Journal of Computational Methods in Molecular Design, 2015, 5 (3):16-24

Scholars Research Library
(http://scholarsresearchlibrary.com/archive.html)

ISSN : 2231- 3176
CODEN (USA): JCMMDA

Acute toxicity of phenol derivatives: Combining DFT and QSAR studies

A. Ousaa[1*], B. Elidrissi[1], M. Ghamali[1], S. Chtita[1], M. Bouachrine[2] and T. Lakhlifi[1]

[1]Molecular Chemistry and Natural Substances Laboratory, Faculty of Science, University Moulay Ismail, Meknes, Morocco
[2]MEM, ESTM, University Moulay Ismail, Meknes, Morocco
**Corresponding E-mail: abdellahousaa@gmail.com*

ABSTRACT

In order to investigate the relationship between activities and structures, a QSAR study is applied to a set of 23 phenol derivatives compounds. This study is conducted using the principal component analysis (PCA) method, the linear multiple regression method (MLR), the non-linear regression (MNLR) and the artificial neural network (ANN). We accordingly propose a quantitative model, and we interpret the activity of the compounds relying on the multivariate statistical analysis. Density functional theory (DFT) and ab-initio molecular orbital calculations have been carried out in order to get insights into the structure, chemical reactivity and property information for the series of study compounds. This study shows that the MRA and MNLR are served also to predict activities, but when compare with the results given by the ANN, we realize that the predictions fulfilled by this latter is more effective. To validate the predictive power of the resulting models, external validation multiple correlation coefficient are 0.80 and 0.78 for the MLR and the MNLR respectively. This model gives statistically significant results and shows good stability to data variation in leave-one-out cross-validation. The obtained results suggested that the proposed combination of several calculated parameters could be useful to predict the biological activity of phenol derivatives over Tetrahymena pyriformis.

Keywords: QSAR model, DFT study, phenol derivatives, *Tetrahymena pyriformis, cross-validation.*

INTRODUCTION

Organic chemicals carrying the structure of phenol have been in production since the 1860s, and include a wide number of applications in various industries such as textile, leather, paper and oil. For example, chlorophenols are utilized in Agriculture to manufacture a range of pesticides; alkylphenols are involved in the production of surfactants and detergents; bisphenol A is used to synthesize epoxy resins for paint coatings and mouldings, and in polycarbonate plastics, familiar in CDs and domestic electrical appliances. They can spread through air and water, with strong carcinogenicity, teratogenicity and mutagenicity, which will cause great damage to environment, plants, animals and human health. Therefore, it is vital to protect the environment and prevent occupational poisoning by studying the acute toxicity of phenols compounds [1-2].

The experiment is a direct way to obtain the toxicity data of organic compounds, which has many deficiencies, such as requirement of myriads of trial organisms, high expense, long time, the difference in measured value between different researchers and so on. Consequently, it would be impossible to gain the toxicity data of all organic compounds by experiment. As new compounds are springing up, other difficulties will follow. So it is necessary to use the theoretical research to make up for disadvantages of the experiment and to predict the toxicity data of compounds quickly and exactly.

QSAR can predict the bioactivity such as toxicity, mutagenicity and carcinogenicity based on structural parameters of compounds and appropriate mathematical models. With the rapid development of computer science and theoretical quantum chemical study, it can speedily and precisely obtain the quantum chemical parameters of compounds by computation.

Moreover, these parameters, which have definite physical meaning, along with the introduction of the QSAR model can increase the interpretability. So quantum chemical theory is extensively applied in establishing QSAR models [3-4].

In this work, we model the toxicity of 23 phenol derivatives compounds to *Tetrahymena pyriformis* using several statistical tools, principal components analysis (PCA), multiple linear regression (MLR), non-linear regression (RNLM) and artificial neural network (ANN) calculations, we accordingly propose a quantitative model, and we try to interpret the activity of these compounds relying on the multivariate statistical analyses.

MATERIAL AND METHODS

Data sources

Acute toxicity data of 23 phenol derivatives to *Tetrahymena pyriformis* are taken from a literature [5]. **IC_{50}** here means the millimolar concentration causing 50% inhibition of growth about phenol derivatives to *tetrahymena pyriformis.* The bigger the value of **$-\log IC_{50}$ (pIC_{50})**, the higher is toxicity of compounds, and vice versa.

The following table shows the studied compounds and the corresponding experimental activities pIC_{50} (Table 1). The experimental toxicity of the studied compounds is collect from recent work [5], 18 molecules are select for the quantitative model (training set), and 5 are select randomly to test the performance of the proposed model (test set).

