https://doi.org/10.15255/KUI.2020.004
KUI-40/2020
Original scientific paper
Received January 18, 2020
Accepted May 13, 2020

Artificial Neural Network and Support Vector Regression Applied in Quantitative Structure-property Relationship Modelling of Solubility of Solid Solutes in Supercritical CO_2

M. Moussaoui,[a,b,*] *M. Laidi*,[a] *S. Hanini*,[a] and *M. Hentabli* [a]

[a] Laboratory of Biomaterials and Transport Phenomena (LBMPT), University of Médéa, Médéa, Algeria
[b] University of Bouira, Bouira, Algeria

Abstract

In this study, the solubility of 145 solid solutes in supercritical CO_2 ($scCO_2$) was correlated using computational intelligence techniques based on Quantitative Structure-Property Relationship (QSPR) models. A database of 3637 solubility values has been collected from previously published papers. Dragon software was used to calculate molecular descriptors of 145 solid systems. The genetic algorithm (GA) was implemented to optimise the subset of the significantly contributed descriptors. The overall average absolute relative deviation MAARD of about 1.345 % between experimental and calculated values by support vector regress SVR-QSPR model was obtained to predict the solubility of 145 solid solutes in supercritical CO_2, which is better than that obtained using ANN-QSPR model of 2.772 %. The results show that the developed SVR-QSPR model is more accurate and can be used as an alternative powerful modelling tool for QSAR studies of the solubility of solid solutes in supercritical carbon dioxide ($scCO_2$). The accuracy of the proposed model was evaluated using statistical analysis by comparing the results with other models reported in the literature.

Keywords

Solubility, solid solutes, supercritical-fluids, computational intelligence techniques, quantitative structure-property relationship

1 Introduction

Supercritical carbon dioxide ($scCO_2$) is generally used in separation processes applied in the food, chemistry, pharmaceutical, and other industries. The design and optimisation of extraction, fractionation, and purification processes are mainly based on the knowledge of the solubility of solid solutes in $scCO_2$.[1] The experimental measurement of the solubility (thermo-physical property) of such compounds in $scCO_2$ is laborious and costly.[2] To avoid expensive and tedious experiments, and to fulfil the lack of solubility data and/or pure component property data required to estimate solubility, a need exists to develop flexible and robust predictive models to estimate the solubility of solid solutes in supercritical solvents using limited information. Literature reports three major modelling approaches have been used to model the solubility of a solid solute in $scCO_2$ using equations of state,[3] semi empirical equations,[2] and computational intelligence techniques such as artificial neural networks and support vector machine.[4] The solid solute solubility in $scCO_2$ is mostly modelled using density-based correlations, such as in [5]. Moreover, MLP-ANN also has been used successfully to model the solid solute solubility in $scCO_2$ based only on thermodynamic parameters.[6] The least square support vector machine LS-SVM may be considered an alternative for classical approaches, such as semi empirical correlations to model the solid solute solubility in $scCO_2$.[7] The most common type of ANN model that is being used nowadays is multilayer perceptron (MLP) feedforward neural network (FNN) trained by back-propagation (BP) algorithm.[8] This ANN type has also been successfully used for mapping complex highly nonlinear input/output relationships between one dependent variable and more independent variables of any systems. But there are numerous limitations, for instance, difficulty in deciding the criterion for the number of hidden neurons and layers, as well as over-fitting.[9] To avoid these problems, SVM is formulated as a quadratic optimisation problem and ensures a global optimal solution. Some studies have proven the superiority of SVM over ANN.[10] SVM method has gained a wide range of engineering applications in forecasting and regression analysis due to its attractive features and remarkable generalization performance.[11] Therefore the present study was carried out to evaluate the predictive performance of computational intelligence techniques including MPL-FNN as the most used neural network type and SVM to foretell the solubility of 145 different solid solutes in $scCO_2$ based on a mixture between thermodynamic properties and molecular descriptors. The novelty of the developed model in contrast to previous models is its ability to extrapolate or interpolate the solubility of several solid solutes in $scCO_2$ with only one combination of the fitted parameters obtained during the training stage. In addition, the independent parameters were a mixture between thermodynamic properties and structural descriptors. Moreover, a simple and convivial graphical user interface was designed to compute the solubility without learning MATLAB software. Finally, a sensitivity analysis was computed by calculating the relative importance of each input parameter on the output.

* Corresponding author: Assist. Professor Mohammed Moussaoui
Email: m.moussaoui@univ-bouira.dz

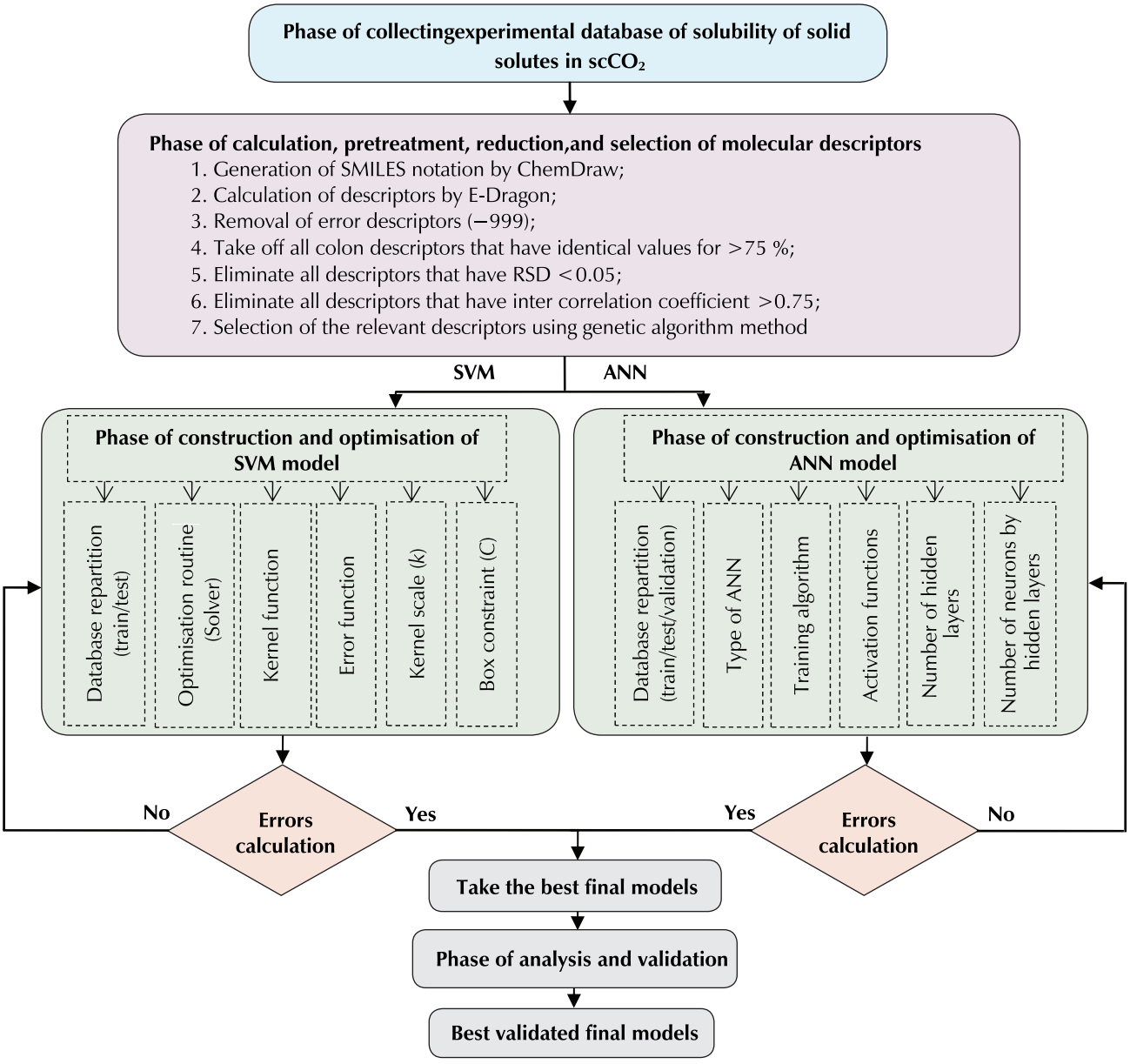

Fig. 1 – Procedure of construction and optimisation of ANN-QSPR and SVR-QSPR models

2 Methodology

In order to develop an ANN-QSPR or SVR-QSPR model, several steps are required, as specified in the following section (Fig. 1).[12] Before sample data were given to one of the machine learning models, relevant descriptors of each system were calculated and optimised, and added to the thermodynamic parameters of each system. The data were then scaled so as to solve the problems of noise and correlated data. Next, splitting of data into two samples training and test, respectively. After settling each model architecture, its parameter optimisation was processed during the training phase. The next phase was to evaluate the performance of the optimised models using other data not used during the training stage. Lastly, statistical parameters had to be calculated to assess the accuracy of each model and validate it to be used for further prediction of solid solute solubility in $scCO_2$. The modelling procedure was run in MATLAB R2018a on the Windows 7 system.

2.1 Data preparation

Reliability and effectiveness of any model highly depends on the availability of accurate experimental data. For that, the required data for training and testing of the computational intelligence techniques were extracted from previous experimental investigations in literature.

Table 1 displays several properties of the 145 selected binary solute-$scCO_2$ systems counting 3637 data points of experimental solubility in $scCO_2$, among these parameters, the name of the corresponding drugs, molecular weight, range of working temperature and pressure, experimental solubility, and references of the 145 selected systems.

Table 1 – Sources and ranges of solubility data of solid drugs in supercritical carbon dioxide

N°	Solid solutes	M / g mol^{-1}	T / K	p / bar	ρ_{CO_2} / kg m^{-3}	$-\log_{10}(y_2)$	N_i	Refs.
1	1,5-Naphthalenediamine (1,5-NDA)	158.204	313.15~333.15	110.00~200.00	357.88~840.66	4.7905~5.7212	27	13
2	4,4-Diaminodiphenylmethane (DADPM)	198.269	313.15~333.15	110.00~200.00	357.88~840.66	3.7940~5.0269	27	14
3	4-Aminoantipyrine	203.240	308.20~328.20	100.30~220.40	338.62~879.62	3.4353~4.9830	21	15
4	Aspirin	180.157	308.15~328.15	120.00~250.00	504.51~901.23	3.4597~4.2007	24	16
5	Cholesteryl acetate	428.701	308.15~328.15	90.00~240.00	385.30~893.89	2.9830~5.3615	24	
6	Cholesteryl benzoate	490.772	308.15~328.15	120.00~270.00	120.00~270.00	4.2774~5.2765	20	
7	Cholesteryl butyrate	456.755	308.15~328.15	100.00~240.00	54.47~893.89	3.0491~4.6497	20	
8	Atorvastatin	558.640	308.00~348.00	121.60~354.60	327.00~955.00	2.8398~6.0000	45	17
9	Simvastatin	418.566	308.00~348.00	121.60~354.60	327.00~955.00	3.2716~5.6990	45	
10	Lovastatin	404.550	308.00~348.00	121.60~354.60	327.00~955.00	3.9431~4.9586	45	
11	Rosuvastatin	481.538	308.00~348.00	121.60~354.60	327.00~955.00	3.6126~5.5229	45	
12	Fluvastatin	411.470	308.00~348.00	121.60~354.60	327.00~955.00	3.2211~5.3010	45	
13	Atropine	289.369	308.00~348.00	122.00~355.00	327.00~955.00	2.7773~4.2218	45	18
14	Carbamazepine	236.269	308.00~348.00	122.00~355.00	477.00~955.00	4.0269~5.5229	39	
15	Codeine	299.364	308.00~348.00	122.00~355.00	327.00~955.00	2.9101~4.3979	45	
16	Diazepam	284.740	308.00~348.00	122.00~355.00	327.00~955.00	2.9547~3.7959	45	
17	Acenaphthene	154.212	308.15~348.15	121.60~354.60	327.00~955.00	1.8540~2.8938	45	
18	Fluoranthene	202.256	308.15~348.15	121.60~354.60	327.00~955.00	2.8195~3.9829	45	
19	Triphenylene	228.294	308.15~348.15	121.60~354.60	327.00~955.00	4.0222~5.6989	45	
20	Azelaic acid	188.221	313.15~333.15	100.00~300.00	289.95~909.89	4.9948~6.3768	14	19
21	Benzocaine	165.189	308.00~348.00	122.00~355.00	327.00~955.00	1.9165~3.5686	40	20
22	Benzoin	212.248	308.15~328.15	121.60~236.10	433.42~897.99	3.3872~4.3872	19	21
23	Mandelic acid	152.150	308.15~328.15	101.00~230.60	356.78~892.63	2.5370~4.5686	21	
24	Propyl 4-hydroxybenzoate	180.203	308.15~328.15	94.10~220.90	361.93~885.02	3.2132~4.7212	21	
25	Bisacodyl	361.400	308.00~348.00	122.00~355.00	396.00~955.00	3.2343~5.0458	39	22
26	Caffeine	194.190	313.00~353.00	199.00~349.00	591.75~934.98	2.9469~3.5482	24	23
27	Naproxen	230.259	313.10~333.10	89.60~193.10	394.16~840.87	4.4976~5.7212	18	24
28	Ketoprofen	254.280	313.15~328.15	90.00~250.00	422.63~889.59	3.7258~5.4815	15	
29	Flurbiprofen	244.261	303.15~323.15	89.00~245.50	521.40~929.10	3.7059~4.7768	27	
30	Ibuprofen	206.281	308.15~318.15	80.00~220.00	290.96~890.04	2.1675~4.5229	29	
31	Gemfibrozil	250.340	308.20~328.20	100.10~220.20	400.28~879.40	2.3778~4.5317	21	25
32	Clofibric acid	214.645	308.2~328.2	100.10~220.20	337.35~879.40	3.0675~4.4934	21	
33	Irgacure® 2959 photoinitiator	224.256	308.20~328.20	101.00~254.00	326.90~901.80	3.5481~5.2840	21	26
34	Dichlone	227.040	313.00~ 333.00	70.70~325.80	160.80~905.80	3.6003~5.1549	23	27
35	Menadione	172.183	313.00~333.00	97.20~306.70	273.50~306.70	2.9974~3.1249	18	
36	Methimazole	114.166	308.00~348.00	122.00–355.00	327.00~955.00	2.7215~4.2676	40	28
37	Phenazopyridine	213.244	308.00~348.00	122.00–355.00	327.00~955.00	2.6944~4.3565	45	
38	Propranolol	259.349	308.00~348.00	122.00~355.00	327.00~955.00	1.6205~3.4461	45	
39	Nicotinamide	122.127	313.15~373.15	54.00~305.00	90.15~905.74	2.5031~5.0000	24	29
40	Nicotinic acid	123.110	313.15~373.15	45.00~302.00	78.13~908.18	4.9825~6.5086	22	

