Int. J. Electrochem. Sci., 9 **(2014)** 3725 - 3735

International Journal of
ELECTROCHEMICAL
SCIENCE
www.electrochemsci.org

Prediction of Polarizability and Absolute Permittivity Values for Hydrocarbon Compounds Using Artificial Neural Networks

Hao Li[1,*,#], *Xifeng Liu*[2,3,#], *Shuangjun Yang*[4], *Tianqi Yi*[5], *Zhi Yang*[6] *and Wen Yuan*[3].

[1] College of Chemistry, Sichuan University, Chengdu, Sichuan 610064, China
[2] College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, China.
[3] Department of Chemistry, Michigan State University, East Lansing, Michigan 48824, United States
[4] College of Light Industry, Sichuan, Chengdu, Sichuan 610064, China
[5] College of Polymer Science and Engineering, Sichuan University, Chengdu, Sichuan 610064, China
[6] College of Software Engineering, Sichuan University, Chengdu, Sichuan 610064, China
#These authors contributed equally to this work.
[*]E-mail: lihao_chem_92@hotmail.com

Received: 19 February 2014 / *Accepted:* 18 March 2014 / *Published:* 14 April 2014

Artificial Neural Networks (ANNs) are computational models used to predict the properties of the organic compounds. In this study, we successfully established a model for predicting polarizability and absolute permittivity of hydrocarbon compounds by analyzing 281 typical hydrocarbon compounds using General Regression Neural Network (GRNN) and Multilayer Feedfoward Neural Network (MLFN) methods. Within permissible error range (30% tolerance) , our results show that GRNN model is an effective model in predicting materials' polarizabilities, 100% tested samples showed accurate results. In addition, MLFN model with nine nodes is demonstrated to be a valid model in predicting the absolute permittivities of hydrocarbon compounds, 97.22% tested samples showed accurate results. Taking together, our results indicated that ANN models can be applied in predicting the polarizability and absolute permittivity values of hydrocarbon compounds.

Keywords: Artificial Neural Networks; Polarizability; Absolute permittivity; General Regression Neural Networks; Multilayer Feedfoward Neural Networks; Prediction;

1. INTRODUCTION

Polarizability is the ability for a molecule to be polarized. As an essential matter property, it determines the dynamical response of a bound system to external fields. The analysis of the polarizability of a material could provide insight into the internal molecular structure[1]. Permittivity is a measure of the resistance that encountered when forming an electric field in a medium. It is a

measure of how an electric field affects, and it is affected by the dielectric medium. The permittivity of a medium describes how much the electric field (more correctly, flux) is 'generated' per unit charge in that medium [2-3].

In the field of electrochemistry, polarizability and permittivity are significant properties of relative materials. However, the value of polarizability and permittivity of electrochemical materials such as hydrocarbon compounds is so incomplete and complex to determine that makes a great obstacle of the related studies. Having realized that a quick and effective model to predict the two properties is needed, we paid our attention to find a suitable method to distinguish the structure characteristics of different hydrocarbons. And at last, we trained the accurate Artificial Neural Networks (ANNs) by means of the module of software in order to obtain the accurate ANN models to predict the polarizabilities and absolute permittivities of hydrocarbons.

Artificial Neural Networks are computational models inspired by animals' central nervous systems that are capable of machine learning and pattern recognition [4-6]. They are usually presented as systems of interconnected "neurons" that can calculate different values from inputs by feeding information through the network. As the development of the algorithm, this method is mature and has been packed into a module of the software. Represented by nonlinear functions, Artificial Neural network analysis is an artificial intelligence (AI) approach to modeling. Polarizability and absolute permittivity are classified nonlinear functions that too complicated to describe by analytical methods or empirical rules. Nevertheless, in this study, we are trying to predict these two properties by applying artificial neural networks into the issue [7-8].

In natural conditions, elements form groups and connect each other as neurons within the discrete layer. Each connection of them has its identified weight coefficient. The multiple layer consisted of the structure of such network. Usually, there are one or more than one layers of the elements followed by an output layer [8-12]. Multiple layers of elements can drive the network to learn nonlinear and linear relationships between input and output vectors [13].

