Regular Article

PHYSICAL CHEMISTRY RESEARCH
Published by the
Iranian Chemical Society
www.physchemres.org
info@physchemres.org

Phys. Chem. Res., Vol. 12, No. 3, 567-578, September 2024
DOI: 10.22036/pcr.2023.391870.2318

Solubility of Pharmaceutical Compounds in Organic Solvents Using Artificial Nural Network and Correlation Model

M. Kiani[a,*], F. Akbari[b], N. Astaneh[a] and R. Behjatmanesh Ardakani[a]

[a]*Department of Chemistry, Payame Noor University, P. O. Box: 19395-3697, Tehran, Iran*

[b]*Department of Chemistry, Eram Institute of Higher Education, P. O. Box: 71957-46733, Shiraz, Iran*

(Received 4 April 2023, Accepted 19 November 2023)

The correlation of solid-liquid solubility using an updated semi-theoretical equation has been reported by our team earlier. This study uses an Artificial Neural Network (ANN) and a refined Apelblat model to predict the solubility of nine pharmaceutical compounds in pure organic solvents over an extensive temperature range. By training the optimized network using the back-propagation method of the Levenberg-Marquardt algorithm, optimal parameters such as neurons, hidden layers and transfer function were established in the ANN through the utilization of training, testing, and validation data. This network was trained with 764 data points and subsequently tested and validated with 164 data points. A satisfactory correlation of 1.33% was obtained from the enhanced Apelblat model across 1092 data points. The findings from thermodynamic analysis and solubility parameters can optimize the purification process in the synthesis of pharmaceutical compounds.

Keywords: Solubility, Pharmaceutical, Apelblat model, Organic solvents, Artificial neural networks

INTRODUCTION

A clear understanding of the solubility in organic and inorganic solvents is essential in the synthesis and purification of pharmaceutical and chemical materials [1-4]. The dissolution properties of medicinal compounds significantly affect their crystallization process. Hence, the selection of the optimal solvent and the right amount of drug as a solute is crucial in the pharmaceutical industry, especially in the manufacturing and development of new pharmaceutical compounds. Due to the diverse nature of organic solvents, they offer various opportunities for managing chemical and physicochemical processes [5]. Consequently, the provision of thermodynamic models, yielding a more precise correlation value to predict solubility, has been the object of extensive research aiming for obtaining the most accurate outcomes.

*Corresponding author. E-mail: M.Kiani@pnu.ac.ir

This study provides a theoretical appraisal of the solubility of nine pharmaceutical compounds in pure organic solvents. One of the compounds, 2,4-dihydro-5-methyl-2-(4-methylphenyl)-3H-pyrazol-3-one (PTMP), is a pharmaceutical intermediate with notable neuroprotective properties. Its therapeutic potential spans anti-tumor, antidepressant, antibacterial, microbicidal, anti-inflammatory, and neuroprotective effects, and it is used in dye production [6]. This drug prevents the oxidative destruction of nerve cells, showcasing its brain-protective abilities [7]. The steroid-sparing compound, 5-chloro-1-methyl-4-nitroimidazole, is a significant therapeutic agent for inflammatory diseases and multiple sclerosis [8,9]. Another compound, 4-nitro-1,2-phenylenediamine, is a vital organic intermediate used in the production of colorants, pesticides, medicine, and dyestuff [10]. The nonsteroidal anti-inflammatory drug, 2-amino-5-methylthiazole, a critical material for synthesizing meloxicam [11], had its solubility in different organic solvents evaluated through the isothermal

saturation method [12]. The compound plays an important role in the production of functional materials such as fluorescent, catalytic, and magnetic materials. It was determined that purity plays a crucial role in the fluorescent properties of synthesized material, which led to the exploration of pure 2-isopropylimidazole [13]. Phthalimides and their derivatives are widely used in dyestuff and agriculture [14]. Valsartan is commonly used in treating high blood pressure [15]. The accurate prediction of drug solubility in both pure and mixed solvents is critical for the development of purification processes in the pharmaceutical industry.

Artificial Neural Networks (ANNs) have widespread applications in the chemical field, *e.g.*, in predicting the thermodynamic characteristics of ionic liquids such as gas removal properties [16-20]. Considering the complexity of the thermodynamic models, alternative estimation methodologies such as ANNs are necessary. ANNs have been successfully deployed in diverse chemical processes, because of their advantages in rapid computations and the ability to use the models without complicated process knowledge. This research explores the potential of using ANNs in predicting the solubility of pharmaceutical compounds in organic solvents.

Twenty organic solvents were selected at various temperatures and pressures. The pharmaceutical compounds were these: 5-chloro-1-methyl-4-nitroimidazole [8], 2-amino-5-methylthiazole [12], 2-isopropylimidazole [13], valsartan [15], 4-nitro-1,2-phenylenediamine [21], phthalimide [22], terephthaldialdehyde [23], 2,4-dihydro-5-methyl-2-(4-methylphenyl)-3H-pyrazol-3-one (PTMP) [24], 2-amino-4-chlorobenzoic [25,26], and sirolimus [27].

Recently, thermodynamic models for predicting the properties and solubility of drugs have been widely studied. In this article, we briefly describe the models used in the field of solubility of compounds. Among the compounds investigated were 5-chloro-1-methyl-4-nitroimidazole, 2-amino-5-methylthiazole, 2-isopropylimidazole, valsartan, 4-nitro-1,2-phenylenediamine, phthalimide, terephthal dialdehyde, 2,4-dihydro-5-methyl-2-(4-methylphenyl)-3H-pyrazol-3-one (PTMP), 2-amino-4-chlorobenzoic and sirolimus. Table 1 provides an overview of these models.

MODELING

ANN Model

Artificial Neural Networks (ANNs) are computational methods wherein information is processed collectively and simultaneously across a network of nodes (wherein the neurons serve as the nodes facilitating information processing). ANNs are designed to mimic how biological neurons transmit and receive information across layers, with each layer comprising several neurons [38,39]. The structure of neurons is typically segmented into three layers: the hidden layer, the input layer, and the output layer. Figure 1 provides a schematic representation of an ANN.

