40.0	193.9032	0.1875	1.0727	1.0149	1.3727	0.9531	13	0.5576	0.6514	2.78	1.482	0.0000201-0.042	[536]
68	Finasteride	308-348	12.2-35.5	175.3167	0.1983	0.8732	1.2519	0.7794	0.7831	3	0.6518	0.3783	3.55	1.942	0.000336-0.00332	[473]
69	Fluoxetine Hydrochloride	308.15-338.15	16-40	182.7174	0.3698	0.9534	1.0135	1.1184	0.4801	12	0.5559	0.3369	2.78	1.236	0.0000265-0.000812	[476]
70	Flurbiprofen	303-323	8.9-24.5	210.3439	0.1535	0.9121	1.019	1.1564	0.8865	14	0.5516	0.6615	2.78	1.388	0.0000167-0.000197	[477]
71	Flutamide	308-348	12.2-35.5	197.4672	0.4344	1.2309	1.1323	1.1207	0.6545	8	0.5392	0.529	1.9	1.863	0.0000046-0.000509	[473]
72	Fluvastatin	308-348	12.16-35.46	192.8681	0.0598	0.6283	0.9066	0.609	1.0013	19	0.5664	1.4356	3.44	2.486	0.000005-0.000601	[33]
73	Fluvoxamine Maleate	308-338	24.0-40.0	165.7628	0.2603	0.7365	0.9666	1.1210	0.5065	7	0.5694	0.5203	2.34	1.068	0.0000123-0.000445	[537]
74	Gabapentin	308-338	16-40	141.7961	0.2009	0.8215	1.2575	0.5609	0.6744	1	0.6152	1.0533	2.12	0.829	0.0000897-0.00736	[478]
75	Galantamine	308-338	12.0-27.0	206.4365	-0.1472	0.8419	1.0653	0.8497	0.9104	7	0.5954	0.4141	2.89	2.151	0.000006-0.00266	[479]
76	Gambogic Acid	308.15-328.15	10.0-30.0	214.8228	0.0998	1.0269	1.0774	1.1301	0.8507	14	0.6086	0.6876	4.76	2.935	0.00000163-0.0000226	[480]
77	Gatifloxacin	313-333	12.0-36.0	203.8424	-0.0339	0.8266	0.8102	0.7465	0.7	9	0.5662	0.8703	2.67	2.313	0.000000106-0.0000016	[461]
78	Gemfibrozil	308.2-328.2	10.01-22.02	173.1881	0.1698	1.2968	1.3523	1.7007	0.8193	7	0.6453	0.6041	2.78	0.992	0.00000294-0.00419	[481]
79	Gliclazide	308.2-328.2	12.0-18.5	163.8889	0.3342	0.9973	1.1953	0.8315	0.7755	9	0.5811	0.9325	2.34	1.758	0.000000126-0.00000501	[531]
80	Ibuprofen	308-318	8.0-13.0	169.2564	0.2781	0.8134	1.1929	0.9692	0.7951	7	0.6351	0.6548	2.67	0.832	0.0000115-0.00326	[434]
81	Imatinib Mesylate	308.2-338.2	12.0-27.0	189.5516	0.2392	0.7186	1.1346	0.5797	0.9800	15	0.5728	1.6131	3.44	6.192	0.0000153-0.0002235	[482]
82	Imipramine Hcl	313-323	30-50	188.3197	-0.1823	1.0557	1.047	1.2625	1.0154	12	0.6213	0	3.33	1.254	0.0000051-0.0000099	[483]

83	Ipriflavone	308.2-328.2	10.0-30	209.4584	-0.0323	0.875	0.8758	0.9364	1.034	14	0.5713	0	3.11	1.921	0.0000054-0.00017	[484]
84	Isoniazid	308-313	13-18.5	188.1176	-0.0365	1.2133	1.0729	1.7202	0.87	7	0.5242	1.3234	1.68	1.447	0.00000053-0.00000605	[485]