Table 1 – (continued)

N°	Solid solutes	M / g mol^{-1}	T / K	p / bar	ρ_{CO_2} / kg m^{-3}	$-\log_{10}(y_2)$	N_i	Refs.
41	Oxymetholone	332.477	308.00~328.00	121.00~305.00	411.00~858.00	3.8268~4.7959	20	30
42	Cefiximetrihydrate	507.489	308.00~328.00	183.00~335.00	634.00~875.00	6.5199~6.7958	18	
43	Progesterone (steroids)	314.460	313.15~338.15	120.00~260.00	387.40~884.80	2.8386~4.8539	18	31
44	Sinapic acid	224.212	313.00~333.00	200.00~400.00	725.00~956.00	6.1524~8.1549	9	32
45	Protocatechuic acid	154.121	313.00~333.00	200.00~400.00	725.00~956.00	5.9731~ 6.8182	9	
46	Chrysin	254.241	313.15~333.15	200.00~400.00	725.00~956.00	6.9469~7.7447	9	
47	Pyrocatechol	110.112	308.15~363.15	100.00~400.00	203.00~975.00	2.3388~3.9136	33 32	33,34,35
48	Phenol	94.113	333.15~363.15	100.00~350.00	208.00~864.00	1.0427~2.9431	33	34
49	Resorcinol	110.112	308.15~338.15	120.00~400.00	296.00~975.00	3.0119~3.9586	32	35
50	Quinine	324.424	308.15~328.15	80.00~240.00	210.62~906.51	5.0915~7.0809	27	36
51	Salicylic acid	138.120	308.15~318.15	92.60~157.90	463.00~824.00	3.4001~4.0132	20	37
52	Uracil	112.080	313.00~333.00	100.00~299.90	289.40~910.00	3.8861~5.6383	12	38
53	Tributylphosphate	266.318	303.00~363.00	150.00~250.00	372.00~923.00	1,0459~1,3152	24	39
54	*N*-phenylacetamide	135.166	308.00~328.00	104.40~225.00	356.70~889.20	3,3408~4,5331	24	40
55	*o*-Phthalic acid	166.132	308.00~328.00	80.00~210.00	206.80~875.60	5,2899~5,9957	15	41
56	1-Aminoanthraquinone	223.231	323.15~383.15	125.00~250.00	236.98~834.01	4.4543~6.2596	18	42
57	1-Nitroanthraquinone	253.213	323.15~383.15	125.00~250.00	236.98~834.01	4.0788~4.6253	18	
58	4-Hydroxycinnamic acid (*p*-coumaric acid)	164.160	313.00~ 333.00	150.00~500.00	607.00~992.00	5.5930~6.8110	24	43
59	3,4-Dihydroxycin-namic acid (caffeic acid)	180.159	313.00~ 333.00	150.00~500.00	607.00~992.00	7.3261~9.0969	24	
60	4-Hydroxy-3-methoxycinnamic acid (ferulic acid)	194.186	313.00~333.00	150.00~500.00	607.00~992.00	4.3633~5.4356	24	
61	α-Tocopherol	430.717	313.00~ 353.00	199.00~349.00	591.60~935.00	2.4473~3.1938	24	44
62	β-Carotene	536.888	313.00~353.00	200.00~350.00	594.20~935.40	3.9281~7.0457	23	
63	2-Methylbenzoic acid	136.150	313.20~333.20	110.00~246.00	369.00~882.00	2.2441~3.6198	18	45
64	3-Methylbenzoic acid	136.150	313.20~333.20	110.00~246.00	369.00~882.00	2.2724~3.7352	18	
65	4-Methylbenzoic acid	136.150	313.20~333.20	110.00~246.00	369.00~882.00	3.1537~4.3665	18	
66	Erythromycin	733.937	308.00~348.00	122.00~355.00	327.00~955.00	3.5058~4.3665	45	46
67	Naphthalene	128.174	308.15~338.00	71.90~321.50	66.95~777.02	0.4898~3.2881	109	47,14
68	Amiodaronehydrochloride	681.778	313.42~343.20	120.00~300.00	347.04~909.94	2.9948~4.6003	28	48
69	1-Amino-4-hydroxyanthraquinone	239.230	323.15~383.15	125.00~250.00	236.79~834.01	4.6117~6.3635	20	49
70	1-Hydroxy-4-nitroanthraquinone	269.212	323.15~383.15	150.00~250.00	303.03~834.00	5.0634~6.2034	15	
71	Bisphenol A	228.291	308.00~328.00	110.00~210.00	417.06~874.40	5.5212~6.9281	15	50
72	*p*-Nitroaniline	138.126	308.00~328.00	110.00~210.00	414.90~873.67	4.3705~5.2055	15	51
73	Methyl salicylate	152.149	343.15~423.15	90.00~310.00	129.10~613.47	1.1230~3.2924	44	52
74	Juglone	174.155	308.20~328.20	92.00~244.00	267.40~895.50	2.7986~4.6989	18	53
75	1-Eicosanol	298.555	308.20~328.20	36.20~412.30	71.30~976.40	2.6556~4.8268	24	54
76	Eicosanoicacid	312.538	308.20~328.20	33.70~211.20	62.90~971.80	2.5654~5.2006	21	
77	1-Octadecanol	270.501	318.00~338.00	139.90~452.80	536.80~955.10	1.6402~2.9830	17	55
78	Stearic acid	284.484	318.00~338.00	145.40~467.50	591.17~935.92	1.8327~3.2069	17	
79	Squalene	410.730	313.00~333.00	100.00~400.00	389.00~956.00	1.1124~3.3010	9	56
80	Carbazole	167.211	308.20	103.00~201.00	706.22~865.71	4.5591~4.8447	5	57

Table 1 – (continued)

N°	Solid solutes	M / g mol^{-1}	T / K	p / bar	ρ_{CO_2} / kg m^{-3}	$-\log_{10}(y_2)$	N_i	Refs.
81	Anthraceneoil	178.234	308.20~333.15	80.00~470.00	206.49~981.24	3.6439~5.4828	5 8 23	57,58,59
82	Ascorbylpalmitate	414.539	308.10~313.10	130.00~200.00	720.15~865.50	5.2358~5.8041	8	60
83	Butyl hydroxy anisole (BHA)	360.494	313.10~333.10	130.00~200.00	604.38~840.15	1.2218~1.8416	7	
84	Dodecyl gallate	338.444	313.50~333.10	150.00~250.00	604.38~878.02	4.6840~5.5086	8	
85	Propyl gallate	212.201	313.10~333.10	150.00~250.00	604.38~878.02	4.1586~5.2840	8	
86	Ascorbic acid	176.124	313.10	130.00~200.00	720.14~840.15	4.4962~4.7328	4	
87	Chrysene	228.294	308.15	84.00~251.00	539.86~901.61	5.0535~6.1373	11	61
88	Triphenylene	228.294	308.15~328.15	85.00~252.00	381.25~898.90	4.3747~5.9957	28	61
89	Lactic acid	90.078	313.00~328.00	58.20~197.10	143.07~836.03	2.7695~4.6197	29	62
90	2-Hydroxyhexanoic acid	132.159	311.00~328.00	74.00~198.30	265.60~848.89	2.5834~4.6383	34	
91	Octacosane	394.772	308.15~318.15	80.00~275.00	485.60~894.80	3.4089~4.9101	28	63
92	Triacontane	422.826	308.15~318.15	90.00~250.00	529.90~250.00	3.5210~5.2924	18	
93	Hexadecanoicacid	256.430	308.00~3018.00	128.00~226.00	639.60~884.56	2.9809~3.5884	10	64
94	Octadecanoicacid	284.484	308.00~318.00	128.00~226.00	639.60~884.56	3.4894~4.1308	10	
95	Tetradecanoic acid	228.376	308.00~318.00	99.00~226.50	641.91~884.92	2.1278~2.9508	11	
96	Artemisinin	282.336	310.10~338.10	100.00~270.00	278.81~904.26	2.5753~4.0088	36	65
97	Ergosterol	396.659	318.15~333.15	120.00~240.00	421.84~848.03	4.5230~5.5324	15	66
98	Cholesterol	386.664	313.15~333.15	100.00~250.00	384.40~879.60	3.8386~5.6383	24	16,67
99	(2,4-Dichlorophenoxyacetic acid) (2,4-D)	221.033	313.10~333.10	104.40~207.90	448.05~846.42	3.9508~4.7852	16	68
100	Methyl gallate	184.147	313.00~333.00	10.00~50.00	295.00~992.00	5.3727~7.7190	27	69
101	Protocatechualdehyde	138.122	313.00~333.00	10.00~50.00	295.00~992.00	5.4270~8.4948	24	
102	1,4-Naphthoquinone	158.156	308.20~328.20	91.00~242.00	267.80~894.93	2.3071~4.3010	18	70
103	*p*-Quinone (1,4-benzoquinone)	108.096	308.15~318.15	86.10~292.10	331.97~923.84	1.5028~2.9730	18	71
104	9,10-Anthraquinone	208.216	273.15~318.15	84.10~306.30	379.83~1055.76	4.1308~6.0545	17	
105	Syringic acid	198.174	313.00~333.00	100.00~500.00	291.41~991.63	4.8959~9.0000	27	72
106	Vanillic acid	168.148	313.00~333.00	85.00~500.00	291.42~991.63	4.2167~6.0283	28	
107	Paracetamol (acetaminophen)	151.165	313.00~353.00	11.00~250.00	19.63~880.24	5.0894~7.1427	10	73
108	Lycopenepe	536.885	323.15~353.15	200.00~400.00	595.40~923.80	5.7277~6.1864	20	74
109	Testosterone (steroids)	288.431	308.15~328.15	87.02~242.82	243.85~895.47	4.1543~6.4089	39	75
110	*o*-Hydroxy benzoic acid	138.122	273.15~328.15	81.10~202.60	263.59~1182.34	3.2426~5.1549	48	76
111	Acetanilide	135.166	308.20~323.20	90.00~400.00	383.93~972.04	0.3003~4.8706	29	77
112	Propanamide	73.095	308.20~323.20	90.00~400.00	284.75~972.04	2.2453~3.6769	30	
113	Butanamide	87.122	308.20~323.20	90.00~400.00	284.75~972.04	0.9780~4.1026	30	
114	Cinnamic acid	148.161	308.20~328.20	12.30~23.61	546.67~889.50	3.3716~4.4559	19	78
115	Geranyl butyrate (10-undecenoic acid)	184.279	308.00~333.00	100.00~180.00	291.42~845.47	1.7594~3.3979	18	79
116	10-Undecenoic acid	184.279	308.00~333.00	100.00~180.00	291.42~845.47	1.7594~3.3979	18	
117	2-Trifluoromethylbenzoic acid	190.121	308.20~323.20	93.40~226.00	370.04~883.63	2.0680~3.5850	21	80
118	3-Trifluoromethylbenzoic acid	190.121	308.20~323.20	94.10~225.40	345.25~883.06	1.3288~3.1487	21	
119	4-Trifluoromethylbenzoic acid	190.121	308.20~323.20	96.80~224.40	364.09~882.48	3.3410~4.5086	21	
120	Cannabinol	310.437	314.00~334.00	130.00~202.00	472.66~836.80	3.3458~3.8996	34	81

Table 1 – (continued)

