2014 3rd International Conference on Environment Energy and Biotechnology
IPCBEE vol.70 (2014) © (2014) IACSIT Press, Singapore
DOI: 10.7763/IPCBEE. 2014. V70. 32

Artificial Neural Networks and the Melting Temperature of Ionic Liquids

José O. Valderrama[1,2+] and Claudio A. Faúndez[3]

[1] Univ. de La Serena, Dpto. de Ingeniería Mecánica, Casilla 554, La Serena-Chile
[2] Centro de Información Tecnológica, La Serena-Chile (jvalderr@userena.cl)
[3] Univ. de Concepción, Fac. de Cs Físicas y Matemáticas, Casilla 160-C, Concepción-Chile

Abstract. The use of artificial neural networks (ANN) for the correlation and prediction of the melting temperature of ionic liquids is analyzed in this paper. Several network architectures and two sets of data were analyzed and results compared with others from the literature. The independent variables considered for training the ANN were: groups forming the molecules, mass of the cation, mass of the anion and mass connectivity index. As a measure of the accuracy of the method the average deviation and the average absolute deviation are evaluated. Results of this work and others from the literature indicate that appropriate selection of data, a good combination of architecture and variables can lead to acceptable correlation of data but accurate prediction is not yet possible.

Keywords: chemical homology, ionic liquids, melting temperature

1. Introduction

In a recent paper I discussed the several methods presented in the literature during the last few years to estimate the melting temperature of ionic liquids and the alleged accuracy of the methods presented by several authors [1]. One of the methods that I mentioned was artificial neural networks (ANN), method that has received some especial attention in the literature without much discussion about its scope and limitations. In fact, the method has been seen as the general and appropriate correlating and predicting tool for any situation, with almost no restrictions. This view is simply wrong.

A complete list of properties that have been analyzed in the literature using different approaches of ANN was reported by Taskinen and Yliruusi [2]. Properties such as boiling point, vapor pressure, melting temperature, critical temperature and pressure, enthalpy of sublimation, heat of vaporization, density, viscosity, thermal conductivity, and acentric factor, among others, were thoroughly reviewed. Melting temperature was not discussed by these authors. However, some works on application of ANN to correlate melting temperature of organic substances and also of ionic liquids have been presented in the literature [1]. This paper discusses the main problems found when correlating and predicting melting temperatures of ionic liquids using ANN. In particular I analyze in this paper two aspects: i) the inaccuracy of the melting temperature data available and that is commonly used for training; and ii) the inadequate or limited knowledge that we have of other properties of ionic liquids that could have influence on the melting temperature.

2. Artificial Neural Networks

Artificial neural networks have several unique characteristics and advantages for applications such as the prediction of physicochemical properties of substances. One of these features is their adaptive nature. This means that the "learning by training" replaces the conventional "programming" commonly used in solving problems. This feature makes ANN very attractive for applications where the relation between variables is not well known, but data are readily available for training [3]. The Table 1 shows some advantages and

[+] Corresponding author. Tel.: +56-51-2551158
E-mail address: jvalderr@userena.cl

disadvantages of ANN. The aspects mentioned in Table 1 indicate that although ANN represent useful mathematical tools, their applications must be taken with care. In particular the requirement of large amount of accurate data (which depends on the complexity of the problem) and the knowledge of which are the most appropriate variables that have most influence on the target variable.

It seems that there is a generalized assumption that ANN can be used for correlating and predicting properties (in particular melting temperature of ionic liquids) if just enough data are available for the variable of interest (for instance the melting temperature). This is general belief is simply not true [1]. As Livingstone et al. [4] wrote: *"data modeling with neural networks is certainly not an answer to the maiden's prayer, but neural networks do offer a number of advantages over some of the more traditional methods of data modeling and should be viewed as a useful adjunct to these techniques*".

Table 1. Some advantages and disadvantages of ANN.

	Advantages:	**Disadvantages:**
1	A neural network can perform tasks that a linear program cannot.	The neural network needs training to operate.
2	If an element of the ANN fails, it can continue without any problem.	The best architecture must be known or determined by trial and error.
3	A neural network learns and does not need to be reprogrammed.	Requires a large amount of accurate data.
4	It can be implemented in any application and it is easy to set up.	Needs to know several variables well related to the target variable.
5	Models do not have to be specified in advance.	Local minima generalization/over fitting hard to interpret.
6	That they are able to fit complex nonlinear models.	Can suffer from over fitting and overtraining.