85	Ketoconazole	308-348	12.2-35.5	196.7504	-0.059	1.1118	0.9759	1.0464	1.1941	18	0.5969	0	3.22	3.481	0.0000005-0.000175	[147,486]
86	Ketoprofen	308.15-338.15	16-40	208.2778	0.19322	1.0397	1.0351	1.42	0.8396	14	0.5545	0.6399	2.89	1.768	0.0000221-0.000712	[487]
87	Ketotifen Fumarate (Ktf)	308.2-338.2	12.0-30.0	218.951	-0.0975	0.5679	0.9644	0.6782	1.009	13	0.6172	0	3.22	1.902	0.0000211-0.00108	[488]
88	Lacidipine	313-333	12.0-36.0	201.0299	0.2281	1.0778	0.9691	0.9463	0.9087	12	0.6183	0.3455	3.55	2.323	0.0000018-0.0000572	[489]
89	Laminarin	308.15-318.15	10.0-16.0	162.1015	-0.2057	1.0874	1.0252	1.0844	0.8183	0	0.5218	4.7297	1.68	3.409	0.00013-0.000341	[538]
90	Lamotrigine	318-348	12.16-35.46	261.4388	0.1705	0.841	0.7891	1.1977	0.7752	13	0.5935	1.2079	1.68	2.856	0.0000003-0.0000059	[463]
91	Lansoprazole	308.2-338.2	12.0-36.0	214.9262	0.2991	1.0793	0.9955	0.7079	0.8571	17	0.5436	0.4243	2.23	3.139	0.000012-0.000736	[490]
92	Letrozole	308-348	12.2-35.5	232.1377	0.0055	0.7349	1.1631	1.0866	0.5422	19	0.5266	0	2.78	2.455	0.000001-0.000083	[439]
93	Levonorgestrel	308-338	12.2-35.5	180.7864	0.0366	0.9999	1.0594	0.9834	0.7736	3	0.6166	0.562	3.55	1.391	0.0000004-0.0000029	[491]
94	Loratadine	308.15-338.15	12.0-27.0	208.6064	0.2117	1.201	0.9043	1.2514	0.6893	14	0.6021	0	3.33	1.68	0.0000045-0.0013	[117]
95	Losartan Potassium	308-338	12.0-27.0	206.5251	0.1536	0.9866	0.9161	0.8590	0.5119	24	0.5657	0.4617	3	3.14	0.0000203-0.0001875	[492]
96	Lovastatine	308-348	12.16-35.46	160.7995	0.1706	0.9191	1.0663	1.048	0.8986	4	0.6266	0.3972	3.55	2.21	0.000016-0.000114	[33]
97	Loxoprofen	308-338	12.0-40.0	167.9278	0.1830	0.9802	1.1247	1.2203	0.9109	8	0.5865	0.6429	2.78	1.453	0.0000135-0.00128	[539]
98	Mandelic Acid	308.15-328.15	10.1-23.06	183.1658	0.1404	1.105	1.1004	1.1027	0.6947	7	0.5242	1.3896	2.01	1.055	0.000027-0.0029	[493]
99	Medroxyprogesterone Acetate	308-348	12.2-35.5	191.7921	0.1513	0.9654	1.119	0.8588	0.8466	4	0.6417	0	3.66	1.783	0.000016-0.009	[467]
100	Mefenamic Acid	308.15-338.15	16-40	213.7418	0.263	0.9038	1.1078	0.9024	0.6543	13	0.5776	1.0147	2.78	1.77	0.0000831-0.00594	[494]
101	Megestrol Acetate	308-338	12.2-35.5	199.9641	0.1541	0.9915	1.1091	0.8883	0.8852	5	0.6343	0	3.66	1.898	0.0000029-0.0000871	[490]
102	Meloxicam Sodium Salt	303-323	14.9-25.5	252.8421	0.229	1.4026	1.1312	1.4267	0.933	15	0.5583	0.4982	2.01	2.973	0.00000441-0.0000128	[495]