N°	Solid solutes	M / g mol^{-1}	T / K	p / bar	ρ_{CO_2} / kg m^{-3}	$-\log_{10}(y_2)$	N_i	Refs.
121	Ethanamide (acetamide)	59.068	308.20~323.20	90.00~400.00	284.75~972.04	2.5000~3.6228	30	82
122	2-Propenamide (acrylamide)	71.079	308.20~323.20	90.00~400.00	284.75~972.04	2.7875~3.9233	28	
123	Hinokitiol	164.204	313.20~333.20	101.40~378.30	523.33~946.73	2.6038~23.2284	30	83
124	Iodopropynylbutylcarbamate	281.093	313.15~333.15	87.60~341.50	417.44~930.67	2.3279~2.9208	27	84
125	Phenanthrene	178.234	308.15	113.00~332.00	729.62~943.89	2.5952~3.1249	11	
126	Levulinicacid	116.116	313.00~342.40	84.10~188.30	268.20~824.70	2.0457~3.3089	34	85
127	*m*-Dinitrobenzene	168.108	308.00~328.00	95.00~145.00	299.78~803.41	2.2565~3.7212	18	86
128	1-Chloro-2,4-dinitrobenzene	202.550	308.00~313.00	95.00~145.00	560.16~803.41	2.2083~2.7495	12	
129	Meloxicam sodium salt	391.392	303.00~323.00	149.00~255.00	704.40~921.60	4.89.41~5.3556	15	87
130	*N*-(4-Ethoxyphenyl) ethanamide	386.665	308.00~328.00	90.00~190.00	401.55~855.74	4.3685~5.0409	16	88
131	Ketotifenfumarate (KTF)	425.50	308.2~338.2	120.00~300.00	384.17~929.68	2.9679~4.6757	28	89
132	Sertraline hydrochloride	342.69	308.00~338.00	120.00~300.00	347.04~909.94	4.0315~6.0000	28	90
133	Aprepitant	534.40	308.15~338.15	120.00~300.00	384.00~944.00	4.1152~5.3468	32	91
134	Coumarin-7	333.38	308.00~338.00	90.00~330.00	220.59~944.01	4.9961~5.3819	20	92
135	Letrozole	285.30	318.20~348.20	120.00~360.00	318.90~922.00	5.0701~5.7959	20	93
136	Imatinibmesylate	589.71	308.20~338.20	120.00~270.00	388.00~914.00	5.3559~7.0000	24	94
137	Esomeprazole	345.417	308.20~338.20	120.00~270.00	382.44~912.88	3.0409~4.9586	24	95
138	Loratadine	382.88	308.15~338.15	120.00~270.00	388.00~914.00	2.8855~5.3468	24	96
139	Lansoprazole	369.363	308.20~338.20	120.00~270.00	382.44~912.88	3.1331~4.9208	27	97
140	Sunitinib malate	532.56	308.00~338.00	120.00~270.00	388.00~914.00	4.0675~4.8894	24	98
141	Azathioprine	277.263	308.00~338.00	120.00~270.00	388.00~914.00	4.7375~5.5686	24	99
142	Sorafenibtosylate	637.03	308.00~338.00	120.00~270.00	388.00~914.00	4.9006~6.1675	24	100
143	Repaglinide	452.595	308.00~338.00	120.00~270.00	384.00~914.00	4.0214~5.5376	24	101
144	Oxcarbazepine	252.273	308.00~338.00	120.00~270.00	384.00~914.00	4.5726~6.9586	24	102
145	Pigment (Phthalocyanine green)	1127.154	308.00~338.00	120.00~270.00	384.00~914.00	3.9165~7.0000	24	103

2.2 Determination of descriptors

Quantitative Structure-Property Relationship (QSPR) models can also be used to perform property estimation since they establish quantitative correlations between diverse molecular properties and the chemical structure.[3] A QSPR model consists of a mathematical relationship between the property or activity and several molecular descriptors, such as structural/topological indices or electronic/quanto-chemical properties. One of the most important steps in the construction of QSPR models is the quantification of structure information of the studied molecules,[104] which are called molecular descriptors. Actually, there exist more than 11145 molecular descriptors that can be used to solve several problems in different fields, such as chemistry, biology, and other related sciences.[105] In this study, 1666 molecular descriptors were calculated online by E-Dragon 1.0 software.[106] The computation of these descriptors was preceded by a very interesting step, *i.e.*, the generating of simplified molecular input-line entry system (SMILES) strings with the use of the "ChemDraw" software. After the calculation of the molecular descriptors, constant and near-constant descriptor values were removed, all missing values or at least one missing value were removed, any descriptor with a relative standard deviation < 0.001 was removed, also descriptors with pair correlation larger or equal to ≥ 0.95 were removed. Finally, a total set of 641 remaining descriptors was reduced as possible using stepwise regression method to select only the relevant subset of descriptors that have significant contributions to the solubility of solid solutes. After this pre-treatment, three pertinent descriptors were obtained for each system to build the ANN and SVR models[107] (Table 2).

2.3 Statistical performance evaluation criteria

The internal predictive capability of the ANN and SVR models was evaluated by root mean square error (RMSE), average absolute relative deviation (AARD), leave-one-out (LOO) cross validation Q^2_{LOO} on the training set, and the coefficient of determination (COD). The definitions of these measures are given in Table 3, where N represents

Table 2 – The three 2D molecular descriptors used in QSPR model

Descriptors	Definition	Type
MPC2	molecular path count of order 2	topological descriptors
WTPT-1	molecular ID	topological descriptors
AATS0i	averaged Moreau-Broto autocorrelation of lag 0 weighted by ionization potential	autocorrelation descriptors

the number of observations, y_2^{exp} and y_2^{cal} are the desired (observed or experimental) and the predicted output values, respectively, $\widehat{y_2^{exp}}$ is the average of experimental values of y_2^{exp}.

The value of RMSE and AARD ranges from 0 to ∞, where the R^2 is used to determine the degree of similarity between the predict value and its observed value. In general, the best value for R^2 and Q^2_{LOO} is 1, and the best value of RMSE and AARD is zero which indicates high performance of the model.

3 Results and discussion

3.1 ANN-QSPR model results

Multi-layer perceptron networks are the most widely used feedforward artificial neural networks, and are considered a powerful nonlinear black-box model for learning complex nonlinear relationships between input and output variables. The multi-layer perceptron neural networks consist of 03 connected layers: input, hidden, and output layers. The first layer receives input data and sends them to the hidden layer. After adjusting weights and biases by interconnected neurons using several non-linear transfer functions and different variables, the results from the hidden layer are sent to the output layer for the outputs.[110]

The weights of the ANN model are modified using the training algorithm until the error criterion is satisfied. In this study, the Bayesian regularisation backpropagation algorithm was used in training the ANN models. An objective function, which is the mean squared error between the outputs of the network and the target (experimental) values is used to calculate the bias and weights of the feedforward neural network. To avoid fitting problems, the number of hidden neurons is usually determined by a trial-error methodology.

In this study, this number was changed from 10 to 30 neurons using one and two hidden layers. Results have shown that ANN-QSPR with two hidden layers performs better in comparison to single layer. This procedure is repeated until the best precision is achieved for each architecture. The input parameters for the proposed model were temperature, pressure, density of $scCO_2$, critical temperature, critical pressure, acentric factor, and the most accurate three selected 2D descriptors {MPC2, WTPT-1, and AATS0i}, while the output was the logarithm of the molar fraction solubility.

In the ANN model, as the inputs have different magnitudes and in order to decrease the error, a normalization must be performed on the inputs and outputs of the artificial neural network (usually between −1 and +1) to achieve fast convergence. The data set was randomly split into three subsets (66 %) for the training phase, (17 %) for the testing phase, and (17 %) for the validation phase of the ANN model.

The best developed three-layer FFNN for the solubility has the structure of 9-25-20-1. This means that the ANN model has nine independent variables (thermodynamic properties and descriptors), 25 numbers of neurons in the first hidden layer, 20 numbers of neurons in the second hidden layer, and one output which is the solubility for each system. Table 4 shows the structure of the optimised ANN model. Hyperbolic tangent sigmoid (tansig) and line-

Table 3 – Definitions of performance measures[108,109]

Measure	Formula
Coefficient of determination (COD) Validation criteria > 0.8	$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_2^{exp}(i) - y_2^{cal}(i)\right)^2}{\sum_{i=1}^{N}\left(y_2^{exp}(i) - \widehat{y_2^{exp}}\right)^2}$
Mean average absolute relative deviation (MAARD/%)	$MAARD/\% = \frac{100}{N}\sum_{i=1}^{N}\left\|\frac{\left(y_2^{exp}(i) - y_2^{cal}(i)\right)}{\left(y_2^{exp}(i)\right)}\right\|$
Root mean square error (RMSE)	$RMSE = \left(\left(\frac{\sum_{i=1}^{N}\left(y_2^{exp}(i) - y_2^{cal}(i)\right)^2}{N}\right)\right)^{1/2}$
Leave-one-out (LOO) cross validation Q^2_{LOO} Validation criteria > 0.5	$Q^2_{LOO} = 1 - \frac{\sum_{i=1}^{N}\left(y_2^{exp\_train}(i) - y_2^{cal\_train}(i)\right)^2}{\sum_{i=1}^{N}\left(y_2^{exp\_train}(i) - \overline{y}_2^{exp\_train}(i)\right)^2}$

Table 4 – Architecture of the optimised ANN model

ANN type	Input layer	Hidden layer				Output layer		Training algorithm
		1st hidden layer		2nd hidden layer				
	Neurons number	Neurons number	Transfer function	Neurons number	Transfer function	Neurons number	Transfer function	
MLP	09	25	hyperbolic tangent sigmoid TANSIG	20	hyperbolic tangent sigmoid TANSIG	1	purelin	Bayesian regularisation backpropagation

ar (purelin) functions were used as the transfer function for the hidden and output layers, respectively.

Fig. 2 displays the scatter plot of all data set of the solid solute solubility calculated using the ANN-QSPR model *vs.* experimental solubility (3637 experimental points). These plots were generated using the postreg function of MATLAB, tracing the calculated solubility as a function of experimental solubility.

In Fig. 2, the first bisector shows the exact fit between the correlated solubilities and experimental data, whereas the cross points demonstrate the real correlated solubility data by the proposed correlation *vs.* experimental data. In detail, the closer the points are to the solid line, the more accurate are the correlated solubility data.

The logarithmic scale of the experimental against calculated solubility is plotted in Fig. 2. On the basis of the obtained results, it can be concluded that the ANN-QSPR model was able to correlate the solubility of solids in $scCO_2$ with an acceptable deviation (Table 5).

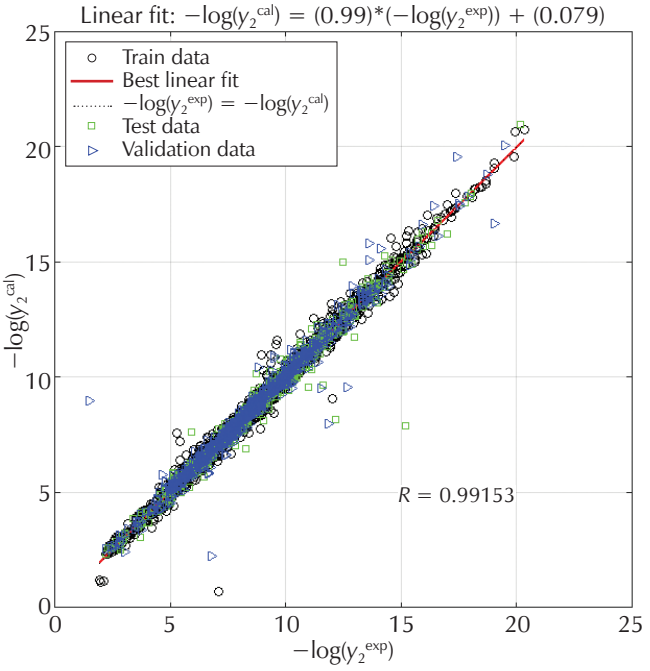

Fig. 2 – Comparison between experimental and predicted values of the solubility of 145 solutes in $scCO_2$ using ANN

The statistical and validation values of the ANN model are listed in Table 5.

Table 5 – Statistical and validation results of the ANN-QSPR model

	Training	Test	Validation	All data set
R^2	0.9895	0.9733	0.9665	0.9830
mean AARD/%	2.637	2.900	3.189	2.772
RMSE/%	0.2932	0.4633	0.5241	0.3726
Q^2_{LOO}	0.9895	0.9731	0.9667	0.9830

3.2 SVR modelling

Support vector regress (SVR) has found many applications in different regression problems.[111] This approach was firstly proposed by *Vapnik*.[112] The quality of SVR models depends on the proper setting of SVR parameters for a given data set. The selection of these parameters (kernel width parameter σ, penalizing parameter C, and error-accuracy parameter ϵ) is investigated in this section.

The same ANN inputs have been adopted to build the SVR model. The data set was randomly split into two subsets (66 %) for the training phase, (33 %) for the testing phase of the SVR model. The optimisation procedure was repeated several times in order to find the most probable global optimum of the fitness function.

In this study, radial basis function (RBF) was selected as the most common kernel function used in literature; details of the other selected parameters are summarized in Table 6.

Table 6 – Details of the proposed SVR model

Type	Value/comment
number of data used for training	2425
number of data used for testing	1212
optimisation routine (solver)	Sequential Minimal Optimisation (SMO)
kernel function	Radial Basis Function (RBF)
penalising parameter "C"	92
kernel width parameter "σ"	1.8
quantity of support vectors	2568

Table 7 – Statistical analyses for the suggested SVR model

	Train	Test	All data set
R^2	0.9952	0.9666	0.9857
mean AARD / %	0.895	2.247	1.345
RMSE	0.1968	0.5177	0.3393
Q^2_{LOO}	0.9953	0.9675	0.9859

Table 7 shows a comparison between experimental and predicted values of solubility for the training, test, and all data set. Results show that the model has captured the features quite accurately with relatively high regression {R^2} and low errors {AARD and RMSE and Q^2_{LOO}}.

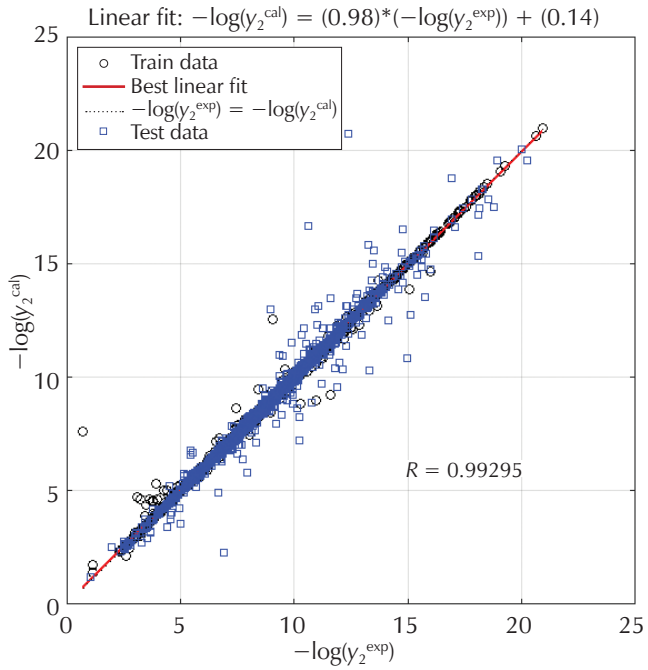

Fig. 3 – Comparison between experimental and predicted values of the solubility of 145 solutes in $scCO_2$ using SVR model

In Fig. 3, the capability of the model was evaluated by plotting experimental values of solubility against the predicted values by the model for all data set. The large correlation coefficient 0.9857 and a small average absolute relative deviation of 1.345 reveal the capability of the model to correlate the solubility of solid solutes in $scCO_2$ by knowing their physical properties.