The present study introduces a method for building neural networks using a combination of feed-forward (backpropagation) neural networks (FNN) and the Levenberg-Marquardt optimization. Weight and bias factors are two parameters that define neurons and are used for\ shifting information about input and output data and their weights. Each input is a real number, and each specific value

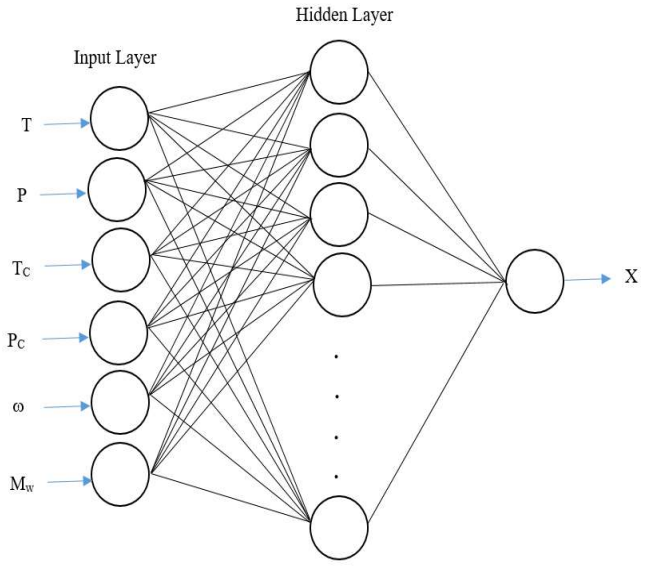

Fig. 1. Schematic diagram of the ANN model. Inputs: T: Temperature (K), P: Pressure (Pa), T_c: Critical temperature (K), P_c: Critical pressure (Pa), ω: Acentric factor, M_w: Molecular Weight (g mol^{-1}). Output: x: Solubility.

Table 1. Overview of Published Articles on Similar Compounds with Different Thermodynamic Models

Authors	Summary points	Pub date	Ref.
Akbari, F.; Didehban, K.; Farhang, M.	A new semi-empirical model for inversely correlating solid and liquid solubility data formulated with temperature	2019	[5]
Wang, J.; Xu, R.; Xu, A.; Cong, Y.	Solubility was ascertained using an isothermal saturation method, in accordance with the thermodynamic model, at temperatures from 283.15 K to 318.15 K under 101.3 kPa in 11 organic solvents	2017	[8]
Chen, G.; Chen, J.; Cheng, C.; Cong, Y.; Du, C.; Zhao, H.	Solubility of 2-Amino-5-methylthiazole was determined in several organic solvents.	2016	[12]
Choi, J. H.; Lee, H. Y.; Towns, A. D.	Solubility of Phthalimide was determined in various organic solvents using high-performance liquid chromatography,	2020	[14]
Wu, Y.; Zhang, X.; Di, Y.; Zhang, Y.	4-Nitro-1,2-phenylenediamine solid-liquid equilibrium was evaluated in eleven organic solvents.	2016	[21]
Xu, R.; Wang, J.; Han, S.; Du, C.; Meng, L.; Zhao, H.	The modified Apelblat equation, the λh equation, and the Wilson and NRTL models to correlate the solubility of phthalimide were utilized in various solvents. They calculated the solubility of terephthalaldehyde in different solvents to improve the purification process and subsequent theoretical investigations.	2012	[22]
Olson, M. E.; Ralston, B.; Burwash, L.	Ssolubility of 5-chloro-1-methyl-4-nitroimidazole was ascertained using an isothermal saturation method, in accordance with the thermodynamic model, at temperatures from 283.15 K to 318.15 K under 101.3 kPa in 11 organic solvents	2016	[28]
Li, M.; Lian, S, Wang, F, Zhou, Y.; Chen, B.; Guan, L.; Wu, Y.	A solubility model was proposed using a hybrid artificial intelligence system and diffusion theory to address the limitations of existing solubility prediction techniques. Their model incorporated empirical results of supercritical carbon dioxide solubility in eight polymers and utilized a modified double-population disordered accelerated particle swarm optimization (APSO) algorithm.	2020	[29]
Fu, L.; Hu, J.; Zhang, Y.; Li, Q	A T-S fuzzy neural network (T-S FNN) was used to predict sulfur solubility in sour gas over temperature and pressure ranges, using sulfur solubility as the target, and temperature, the mole fraction of methane, carbon dioxide, and hydrogen sulfide as the input parameters.	2017	[30]
Faúndez,, C. A.; Campusano, R. A.; Valderrama, J. O	An Artificial Neural Network (ANN) was used to correlate and predict the solubility of difluoromethane (R-32) in 17 ionic liquids, with R-32 solubility as the dependent variable and temperature and pressure as the independent variables. Additional properties such as the mass connectivity index (λ), acentric factor (ω), critical pressure (Pc), critical temperature (Tc), the mass of the anion (M^-), and the mass of the cation (M^+) were identified for each ionic liquid.	2010	[31]
Aminian, A.	The solubility of fifteen solutes was investigated in supercritical carbon dioxide (CO_2) using four neural network models at different temperatures and pressures.	2006	[32]

Table 1. Continued

Matsuda, H.; Kaburagi, K.; Kurihara, K	The one-parameter Wilson activity model via the co-solvent method were employed to predict the solubility of various medicines in solvents and water. Their study correlated the solubility data for acetaminophen, salicylic acid, benzocaine, phenacetin, and acetanilide.	2006	[33]
Mirmehrabi, M.; Rohani, S.; Perry, L	The NRTL and UNIQUAC models was used to assess the solubility of ranitidine hydrochloride, stearic acid, and stavudine in organic solvents.	2016	[34,35]
Boobier, S.; Hose, D. R. J.; Blacker, A. J.; Nguyen B. N.	A combination of machine learning (ANN, SVM, RF, ExtraTrees, Bagging and GP) and computational chemistry was used to assess the solubility prediction	2020	[36]
Tosca E. M.; Bartolucci R.; Magni P.	A QSPR model based on artificial neural networks (ANNs) was built to predict the intrinsic solubility (logS0) of the 100-compound low-variance tight set and the 32-compound high-variance loose set provided by Second Solubility Challenge (SC-2) as test datasets.	2021	[37]

is affected by coefficients, representing the weights of the mutual connections. So, a neuron with input *j* has a potential calculated as follows.