103	Metaxalone	308.2-328.2	11.95-24	172.6197	0.2069	0.8954	1.2285	0.9023	0.9941	7	0.5905	0.5441	2.34	1.292	0.00000058-0.000148	[454]
104	Methimazole	308-348	12.2-35.5	183.7796	0.0817	0.6811	1.2079	1.1281	1.0643	2	0.6895	0.4563	1.57	1.609	0.0000054-0.00019	[496]
105	Methocarbamol	308.2-328.2	11.99-22.05	162.5395	-0.0777	1.2056	1.0348	1.2422	1.2814	7	0.5739	1.1066	2.23	2.348	0.000028-0.000489	[454]
106	Methylparaben	308-348	12.2-35.5	206.4829	0.0594	1.1821	1.1733	1.6055	0.939	7	0.5521	0.6907	2.01	1.282	0.000108-0.00121	[446]
107	Metronidazole Benzoate	308-348	12.2-35.5	196.6955	0.1749	1.2717	0.9401	1.264	0.6216	13	0.5352	0	2.12	1.87	0.00007-0.00455	[445]
108	Minoxidil	308-338	12.0-27.0	159.2087	0.1781	0.8148	1.0014	0.9025	0.4671	4	0.5653	1.870	1.79	1.855	0.0000027-0.0000339	[497]
109	Nabumetone	308.2-328.2	10.0-22.0	200.5738	-0.0591	0.7723	1.0898	0.7202	1.445	12	0.5996	0	2.89	1.527	0.0000393-0.00268	[140]
110	Naproxen	308-348	12.2-35.5	216.5803	0.1069	0.8528	1.0591	1.012	0.8845	12	0.5767	0.6465	2.67	1.398	0.00001-0.0002	[445]

111	Nicotinic Acid	313.15-373.15	8.0-30.2	195.7393	0.187	0.8557	1.1989	1.3406	0.8178	7	0.5124	0.8023	1.79	1.107	0.00000031-0.0000104	[498]
112	Nifedipine	313-333	12.0-36.0	244.3145	0.1565	0.9267	1.0163	1.0799	0.6539	11	0.5654	0.4034	2.45	2.085	0.0000023-0.0000705	[488]
113	Niflumic Acid	313-333	12.0-36.0	222.924	0.4391	0.7966	1.131	1.0907	0.6016	13	0.4989	1.2355	2.12	2.075	0.00000163-0.000021	[451]
114	Nimodipine	313-333	10.0-36.0	208.5713	0.1133	1.0184	0.9741	0.9504	0.704	11	0.5844	0.3696	2.78	2.231	0.0000006-0.0000423	[499]
115	Nitrendipine	308-328	8.0-20.0	227.3893	0.1785	1.0325	0.944	1.0861	0.6581	11	0.5708	0.3883	2.56	2.085	0.000001-0.0000632	[500]
116	Norfloxacin	313.15-323.15	9.94-30.33	197.5117	-0.0077	0.6611	0.8374	0.7473	0.6801	9	0.5566	0.9682	2.45	2.204	0.00000122-0.0000244	[501]
117	Ofloxacin	318.15-328.15	10.11-30.33	205.1238	-0.0528	0.7616	0.851	0.7801	0.667	9	0.5484	0.9501	2.45	2.352	0.00000039-0.00000134	[501]
118	O-Nitrobenzoic Acid	308-328	12.0-36.0	222.9693	0.2922	1.0443	1.0917	1.2823	0.7149	8	0.4816	0.7762	1.68	1.124	0.00000152-0.0000356	[502]
119	Oxcarbazepine	313-333	12.0-27.0	235.6584	0.2519	0.9417	0.8755	0.988	0.9602	14	0.5351	0.606	2.67	1.82	0.00000011-0.0000268	[503]