4 Comparison

The comparative study between ANN-QSPR and SVR-QSPR models in terms of regression and different errors using all data set was conducted (Fig. 4). The results show that the solubility data of solid solutes in $scCO_2$ are better correlated by SVR-QSPR model than with the ANN-QSPR model. Statistical analyses show that support vector regress (SVR) predictions have an excellent agreement with the experimental data set.

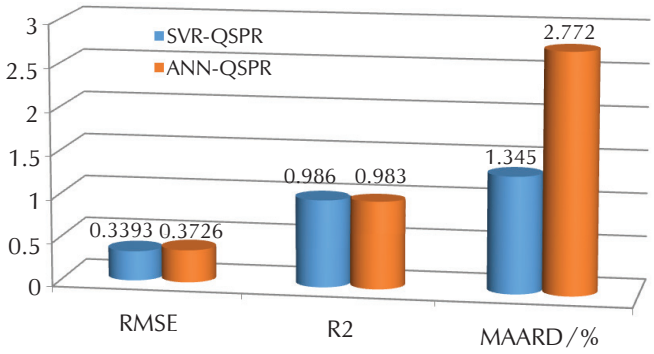

Fig. 4 – Statistical comparison between ANN and SVR models

In addition to the statistical validation of the developed models, further evaluation of the accuracy and the regression was obtained from the most relevant papers developed by other researchers (Table 8 and Fig 5). Despite numerous previously published papers in literature, the developed models in this study showed better performance in cor-

Table 8 – Comparison between the developed SVR-QSPR model in this study and other smart models reported in literature for all data set in binary solute-$scCO_2$ systems

Studies	Methods	Compounds number	Experimental data	Model inputs	R^2	RMSE	MAARD / %	Q^2_{LOO}	Year
This work	SVR-QSPR ANN-QSPR	145	3637	9	0.9857 0.9830	0.3393 0.3726	1.3450 2.7720	0.9859 0.9830	2020
[117]	LSSVR	33	1162	5	0.9975	–	5.6100	–	2018
[118]	ANN	6	155	4	0.91999	–	5.9300	–	2018
[119]	ANN	20	439	6	0.9955	–	5.4200	–	2017
[120]	ANFIS	29	795	6	0.9830	0.156	4.6617	–	2017
[121]	GWO-SVR	18	1148	3	0.9936	–	3.20	–	2017
[122]	ANN	8	198	5	0.99699	–	4.99	–	2013
[123]	FFNN	21	795	6	0.9533	–	14.000	–	2011

relating the solubility of solid solutes in $scCO_2$ based on some relevant descriptors. The obtained models were able to correlate 145 systems (3637 data points) with lowest errors. The QSPR-SVM model was obtained with very acceptable statistical parameters, such as R^2 of 0.9857, RMSE of 0.3393, and MAARD of 0.9859, as presented in Table 8. However, the global comparison in terms of the correlated systems indicated that the proposed model performed 145 systems with low errors and high collection coefficient in contrast to the other correlations, as depicted in Fig. 5.

The performance of our model (SVR-QSPR) was compared again to six density-based models previously published (Table 9).

The parameters of the selected literature models were adjusted using genetic algorithms (ga MATLAB function).

The global comparison for 145 compounds was calculated for each of the 07 studied correlations. The correlation results of the experimental data from this study are presented in terms of a comparative table of (AARD) and R^2 grouped in Table 10 and Fig 6.

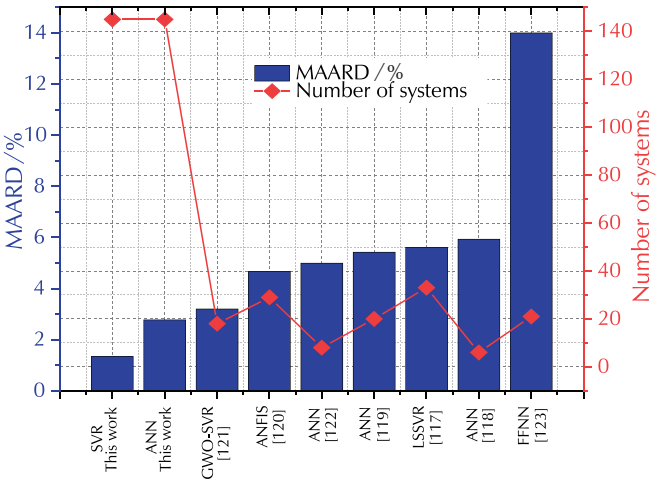

Fig. 5 – Comparison between the developed models in this study and those previously published in literature in terms of mean average absolute relative deviation

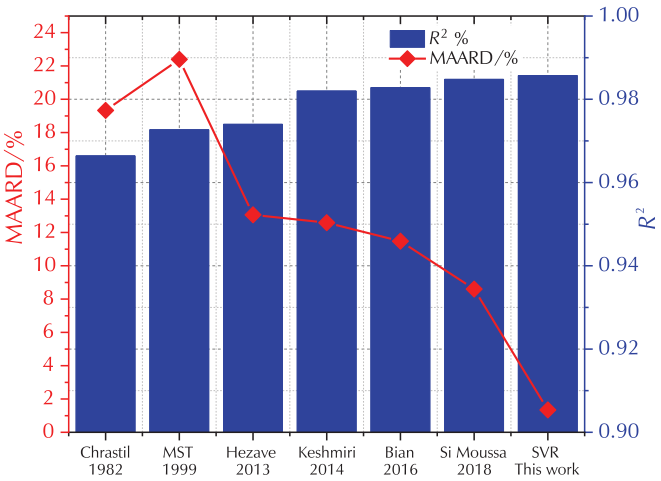

Fig. 6 – Global graphical comparison for each of the 06 density-based models and our SVR-QSPR model

Table 10 – Global comparison for each of the 06 density-based models and our SVR-QSPR model

Errors	*Chrastil* (1982)[113]	*Mèndez-Santiago and Teja* (1999)[114]	*Hezave and Lashkarbolooki* (2013)[5]	*Keshmiri* (2014)[115]	*Bian* (2016)[116]	*Si Moussa* (2018)[2]	Our model SVR-QSPR
mean AARD/%	19.3312	22.4011	13.0631	12.5889	11.4792	8.6026	**1.345**
R^2	0.9664	0.9727	0.9740	0.9820	0.9828	0.9848	**0.9857**

Table 9 – Density-based models for the correlation of solid solutes in $scCO_2$

Models	Equations	Refs.
Chrastil, 1982	$\ln(y_2) = a_0 + a_1 \ln(\rho_1)$	113
Mèndez-Santiago and Teja (MST), 1999	$\ln(y_2) = a_0 + \frac{a_1}{T} + a_2 P^2 + \left(a_3 + \frac{a_4}{T}\right) ln(\rho_1)$	114
Hezave and Lashkarbolooki, 2013	$\ln\left(\frac{P(y_2 - dxT)}{P^{ref}}\right) = a + \frac{b}{T} + c\left(\rho_1 - \rho_{ref}\right)$	5
Keshmiri et al., 2014	$\ln(y_2) = a_0 + \frac{a_1}{T} + a_2 P^2 + \left(a_3 + \frac{a_4}{T}\right) \ln(\rho_1)$	115
Bian et al., 2016	$\ln(y_2) = a_0 + \frac{a_1}{T} + \frac{a_2 \rho_1}{T} + \left(a_3 + a_4 \rho_1\right) \ln(\rho_1)$	116
Si Moussa et al., 2018	$\ln(y_2) = a_0 + a_1 \rho_1 + a_2 \rho_1^2 + a_3 \rho_1 T + a_4 T + a_5 T^2 + a_6 \ln(\rho_1) + \frac{a_7}{T}$	2

Sharing the obtained results by SVR model with users might be interesting using a simple interface. In Fig. 7, a convivial graphical user interface based on MATLAB software and using the trained SVR model was developed to evaluate the solubility of 145 solid solutes in $scCO_2$.

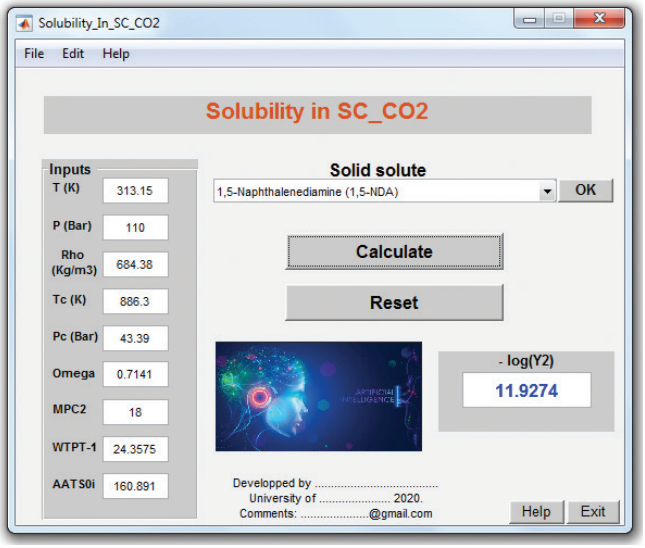

Fig. 7 – A MATLAB calculator to evaluate the solubility of 145 solid solutes in $scCO_2$ using support vector machine model

where i and k are the dimensions of the input matrix (i = 1 to 3637, k = 1 to m), j is the dimension of the output vector (j = 1 to 3637), and m is the number of input parameters (m = 9).

The importance of input parameters in QSPR-SVM for solid solute solubility in $scCO_2$ is shown in Fig. 8. As observed, all parameters have approximately the same effects as well. The results prove that all of the designated input parameters in this study have crucial effects on the solid solute solubility in $scCO_2$, hence, they have been appropriately selected.

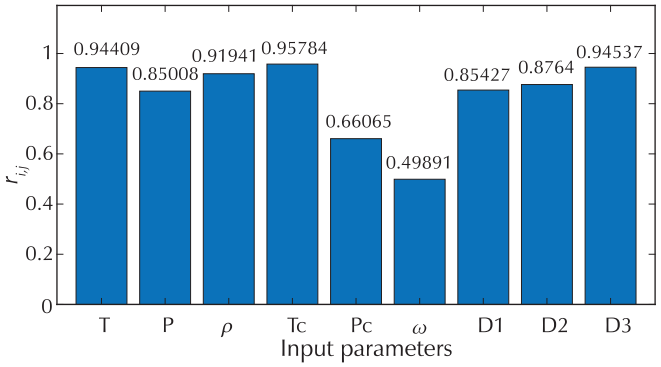

Fig. 8 – Sensitivity analysis between the input variables and the calculated $-\log(y_2)$ in $scCO_2$

5 Global Sensitivity analysis

The objective of the global sensitivity analysis is a mathematical technique applied to find which input variable is considered more important for prediction of solubility values.[124,125] In this work, the cosine amplitude method (CAM) was employed to assess the effect of each input on the output.[126] The used data pairs build a data array X which are defined as $X = \{X_1, X_2, X_3,...,X_m\}$, where X_i is a vector of lengths m, expressed as: $X_i = \{x_{i1},x_{i2},x_{i3},...,x_{im}\}$. The r_{ij} values are defined as the strength relationships between the output x_j and input x_i parameters, and can be calculated using Eq. (1).

$$r_{ij} = \frac{\sum_{k=1}^{m} x_{ik} x_{jk}}{\sqrt{\sum_{k=1}^{m} x_{ij}^2 \sum_{k=1}^{m} x_{jk}^2}} \quad (1)$$

6 Extrapolation capability

In this section, the extrapolation capability of the proposed model was discussed using the experimental measurements reported in the following papers.[127–134] The predicted solubility values obtained by the SVR-QSPR model were correlated to the experimental ones of eight systems and also depicted clearly in terms of $scCO_2$ density. The results of the comparison, shown in Figs. (9–16), indicated that the model performed well with a high goodness of fit R during the prediction of the solubility of eight systems, and almost all predicted solubility values of the eight systems fall inside the experimental ones. These results demonstrated clearly that the performance of SVR-QSPR based forecasting model is quite satisfactory.