$$n = \sum_{j=1}^{J} x_j w_j + b_j \tag{1}$$

Here, x_j represents the *j*th input, w_j is the corresponding weight, and *b* is the bias term. In this research, the transfer function of the first layer is the hyperbolic tangent sigmoid, and the second layer is the linear function on the other side. Then the following functions are defined.

$$f(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \tag{2}$$

These independent variables are fed into the input layer, transferred to artificial neurons in the hidden layers, and then transferred into the output layer. pharmaceutical compound solubility is a function of pressure, temperature, acentric factor, critical pressure, critical temperature, and molecular weight, as shown in the following:

$$\text{Solubility} = f(T, P, T_c, P_c, w, Mw) \tag{3}$$

In Fig. 1, a typical diagram illustrates a one-hidden-layer network assuming acentric factor, critical pressure, critical temperature, pressure, temperature, and molecular weight. Absolute solubility is the output, and parameters in the network are represented by arrows.

As part of this study, two different error functions were used to determine neural network accuracy: the percent of average absolute relative deviations (AARD%) and the root mean square error (RMSE). These equations are arranged as below:

$$AARD\ \% = \frac{100}{NP} \sum_{i=1}^{NP} \frac{\left| x_i^{calc} - x_i^{\exp} \right|}{x_i^{\exp}} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{NP} \sum_{i=1}^{NP} (x_i^{calc} - x_i^{\exp})^2} \tag{5}$$

Here, NP is the number of experimental data points, $x_i^{\exp}$ is the experimental solubility, and x_i^{calc} represents the calculated solubility. The solubility in organic solvents is calculated by pharmaceutical compounds. Table 2 shows the

Table 2. Calculating Neural Network Accuracy Overall Values of AARD% and RMSE for Solid Components in Different Solvents. The Experimental Values Taken from Literature [8,12,13,15,21-26]

Compound	T_{min}-T_{max}	NP[a]	NN[b]	AARD%[c]	RMSE[d]
5-Chloro-1-methyl-4-nitroimidazole					
	283.15-318.15	150	30	1.904	0.000286
2-Amino-5-methylthiazole					
	278.15-313.15	180	40	0.648	0.000610
2-Isopropylimidazole					
	278.15-288.15	40	55	0.738	0.001382
Valsartan					
	278.15-313.15	48	10	0.226	0.008765
4-Nitro-1,2-phenylenediamine					
	283.15-318.15	150	30	7.960	0.001760
Phthalimide					
	283.15-318.15	135	10	6.328	0.000362
Terephthaldialdehyde					
	273.15-318.15	80	60	3.031	0.004526
PTMP					
	278.15-313-15	161	40	1.0821	0.000250
2-Amino-4-chlorobenzoic acid					
	278.15-313.15	148	40	2.834	0.000130

[a]NP represents the number of data points examined. [b]NN represents the number of neurons in the hidden layer.

details of the Overall values of *AARD*% and *RMSE*. The experimental values were taken from literature [8,12,13,15, 21-26].

$$^{c}\,AARD\ \% = \frac{100}{NP}\sum_{i=1}^{NP}\frac{\left|x_i^{calc} - x_i^{\exp}\right|}{x_i^{\exp}}$$

$$^{d}\,RMSE = \sqrt{\frac{1}{NP}\sum_{i=1}^{NP}\left(x_i^{calc} - x_i^{\exp}\right)^2}$$

It can be seen that ANN*s* are highly accurate in predicting the solubility of pharmaceutical compounds in organic solvents.

The Modified Apelblat Model

Apelblat is a semi-empiric model expressed as [38,39]:

$$\ln(x) = A + \frac{B}{T} + C\ln T \qquad (6)$$

In this work, the equation parameters that can be adjusted are *A*, *B*, and *C*, which represent the effect of the non-ideality solution on the coefficient activity, solubility, and the effect of temperature on the enthalpy of fusion, respectively.

RESULTS AND DISCUSSION

The data were randomly divided into two sub-data. The sub-data occupied for ANN training entailed approximately 70% of data points, and the remaining 30% were used for the verification of the network. This research uses six parameters as inputs for the presented network: critical temperature T_c (K), temperature (K), critical pressure P_c (MPa), pressure (MPa), acentric factor (ω), and molecular weight (Mw) of the organic solvents. The solubility of pharmaceutical compounds is the output parameter calculated in various organic solvents. In a trial-and-error process, the optimum number of neurons was determined. It was necessary as an objective for choosing the most suitable model. Table 3 displays the references between all information about the

Table 3. Molecular Weight, Critical Temperature, Critical Pressure, and Acentric Factor Used in this Work

Compound	*Mw* (g mol^{-1})	T_c (K)	P_c (MPa)	ω
1,4-Dioxane	88.106	587.30	5.47	0.280
2-Butanone	72.107	535.50	4.17	0.320
Acetone	58.080	508.20	4.71	0.307
Acetonitrile	41.053	545.50	4.85	0.340
Chloroform	119.370	536.40	5.37	0.218
Cyclohexane	84.161	553.58	4.10	0.212
Dichloromethane	84.930	236.68	6.10	0.199
Ethanol	46.069	513.92	6.12	0.643
Ethyl acetate	88.106	523.30	3.85	0.363
Ethyl benzene	106.165	617.12	3.62	0.304
Isopropanol	60.096	508.30	4.79	0.670
Isobutanol	74.123	536.05	4.2	0.574
Isopentanol	88.150	565.00	3.87	0.678
Methanol	32.040	512.60	8.14	0.566
Methyl acetate	74.079	506.55	4.69	0.326
N,N-Dimethylformamide	73.095	649.60	4.37	0.312
n-Butanol	74.123	563.05	4.34	0.585
n-Butyl acetate	116.160	579.15	3.11	0.410
n-Propanol	60.097	536.78	5.12	0.617
Toluene	92.141	591.80	4.10	0.262

solubility of pharmaceutical compounds and the related references in the expanded range of temperature. The molecular weights, critical properties, and acentric factors of each organic solvent are provided in Table 3.