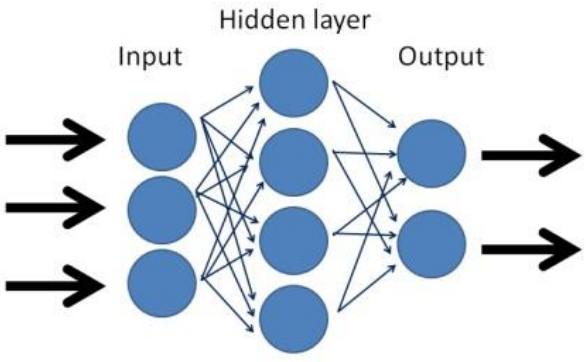

Figure 1. A schematic view of artificial neural network structure.

The figure above shows the main structure of the ANN. It is mainly made up of the input layer and the output layer. The input layer introduces the input variables to the network [14]. The output of the nodes in this layer represents the predictions made by the network for the response variables. In addition, it contains hidden layers. The optimal number of neurons in the hidden layers depends on the type and complexity of the process or experimentation and it's usually iteratively determined [15].

2. COMPUTATIONAL METHODS

2.1. A new distinction method of hydrocarbon compounds

To ensure the ANN model can recognize the different structure of different hydrocarbons, we designed a new distinction method of hydrocarbon compounds. The regulation of definition is show as follows:

Table 1. The description of the distinction method

Element	Description
Type	Identify different types of hydrocarbons
Molecular Mass	Relative molecular mass
Total Carbon Number	The overall carbon number of a single molecule
Total Hydrogen Number	The overall hydrogen number of a single molecule
Backbone's Carbon Number	The overall carbon number of backbone [a]
Number of Side Chain	The quantity of the side chain [b]
Identifier of Side Chain's Position（I）	The position number in accordance with the defined rule [c]
Carbon Number of Side Chain I	The carbon number of Side Chain I
Identifier of Side Chain's Position（II）	The position number in accordance with the defined rule c
Carbon number of Side Chain II	The carbon number of Side Chain II
Identifier of Side Chain's Position（III）	The position number in accordance with the defined rule c
Carbon Number of Side Chain III	The carbon number of Side Chain III
Double Bond‘s Position	The position number in accordance with the defined rule [c]
Number of Double Bond(s)	The position number in accordance with the defined rule [c]
Triple Bond’s Position	The position number in accordance with the defined rule [c]
Number of Triple Bond(s)	The total number of triple bond(s)
Carbon Numbers of Ring(s)	The total carbon number of the ring(s)
Number of Ring(s)	The quantity of the ring(s)

a: If there exists ring(s), the backbone should be considered with rings as possible, if there exists double/triple bond(s) together with rings, the chain with more double/triple bonds would be the backbone;
b: The definition of the side chains is based on the rule ***a***;
c: The location method is based on rule ***a*** and the nomenclature of organic defined by IUPAC[16];

Parameters of hydrocarbon molecules can be input into the form below, which constructs a database as the evidence of the ANN prediction model. The distinction rules above bases on the repeated trials of the training. Rules have been simplified in order to be easy understanding and more applicable.

2.2. Training process of the neural network

The ANN prediction model is constructed by the NeuralTools® software (Trial Version, Palisade Corporation, NY, USA)[17]. We chose the General Regression Neural Networks [18-21] (GRNN) module and Multilayer Feedfoward Neural Networks [22-26] (MLFN) module as the training modules. The training results are shown as follows (Data source: *CRC Handbook of Chemistry and Physics*2012-2013[27] and Wingch Database [28]):

Table 2. The training process of the samples

Physical Properties	Trained Cases	Number of Trials	Tested Cases
Polarizability	211	680000	53
Absolute Permittivity	145	1000000	36

The data of training process, including the number of trained cases and tested cases, as well as the number of trials, were presented in Table 2. There were 281 hydrocarbon compounds were taken into consideration for the training process.