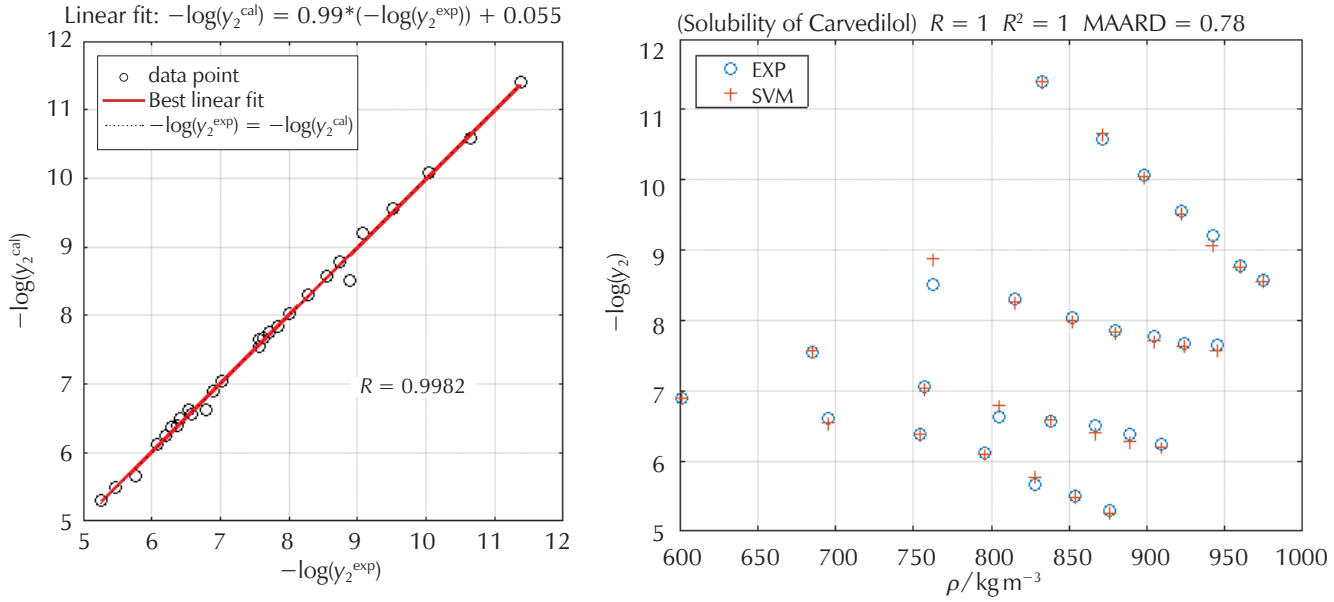

Fig. 9 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Carvedilol system, b) Comparison between experimental data[127] and SVR-QSPR predicted results for Carvedilol system *vs* $scCO_2$ density

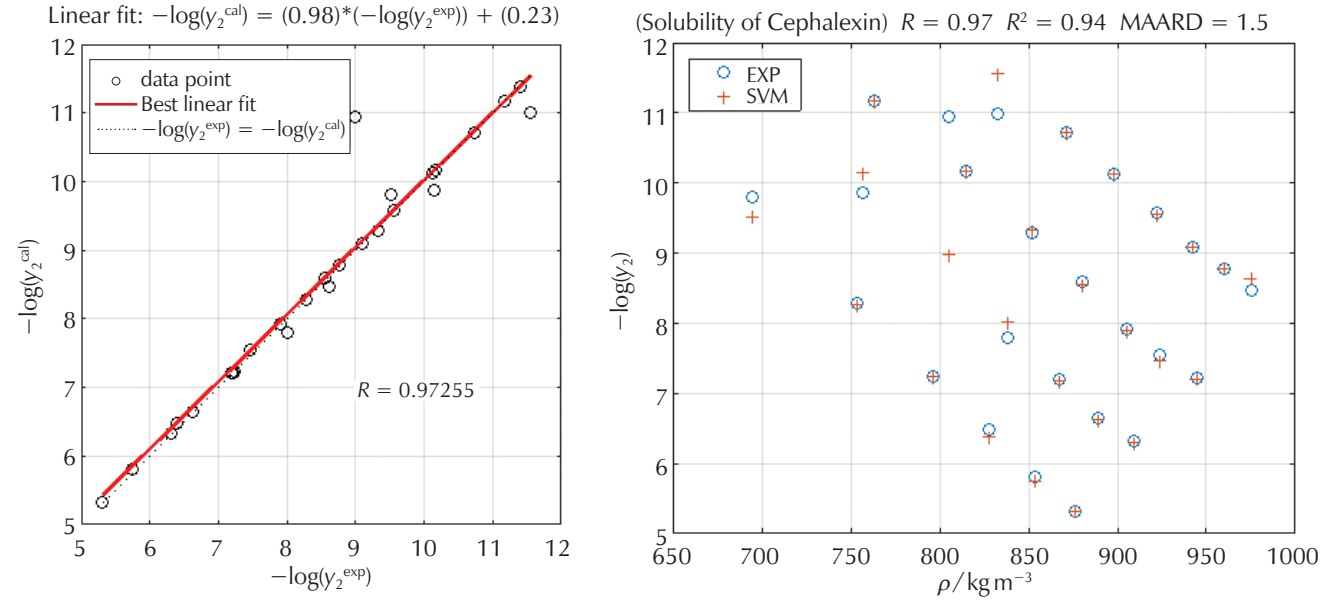

Fig. 10 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Cephalexin system, b) Comparison between experimental data[128] and SVR-QSPR predicted results for Cephalexin system *vs* $scCO_2$ density

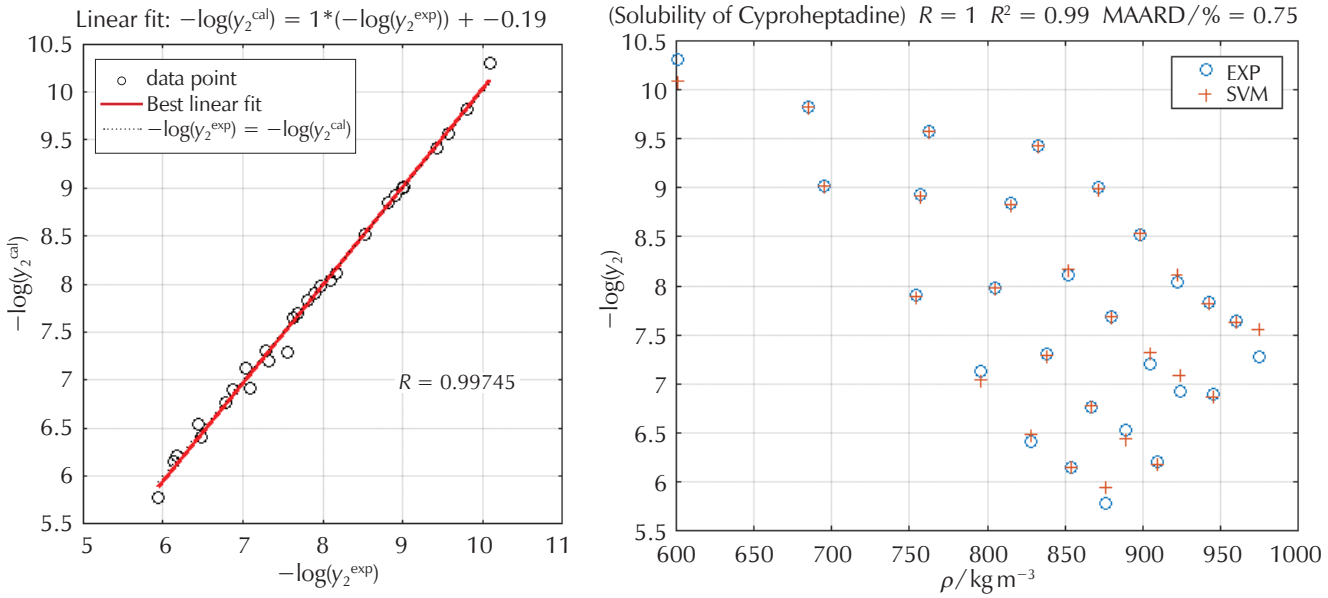

Fig. 11 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Cyproheptadine system, b) Comparison between experimental data[129] and SVR-QSPR predicted results for Cyproheptadine system *vs* $scCO_2$ density

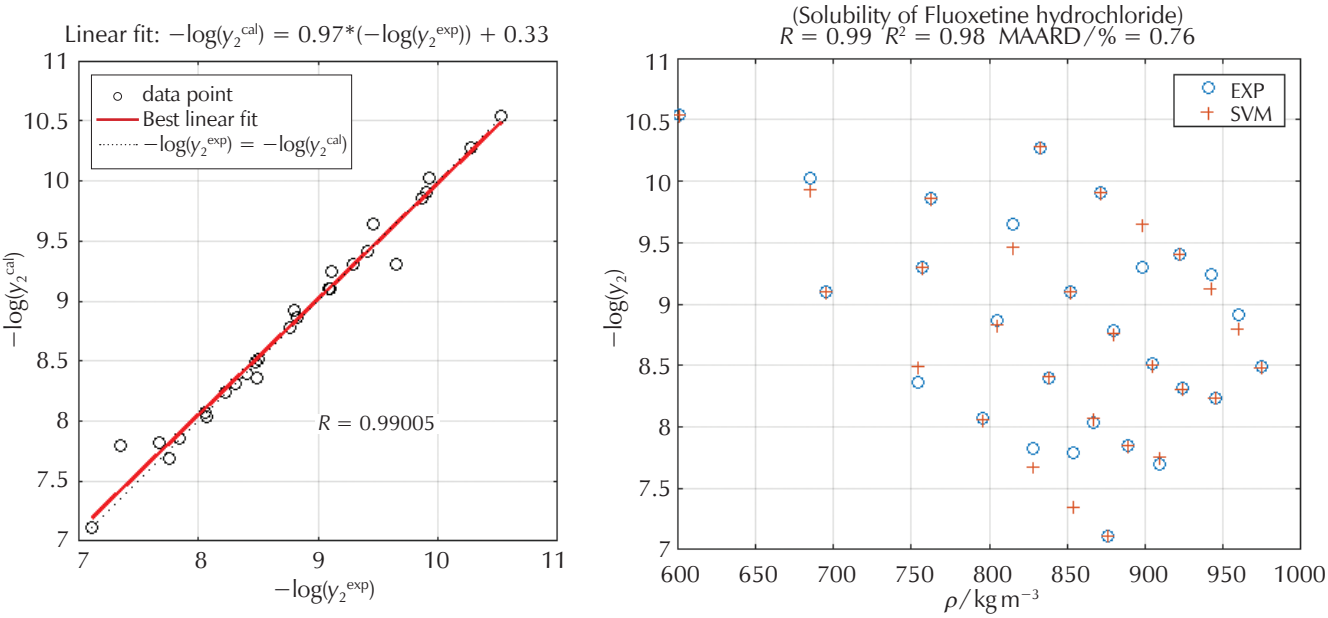

Fig. 12 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Fluoxetine hydrochloride system, b) Comparison between experimental data[130] and SVR-QSPR predicted results for Fluoxetine hydrochloride system *vs* $scCO_2$ density

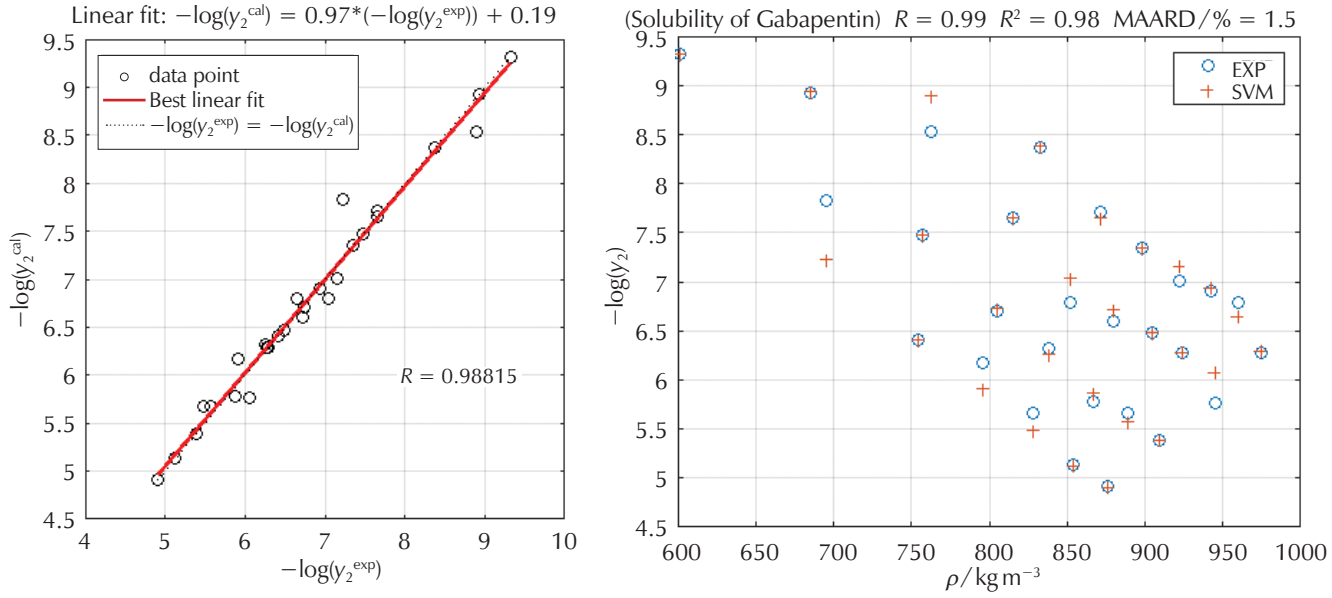

Fig. 13 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Gabapentin system, b) Comparison between experimental data[131] and SVR-QSPR predicted results for Gabapentin system *vs* $scCO_2$ density