120	Oxymetholone	308-328	12.1-30.5	182.579	0.0218	0.7743	1.2362	0.7991	0.7801	2	0.6394	0.6042	3.44	1.525	0.000000016-0.000000149	[107]
121	Paclitaxel	310.2-329	14.09-34.69	212.3053	0.0987	1.2584	1.0523	1.1885	0.9297	25	0.5682	0.1712	4.98	5.288	0.00000206-0.00000632	[434]
122	Paeonol	313.2-333.2	8.6-14.3	215.1309	-0.0969	0.7256	1.0566	0.7984	1.3742	7	0.5829	0.6174	2.12	1.101	0.00388-0.0132	[504][504]
123	Penicillin G	313.15-333.15	10.0-35.0	199.3197	0.2017	1.1924	1.2814	1.0114	0.9273	9	0.5943	0.9122	2.45	2.761	0.0000042-0.0000633	[142]
124	Penicillin V	314.85-334.85	8.0-28.05	198.3949	0.1103	1.2455	1.2529	1.0672	0.9342	9	0.5865	0.9655	2.34	2.844	0.0000545-0.000576	[505]
125	Pentoxifylline	308.15-328.15	11.98-24	186.6443	0.1796	1.1029	1.0849	1.2261	1.0568	8	0.5959	0	2.12	1.467	0.00003-0.00137	[459]
126	Phenazopyridine	308-348	12.2-35.5	191.539	0.2091	0.9277	1.0141	1.3192	0.989	13	0.5336	1.1849	2.12	2.298	0.0000044-0.000202	[496]
127	Phenylephrine Hydrochloride	308.15-338.15	16-40	162.8124	-0.1167	1.0426	1.1078	1.0227	1.0123	6	0.5823	1.5194	2.12	1.23	0.000101-0.00289	[506]
128	Pholcodine	308-338	12.0-27.0	191.6056	-0.1771	0.7784	1.0073	0.8155	0.8385	7	0.5879	0.2687	3.33	2.739	0.000206-0.000593	[507]
129	Pindolol	298-318	8.0-27.5	156.2058	-0.1023	1.0287	1.003	1.0614	1.1802	10	0.6006	1.1534	2.56	1.474	0.000027-0.000224	[508]
130	Piracetam	308.15-328.15	12.0-24.0	137.1858	0.1394	1.173	1.1344	1.227	1.1751	2	0.5669	0.6283	1.68	1.69	0.0000075-0.0000373	[459]
131	Piroxicam	308.15-338.15	16-40	239.6693	0.204	1.3062	0.9891	1.4006	0.9136	16	0.5346	0.9707	2.23	3.344	0.0000117-0.000512	[509]
132	Prednisolone	308.15-373.15	8.24-22.65	195.318	-0.0136	0.797	1.1443	0.8253	0.7194	4	0.5939	0.7067	3.22	2.228	0.000000001-0.000000728	[465]
133	Prednisone	308.15-373.15	8.24-22.65	201.6644	0.0017	0.8339	1.1362	0.848	0.7567	5	0.5894	0.6249	3.22	2.416	0.000000001-0.000000893	[465]
134	Propranolol	308-348	12.2-35.46	165.2954	-0.0852	1.123	0.9722	1.1398	1.1646	11	0.6034	0.7016	2.89	1.353	0.0000358-0.0024	[495]
135	Protocatechuic Acid	313-333	20-40	210.4681	0.117	1.2791	1.2192	1.4321	0.848	7	0.4983	2.1877	1.79	1.63	0.000000152-0.000000723	[457]
136	Pyridin-4-Amine	308.2-328.2	10.0-24.0	160.3554	-0.1414	0.6104	1.2343	0.5937	1.2442	6	0.5477	0.5606	1.79	1.184	0.0000209-0.000207	[510]
137	Pyrocatechol	308.15-338.15	12.0-40.0	173.2791	-0.021	1.5527	1.1495	1.6125	0.8387	6	0.5254	1.4441	1.9	1.273	0.00066-0.00394	[511]