Table 3. The training result of polarizability in different ANN models

ANN Model	RMS Error	Training Time	Finished Reason
Linear Prediction	13.53	0:00:00	Auto
GRNN	13.35	0:00:00	Auto
MLFN: 2 Nodes	16.38	0:00:59	Auto
MLFN: 3 Nodes	14.58	0:00:49	Auto
MLFN: 4 Nodes	14.98	0:00:50	Auto
MLFN: 5 Nodes	14.78	0:00:59	Auto
MLFN: 6 Nodes	13.89	0:01:24	Auto
MLFN: 7 Nodes	14.79	0:01:36	Auto
MLFN: 8 Nodes	14.87	0:01:41	Auto
MLFN: 9 Nodes	14.42	0:01:55	Auto
MLFN: 10 Nodes	15.34	0:02:13	Auto
MLFN:11 Nodes	14.50	0:02:52	Auto
MLFN: 12 Nodes	13.61	0:02:36	Auto
MLFN:13 Nodes	15.70	0:03:21	Auto
MLFN: 14 Nodes	13.89	0:03:46	Auto
MLFN: 15 Nodes	14.18	0:04:32	Auto
MLFN: 16 Nodes	14.79	0:05:41	Auto
MLFN: 17 Nodes	14.29	0:06:22	Auto
MLFN: 18 Nodes	16.11	0:06:21	Auto
MLFN: 19 Nodes	14.69	0:06:17	Auto
MLFN: 20 Nodes	14.55	0:07:22	Auto

As displayed in Table 3, the prediction model developed by GRNN method has the minimum RMS errors, which are lower than the linear prediction methods. These phenomena indicate that the polarizabilities of hydrocarbons can be well predicted by the non-linear model.

Table 4. The training result of absolute permittivity in different ANN models

ANN Model	RMS Error	Training Time	Finished Reason
Linear Prediction	10.21	0:00:00	Auto
GRNN	0.29	0:00:00	Auto
MLFN: 2 Nodes	0.34	0:00:39	Auto
MLFN: 3 Nodes	0.34	0:00:38	Auto
MLFN: 4 Nodes	0.32	0:00:49	Auto
MLFN: 5 Nodes	0.40	0:00:53	Auto
MLFN: 6 Nodes	0.27	0:01:24	Auto
MLFN: 7 Nodes	0.33	0:01:32	Auto
MLFN: 8 Nodes	0.30	0:02:29	Auto
MLFN: 9 Nodes	0.26	0:02:30	Auto
MLFN: 10 Nodes	0.32	0:03:11	Auto
MLFN: 11 Nodes	0.33	0:02:49	Auto
MLFN: 12 Nodes	0.37	0:04:18	Auto
MLFN: 13 Nodes	0.34	0:03:52	Auto
MLFN: 14 Nodes	0.50	0:05:16	Auto
MLFN: 15 Nodes	0.40	0:06:56	Auto
MLFN: 16 Nodes	0.34	0:08:50	Auto
MLFN: 17 Nodes	0.42	0:13:42	Auto
MLFN: 18 Nodes	0.42	0:33:27	Auto
MLFN: 19 Nodes	6.40	1:45:49	Auto
MLFN: 20 Nodes	0.42	1:45:30	Auto

Table 4 summarizes the MLFN model testing results, which indicate that 9 nodes has the minimum RMS error. In addition, we found that the RMS error of linear prediction is much greater in value than those obtained using ANN models. Thus, the absolute permittivities of hydrocarbon compounds are not fitting linear laws in these cases. Therefore, non-linear methods are supposed to be superior in predicting the values of absolute permittivity.

3. RESULTS AND DISCUSSION

3.1 Multiple model training

According to training and testing results, the best prediction model of polarizability is the GRNN model and the best prediction model of absolute permittivity is the MLFN model with 9 nodes (MLFN-9). As special distinction method for hydrocarbon compounds, the trained results are acknowledged to be as important as accurate tested results. Therefore, in this article, we presented all the training results to show comparison between predicted values and actual values, the comparison between residual errors and actual values, as well as the relationship between residual values and predicted values.

The outcomes of GRNN model of polarizability values were shown as follow (Figure 2 to Figure 4):

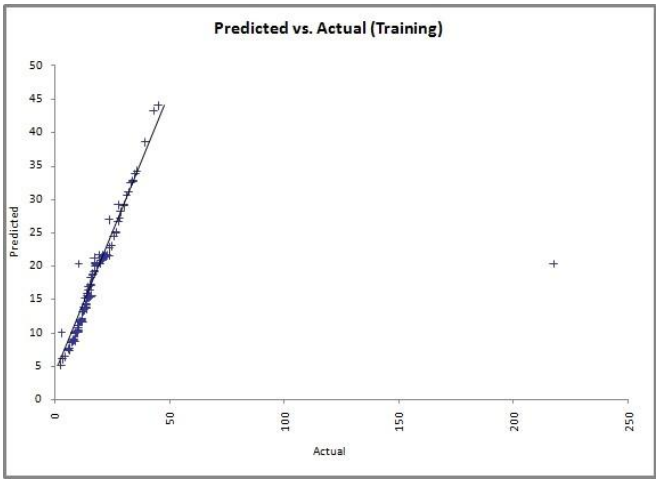

Figure 2. A comparison of the predicted values with actual values of polarizability of hydrocarbon compounds using GRNN model.