It is necessary to expand the multilayer of ANN to detect the number of neurons in the hidden layer, as it has a considerable effect on the network yield. The percent of average absolute relative deviation (AARD%) is the basis for controlling the criterion as a network performance function used for training data sets. The collected number of neurons in the hidden layer was selected from 10 to 60 for optimizing neuron numbers and achieving the best ANN structure. Figure 2 demonstrates the number of neuron functions based on the variation of AARD%. As a result, a network with the best ANN structure was formed on the lowest value of AARD%. The optimum neuron numbers in the hidden layer are displayed in Table 2.

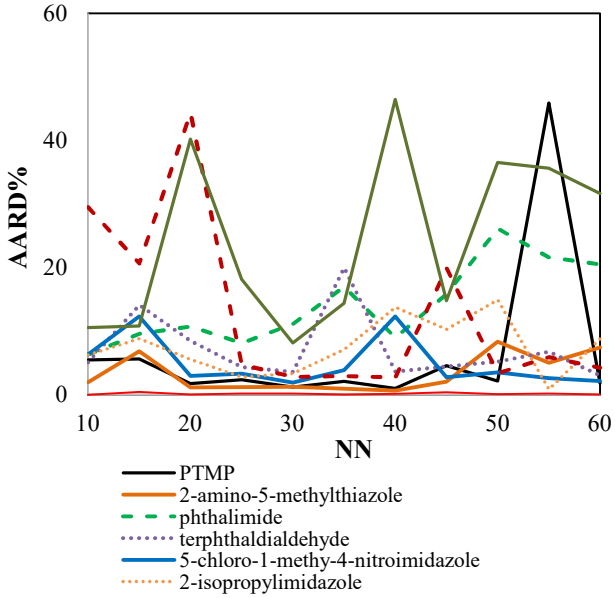

Fig. 2. AARD% between data collected and ANN results *versus* number of neurons in the hidden layer.

The values of parameters *A*, *B,* and *C* in the rectified Apelblat method are shown in Table 4, together with the AARD% values. The experimental values are taken from literature [13,15,22,25,27].

Based on the model presented in Fig. 3, it is evident that the mole fraction solubility of 2-Isopropylimidazole grew by increasing the temperature. At a fixed temperature, the mole fraction solubility of 2-Isopropylimidazole was the largest in methanol and lowest in cyclohexane. Figure 3 also showed that the solubility of 2-Isopropylimidazole in the mole fraction in various solvents was: methanol > ethyl acetate > acetone > ethanol > 1,4-dioxane > 2-butanone > n-propanol > isopropanol > acetonitrile > toluene > cyclohexane. These results are satisfactorily and consistent with previous research that used other methods.

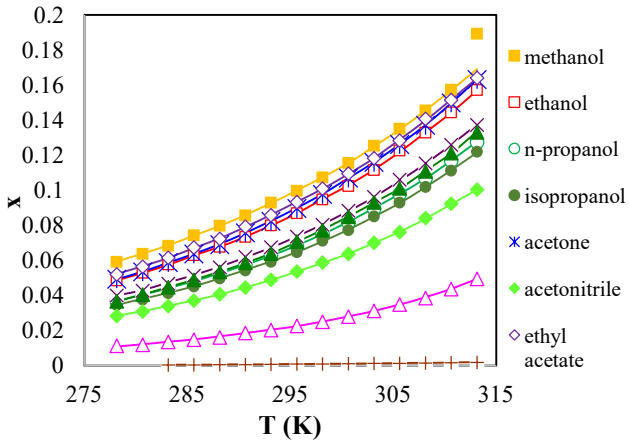

Fig. 3. Mole fraction solubility (*x*) of 2-Isopropylimidazole in the selected solvents at different temperature. The markers represent the experimental data [13] and solid lines are those correlated by the ANN.

In Fig. 4, the solubility of valsartan based on different parameters was investigated in organic solvents such as n-butyl acetate, acetonitrile, methyl acetate, dichloromethane, chloroform, and N,N-dimethylformamide at certain temperatures. The results of ANN method in this research in the range of selected temperature showed that the solubility of valsartan raised by increasing the temperature in all six pure solvents, as expected.

As shown in Fig. 5, the solubility of phthalimide increased with an increase in the temperature. The mole fraction solubility of phthalimide was maximum in the acetone and minimum in the ethanol and n-propanol. Figure 5 shows the solubility of phthalimide by mole fraction in various solvents, which is in this order: acetone > ethyl acetate > isopropanol > n-butanol > methanol > n-propanol. These results are satisfactorily and consistent with previous research that used other methods.

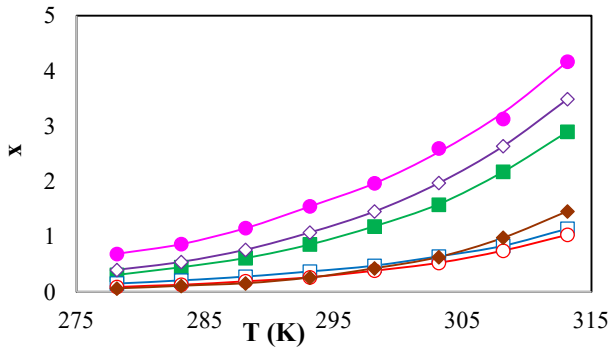

Fig. 4. Mole fraction solubility (*x*) of valsartan in the selected solvents at different temperature. The markers represent the experimental data [15,22] and solid lines are those correlated by the ANN.