138	Quetiapine Hemifumarate	308-338	12.0-27.0	193.8001	-0.1956	0.7809	0.8693	0.7208	0.7673	13	0.5999	0.6394	3.11	2.593	0.000003-0.00000903	[512]

139	Quinine	323.15-343.15	12.5-27.5	187.7316	-0.1334	1.0275	1.042	0.9547	0.7903	12	0.5857	0.1959	3.22	1.916	0.000012-0.0000504	[513]
140	Repaglinide	308-338	12.0-27.0	171.4509	0.1245	0.7644	0.9959	0.7364	0.7723	14	0.6052	0.8312	3.77	2.676	0.0000029-0.0000952	[514]
141	Rosuvastatin	308-348	12.16-35.46	190.1097	0.1763	0.8297	1.0194	0.8149	1.1091	17	0.576	1.4526	2.67	3.247	0.000003-0.000244	[33]
142	Salicylamide	308.2-328.2	10.0-22.0	196.9071	0.1322	0.8218	1.0996	1.0555	1.0942	7	0.5242	1.32	1.9	1.403	0.00000845-0.00021	[140]
143	Salicylic Acid	313-333	10.0-35.0	202.4417	0.1948	0.8128	1.1015	1.0632	0.883	7	0.5112	1.4675	1.9	0.888	0.0000855-0.000659	[515]
144	Salsalate	308-338	12.0-40.0	222.2185	0.2053	0.9191	1.0095	1.1943	0.8228	14	0.5125	1.2943	2.45	1.721	0.0000377-0.00388	[540]
145	Sertraline Hydrochloride	308-338	12.0-30.0	209.0308	-0.0491	0.9301	0.9282	0.7247	0.8609	12	0.6439	0.2174	3	1.515	0.0000053-0.000093	[516]
146	Simvastatine	308-348	12.16-35.46	164.5879	0.1765	0.978	1.1563	1.1741	0.8803	4	0.6365	0.3912	3.66	2.17	0.000002-0.000535	[33]
147	Sinapic Acid	313-333	20-20	224.6376	-0.0717	0.8698	1.0916	0.8907	0.8611	8	0.558	1.3686	2.12	1.816	0.000000007-0.000000704	[457]
148	Sodium Valproate	308.15-338.15	12.0-27.0	123.5955	0.2823	1.2126	0.9490	1.3502	0.6349	1	0.6314	0	2.12	0.222	0.000005-0.000371	[517]
149	Sorafenib Tosylate	308-338	12.0-27.0	224.3504	0.3286	1.079	0.9818	1.0898	0.727	20	0.5379	1.6217	2.56	3.495	0.00000068-0.0000126	[518]
150	Spironolactone	308-338	16-40	192.5641	0.1583	0.7561	1.1411	0.7154	0.7552	4	0.6416	0	3.55	2.707	0.0000625-0.00501	[519]
151	Sulfabenzamide	308-338	12.0-27.0	210.4415	0.34804	1.1464	1.1258	1.3454	1.0015	15	0.5321	1.0364	2.23	2.873	0.000001-0.00004406	[242]
152	Sulindac	308.15-338.15	16-40	237.552	0.116	0.6604	1.019	0.6298	1.0402	16	0.5733	0.6175	3.11	3.076	0.0000365-0.00869	[520]
153	Sunitinib Malate	308-338	12.0-27.0	194.3161	0.0706	1.0411	0.9237	0.8549	0.9051	14	0.5958	1.1794	3.11	3.02	0.000005-0.0000856	[521]
154	Tamsulosin	308-338	12.0-27.0	180.7768	0.05181	0.8715	0.9336	0.8792	0.9643	14	0.5956	0.7457	2.78	2.769	0.0000018-0.0001013	[522]