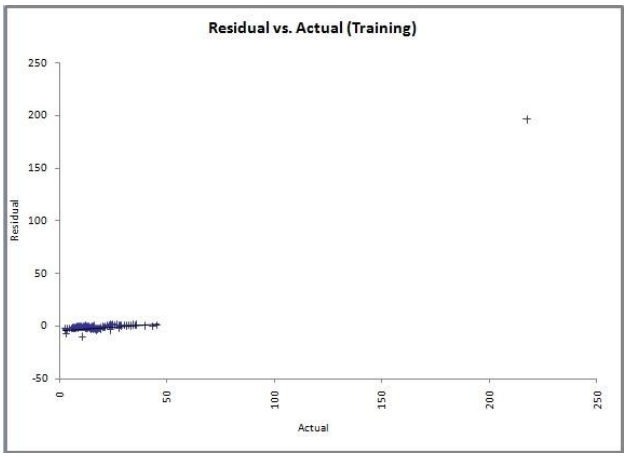

Figure 3. A comparison of the residual errors with actual values of polarizability of hydrocarbon compounds using GRNN model.

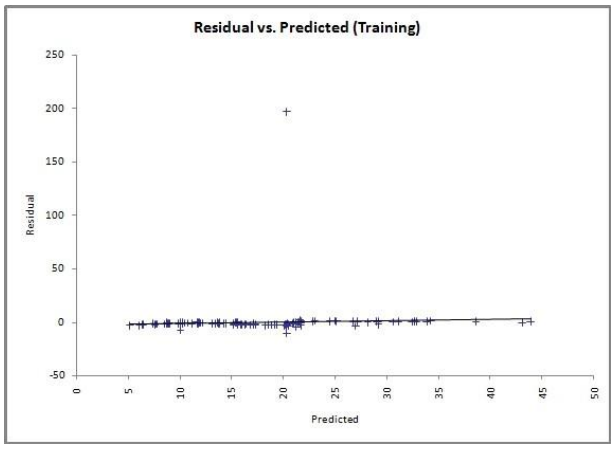

Figure 4. The relationship between the residual values and predicted values of polarizability of hydrocarbon compounds using GRNN model.

As displayed in Figure 2 to Figure 4, the trained result of polarizability values by GRNN model exhibited a linear trend. All the three groups of data fit well with the linear law consistently, indicating the training results of the GRNN model became a precise model after implanting our new distinction method. Likewise, the results of MLFN-9 model of absolute permittivity are shown in Figure 5 to Figure 7. As can be seen in these three figures, the three groups of data follow the linear law excellently with dispersion to small extent. These training results indicate that MLFN-9 model is precise by using our new distinction method.

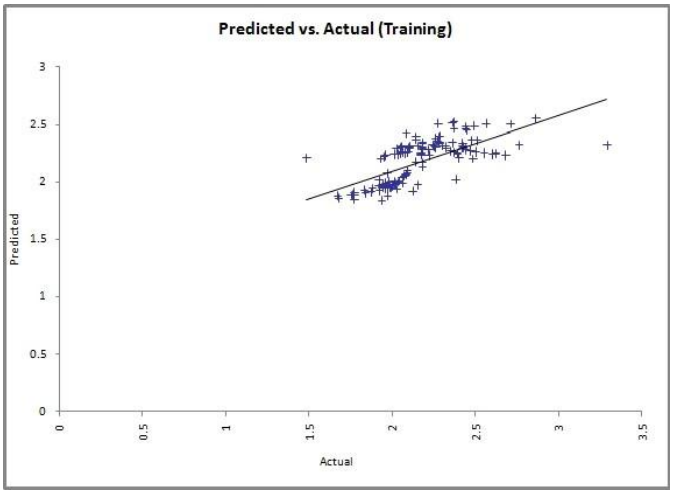

Figure 5. The relationship between the predicted value and actual value of absolute permittivity of hydrocarbon compounds in MLFN-9 model.