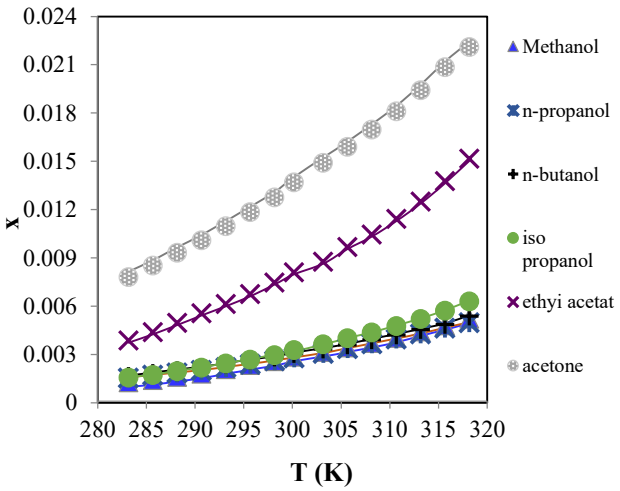

Fig. 5. Mole fraction solubility (*x*) of phthalimide in the selected solvents at different temperature. The markers represent the experimental data [22] and solid lines are those correlated by the ANN.

Table 4. Parameters of the Modified Apelblat Model and AARD% Values for Solid Components in Different Solvents. The Experimental Values Taken from Literature [13,15,22,25,27]

Compound	T_{min}-T_{max}	*AARD*%	*A*	*B*	*C*
2-Isopropylimidazole					
Methanol	278.15-288.15	0.981850913	-3.22E+01	3.07E-01	5.46E+00
Ethanol	278.15-288.15	0.376226238	-2.58E+01	3.80E-01	4.28E+00
Toluene	278.15-288.15	0.351039057	-3.23E+01	3.07E-01	5.06E+00
n-Propanol	278.15-288.15	0.379159432	-2.88E+01	3.07E-01	4.71E+00
Isopropanol	278.15-288.15	0.173015062	-2.58E+01	3.79E-01	4.12E+00
Acetone	278.15-288.15	0.387329115	-3.23E+01	3.07E-01	5.30E+00
2-Butanone	278.15-288.15	0.124622726	-3.22E+01	3.07E-01	5.34E+00
Acetonitrile	278.15-288.15	1.383952805	-3.23E+01	3.07E-01	5.09E+00
Average		0.519649			
Valsartan					
Methyl acetate	278.15-313.15	1.137659946	-1.07E+02	9.10E+00	1.87E+01
n-Butyl acetate	278.15-313.15	1.136752894	-9.66E+01	1.13E+01	1.68E+01
N,N-Dimethylformamide	278.15-313.15	1.009556332	-8.65E+01	-1.79E+00	1.53E+01
Acetonitrile	278.15-313.15	1.092586572	-1.05E+02	9.12E+00	1.85E+01
Dichloromethane	278.15-313.15	1.116504092	-1.19E+02	1.08E+01	2.08E+01
Chloroform	278.15-313.15	3.2296435	-1.55E+02	9.96E+00	2.70E+01
Average		1.453784			
Phthalimide					
Methanol	283.15-318.15	2.86455935	-7.37E+01	-6.89E+00	1.19E+01
Isopropanol	283.15-318.15	0.991617164	-7.26E+01	2.63E-01	1.17E+01
n-Propanol	283.15-318.15	0.40062833	-6.30E+01	4.90E-01	1.00E+01
Ethyl acetate	283.15-318.15	1.561502551	-6.88E+01	3.56E-01	1.12E+01
Acetonitrile	283.15-318.15	2.211627807	-8.37E+01	1.32E-02	1.37E+01
i-Butanol	283.15-318.15	0.765106369	-6.40E+01	4.67E-01	1.02E+01
n-Butanol	283.15-318.15	0.836152052	-6.19E+01	5.12E-01	9.83E+00
Toluene	283.15-318.15	4.302341864	-7.40E+01	-1.20E02	1.17E+01
Acetone	283.15-318.15	1.022162268	-5.38E+01	9.89E-04	8.68E+00
Ethanol	283.15-318.15	0.882739957	-8.41E+01	1.11E-03	1.37E+01
Average		1.583844			
2-Amino-4-chlorobenzoic					
Ethanol	278.15-313.15	4.476923928	-5.44E+01	-1.28E+00	8.85E+00
n-Propanol	278.15-313.15	0.475658961	-5.76E+01	-3.91E+00	9.39E+00
Iso propanol	278.15-313.15	0.544629848	-5.73E+01	-3.92E+00	9.35E+00
n-Butanol	278.15-313.15	0.703088698	-6.08E+01	-1.41E+00	9.94E+00
Ethyl acetate	278.15-313.15	0.690223944	-5.57E+01	-6.04E+00	9.13E+00
Acetonitrile	278.15-313.15	0.65046545	-6.80E+01	-6.32E+00	1.09E+01
Toluene	278.15-313.15	1.17125963	-8.84E+01	-8.99E+00	1.41E+01
Ethyl benzene	278.15-313.15	1.058168511	-9.30E+01	-6.90E+00	1.49E+01
Acetone	278.15-313.15	0.986788011	-5.29E+01	-1.41E+00	8.68E+00
1,4-Dioxan	278.15-313.15	0.569540423	-6.08E+01	-1.41E+00	1.00E+01
Methanol	278.15-313.15	0.331589794	-5.73E+01	-1.33E+00	9.35E+00
Average		1.059849			
Sirolimus					
Acetone	295-345	3.265354953	-3.45E+01	-4.26E+01	5.07E+00
Chloroform	295-345	3.247481762	-6.22E+01	-4.39E+01	9.82E+00
Dichloromethane	295-345	1.716318949	-3.60E+01	-1.11E+01	5.23E+00
Ethanol	295-345	1.713357515	-3.77E+01	-9.36E-01	5.48E+00
Methanol	295-345	2.950515498	-3.47E+01	-1.50E+01	4.89E+00
Average		2.578606			
Overall		1.33174			

As shown in Fig. 6, the solubility of 2-amino-4-chlorobenzoic acid increased with an increase in the temperature. At a specific temperature, the mole fraction solubility of 2-amino-4-chlorobenzoic acid was maximum in the acetone and minimum in the ethyl benzene.