155	Teriflunomide	308-338	12.0-27.0	225.4198	0.4297	0.6198	1.1428	0.9912	0.7796	6	0.5129	1.2336	2.01	2.533	0.0000804-0.0005428	[523]
156	Teriflunomide	308-338	12.0-27.0	225.4198	0.4297	0.6198	1.1428	0.9912	0.7796	6	0.5129	1.2336	2.01	2.533	0.000884-0.005428	
157	Tetramethylpyrazine	318-338	10.0-30.0	198.2682	0.2432	1.6602	1.0242	1.9067	1.9052	6	0.6894	0	2.12	0.725	0.0022-0.131	[524]
158	Theobromine	313-353	19.3-34.5	225.0545	0.1779	0.9163	1.2139	1.102	1.0652	7	0.5565	0.647	1.57	1.013	0.0000088-0.0000471	[448]
159	Theophylline	313-353	19.9-34.9	226.2795	0.1779	1.1128	1.1859	1.7682	0.9489	7	0.5565	0.5649	1.57	1.027	0.0000103-0.0000332	[448]
160	Thymidine	308-328	10.0-30.0	146.6929	0.3309	0.8811	1.1562	1.372	1.0501	2	0.5667	1.147	1.57	0.347	0.0000012-0.000008	[434]
161	Tolbutamide	313.15-353.15	10.0-30.0	167.1545	0.4476	1.0151	1.0224	1.0788	0.9488	9	0.5981	0.9488	2.12	1.876	0.0000166-0.000405	[456]
162	Tolfenamic Acid	313-333	12.0-36.0	225.3529	0.2001	0.9252	1.0254	1.0827	0.6526	13	0.585	1.0134	2.56	1.889	0.00000161-0.0000181	[238]
163	Tolmetin	308-338	12.0-40.0	211.9995	0.1457	0.6473	1.1113	0.7509	0.9596	13	0.5727	0.6931	2.67	1.873	0.00005-0.00259	[541]
164	Triamterene	308-338	12.0-27.0	227.9452	0.4132	0.8020	0.8863	0.7956	0.8199	18	0.5214	1.8105	2.01	3.463	0.000003-0.000289	[525]
165	Triclocarban	313.2-333.2	10.93-38.96	230.821	0.1677	1.1382	1.0555	1.0755	1.0864	13	0.6237	0.9365	2.23	2.187	0.00009-0.00087	[462]
166	Vitamin E Acetate	308.15-328.15	8.41-14.19	159.6518	0.2274	1.1469	1.136	1.1251	0.992	7	0.698	0	4.54	1.003	0.000139-0.000726	[526]

167	Warfarin	308.2-328.2	10.0-18.0	218.4587	0.1384	1.0262	0.9961	1.0387	1.045	15	0.5501	0.3817	3.11	2.927	0.000000148-0.00000539	[527]
168	Zopiclone	313-333	10.0-25.0	205.2551	0.0835	1.0419	1.0251	1.1723	0.7797	14	0.5679	0	2.23	2.794	0.0000015-0.0000219	[499]

Table A3. Summary of API and ILs, systems, rows and their references contained in the database used in **chapter IV**.

N°	Apis	Ils	T(K)	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	$N^{\circ}_{systems}$	N°_{rows}	Ref°
1	Lamotrigine	[OMIm][1] [BR][2]	293.2-313.2	-7,80E-06	2,63E+00	2,55E+01	8,06E-09	-1,43E-08	-7,47E-04	5,86E+00	-2,08E-04	2,71E+00	8,18E+00	1,13E+01	2,25E+00	1	5	[542]
2	Delafloxacin	[BMIm][3] [PF_6][4] [HMIm][5] [PF_6][4] [OMIm][1] [PF_6][4]	298.2-318.2	1,65E-06 -0.0827	2,5849-2.6310	17.5744-25.5105	4.31E-10-8.74E-5	9,6771	1,7430	1,37E-08	1,35E-13	3,0027	16,7270	12,2931	4,6922	3,0027	15	[543]