For the case of interest here, that is correlating and predicting the melting temperature of ILs, there is not enough information (for instance other properties of ILs) to find an appropriate relation between T_m and those other variables. Also, data reported in the literature which are used to train the ANN are of unknown accuracy and in many cases very different values are reported for the same ionic liquid [5]. The true value of T_m (if there is any) is unknown. These are the main reasons for the little progress done on this area.

However, even if accurate data were available, ANN cannot give a clear answer to the problem of predicting the melting temperature of ionic liquids in an acceptable way. This view agrees with that of Preiss et al. [6] who wrote: "*But, even if a pure substance was synthesized, and the melting temperature was unequivocally determined, there might be a multitude of unfavorable factors*", for developing a model for the prediction of the melting temperature of ionic liquids.

3. Literature data of T_m for ILs

There are some databases, handbooks and compilations of data for the melting temperature of ionic liquids [7]-[9]. The most complete compilation is that of Zhang et al. [9] which includes 1103 data of melting temperature, for 953 ionic liquids. New data frequently appear in journals and monographs. However, values reported in the literature for the same ionic liquid may show great differences. The thermal behavior of many ionic liquids is relatively complex which may explain the great differences in melting temperature values found in the literature. Valderrama and Rojas [5] discussed the deviations between experimental melting temperature data published by different authors for some ionic liquids.

Our group has been working for about three years on applications of ANN to property estimation and it seems that good data selection, good classification of data (types of ionic liquids) and a reasonable amount of data are the key factors for obtaining good correlating and predicting models. Chemical descriptors from computational chemistry (QSRP) have been proposed to be included in ANN models [10].

4. Applications of ANN models

For an ANN being capable of correlating data it is necessary that a series of requirements are fulfilled for that specific property and for the variables on which such a property depends. One approach that we have explored has been combining group contribution and ANN to correlate and predict melting temperature of ILs. Among the several independent variable that we have used in ANN studies are: mass of the cation, mass of the anion, density, mass connectivity index and groups forming the molecule ($M+$, $M-$, ρ, λ, G_i). Table 2 shows the groups considered in our studies.

Table 2. Groups Gi considered as independent variables in ANN models

N º	Groups	N º	Groups	N º	Groups	N º	Groups
	without rings	12	-O- [-O]$^-$	23	-CN	with rings	
1	-CH_3	13	>C=O	24	-NO_2	35	-CH_2-
2	-CH_2-	14	-CHO	25	-F [F^-]	36	>CH-
3	>CH-	15	-COOH	26	-Cl [Cl^-]	37	=CH-
4	>C< [>C-]$^-$	16	-COO-	27	-Br	38	>C<
5	=CH_2	17	-HCOO-	28	-I [I^-]	39	=C<
6	=CH-	18	=O	29	-P	40	-O-
7	=C<	19	-NH_2	30	-B	41	-OH
8	=C=	20	-NH-	31	-S- [>S-]$^+$	42	>C=O
9	=(-)CH	21	>N- [>N<]$^+$	32	-SO_2-	43	-NH-
10	=(-)C-	22	-N=	33	[-Sb^-]	44	>N- [>N<]$^+$
11	-OH			34	[-Al^-]	45	-N= [>N=]$^+$

The most basic architecture normally used for this type of applications involves a back propagation feed-forward neural network containing three or four layers: the array input layer, one or two hidden layers and the output layer [3]. According to other studies four layer architectures with 5 to 25 neurons in the inner layers are appropriate for correlating properties of fluids. The accuracy of the network was checked by determining the deviations between the calculated values of melting temperature after training and data from the literature. The optimum model is that providing the lowest deviations in the testing section. The maximum absolute deviation is also a distinctive parameter to define the goodness and accuracy of a model and of its predictive capability. A low maximum deviation would guarantee that the ANN not only learned but also has the capability of acceptably predicting Tm for cases not considered during training.