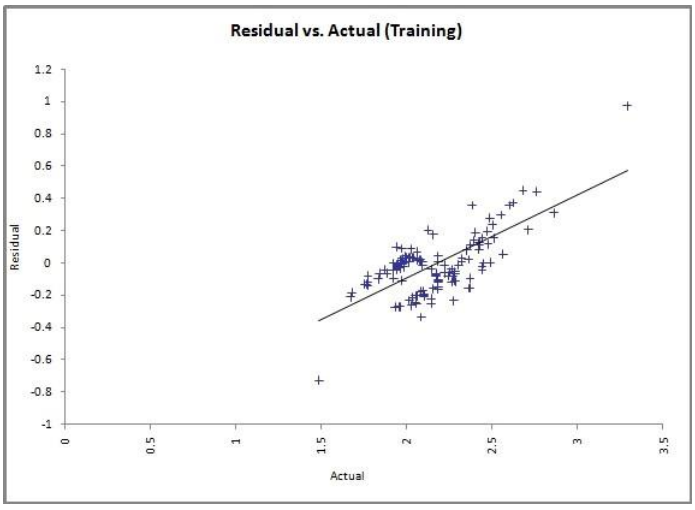

Figure 6. The relationship between the residual errors and actual values of absolute permittivity of hydrocarbon compounds in MLFN-9 model.

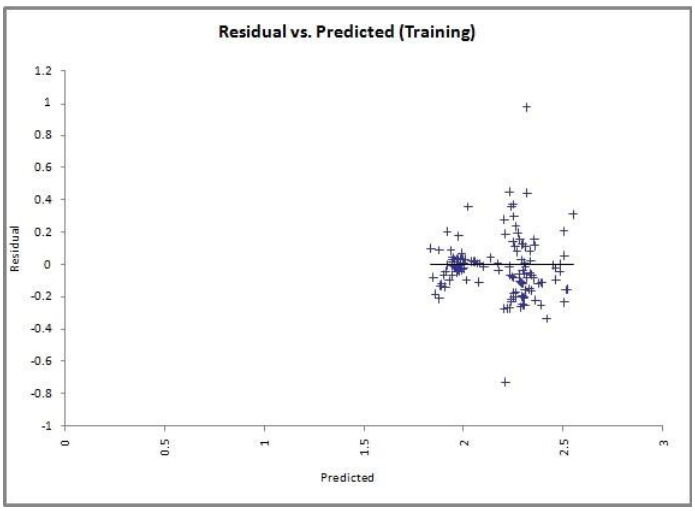

Figure 7. The relationship between the residual errors and predicted values of absolute permittivity of hydrocarbon compounds in MLFN-9 model.

According to the training results above, we considered the GRNN module is a better module to predict the values of polarizability (standard error: 13.35). 100% tested samples showed accurate results within permissible error range (30% tolerance). Besides, the MLFN module with 9 nodes is a better module to predict the values of absolute permittivity (standard error: 0.2563). 97.22% tested samples showed accurate results. Based on this analysis, we conducted the studies to obtain and analysis the prediction model using these two modules.

3.2 Accuracy and serviceability of hydrocarbon compounds testing

The GRNN and MLFN-9 prediction models were developed by the operation above. According to the generated report, the average relative errors of the two models are 4.7103% (polarizability, GRNN model) and 3.7736% (absolute permittivity, MLFN-9 model) respectively. To estimate the accuracy and serviceability, new data of the hydrocarbon compounds were inputted into the model. The testing results show as follows:

Table 5. The test result of the prediction on polarizability in GRNN model

Sample	Experimental Value	Predicted Value	Absolute Error	Relative Error
2,3-Dimethylbutane	11.79	11.75	0.04	0.339%
3,3-Dimethylpentane	13.65	13.76	0.10	0.733%
Amylcyclohexane	20.17	20.17	0.00	0.000%
2-Pentene	10.01	10.13	0.12	1.20%
Cyclooctene	14.46	14.44	0.02	0.138%
2-Methylbutadiene	9.71	8.98	0.27	2.78%
1,4-Decadiyne	17.52	17.44	0.08	0.457%
3,9-Dodecadiyne	21.26	21.11	0.15	0.706%
2-Methyl-1-buten-3-yen	8.96	9.10	0.14	1.56%

Table 6. The test result of the prediction on absolute permittivity in MLFN-9 model