3	N-Acetyl-L-Cysteine	[EMIm][6] [OTF][7] [BMIm][3] [OTF][7] [BMIm][3] [NTF_2][8] [HMIm][5] [NTF_2][8] [DMIm][9] [NTF_2][8]	283.34-381.04	9,61E-06-0.0827	2.8549-2.6618	13.5141-29.5087	4,30E-10-8.74E-05	6,8590-15,0938	4,5427-5,2186	0,0244-1,3087	4.73E-09 -1,78E-06	1,0851	9,3299	4,9434	2,8335	5	56	[544]
4	Isoniazid	[BMIm][3] [BF_4][10] [BMIm][3] [OTF][7] [BMIm][3] [PF_6][4] [OMIm][1] [OTF][7] [DMIm][9] [OTF][7] [DDA][11] [NO_3][12] [C_2][13] [NO_3][12] [BA][14] [NTF_2][8] [EMIm][6] [NTF_2][8] [BMIm][3] [NTF_2][8] [HMIm][5] [NTF_2][8] [OMIm][1] [NTF_2][8] [DMIm][9] [NTF_2][8]	290.37-418.48	9.61E-06-0.1262	1.7862-2.7157	13.2625-44.0115	1.13E-06-0.6485	1.2885-15.0938	1.7409-7.2302	0.0029-2.7758	1.78E-06-0.0726	1,9326	6,9550	4,9696	2,9294	13	246	[545] [546] [482]
5	Pyrazine Carboxamide	[BMIm][3] [BF_4][10] [BMIm][3] [OTF][7] [BMIm][3] [PF6][4] [OMIm][1] [OTF][7] [DMIm][9] [OTF][7] [DDA][11] [NO_3][12] [C_2][13] [NO_3][12] [BA][14] [NTF_2][8] [EMIm][6] [NTF_2][8]	248.92-423.73	7.80E-06-0.1262	1.7862-2.7157	13.2625-44.0115	11.31E-10-0.6485	1.2385-15.0938	1.7431-7.2302	0.0029-2.7758	1.35E-13-0.0729	1,3389	7,3768	3,6761	2,4562	13	250	[545] [546]

[1] 1-octyl-3-methylimidazolium
[2] bromide
[3] 1-butyl-3-methylimidazolium
[4] Hexafluorophosphate
[5] 1-hexyl-3-methylimidazolium
[6] 1-ethyl-3-methylimidazolium
[7] trifluoromethanesulfonate
[8] bis(trifluoromethylsulfonyl)amide
[9] 1-decyl-3-methylimidazolium
[10] tetrafluoroborate
[11] Didecyldimethylammonium
[12] nitrate
[13] N-benzyl-N,N-dimethyldodecan-1-aminium
[14] N-ethyl-2-hydroxy-N,N-dimethylethanaminium

		[BMIm][3] [NTF_2][8]11. [HMIm][5] [NTF_2][8] [OMIm][1] [NTF_2][8] [DMIm][9] [NTF_2][8]																[547]
6	Coumarin	[EMIm][6] [OTF][7] [BMIm][3] [OTF][7] [EMIm][6] [NTF_2][8] [BMIm][3] [NTF_2][8] [HMIm][5] [NTF_2][8] [DMIm][9] [NTF_2][8]	263.58-338.77	1.65E-06-0.0827	2.5849-2.6618	13.5141-29.5087	4.31E-10-8.74E-05	6.8591-15.0938	4.5427-5.2186	0.0024-1.3087	4.73E-09-1.78E-06	0,1914	8,6955	6,3308	1,8810	6	43	[544]
7	4-Hydroxycoumarin	[EMIm][6] [OTF][7] [BMIm][3] [OTF][7] [BMIm][3] [NTF_2][8]	283.02-415.2	9.61E-06-0.0827	2.5849-2.6618	13.5141-17.5745	1.12E-08-8.74E-05	6.8591-15.0938	4.5427-5.2186	0.0244-1.3087	4.73E-09-1.78E-06	1,2222	7,4744	7,0979	2,1185	3	24	[544]