5. Results

The results are organized according to the two different studies presented in this paper and data analyzed using statistical parameters that according to the authors are the most representative of the accuracy of the method, as previously discussed in the literature: the average relative percent deviation (%T_m) and the average absolute percent deviation %|T_m| between the calculated values of melting temperature (T_m^{cal}), and data from the literature (T_m^{lit}) [11]. These are also the statistical parameters that are usually reported in the literature, so comparison with reported studies can be done. These deviations are the average deviation and the absolute average deviation.

5.1. Study 1: Several architectures for different types of ILs

In this study 667 data of melting temperature for all kind of ionic liquids were considered. These data were those reported by the authors as reasonable consistent in a previous paper [5]. The groups considered in this Study 1 were those used by the authors for predicting other properties. These groups were set as an extension of group contribution methods used for organic substances and are shown in Table 2. In this study, two cases, using the same 667 data points, were analyzed. The ANN was trained with the 45 groups, the mass connectivity index, the mass of the cation and the mass of the anion as the independent variables. The Table 3 presents results of these studies for ten four-layer architectures, starting from a simple network of 16 neurons (5,5,5,1) to a more complex of 61 neurons (10,25,25,1).

Table 3. Deviations in calculating T_m for 667 ionic liquids (567 for training and 100 for testing) [12]

	Training (567 data)		*Testing (100 data)*	
Architecture	*%ΔT_m*	*/%ΔT_m/*	*%ΔT_m*	*/%ΔT_m/*
5,5,5,1	0.5	9.1	3.6	15.7
5,10,10,1	0.4	6.8	3.5	16.4
5,15,15,1	0.2	3.8	-0.9	15.2
5,20,20,1	0.2	3.7	0.6	14.6
5,25,25,1	0.4	3.7	0.2	14.9
10,5,5,1	0.4	6.9	0.9	14.9
10,10,10,1	0.3	6.9	0.9	14.8
10,15,15,1	0.4	5.3	-1.8	13.5
10,20,20,1	0.3	3.7	1.9	15.7
10,25,25,1	0.5	3.7	-1.1	14.6

As seen in Table 3, results are variable, but average deviations are relatively low (<1%) and the average absolute deviation is below 10%. The increase in complexity of the network (more neurons) improves

correlation. As seen in the table the number of neurons in the entrance layer did not have much influence in the results.

5.2. Study 2: Several architectures for imidazolium-type ILs

In this study 297 selected data of imidazolium-type ionic liquids were considered. Also the groups forming the molecules were reconstructed and a new set of groups is proposed. This new set takes into account the particular structure of ionic liquids themselves more than being an extension of organic chemistry group contribution methods, as done in the Study 1. Similar to Study 1 above, besides the groups the ANN also considered the mass connectivity index, the mass of the cation and the mass of the anion as independent variables.

The Table 4 summarizes some statistical values such as average and absolute average deviations found between correlated and experimental T_m for all cases of the Study 2. The Table presents results for four architectures of type (5,N,N,1) for five cases, named 2a, 2b, 2c, 2d and 2e. The first case considers all 270 data for training and the other four cases with a reduced data set in which those showing high deviation were eliminated. In this study the variables considered for training are connectivity, groups, mass of the cation and mass of the anion.

Table 4. Deviations in calculating T_m for 297 imidazolium-type ionic liquid (270 for training and 27 for testing).

Architecture	*Training*		*Testing*	
	$\%\Delta T_m$	$\%\lvert\Delta T_m\rvert$	$\%\Delta T_m$	$\%\lvert\Delta T_m\rvert$
	(270 values)		*(27 values)*	
5-5-5-1	0.38	5.1	0.1	8.0
5-10-5-1	0.71	5.7	0.1	7.6
5-10-10-1	0.49	4.6	1.8	7.3
5-15-10-1	0.66	5.9	-3.6	7.6
	(239 values)		*(27 values)*	
5-5-5-1	0.41	5.1	-0.4	6.4
5-10-5-1	0.31	4.6	0.1	7.6
5-10-10-1	0.38	4.6	0.7	7.4
5-15-10-1	0.44	5.3	0.8	5.8
	(212 values)		*(27 values)*	
5-5-5-1	0.10	3.5	1.5	7.1
5-10-5-1	0.07	3.1	0.6	5.3
5-10-10-1	0.15	3.8	1.3	7.2
5-15-10-1	0.07	3.3	0.3	6.9
	(171 values)		*(27 values)*	
5-5-5-1	0.05	2.4	0.3	7.4
5-10-5-1	0.03	2.4	-0.2	7.4
5-10-10-1	0.06	2.7	-1.7	9.4
5-15-10-1	0.02	2.3	0.9	8.4
	(124 values)		*(27 values)*	
5-5-5-1	0.08	1.9	-0.9	7.4
5-10-5-1	0.07	1.5	3.1	8.2
5-10-10-1	0.06	1.9	-0.4	7.6
5-15-10-1	0.10	2.2	2.2	8.3