Figure 6 showed that the mole fraction solubility was reduced according to the following arrangement, which was checked out in organic solvents: acetone > 1, 4-dioxane > ethyl acetate > ethanol> isopropanol > n-propanol > n-butanol > acetonitrile > toluene > ethyl benzene.

From Fig. 7, the solubility of sirolimus increased with temperature but not in liner form. At a specific temperature, the mole fraction solubility of sirolimus was maximum in the acetone and minimum in the methanol.

Figure 7 shows that the mole fraction solubility was in the following order, which was checked out in organic solvents: acetone > chloroform > dichloromethane > ethanol > methanol.

The outcome of the compared results with the other literature results are shown in Table 5, showing the average absolute relative deviation (AARD) of the calculated solubility of pharmaceutical compounds in organic solvents using the Apelblat ANN model from the experimental values [8,12,13,15,21-26]. Based on the data presented in Table 5, it is evident that the model we employed can precisely anticipate the solubility of the compounds in organic solvents.

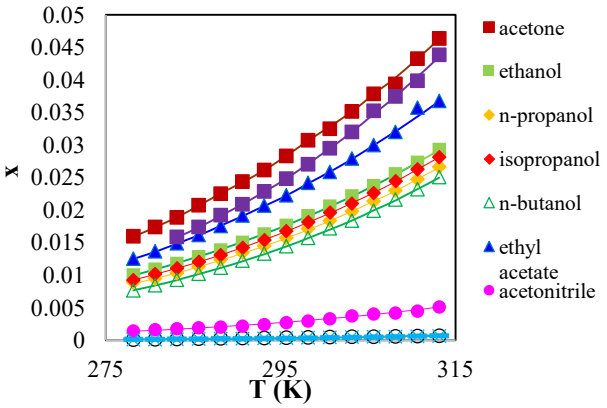

Fig. 6. Mole fraction solubility (x) of 2-amino-4-chlorobenzoic acid in the selected solvents at different temperature. The markers represent the experimental data [15,25,26] and solid lines are those correlated by the ANN.

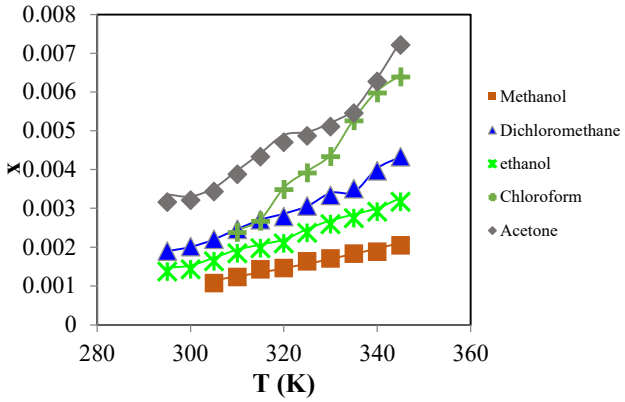

Fig. 7. Mole fraction solubility (x) of sirolimus in the selected solvents at different temperature. The markers represent the experimental data [27] and solid lines are those correlated by the ANN.

CONCLUSIONS

This study focused on assessing the solubility of pharmaceutical compounds in twenty candidate organic solvents under various temperatures and pressures conditions. These compounds included 2-amino-5-methylthiazole [12], 2-isopropylimidazole [13], valsartan [15], 4-Nitro-1,2-phenylenediamine [21], phthalimide [22], terephthaldialdehyde [23], 2,4-dihydro-5-methyl-2-(4-methylphenyl)-3H-pyrazol-3-one (PTMP) [24], 5-chloro-1-methyl-4-nitroimidazole [8], 2-amino-4-chlorobenzoic [25,26], and sirolimus [27]. Here, a modified Apelblat equation and artificial neural network model were expanded to equivalent and predict the solubility of pharmaceutical compounds in organic solvents for 20 binary pharmaceutical compounds and organic solvents systems. An extra equipoise existed between the calculated data of this study and the information extracted from the literature. The obtained data can contribute to understanding the connection between the solubility, solvent composition, and temperature. The conclusions can be used as support indexes in the purification of pharmaceutical compounds. These findings have two significant inferences: 1) the solubility of pharmaceutical compounds in organic solvents can be obtained with acceptable accuracy, with overall absolute average relative deviations of 2.75%, and 2) all deviation points are lower than 13.041%.

Table 5. The Overall Value Results of AARD% and RMSE of the Literature of Several Compounds Compared with the Present Work

Compound	From the present work		from the literature data		Ref. number of literature data
	RMSE	AARD%	RMSE	AARD%	
5-Chloro-1-methyl-4-nitroimidazole	2.86×10^{-4}	1.904	9.20×10^{-3}	2.86	[8]
2-Amino-5-methylthiazole	6.10×10^{-4}	0.65	10.6×10^{-4}	1.74	[12]
2-Isopropylimidazole	1.38×10^{-3}	0.73	19.3×10^{-3}	1.07	[13]
Valsartan	8.76×10^{-3}	0.226	-	0.5	[15]
4-Nitro-1,2-phenylenediamine	1.76×10^{-3}	7.96	0.83×10^{-4}	1.53	[21]
Phthalimide	3.62×10^{-4}	6.328	1.08×10^{-4}	1.77	[22]
Terephthaldialdehyde	4.5×10^{-3}	3.031	8.7×10^{-4}	1.2	[23]
PTMP	4.52×10^{-3}	3.031	1.35×10^{-3}	0.32	[24]
2-Amino-4-chlorobenzoic acid	1.3×10^{-4}	2.834	8.25×10^{-4}	3.35	[25,26]

REFERENCES

[1] Gupta, P.; Gupta, J. K.; Halve, A. K., Synthesis and biological significance of pyrazolones: A review. *Int. J. Pharm. Sci.* **2016**, *47* (13), 13-41, DOI: 10.1002/chin.201613221.

[2] Liu, B.; Zhou, Y.; Liang, Y.; Miao, C. Q., Synthesis method of pyrazolone. *CN Patent.* **2013**, *103* (396), 365-374.