8	Diosgenin	[BMIm][3] [NO_3][12] [EMIm][6] [NTF_2][8] [BMIm][3] [NTF_2][8] [OMIm][1] [NTF_2][8]	312.85-363.95	7.80E-06-0.0827	2.5849-2.6618	13.5141-25.5103	8.06E-09-8.74E-05	1.2385-15.0938	3.5645-5.2186	0.0244-2.7758	4.73E-09-0.0726	0,6811	24,3218	13,7444	2,6303	4	20	[548]
9	P-Coumaric	[BMIm][3] [PF_6][4] [OMIm][1] [PF_6][4] [BMIm][3] [BF_4][10] [OMIm][1] [BF_4][10] [BMIm][3] [NTF_2][8] [BMIm][3] [OTF]	303.1-317.4	7.80E-06-0.0827	2.5849-2.6310	17.5745-25.5163	8.06E-09-8.74E-05	1.8168-15.0938	1.7431-7.2302	1.37E-08-1.3087	1.35E-13-7.75E-09	1,8564	6,9428	8,4971	2,4889	6	24	[549]
10	Caffeic Acid	[BMIm][3] [PF_6][4] [OMIm][1] [PF_6][4] [BMIm][3] [BF_4][10] [OMIm][1] [BF_4][10] [BMIm][3] [NTF_2][8] [BMIm][3] [OTF][7]	303.1-317.4	7.80E-06-0.0827	2.5849-2.6310	17.5745-25.5163	8.06E-09-8.74E-05	1.8168-15.0938	1.7431-7.2302	1.37E-08-1.3087	1.35E-13-7.75E-09	2,5131	6,3735	9,1127	2,6058	6	24	[549]
11	Aliskiren Hemifumarate	[BMIm][3] [PF_6][4] [HMIm][5] [PF_6][4] [OMIm][1] [PF_6][4]	298.2-318.2	7.80E-06-0.0827	2.5849-2.6310	17.5745-25.5163	8.06E-09-8.74E-05	9,6772	1,7431	0,0000	0,0000	6,4366	69,6853	38,5946	14,7455	3	15	[550]
12	Erythromycin	[BMIm][3] [NTF_2][8] [DMIm][9] [NTF_2][8] [$P_{6,6,6,14}$][15] [CL][16] [$N_{4,1,1,1}$][17] [NTF_2][8] [$N_{1,8,8,8}$][18] [NTF_2][8] [$Pyrr_{4,1}$][19] [NTF_2][8]	284.4-358	0.0005-0.0827	0.9058-2.5977	16.1441-65.4836	8.74E-09-0.0003	1.98E-09-15.0938	0.0001-5.2186	0.0244-4.6548	4.73E-09-0.6128	1,3933	35,1450	17,8496	6,6067	6	44	[551]
13	Theobromine	[BMIm][3] [BF_4][10] [BMIm][3] [PF_6][4] [BMIm][3] [NO_3][12]	303.15-353.15	1.65E-06-2.6310	2.5849-2.6310	17.5744-25.5105	4.31E-10-8.74E-05	1.2385-9.6771	1.7430-7.2302	0.0029-2.7758	1.35E-13-0.0273	1,2512	11,5490	2,8743	3,7362	7	42	[552]

[15] Trihexyltetradecylphophonium
[16] chloride
[17] Butyltrimethylammonium
[18] Methyltrioctylammonium
[19] 1-Butyl-1-methylpyrrolidinium

		[HMIm]^{5} [BF_4]^{10} [HMIm]^{5} [PF6]^{4} [HMIm]^{5} [NO_3]^{12} [OMIm]^{1} [BF_4]^{10}																
14	Theophylline	[BMIm]^{3} [BF_4]^{10} [BMIm]^{3} [PF6]^{4} [BMIm]^{3} nitrate [HMIm]^{5} [BF_4]^{10} [HMIm]^{5} [PF6]^{4} [HMIm]^{5} nitrate [OMIm]^{1} [BF_4]^{10}	303.15-353.15	1.65E-06-2.6310	2.5849-2.6310	17.5744-25.5105	4.31E-10-8.74E-05	1.2385-9.6771	1.7430-7.2302	0.0029-2.7758	1.35E-13-0.0273	1,5669	10,9628	3,0172	3,6706	7	42	[552]