Sample	Experimental Value	Predicted Value	Absolute Error	Relative Error
2,3-Dimethylbutane	1.88	1.81	0.07	3.72%
3,3-Dimethylpentane	1.92	1.86	0.06	3.13%
Amylcyclohexane	2.06	2.11	0.05	2.43%
2-Pentene	2.60	2.59	0.01	0.385%
Cyclooctene	2.14	2.13	0.01	0.467%
2-Methylbutadiene	2.18	2.18	0.00	0.000%
1,4-Decadiyne	2.57	2.57	0.00	0.000%
3,9-Dodecadiyne	2.37	2.38	0.01	0.422%
2-Methyl-1-buten-3-yen	2.63	2.61	0.02	0.760%

A comparison of polarizabilities between experimental values and predicted values for these 9 new samples were summarized in Table 5. At the same time, the absolute error and relative error were we calculated from these data also listed in the table. On the other hand, the comparison of absolute permittivities between experimental values and predicted values for these 9 new samples were summarized in Table 6. On the basis of these testing data above, the overall relative error of the prediction on polarizability in GRNN model is 0.8792%. While the overall relative error of the prediction on absolute permittivity in MLFN-9 model is 1.2571% , and they're proved to be applicable.

According to previous studies [29-39], we found out the similar researches on the prediction models of polarizabilities and permittivities. In the field of polarizability. Stout and his co-workers [29] have developed a model to predict the static dipole polarizabilities, using electronic structure calculations, with 30 organic molecules, which average error was around 3%. Another study shows the further research on the calculation of static dipole polarizabilities of polyene by abinitio coupled - perturbed Hartree–Fock theory [30], and Thole [31] has created a modified dipole interaction to calculate the polarizabilities of different molecules. Some references[32-33] reported empirical methods to calculate the molecular polarizabilities. In addition, some scientists focused on the studies of calculating the polarizabilities of inorganic substances [34-36]. Apropos of permittivity, Smith and his co-workers [37] have developed three different methods to calculate the permittivity of left-handed metamaterial, which is based on finite-difference simulations. Stern and Feller [38] developed a calculation method of the dielectric permittivity profile for a nonuniform system as an application to a lipid bilayer simulation. Besides, Paddison and his co-workers [39] have established an equilibrium statistical mechanical model for the calculation of the permittivity of water in hydrated polymer electrolyte membrane pores.

These previous studies were successful and can be used for reference to our study, which have different advantages respectively by using different prediction models or calculation methods. However, there is still no report on prediction of the two properties by using Artificial Neural Networks. Compared to the previous researches above, using ANN models to predict the polarizabilities and absolute permittivities is a more convenient and easy-operated method, without complex calculations. Interestingly, we found that ANN models of hydrocarbon compounds are more precise than other similar studies reported before. What's worth mentioning is that using ANN models

to predict such properties with the new distinction method of hydrocarbon compounds is easy to understand for operators.

In the field of electrochemistry, polarizability and permittivity are significant properties of hydrocarbon compound, playing important roles in every branch. Different from other studies, we concentrated our research on the hydrocarbon compounds, which is difficult to be described by the existing equations or calculation methods. The great error of linear fitting suggested that the polarizability and absolute permittivity of hydrocarbon compounds cannot be predicted by linear prediction. Our study shows that ANN model is a precise model to predict the polarizabilities and absolute permittivitis of hydrocarbon compounds.

4. CONCLUSION

Instead of measuring the values of polarizabilities and absolute permittivities of hydrocarbons from experiments in the lab, it is now possible to use the artificial neural networks with known experimental data to predict such properties of organic compounds. Our study has proved that the neural network can effectively generalize correct responses that only broadly resemble the data in the training set. The neural network can now be put to use with the actual data, which involves the values of polarizability and absolute permittivity.