6. Conclusions

It is essential for making progress in this area of application of ANN for estimating T_m of ionic liquids to standardize the experimental determination of this property so good experiments guarantee reproducibility, and elimination of wrong data could be done.

It is necessary to know better the structure, the relation between groups, and the characteristics of the anion and cation that form the ionic liquid to find the most appropriate variables for ANN training.

With the present data one cannot obtain good results, especially in predicting the melting temperature of ILs using ANN, at least in the way done until now.

Other independent variables and better selection and analysis of data are necessary to improve correlation and prediction. Literature information and our own studies indicate that chemical descriptors seem to be the way for making progress in this area of correlating and predicting the melting temperature of ionic liquids

7. Acknowledgement

The authors thank the support of the National Council for Scientific and Technological Research

(CONICYT), through the research grant FONDECYT 1120162 and the Center for Technological Information (CIT, La Serena-Chile) for computer and library facilities. JOV thanks the Direction of Research of the University of La Serena and CAF thanks the Direction of Research of the University of Concepción, for permanent support.

8. References

[1] J.O. Valderrama, Myths and realities about existing methods for calculating the melting temperature of ionic liquids, Ind. *Eng. Chem. Res.*, 2014, 53, 1004-1014.

[2] J. Taskinen and J. Yliruusi, Prediction of physico-chemical properties based on neural network modeling, *Adv. Drug Delivery Rev.*, 2003, 55, 1163- 1183.

[3] N.K. Bose and P. Liang, *Neural Networks Fundamentals with Graphs, Algorithms, and Applications, in Electrical and Computer Engineering*, McGraw-Hill Series, The McGraw-Hill Companies, Inc., USA, 1996.

[4] D.J. Livingstone, D.T. Manallack and I.V. Tetko, Data Modeling With Neural Networks - An Answer to the Maiden's Prayer ?, *J. Comp. Aid. Mol. Design*, 1996, 11, 135-142..

[5] J.O Valderrama and R.E. Rojas, Data selection and estimation of the normal melting temperature of ionic liquids using a method based on homologous cations, *C. R. Chimie*, 2012, 15, 693–699.

[6] U. Preiss, S. Bulut and I. Krossing, In Silico Prediction of the Melting Points of Ionic Liquids from Thermodynamic Considerations: A Case Study on 67 Salts with a Melting Point Range of 337 C, The *Journal of Physical Chemistry B.* 2010, 114, 34, 11133-11140.

[7] NIST (National Institute of Standars and Techology). (July 2007). Ionic Liquids Database-ILThermo, Available: http://ilthermo.boulder.nist.gov/IL Thermo/mainmenu.uix

[8] Dortmund Data Bank, (March. 2011). Available: http://www.ddbst.com/en/ddbst/index.php

[9] S. Zhang, X. Lu, Q. Zhou, X. Li, X. Zhang, S. Li, *Ionic Liquids, Physicochemical Properties*; 1st. ed. Elsevier B.V. Jordan, Oxford 2009.

[10] J.O. Valderrama J.O., P.F. Arce and R.E. Rojas, Group Contribution Method for Estimating the Melting Temperature. Application to ionic Liquids, presented at the 11th Interamerican Congress CAIP 2013, Lima-Perú, October 2013.

[11] J.O. Valderrama and V.H. Alvarez, Correct way of Reporting Results when Modeling Supercritical Phase Equilibria using Equations of State. *Can. J. Chem. Eng*. 2005, 83, 1-4.

[12] V. Vicencio, Correlación y Predicción de la Temperatura de Fusión de Líquidos Iónicos usando Redes Neuronales, B.Sc. Thesis, Dept. of Chemistry, University of La Serena, Chile, 2012.