[3] Amborabé, B. E.; Fleurat-Lessard, P.; Chollet, J. F.; Roblin, G., Antifungal effects of salicylic acid and other benzoic acid derivatives towards Eutypa lata: structure-activity relationship. *Plant. Physiol. Bioch.* **2002**, 40, 1051-1060, DOI: 10.1016/S0981-9428(02)01470-5.

[4] Evans, W. E., Pharmacogenetics of thiopurine S-methyltransferase and thiopurine therapy. *Therapeutic. Drug. Monitoring,* **2004**, *26*, 186-191, DOI: 10.1097/00007691-200404000-00018.

[5] Akbari, F.; Didehban, K.; Farhang, M., Solubility of solid intermediate of pharmaceutical compounds in pure organic solvents using semi-empirical models. *Eur. J. Pharm. Sci.* **2019**, *143*, 1-13, DOI: 10.1016/j.ejps.2019.105209.

[6] Klimek, J.; Wolski, T.; Wawrzycki, S.; Wronska, J., Determination of cysteine with 1-methyl-4-nitro-5-chloroimidazole. *Chem. Anal. (Warsaw).* **1978**, 317-323.

[7] Palomar, J.; Torrecilla, J. S.; Ferro, V. R.; Rodríguez, F., Development of an a priori ionic liquid design tool. 2. Ionic liquid selection through the prediction of COSMO-RS molecular descriptor by inverse neural network. *Ind. Eng. Chem. Res.* **2009**, *48*, 2257-2265, DOI: 10.1021/ie8009507.

[8] Wang, J.; Xu, R.; Xu, A.; Cong, Y., Thermodynamic Modelling for Solubility of 5-Chloro-1-methyl-4-nitroimidazole in Eleven Organic Solvents from T = (283.15 to 318.15) K. *J. Chem. Thermodynamic.* **2017**, *105*, 58-70 , DOI: 10.1016/j.jct.2016.10.006.

[9] Ge, Z. M.; Wang, C.; Hu, Y.; Hu, H., Synthesis of brimonidine tartrate. *Chin. J. Pharmaceuticals.* **2002**, *33* (1), 1-3. ISSN:1001-8255.

[10] Toering, S. J.; Gentile, G. J.; Gentile, J. M., Mechanism of antimutagenic action of (+)-catechin against the plant-activated aromatic amine 4-nitro-o-phenylenediamine. *Mutat. Res.* **1996**, *361*, 81-87, DOI: 10.1016/S0165-1161(96)90242-3.

[11] Zhou, W.; Zhao, L.; An, Z.; Li, G., Three metal-organic frameworks constructed from imidazole-based multi-carboxylate ligands: Syntheses, structures and photoluminescent properties. *Polyhedron.* **2016**, *117*,

202-208, DOI: 10.1016/j.poly.2016.05.059.

[12] Chen, G.; Chen, J.; Cheng, C.; Cong, Y.; Du, C.; Zhao, H., Solubility determination and thermodynamic modelling of 2-amino-5-methylthiazole in eleven organic solvents from T = (278.15 to 313.15) K and mixing properties of solutions. *J. Mol. Liq.* **2017**, *232*, 226-235, DOI: 10.1016/j.molliq.2017.02.084.

[13] Chen, J.; Chen, G.; Cong, Y.; Du, C.; Zhao, H., Solubility of 2-isopropylimidazole in nine pure organic solvents and liquid mixture of (methanol + ethyl acetate) from T = (278.15 to 313.15) K: Experimental measurement and thermodynamic modelling. *J. Chem. Thermodynamics.* **2017**, *10*, 133-140, DOI: 10.1016/j.jct.2016.12.028.

[14] Choi, J. H.; Lee, H. Y.; Towns, A. D., Dyeing properties of novel azo disperse dyes derived from phthalimide and color fastness on poly (lactic acid) fiber. *Polym.* **2010**, *11*, 199-204, DOI: 10.1007/s12221-010-0199-1.

[15] See, S., Angiotensin II receptor blockers for the treatment of hypertension. *J. Ezpert. Opin. Pharmacother.* **2001**, 2, 1795-1802, DOI: 10.1517/14656566.2.11.1795.

[16] Eslamimanesh, A.; Gharagheizi, F.; Mohammadi, A. H.; Richon, D., Artificial neural network modeling of solubility of supercritical carbon dioxide in 24 commonly used ionic liquids. *Chem. Eng. Sci.* **2011**, *66*, 3039-3044, DOI: 10.1016/j.ces.2011.03.016.

[17] Safamirzaei, M.; Modarress, H., Application of neural network molecular modeling for correlating and predicting Henry's law constants of gases in [bmim][PF_6] at low pressures. *Fluid. Phase. equilibria*. **2012**, *332*, 165-172, DOI: 10.1016/j.fluid.2012.07.009.

[18] Safamirzaei, M.; Modarress, H., Correlating and predicting low pressure solubility of gases in [bmim][BF_4] by neural network molecular modeling. *Thermochim. Acta.* **2012**, *545*, 125-130, DOI: 10.1016/j.tca.2012.07.005.

[19] Abdollahi, Y.; Sairi, N. A.; Aroua, M. K.; Masoumi, H. R. F.; Jahangirian, H.; Alias, Y., Fabrication modeling of industrial CO_2 ionic liquids absorber by artificial neural networks. *J. Ind. Eng. Chem*. **2015**, *25*, 168-175, DOI: 10.1016/j.jiec.2014.10.029.

[20] Liang, P.; Bose, N. K., *Neural Networks Fundamentals with Graphs, Algorithms, and Applications, in Electrical and Computer Engineering*. New York McGraw-Hill Inc: **1996**, pp.155-212.

[21] Wu, Y.; Zhang, X.; Di, Y.; Zhang, Y., Solubility determination and modelling of 4-nitro-1,2-phenylenediamine in eleven organic solvents from T = (283.15 to 318.15) K and thermodynamic properties of solutions. *J. Chem. Thermodyn.* **2016**, *106*, 22-35, DOI: 10.1016/j.jct.2016.11.014.