References

1. L. Zhou, F. X. Lee, W. Wilcox, J. Christensen, *Nucl. Phys. B-Proceedings Supplements*, 119 (2003) 272-274.
2. Y. Y. Peter, M. Cardona, *Fundamentals of semiconductors: physics and materials properties*, Springer, Berlin,2010.
3. John Moore, Nicholas Spencer. *Encyclopedia of chemical physics and physical chemistry: Applications*, Taylor & Francis, Abingdon,2001.
4. Zeidenberg, Matthew, *Neural networks in artificial intelligence*, Ellis Horwood, Copenhagen, 1990.
5. Hagan, T. Martin, B. Howard Demuth, H. Mark Beale, *Neural network design*, Boston. Pws. Pub., 1996.
6. H. T. Siegelmann, E. D. Sontag, *Appl. Math. Lett.*,4 (1991) 77-80.
7. K. F. Khaled, N. A. Al-Mobarak, *Int. J. Electrochem. Sci*, 7 (2012) 1045-1059.
8. K. F. Khaled, A. M. El-Sherik, *Int. J. Electrochem. Sci*, 8 (2013) 10022-10043.
9. G. N. Vanderplaats, *Numerical Optimization Techniques for Engineering Design: with Applications.* McGraw-Hill, New York,1984.
10. D. Colorado-Garrido, S. Serna, M. Cruz-Chávez, J. A. Hernandez, B. Campillo, *CERMA2010,. IEEE*, 4 (2010) 185-190.
11. Carpintero, E. M., Malo, J. M., J. U. Chavarin , J. A. Hernández, *J. Solid State Electrochem,* 13 (2009) 1715-1722.
12. D. M. Colorado-Garrido , J. A. Ortega-Toledo, J. G. Hernández, González-Rodríguez, *CERMA ,IEEE*, 200 (2007) 213-218.
13. K. Hornik, M. Stinchcombe, H. White, *Neural networks*,2 (1989) 359-366.
14. J. A. Ciezak, S. F. Trevino, *J. Phys. Chem. A*, 110 (2006) 5149-5155.

15. D. Colorado-Garrido, D.M. Ortega-Toledo, J.A. Hernández, J.G. González-Rodríguez, J. Uruchurtu, *J. Solid State Electrochem*, 13 (2009) 1715-1722.
16. R. Panico, *A guide to IUPAC Nomenclature of Organic Compounds.* Blackwell Scientific Publ., Zurich, 1993.
17. Mostafa, M. Mohamed, Nedret Billor, *Expert. Syst. Appl.*, 36 (2009) 11378-11389.
18. D. F. Specht, *IEEE. T. Neural. Networ.*,2 (1991) 568-576.
19. M. T. Leung, A. S. Chen, H. Daouk, *Comput. Oper. Res.*, 27 (2000) 1093-1110.
20. Lee Won-Yong, M. John House, Nam-Ho Kyong, *Applied Energy*, 77 (2004) 153-170.
21. S. M. Lin, *Neural. Comput. Appl.*,22 (2013) 783-791.
22. K. Hornik, M. Stinchcombe, *Neural networks*, 2 (1989) 359-366.
23. J. S. Pei, E. C. Mai, J. P. Wright, *Nonlinear. Dynam.*,71 (2013) 371-399.
24. C. H. Chen, T. K. Yao, C. M. Kuo, C. Y. Chen, *J. Vib. Control.*, 19 (2013) 2413-2420.
25. R. B. Santos, M. Ruppb, S. J. Bonzi, *Chem. Eng. Trans*, 32 (2013) 1375-1380.
26. X. Glorot, Y. Bengio, *AISTATS2010*,4 (2010) 249-256.
27. Haynes, M. William, *CRC handbook of chemistry and physics*, CRC press, Boca Raton, 2012.
28. Wingch Database. *http://www.basechem.org/*
29. Stout, M. Joyce, Clifford E. Dykstra, *J. Am. Chem. Soc.* , 117 (1995) 5127-5132.
30. Hurst, J. B. Graham, Michel Dupuis, Enrico Clementi, *J.Phys.Chem*., 89 (1988) 385-395.
31. B. T. Thole, *Chemical Physics*, 59 (1981) 341-350.
32. Sadlej, J. Andrzej, *Collect. Czech. Chem. C.*,53 (1988) 1995-2016.
33. K. J. Miller, J. Savchik, *J. Am. Chem. Soc*.,101(1979) 7206-7213.
34. A. W. R. J. Derevianko, W. R. Johnson, M. S. Safronova, J. F. Babb, *Phys. Rev. Lett.*, 82 (1999) 3589.
35. R. D. Amos , *Chem. Phys. Lett.* ,124 (1986) 376-381.
36. J. Deiglmayr, M. Aymar, R. Wester, M. Weidemüller, O. Dulieu, *J. Chem. Phys.*, 129 (2008) 064-309.
37. D. R. Smith, D. C. Vier, N. Kroll, S. Schultz, *Appl. Phys. Lett*., 77 (2000) 2246-2248.
38. Stern, A. Harry, Scott E. Feller., *J. Chem. Phys.*, 118 (2003) 3401-3412.
39. Paul, Reginald, Stephen J. Paddison, *J. Chem. Phys.*, 115 (2001) 7762-7771.