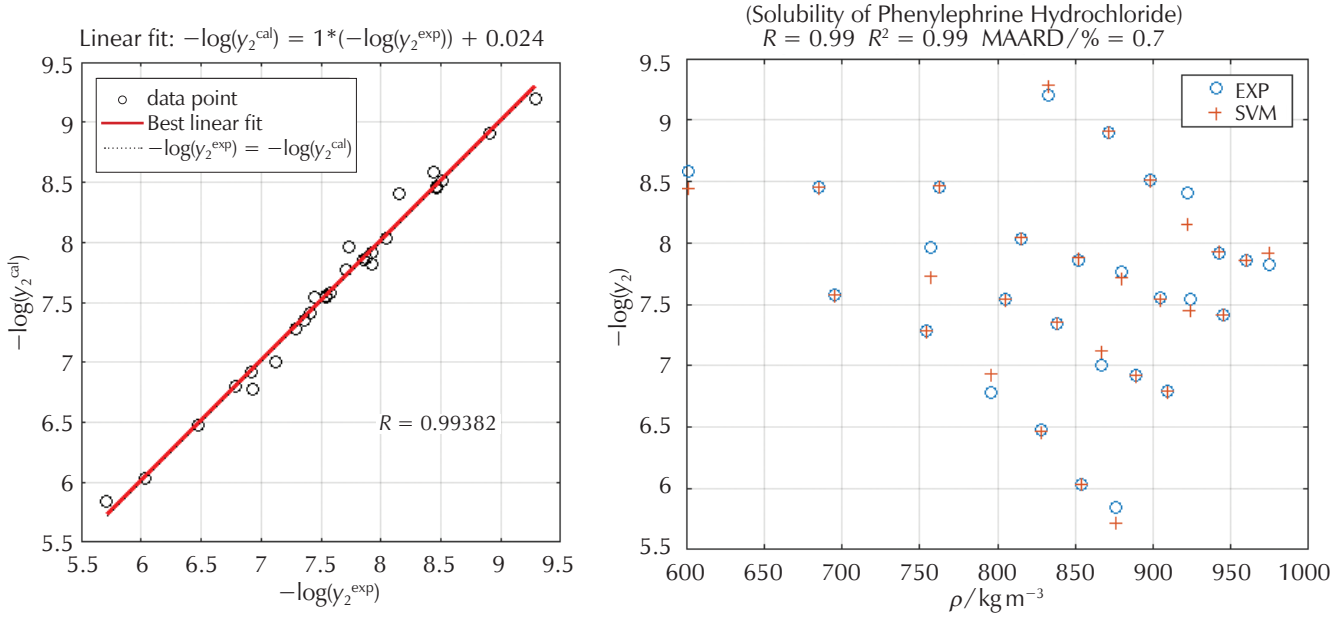

Fig. 14 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Phenylephrine Hydrochloride system, b) Comparison between experimental data[132] and SVR-QSPR predicted results for Phenylephrine Hydrochloride system *vs* $scCO_2$ density

7 Conclusion

In this study, a comparison was conducted between support vector machine (SVM) and artificial neural network (ANN) to predict non-linear variation of the solubility of 145 different compounds in supercritical carbon dioxide. Results show clearly that coupling the SVR approach with QSPR is more accurate with R^2 and AARD% of 0.9857 and 1.345 %, respectively, in comparison with ANN-QSPR for estimating the drug solubility in $scCO_2$. The predictive ability index of the optimal model was estimated with a leave-one-out cross-validated coefficient (Q^2_{LOO}) of 0.9859.

Based on the calculated statistical parameters, the QSPR-SVM has better performance than QSPR-ANN. Therefore, advantageously, the proposed QSPR-SVM as a simple approach to forecasting solid solute solubility in $scCO_2$ can be used in commercial software in the absence of experimental data. Also, QSPR-SVM modelling for the prediction of solubility of solid solute in $scCO_2$ provided in this study shows better performance in comparison with the other previous correlations. A sensitivity analysis has been conducted to show the influence of each input parameter on the output variable. All parameters have approximately the same effects as well.

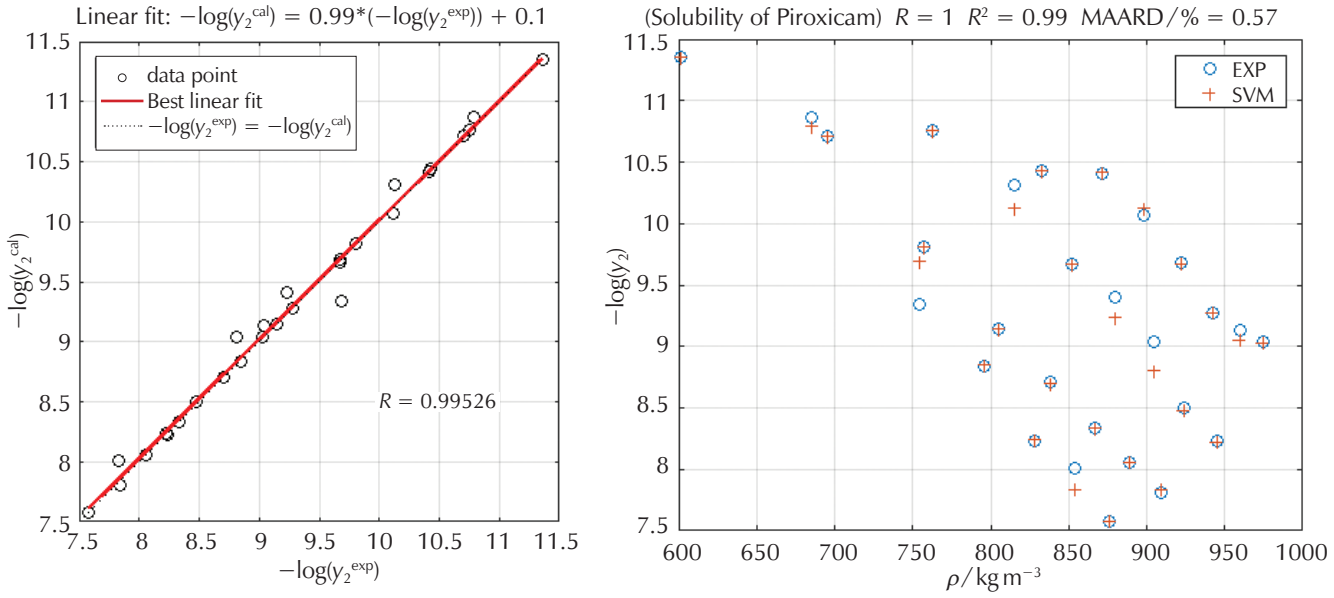

Fig. 15 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Piroxicam system, b) Comparison between experimental data[133] and SVR-QSPR predicted results for Piroxicam system *vs* $scCO_2$ density

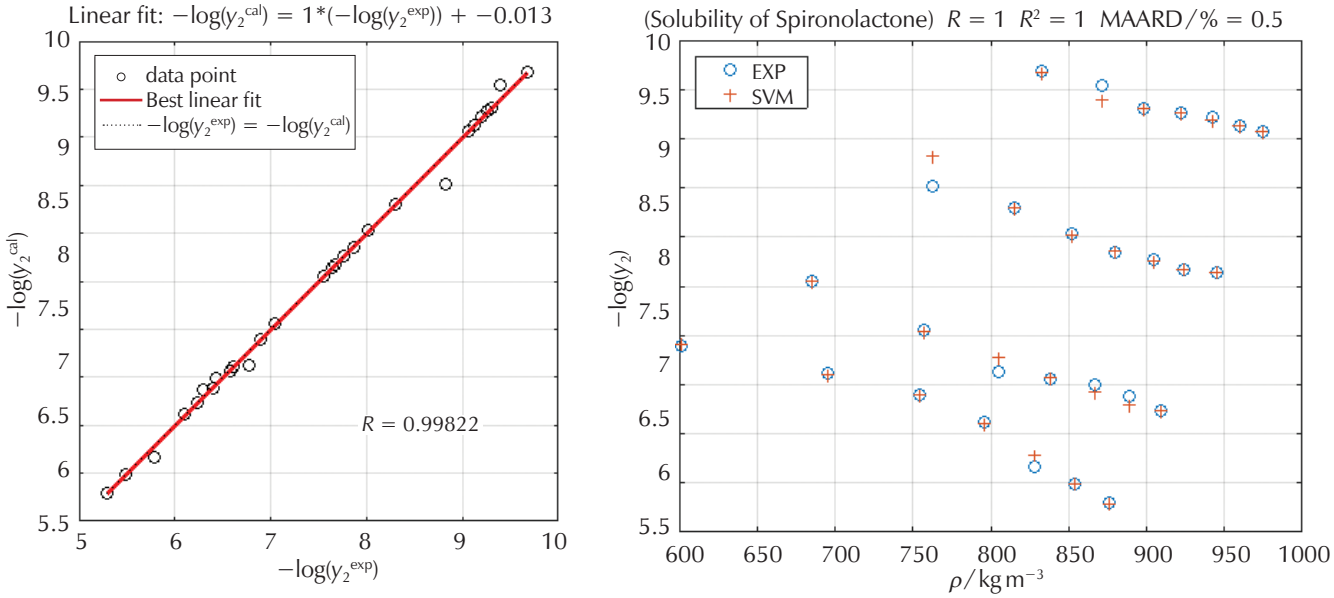

Fig. 16 – a) Plot of predicted solute solubility values in $scCO_2$ *vs* experimental ones for Spironolactone system, b) Comparison between experimental data[134] and SVR-QSPR predicted results for Spironolactone system *vs* $scCO_2$ density

List of abbreviations

ANFIS – Adaptive Neuro-Fuzzy Inference System
ANN – Artificial Neural Network
C – Penalising Parameter
FFNN – Feed Forward Neural Network
GA – Genetic Algorithm
GWO – Grey Wolf Optimiser
LS-SVR – Least Square Support Vector Regress
MAARD – Mean Average Absolute Relative Deviation, %
MSE – Mean Squared Error, %
QSPR – Quantitative Structure-Property Relationship
R – Correlation Coefficient
RBF – Radial Basis Function
RMSE – Root Mean Square Error, %
s – Kernel Width Parameter
sc – Supercritical
SVR – Support Vector Regress
ϵ – Error-Accuracy Parameter

References

Literatura

1. *A. J. A. Meirelles*, Prediction of solid solute solubility in supercritical CO_2 with cosolvents using the CPAEoS, Fluid Phase Equilib. **482** (2019) 1–10, doi: https://doi.org/10.1016/j.fluid.2018.10.020.
2. *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, doi: https://doi.org/10.1016/j.crci.2018.02.006.
3. *A. R. Katritzky, M. Kuanar, S. Slavov, C. D. Hall, M. Karelson, I. Kahn, D. A. Dobchev*, Quantitative correlation of physical and chemical properties with chemical structure: utility for prediction, Chem. Rev. **110** (2010) 5714–5789, doi: https://doi.org/10.1021/cr900238d.
4. *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, doi: https://doi.org/10.1007/s00521-015-2038-1.
5. *A. Z. Hezave, M. Lashkarbolooki*, A new simple correlation for calculating solubility of drugs in supercritical carbon dioxide, J. Theor. Comput. Chem. **12** (2013) 1350062, doi: https://doi.org/10.1142/S0219633613500624.
6. *M. Abdi-Khanghah, A. Bemani, Z. Naserzadeh, Z. Zhang*, Prediction of solubility of *N*-alkanes in supercritical CO_2 using RBF-ANN and MLP-ANN, J. CO2 Util. **25** (2018) 108–119, doi: https://doi.org/10.1016/j.jcou.2018.03.008.
7. *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, doi: https://doi.org/10.1016/j.cherd.2011.03.017.
8. *A. Ray, T. Halder, S. Jena, A. Sahoo, B. Ghosh, S. Mohanty, N. Mahapatra, S. Nayak*, Application of artificial neural network (ANN) model for prediction and optimization of coronarin D content in Hedychium coronarium, Ind. Crops Prod. **146** (2020) 112186, doi: https://doi.org/10.1016/j.indcrop.2020.112186.
9. *A. Oppio, S. Corsi*, Territorial vulnerability and local conflicts perspectives for waste disposals siting. A case study in Lombardy region (Italy), J. Clean. Prod. **141** (2017) 1528–1538, doi: https://doi.org/10.1016/j.jclepro.2016.09.203.
10. *W. H. Chen, S. H. Hsu, H. P. Shen*, Application of SVM and ANN for intrusion detection, Comput. Oper. Res. **32** (2005) 2617–2634, doi: https://doi.org/10.1016/j.cor.2004.03.019.
11. *D. K. Agrafiotis, W. Cedeno, V. S. Lobanov*, On the use of neural network ensembles in QSAR and QSPR, J. Chem. Inf. Comput. Sci. **42** (2002) 903–911, doi: https://doi.org/10.1021/ci0203702.
12. *T. N. G. Borhani, A. Afzali, M. Bagheri*, QSPR estimation of the auto-ignition temperature for pure hydrocarbons, Process Saf. Environ. Prot. **103** (2016) 115–125, doi: https://doi.org/10.1016/j.psep.2016.07.004.
13. *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, doi: https://doi.org/10.1016/j.ces.2008.05.031.
14. *K. Khimeche, P. Alessi, I. Kikic, A. Dahmani*, Solubility of diamines in supercritical carbon dioxide. Experimental determination and correlation, J. Supercrit. Fluids **41** (2007) 10–19, doi: https://doi.org/10.1016/j.supflu.2006.09.004.
15. *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, doi: https://doi.org/10.1016/j.fluid.2009.04.019.
16. *Z. Huang, S. Kawi, Y. C. Chiew*, Solubility of cholesterol and its esters in supercritical carbon dioxide with and without cosolvents, J. Supercrit. Fluids **30** (2004) 25–39, doi: https://doi.org/10.1016/S0896-8446(03)00116-5.
17. *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, doi: https://doi.org/10.1016/j.supflu.2006.10.006.
18. *Y. Yamini, N. Bahramifar*, Solubility of polycyclic aromatic hydrocarbons in supercritical carbon dioxide, J. Chem. Eng. Data **45** (2000) 53–56, doi: https://doi.org/10.1021/je990129s.
19. *D. L. Sparks, R. Hernandez, L. A. Estévez, N. Meyer, T. French*, Solubility of Azelaic Acid in Supercritical Carbon Dioxide, J. Chem. Eng. Data **52** (2007) 1246–1249, doi: https://doi.org/10.1021/je600572z.
20. *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, doi: https://doi.org/10.1021/je020218w.
21. *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, doi: https://doi.org/10.1016/S0378-3812(02)00070-5.
22. *M. Asghari-Khiavi, Y. Yamini*, Solubility of the Drugs Bisacodyl, Methimazole, Methylparaben, and Iodoquinol in Supercritical Carbon Dioxide, J. Chem. Eng. Data **1** (2003) 61–65, doi: https://doi.org/10.1021/JE020080H.
23. *M. Johannsen, G. Brunner*, Solubilities of the xanthines caffeine, theophylline and theobromine in supercritical carbon dioxide, Fluid Phase Equilib. **95** (1994) 215–226, doi: https://doi.org/10.1016/0378-3812(94)80070-7.
24. *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, doi: https://doi.org/10.1016/j.fluid.2005.11.028.
25. *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, doi: https://doi.org/10.1016/j.supflu.2009.12.012.
26. *P. Coimbra, D. Fernandes, P. Ferreira, M. H. Gil, H. C. de Sousa*, Solubility of Irgacure® 2959 photoinitiator in supercritical carbon dioxide: Experimental determination and correlation, J. Supercrit. Fluids **45** (2008) 272–281, doi: https://doi.org/10.1016/j.supflu.2008.01.014.
27. *A. G. Reveco-Chilla, A. L. Cabrera, J. C. de la Fuente, F. C. Zacconi, J. M. del Valle, L. M. Valenzuela*, Solubility of menadione and dichlone in supercritical carbon dioxide, Fluid Phase Equilib. **423** (2016) 84–92, doi: https://doi.org/10.1016/j.fluid.2016.04.001.
28. *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, doi: https://doi.org/10.1016/S0731-7085(03)00016-5.
29. *P. Kotnik, M. Škerget, Z. 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, doi: https://doi.org/10.1021/je100697a.
30. *M. Khamda, M. H. Hosseini, M. Rezaee*, Measurement and

correlation solubility of cefixime trihydrate and oxymetholone in supercritical carbon dioxide (CO_2), J. Supercrit. Fluids **73** (2013) 130–137, doi:https://doi.org/10.1016/j.supflu.2012.09.006.