[22] Xu, R.; Wang, J.; Han, S.; Du, C.; Meng, L.; Zhao, H., Solubility modelling and thermodynamic dissolution functions of phthalimide in ten organic solvents. *J. Chem. Thermodyn.* **2016**, *94*, 160-168, DOI: 10.1016/j.jct.2015.10.024.

[23] Xu, A.; Xu, R.; Wang, J., Solubility determination and thermodynamic modelling of terephthaldialdehyde in ten organic solvents from T = (273.15 to 318.15) K and mixing properties of solutions. *J. Chem. Thermodyn.* **2016**, *102*, 188-198, DOI: 10.1016/j.jct.2016.07.013.

[24] Xie, Y.; Shi, H, Cong; Y.; Du, C.; Wang, J.; Zhao, H., Solubility determination and thermodynamic modelling of 2,4-dihydro-5-methyl-2-(4-methylphenyl)-3H-pyrazol-3-one in twelve organic solvents from T = (278.15 to 313.15) K and mixing properties of solutions. *J. Chem. Thermodynamic.* **2016**, *103*, 86-98, DOI: 10.1016/j.jct.2016.08.007.

[25] Yongjie, C.; Zhenfang, X.; Shuang, C. A. O.; Ying, W., Synthesis and luminescent properties of terbium complex with 2-amino-4-chlorobenzoic acid. *J. Rare. Earths.* **2016,** *34* (3), 240-244, DOI: 10.1016/S1002-0721(16)60020-0.

[26] Li, X.; Wang, M.; Cong, Y.; Du, C.; Zhao, H., Solubility determination and thermodynamic modelling for 2-amino-4-chlorobenzoic acid in eleven organic solvents from T = (278.15 to 313.15) K and mixing properties of solutions. *J. Chem. Thermodyn.* **2017**, *106*, 71-83, DOI: 10.1016/j.jct.2016.11.019.

[27] Gandhi, P. J.; Murthy, Z. V. P., Solubility and crystal size of sirolimus in different organic solvents. *J. Chem. Eng. Data.* **2010**, *55*, 5050-5054, DOI: 10.1021/je100626x.

[28] Olson, M. E.; Ralston, B.; Burwash, L.; Matheson-Bird, H.; Allan, N. D., Efficacy of oral meloxicam suspension for prevention of pain and inflammation following band and surgical castration in calves. *BMC. Vet. Res.* **2016**,

12, 1-11, DOI: 10.1186/s12917-016-0735-3.

[29] Li, M.; Lian, S, Wang, F, Zhou, Y.; Chen, B.; Guan, L.; Wu, Y., Neural network modeling based double-population chaotic accelerated particle swarm optimization and diffusion theory for solubility prediction. *Chem. Eng. Res. Des.* **2020**, *155*, 98-107, DOI: 10.1016/j.cherd.2020.01.003.

[30] Fu, L.; Hu, J.; Zhang, Y.; Li, Q., Investigation on sulfur solubility in sour gas at elevated temperatures and pressures with an artificial neural network algorithm. *Fuel.* **2020**, *262*, 116541-116550, DOI: 10.1016/j.fuel.2019.116541.

[31] Faúndez, C. A.; Campusano, R. A.; Valderrama, J. O., Misleading results on the use of artificial neural networks for correlating and predicting properties of fluids. A case on the solubility of refrigerant R-32 in ionic liquids. *J. Mol. Liq.* **2020**, *298*, 112009-112021, DOI: 10.1016/j.molliq.2019.112009.

[32] Aminian, A., Estimating the solubility of different solutes in supercritical CO_2 covering a wide range of operating conditions by using neural network models. *J. Super. Fluids.* **2017**, *125*, 79-87, DOI: 10.1016/j.supflu.2017.02.007.

[33] Matsuda, H.; Kaburagi, K.; Kurihara, K.; Tochigi, K.; Tomono, K., Prediction of solubilities of pharmaceutical compounds in water + co-solvent systems using an activity coefficient model. *Fluid. Phase. Equilibria.* **2010**, *290*, 153-157, DOI: 10.1016/j.fluid.2009.08.021.

[34] Mirmehrabi, M.; Rohani, S.; Perry, L., Thermodynamic modeling of activity coefficient and prediction of solubility: Part 1 predictive models. *J. Pharm. Sci.* **2006**, *95*, 790-797, DOI: 10.1002/jps.20560.

[35] Mirmehrabi, M.; Rohani, S.; Perry, L., Thermodynamic modeling of activity coefficient and prediction of solubility: Part 2. Semipredictive or semiempirical models. *J. pharm. Sci.* **2006**, *95*, 798-809, DOI: 10.1002/jps.20576.

[36] Boobier, S.; Hose, D. R. J.; Blacker, A. J.; Nguyen B. N., Machine learning with physicochemical relationships: solubility prediction in organic solvents and water. *Nat. Commun.* **2020**, *11*, 1-10, DOI: 10.1038/s41467-020-19594-z.

[37] Tosca E. M.; Bartolucci R.; Magni P., Application of Artificial Neural Networks to Predict the Intrinsic Solubility of Drug-Like Molecules. *Pharmacutics*. **2021**, *13*, 1-20, DOI: 10.3390/pharmaceutics13071101.

[38] Apelblat, A.; Manzurola, E., Solubilities of o-acetylsalicylic, 4-aminosalicylic, 3,5-dinitrosalicylic, and p-toluic acid, and magnesium-DL-aspartate in water from T = (278 to 348) K. *J. Chem. Thermodyn*. **1999**, *31*, 85-91, DOI: 10.1006/jcht.1998.0424.

[39] Apelblat, A.; Manzurola, E., Solubilities of L-glutamic acid, 3-nitrobenzoic acid, p-toluic acid, calcium-L-lactate, calcium gluconate, magnesium-DL-aspartate, and magnesium-L-lactate in water. *J. Chem. Thermodyn*. **2002**, *34*, 1127-1136, DOI: 10.1006/jcht.2002.0975.