31. *Z. Huang, Y.-H. Guo, H. Miao, L.-J. Teng*, Solubility of progesterone in supercritical carbon dioxide and its micronization through RESS, Powder Technol. **258** (2014) 66–77, doi: https://doi.org/10.1016/j.powtec.2014.03.009.
32. *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, doi: https://doi.org/10.1016/j.supflu.2016.02.014.
33. *J. García-González, M. J. Molina, F. Rodríguez, F. Mirada*, Solubilities of Phenol and Pyrocatechol in Supercritical Carbon Dioxide, J. Chem. Eng. Data **46** (4) (2001) 918–921, doi: https://doi.org/10.1021/JE0003795.
34. *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, doi: https://doi.org/10.1021/je0003795.
35. *Y. Yamini, M. R. Fat'Hi, N. Alizadeh, M. Shamsipur*, Solubility of dihydroxybenzene isomers in supercritical carbon dioxide, Fluid Phase Equilib. **152** (1998) 299–305, doi: https://doi.org/10.1016/S0378-3812(98)00385-9.
36. *F. Zabihi, M. Mirzajanzadeh, J. Jia, Y. Zhao*, Measurement and calculation of solubility of quinine in supercritical carbon dioxide, Chinese J. Chem. Eng. **25** (2017) 641–645, doi: https://doi.org/10.1016/j.cjche.2016.10.003.
37. *J. Ke, C. Mao, M. Zhong, B. Han, H. Yan*, Solubilities of salicylic acid in supercritical carbon dioxide with ethanol cosolvent, J. Supercrit. Fluids **9** (1996) 82–87, doi: https://doi.org/10.1016/S0896-8446(96)90002-9.
38. *G. I. Burgos-Solórzano, J. F. Brennecke, M. A. Stadtherr*, Solubility measurements and modeling of molecules of biological and pharmaceutical interest with supercritical CO_2, Fluid Phase Equilib. **220** (2004) 57–69, doi: https://doi.org/10.1016/j.fluid.2004.01.036.
39. *S. J. Nejad, R. Mohammadikhah, H. Abolghasemi, M. A. Moosavian, M. G. Maragheh*, A Novel equation of state (EOS) for prediction of solute solubility in supercritical carbon dioxide: Experimental determination and correlation, Can. J. Chem. Eng. **87** (2009) 930–938, doi: https://doi.org/10.1002/cjce.20232.
40. *S. Y. Huang, M. Tang, S. L. Ho, Y. P. Chen*, Solubilities of N-phenylacetamide, 2-methyl-*N*-phenylacetamide and 4-methyl-*N*-phenylacetamide in supercritical carbon dioxide, J. Supercrit. Fluids **42** (2007) 165–171, doi: https://doi.org/10.1016/j.supflu.2007.04.001.
41. *G. Tian, J. Jin, Z. Zhang, J. Guo*, Solubility of mixed solids in supercritical carbon dioxide, Fluid Phase Equilib. **251** (2007) 47–51, doi: https://doi.org/10.1016/j.fluid.2006.11.002.
42. *K. Tamura, R. S. Alwi, T. Tanaka, K. Shimizu*, Solubility of 1-aminoanthraquinone and 1-nitroanthraquinone in supercritical carbon dioxide, J. Chem. Thermodyn. **104** (2017) 162–168, doi: https://doi.org/10.1016/j.jct.2016.09.032.
43. *R. Murga, M. T. Sanz, S. Beltrán, J. L. Cabezas*, Solubility of three hydroxycinnamic acids in supercritical carbon dioxide, J. Supercrit. Fluids **27** (2003) 239–245, doi: https://doi.org/10.1016/S0896-8446(02)00265-6.
44. *M. Johannsen, G. Brunner*, Solubilities of the Fat-Soluble Vitamins A, D, E, and K in Supercritical Carbon Dioxide, J. Chem. Eng. Data **42** (1997) 106–111, doi: https://doi.org/10.1021/je960219m.
45. *K.-L. Tsai, F.-N. Tsai*, Solubilities of Methylbenzoic Acid Isomers in Supercritical Carbon Dioxide, J. Chem. Eng. Data **40** (1995) 264–266, doi: https://doi.org/10.1021/je00017a057.
46. *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, doi: https://doi.org/10.1016/j.supflu.2015.05.018.
47. *V. Pauchon, Z. Cissé, M. Chavret, J. Jose*, A new apparatus for the dynamic determination of solid compounds solubility in supercritical carbon dioxide: Solubility determination of triphenylmethane, J. Supercrit. Fluids **32** (2004) 115–121, doi: https://doi.org/10.1016/j.supflu.2004.03.003.
48. *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, doi: https://doi.org/10.1016/j.fluid.2017.07.015.
49. *K. Tamura, R. S. Alwi*, Solubility of anthraquinone derivatives in supercritical carbon dioxide, Dye. Pigment. **113** (2015) 351–356, doi: https://doi.org/10.1016/j.dyepig.2014.09.003.
50. *J. Jin, Y. Wang, H. Liu, Z. Zhang*, Determination and calculation of solubility of bisphenol A in supercritical carbon dioxide, Chem. Eng. Res. Des. **91** (2013) 158–164, doi: https://doi.org/10.1016/j.cherd.2012.06.013.
51. *J. Jin, Y. Ning, K. Hu, H. Wu, Z. Zhang*, Solubility of p -Nitroaniline in Supercritical Carbon Dioxide with and without Mixed Cosolvents, J. Chem. Eng. Data **58** (2013) 1464–1469, doi: https://doi.org/10.1021/je300987d.
52. *S. Ismadji*, Solubility of methyl salicylate in supercritical carbon dioxide at several temperatures, J. Chem. Eng. Data **53** (2008) 2207–2210, doi: https://doi.org/10.1021/je800476n.
53. S. *Marceneiro, P. Coimbra, M. E. M. Braga, A. M. A. Dias, H. C. De Sousa*, Measurement and correlation of the solubility of juglone in supercritical carbon dioxide, Fluid Phase Equilib. **311** (2011) 1–8, doi: https://doi.org/10.1016/j.fluid.2011.08.024.
54. *J. Yau, F. Tsai*, Solubilities of 1-Eicosanol and Eicosanoic Acid in Supercritical Carbon Dioxide from 308.2 to 328.2 K at Pressures to 21.26 MPa, J. Chem. Eng. Data **39** (1994) 827–829, doi: https://doi.org/10.1021/je00016a042.
55. *A. Kramer, G. Thodos,* Solubility of *I*-Octadecanol and Stearic Acid in Supercritical Carbon Dioxide, J. Chem. Eng. Data **34** (2) (1989) 184–187, doi: https://doi.org/10.1021/je00056a011.
56. *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, doi: https://doi.org/10.1016/j.jfoodeng.2009.06.041.
57. *I. Goodarznia, F. Esmaeilzadeh*, Solubility of an anthracene, phenanthrene, and carbazole mixture in supercritical carbon dioxide, J. Chem. Eng. Data **47** (2002) 333–338, doi: https://doi.org/10.1021/je010093f.
58. *M. Ashraf-Khorassani, M. T. Combs, L. T. Taylor, F. K. Schweighardt, P. S. Mathias*, Solubility study of sulfamethazine and sulfadimethoxine in supercritical carbon dioxide, fluoroform, and subcritical freon 134A, J. Chem. Eng. Data **42** (1997) 636–640, doi: https://doi.org/10.1021/je960402f.
59. *J. W. Hampson*, A Recirculating Equilibrium Procedure for Determining Organic Compound Solubility in Supercritical Fluids. Anthracene in Carbon Dioxide, J. Chem. Eng. Data **41** (1996) 97–100, doi: https://doi.org/10.1021/je950167l.
60. *A. Cortesi, I. Kikic, P. Alessi, G. Turtoi, S. Garnier*, Effect of chemical structure on the solubility of antioxidants in supercritical carbon dioxide : experimental data and correlation, J. Supercritic. Fluid. **14** (2) (1999) 139–144, doi: https://doi.org/10.1016/S0896-8446(98)00119-3.
61. *L. Barna, J. Blanchard, E. Rauzy, C. Berro, D. Lyon*, Solubility of Flouranthene , Chrysene , and Triphenylene in Super-

critical Carbon Dioxide, J. Chem. Eng. Data **41** (6) (1996) 1466–1469, doi: https://doi.org/10.1021/je960189n.

62. *J. Gregorowicz*, Solubilities of lactic acid and 2-hydroxyhexanoic acid in supercritical CO_2, Fluid Phase Equilib. **166** (1999) 39–46, doi: https://doi.org/10.1016/S0378-3812(99)00283-6.
63. *E. Reverchon, P. Russo, A. Stassi*, Solubilities of Solid Octacosane and Triacontane in Supercritical Carbon Dioxide, J. Chem. Eng. Data **38** (1993) 458–460, doi: https://doi.org/10.1021/je00011a034.
64. *C. Garlapati, G. Madras*, Solubilities of hexadecanoic and octadecanoic acids in supercritical CO with and without cosolvents, J. Chem. Eng. Data **53** (2008) 2913–2917, doi: https://doi.org/10.1021/je8007149.
65. *H. Xing, Y. Yang, B. Su, M. Huang, Q. Ren*, Solubility of Artemisinin in Supercritical Carbon Dioxide, J. Chem. Eng. Data **48** (2) (2003) 330–332, doi: https://doi.org/10.1021/je025575l.
66. *S. H. Cheng, F. C. Yang, Y. H. Yang, C. C. Hu, W. T. Chang*, Measurements and modeling of the solubility of ergosterol in supercritical carbon dioxide, J. Taiwan Inst. Chem. Eng. **44** (2013) 19–26, doi: https://doi.org/10.1016/j.jtice.2012.09.001.
67. *S. L. J. Yun, K. K. Liong, G. S. Gurdial, N. R. Foster*, Solubility of cholesterol in supercritical carbon dioxide, Ind. Eng. Chem. Res. **30** (1991) 2476–2482, doi: https://doi.org/10.1021/ie00059a018.
68. *S. J. Macnaughton, N. R. Foster*, Solubility of DDT and 2,4-D in Supercritical Carbon Dioxide and Supercritical Carbon Dioxide Saturated with Water, Ind. Eng. Chem. Res. **33** (1994) 2757–2763, doi: https://doi.org/10.1021/ie00035a027.
69. *R. Murga, M. T. Sanz, S. Beltrán, J. L. Cabezas*, Solubility of some phenolic compounds contained in grape seeds, in supercritical carbon dioxide, J. Supercrit. Fluids **23** (2002) 113–121, doi: https://doi.org/10.1016/S0896-8446(02)00033-5.
70. *S. Marceneiro, M. E. M. Braga, A. M. A. Dias, H. C. De Sousa*, Measurement and Correlation of 1,4-Naphthoquinone and of Plumbagin Solubilities in Supercritical Carbon Dioxide, J. Chem. Eng. Data **56** (11) (2011) 4173–4182, doi: https://doi.org/10.1021/je200675g.
71. *P. Coutsikos, K. Magoulas, D. Tassios*, Solubilities of *p*-Quinone and 9 , 10-Anthraquinone in Supercritical Carbon Dioxide, J. Chem. Eng. Data **42** (1997) 463–466, doi: https://doi.org/10.1021/je960309r.
72. *R. Murga, M. T. Sanz, S. Beltrán, J. L. Cabezas*, Solubility of syringic and vanillic acids in supercritical carbon dioxide, J. Chem. Eng. Data **49** (2004) 779–782, doi: https://doi.org/10.1021/je034129a.
73. *S. Bristow, B. Y. Shekunov, P. York*, Solubility Analysis of Drug Compounds in Supercritical Carbon Dioxide Using Static and Dynamic Extraction Systems, Ind. Eng. Chem. Res. **40** (2001) 1732–1739, doi: https://doi.org/10.1021/ie0002834.
74. *J. Shi, M. Khatri, S. J. Xue, G. S. Mittal, Y. Ma, D. Li*, Solubility of lycopene in supercritical CO2 fluid as affected by temperature and pressure, Sep. Purif. Technol. **66** (2009) 322–328, doi: https://doi.org/10.1016/j.seppur.2008.12.012..
75. *E. Kosal, C. H. Lee, G. D. Holder*, Solubility of progesterone, testosterone, and cholesterol in supercritical fluids, J. Supercrit. Fluids **5** (1992) 169–179, doi: https://doi.org/10.1016/0896-8446(92)90004-4.
76. *N. R. Foster, S. L. J. Yun, S. S. T. Ting*, Solubility of oleic acid in supercritical carbon dioxide, J. Supercrit. Fluids **4** (1991) 127–130, doi: https://doi.org/10.1016/0896-8446(91)90041-4.
77. *J. P. Coelho, G. P. Naydenov, D. S. Yankov, R. P. Stateva*, Experimental measurements and correlation of the solubility of three primary amides in supercritical CO_2: Acetanilide, propanamide, and butanamide, J. Chem. Eng. Data **58** (2013) 2110–2115, doi: https://doi.org/10.1021/je400357t.
78. *Y. P. Chen, Y. M. Chen, M. Tang*, Solubilities of cinnamic acid, phenoxyacetic acid and 4-methoxyphenylacetic acid in supercritical carbon dioxide, Fluid Phase Equilib. **275** (2009) 33–38, doi: https://doi.org/10.1016/j.fluid.2008.09.009.
79. *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, doi: https://doi.org/10.1016/j.supflu.2015.09.026.
80. *H. Higashi, Y. Iwai, K. Miyazaki, Y. Ogino, M. Oki, Y. Arai*, Measurement and correlation of solubilities for trifluoromethylbenzoic acid isomers in supercritical carbon dioxide, J. Supercrit. Fluids **33** (2005) 15–20, doi: https://doi.org/10.1016/j.supflu.2004.03.006.
81. *H. Perrotin-Brunel, M. J. E. Van Roosmalen, M. C. Kroon, J. Van Spronsen, G. J. Witkamp, C. J. Peters*, Solubility of cannabinol in supercritical carbon dioxide, J. Chem. Eng. Data **55** (2010) 3704–3707, doi: https://doi.org/10.1021/je100245n.
82. *J. P. Coelho, K. Bernotaityte, M. A. Miraldes, A. F. Mendonça, R. P. Stateva*, Solubility of ethanamide and 2-propenamide in supercritical carbon dioxide. Measurements and correlation, J. Chem. Eng. Data **54** (2009) 2546–2549, doi: https://doi.org/10.1021/je900109b.
83. *J. Lim, H. Kim, H. K. Cho, M. S. Shin*, Solubility of hinokitiol in supercritical fluids; measurement and correlation, Korean J. Chem. Eng. **28** (2011) 2319–2323, doi: https://doi.org/10.1007/s11814-011-0112-7.
84. *M. S. Shin, H. Kim*, Solubility of iodopropynyl butylcarbamate in supercritical carbon dioxide, Fluid Phase Equilib. **270** (2008) 45–49, doi: https://doi.org/10.1016/j.fluid.2008.05.010.
85. *J. Fan, Y. Hou, W. Wu, J. Zhang, S. Ren, X. Chen*, Levulinic acid solubility in supercritical carbon dioxide with and without ethanol as cosolvent at different temperatures, J. Chem. Eng. Data **55** (2010) 2316–2321, doi: https://doi.org/10.1021/je900727r.
86. *S. N. Reddy, G. Madras*, Solubilities of Benzene Derivatives in Supercritical Carbon Dioxide, J. Chem. Eng. Data **56** (2011) 1695–1699, doi: https://doi.org/10.1021/je100863p.
87. *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, doi: https://doi.org/10.1016/j.supflu.2011.12.004.
88. *R. Ch, C. Garlapati, G. Madras*, Solubility of n-(4-ethoxyphenyl)ethanamide in supercritical carbon dioxide, J. Chem. Eng. Data **55** (2010) 1437–1440, doi: https://doi.org/10.1021/je900614f.
89. *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, doi: https://doi.org/10.1016/j.fluid.2017.11.016.
90. *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, doi: https://doi.org/10.1016/j.supflu.2019.03.020.
91. *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, doi: https://doi.org/10.1016/j.supflu.2017.05.019.
92. *G. Sodeifian, N. Saadati Ardestani, S. A. Sajadian, H. Soltani*

Panah, Experimental measurements and thermodynamic modeling of Coumarin-7 solid solubility in supercritical carbon dioxide: Production of nanoparticles *via* RESS method, Fluid Phase Equilib. **483** (2019) 122–143, doi: https://doi.org/10.1016/j.fluid.2018.11.006.

93. *G. Sodeifian, S. A. Sajadian*, Solubility measurement and preparation of nanoparticles of an anticancer drug (Letrozole) using rapid expansion of supercritical solutions with solid cosolvent (RESS-SC), J. Supercrit. Fluids **133** (2018) 239–252, doi: https://doi.org/10.1016/j.supflu.2017.10.015.
94. *G. Sodeifian, F. Razmimanesh, S. A. 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, doi: https://doi.org/10.1016/j.supflu.2019.01.006.
95. *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, doi: https://doi.org/10.1016/j.supflu.2019.104606.
96. *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, doi: https://doi.org/10.1016/j.fluid.2018.05.018.
97. *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, doi: https://doi.org/10.1016/j.fluid.2019.112422.
98. *G. Sodeifian, F. Razmimanesh, S. A. Sajadian*, Prediction of solubility of sunitinib malate (an anti-cancer drug) in supercritical carbon dioxide (SC–CO_2): Experimental correlations and thermodynamic modeling, J. Mol. Liq. **297** (2020) 111740, doi: https://doi.org/10.1016/j.molliq.2019.111740.
99. *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, J. Mol. Liq. **299** (2020), doi: https://doi.org/10.1016/j.molliq.2019.112179.
100. *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, doi: https://doi.org/10.1016/j.jct.2019.105998.
101. *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 (2019), doi: https://doi.org/10.1021/acs.jced.9b00550.
102. *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, doi: https://doi.org/10.1016/j.fluid.2019.04.013.
103. *G. Sodeifian, N. Saadati Ardestani, S. A. Sajadian*, Solubility measurement of a pigment (Phthalocyanine green) in supercritical carbon dioxide: Experimental correlations and thermodynamic modeling, Fluid Phase Equilib. **494** (2019) 61–73, doi: https://doi.org/10.1016/j.fluid.2019.04.024.
104. *M. Hamadache, O. Benkortbi, S. Hanini, A. Amrane, L. Khaouane, C. Si Moussa*, A Quantitative Structure Activity Relationship for acute oral toxicity of pesticides on rats: Validation, domain of application and prediction, J. Hazard. Mater. **303** (2016) 28–40, doi: https://doi.org/10.1016/j.jhazmat.2015.09.021.
105. *V. H. Masand, V. Rastija*, PyDescriptor: A new PyMOL plugin for calculating thousands of easily understandable molecular descriptors, *Chemom. Intell. Lab. Syst.* **169** (2017) 12–18, doi: https://doi.org/10.1016/j.chemolab.2017.08.003.
106. *I. V Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V Radchenko, N. S. Zefirov, A. S. Makarenko, others*, Virtual computational chemistry laboratory--design and description, J. Comput. Aided. Mol. Des. **19** (2005) 453–463, doi: https://doi.org/10.1007/s10822-005-8694-y.
107. *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, doi: https://doi.org/10.1016/j.chemolab.2016.03.007.
108. *A. H. Elsheikh, S. W. Sharshir, M. Abd Elaziz, A. E. Kabeel, W. Guilan, Z. Haiou*, Modeling of solar energy systems using artificial neural network: A comprehensive review, Sol. Energy **180** (2019) 622–639, doi: https://doi.org/10.1016/j.solener.2019.01.037.
109. *A. M. Ghaedi, A. Vafaei*, Applications of artificial neural networks for adsorption removal of dyes from aqueous solution: A review, Adv. Colloid Interface Sci. **245** (2017) 20–39, doi: https://doi.org/10.1016/j.cis.2017.04.015.
110. *J. Zupan, J. Gasteiger*, Neural networks for chemists: an introduction (John Wiley & Sons, Inc., 1993).
111. *A. J. Smola, B. Schölkopf*, A tutorial on support vector regression, Stat. Comput. **14** (2004) 199–222, doi: https://doi.org/10.1023/B:STCO.0000035301.49549.88.
112. *V. Vapnik*, Statistical Learning Theory/Vapnik V.--NY, (1998).
113. *J. Chrastil*, Solubility of solids and liquids in supercritical gases, J. Phys. Chem. **86** (1982) 3016–3021, doi: https://doi.org/10.1021/j100212a041.
114. *J. Méndez-Santiago, A. S. Teja*, The solubility of solids in supercritical fluids, Fluid Phase Equilib. **158–160** (1999) 501–510, doi: https://doi.org/10.1016/s0378-3812(99)00154-5.
115. *K. Keshmiri, A. Vatanara, Y. Yamini*, Development and evaluation of a new semi-empirical model for correlation of drug solubility in supercritical CO_2, Fluid Phase Equilib. **363** (2014) 18–26, doi: https://doi.org/10.1016/j.fluid.2013.11.013.
116. *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, doi: https://doi.org/10.1016/j.fluid.2015.12.017.
117. *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, doi: https://doi.org/10.1016/j.supflu.2017.10.032.
118. *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, doi: https://doi.org/10.1007/s11814-018-0125-6.
119. *A. Aminian*, Estimating the solubility of different solutes in supercritical CO_2 covering a wide range of operating conditions by using neural network models, J. Supercrit. Fluids **125** (2017) 79–87, doi: https://doi.org/10.1016/j.supflu.2017.02.007.
120. *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 CO_2 using a connectionist smart technique,

J. Supercrit. Fluids **120** (2017) 181–190, doi: https://doi.org/10.1016/j.supflu.2016.06.006.

121. *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, doi: https://doi.org/10.1016/j.cherd.2017.05.008.
122. *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, doi: https://doi.org/10.1016/j.supflu.2013.02.027.
123. *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, doi: https://doi.org/10.1016/j.ces.2011.03.016.
124. *H. Ahmadi, M. Rodehutscord*, Application of Artificial Neural Network and Support Vector Machines in Predicting Metabolizable Energy in Compound Feeds for Pigs, Front. Nutr. **4** (2017) 1–8, doi: https://doi.org/10.3389/fnut.2017.00027.
125. *S. Andrea, T. Stefano, C. Francesca, R. Marco*, Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models (John Wiley & Sons, Ltd, 2004), doi: https://doi.org/10.1002/0470870958.
126. *E. Uzlu, A. Akpinar, H. T. Özturk, S. Nacar, M. Kankal*, Estimates of hydroelectric generation using neural networks with the artificial bee colony algorithm for Turkey, Energy **69** (2014) 638–647, doi: https://doi.org/10.1016/j.energy.2014.03.059.
127. *S. A. Shojaee, H. Rajaei, A. Z. Hezave, M. Lashkarbolooki, F. Esmaeilzadeh*, Experimental investigation and modeling of the solubility of carvedilol in supercritical carbon dioxide, J. Supercrit. Fluids **81** (2013) 42–47, doi: https://doi.org/10.1016/j.supflu.2013.04.013.
128. *S. A. L. I. Shojaee, H. Rajaei, A. L. I. Z. Hezave*, Experimental solubility measurement of cephalexin in supercritical carbon dioxide, Chem. Ind. Chem. Eng. Q. **20** (3) (2014) 387–396, doi: https://doi.org/10.2298/CICEQ121128021S.
129. *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, doi: https://doi.org/10.1016/j.supflu.2013.09.004.
130. *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, doi: https://doi.org/10.1016/j.supflu.2012.11.005.
131. *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, doi: https://doi.org/10.1016/j.supflu.2013.02.003.
132. *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, doi: https://doi.org/10.1016/j.supflu.2012.11.014.
133. *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, doi: https://doi.org/10.1016/j.supflu.2013.03.015.
134. *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, doi: https://doi.org/10.1016/j.fluid.2013.07.003.

SAŽETAK

Primjena umjetne neuronske mreže i regresije potpornih vektora u modeliranju kvantitativnog odnosa strukture-svojstva i topljivosti otopljenih čvrstih tvari u superkritičnom CO_2

Mohammed Moussaoui,[a,b,*] *Maamar Laidi*,[a] *Salah Hanini* [a] i *Mohamed Hentabli* [a]

U ovom je istraživanju korelirana topljivost 145 čvrstih otopljenih tvari u superkritičnom CO_2 ($scCO_2$) primjenom tehnika računalne inteligencije zasnovanim na modelima kvantitativne strukture i svojstva (QSPR). Baza podataka 3637 topljivosti prikupljena je iz prethodno objavljenih radova. Program Dragon primijenjen je za izračunavanje molekularnih deskriptora 145 čvrstih sustava. Genetski algoritam (GA) implementiran je kako bi se optimizirao podskup deskriptora sa značajnim doprinosom. Ukupno prosječno apsolutno relativno odstupanje MAARD od oko 1,345 % između eksperimentalnih i izračunatih vrijednosti pomoću regresije potpornih vektora modelom SVR-QSPR dobiveno je za predviđanje topljivosti 145 čvrstih otopljenih tvari u superkritičnom CO_2, što je bolje od onog dobivenog primjenom modela ANN-QSPR (2,772 %). Rezultati pokazuju da je razvijeni model SVR-QSPR precizniji i da se može primijeniti kao alternativni alat za modeliranje QSAR studija topljivosti otopljenih čvrstih tvari u superkritičnom ugljikovu dioksidu ($scCO_2$). Točnost predloženog modela procijenjena je statističkom analizom uspoređivanjem rezultata s ostalim modelima zabilježenim u literaturi.

Ključne riječi
Topljivost, otopljene čvrste tvari, superkritične tekućine, tehnike računalne inteligencije, kvantitativni odnos struktura-svojstvo

[a] *Laboratory of Biomaterials and Transport Phenomena (LBMPT), University of Médéa, Médéa, Alžir*
[b] *University of Bouira, Bouira, Alžir*

Izvorni znanstveni rad
Prispjelo 18. siječnja 2020.
Prihvaćeno 13. svibnja 2020.