Table 1: Phenol derivatives and their observed toxicities against *Tetrahymena pyriformis* (Training and test set)

Training Set					
N°	**Name (IUPAC)**	**pIC_{50}**	**N°**	**Name (IUPAC)**	**pIC_{50}**
1	Phenol	-0.431	**10**	3,5-Dimethylphenol	0.113
2	4-Methylphenol	-0.192	**11**	2-Isopropylphenol	0.803
3	3-Methylphenol	-0.062	**12**	3-Isopropylphenol	0.609
4	3-Ethylphenol	0.229	**13**	4-Isopropylphenol	0.473
5	2-Ethylphenol	0.176	**14**	3-tert-Butylphenol	0.730
6	2,3-Dimethylphenol	0.122	**15**	4-tert-Butylphenol	0.913
7	2,4-Dimethylphenol	0.128	**16**	2-Phenylphenol	1.094
8	2,5-Dimethylphenol	0.009	**17**	2,3,6-Trimethylphenol	0.418
9	3,4-Dimethylphenol	0.122	**18**	3,4,5-Trimethylphenol	0.930
Test Set					
N°	**Name (IUPAC)**	**pIC_{50}**	**N°**	**Name (IUPAC)**	**pIC_{50}**
19	2,4,6-Trimethylphenol	1.695	**22**	2,6-Diphenylphenol	2.113
20	2-tert-Butyl-4-methylphenol	1.297	**23**	2,6–Di-tert-butyl-4-methylphenol	1.788
21	6-tert-Butyl-2,4-dimethylphenol	1.245			

Molecular descriptors

All computations are performing by using Gaussian 03W program [6]. The geometries of all 23 theoretically possible phenol derivatives are fully optimize with DFT method at the B3LYP/6-31G (d) level and frequency calculations are performing to calculate at the same level for all of the possible geometries to ensure they are minimal on the potential energy surface. Then we choose some related structural parameters from the results of quantum chemical computations as the highest occupied molecular orbital energy E_{HOMO}, the lowest unoccupied molecular orbital energy E_{LUMO}, energy gap ΔE, dipole moment μ, the total energy E_T, the activation energy E_a, the absorption maximum λ_{max}, the factor of oscillation $f_{(SO)}$.

Statistical analysis

The structures of 23 phenol derivatives to *tetrahymena pyriformis* are stud by statistical methods based on the principal component analysis (PCA) [7] using the software XLSTAT version 2013 [8]. PCA is a statistical technique useful for summarizing all the information encoded in the structures of the compounds. It is also very helpful for understanding the distribution of the compounds [9]. This is an essentially descriptive statistical method which aims to present, in graphic form, the maximum of information contained in the data table 1 and table 2.

The linear multiple regression (MLR) analysis with descendent selection and elimination of variables was employed to model the structure with activity. It is a statistical technique that minimizes differences between actual and

predicted values. It has served also to select the descriptors used as the input parameters in the nonlinear multiple regression (MNLR) and artificial neural network (ANN) methods.

The (MLR) and the (MNLR) were generated using the software XLSTAT version 2013 [8], to predict cytotoxic effects IC_{50}. Equations were justified by the correlation coefficient (R) and mean squared error (MSE) values [8]. ANN is an artificial system simulating the function of the human brain. Three components constitute a neural network: the processing elements or nodes, the topology of the connections between the nodes, and the learning rule by which new information is encoded in the network. While there are a number of different ANN models, the most frequently used type of ANN in QSAR is the three-layered feed forward network [10]. In this type of networks, the neurons are arranged in layers (an input layer, one hidden layer and an output layer). Each neuron in any layer is fully connected with the neurons of a succeeding layer and no connections are between neurons belonging to the same layer.

According to the supervised learning adopted, the networks are taught by giving them examples of input patterns and the corresponding target outputs. Through an iterative process, the connection weights are modified until the network gives the desired results for the training set of data. A backpropagation algorithm is used to minimize the error function. This algorithm has been described previously with a simple example of application [11] and a detail of this algorithm is given elsewhere [12].

The ANNs analysis was performed with the use of Matlab software version 2009a Neural Fitting tool (nftool) toolbox [13].

RESULTS AND DISCUSSION

QSAR models and analysis

The quantitative structure–activity relationship analysis is perform using the pIC_{50} of the 23 phenol derivatives to *tetrahymena pyriformis* as reported in [14], the values of the 8 chemical descriptors as shown in table 2.

The principle (for the two studies) is to perform in the first time, a main component analysis (PCA), which allows us to eliminate descriptors that are highly correlated (dependent), then perform a decreasing study of MLR based on the elimination of descriptors (one by one) aberrant until a valid model (including the critical probability: p-value<0.05 for all descriptors and the complete model).

Table 2: The values of the eight chemical descriptors

N°	E_T (Ua)	E_{HOMO} (ev)	E_{LUMO} (ev)	ΔE (ev)	μ (debye)	Ea (ev)	λmax (nm)	$f_{(SO)}$
1	-8372.09	-6.48	0.01	6.49	1.60	5.26	235.93	0.025
2	-9442.71	-6.21	0.03	6.25	1.70	5.34	232.33	0.011
3	-9442.90	-5.86	0.13	5.99	1.09	4.83	256.56	0.015
4	-10513.41	-5.87	0.09	5.96	1.11	4.50	275.72	0.013
5	-10509.65	-5.20	-2.36	2.84	2.43	4.95	250.51	0.174
6	-10513.45	-5.75	0.15	5.90	1.91	3.03	409.87	0.039
7	-10509.79	-5.08	-2.26	2.82	3.08	3.00	413.86	0.040
8	-10509.79	-5.19	-2.24	2.94	2.74	3.50	354.10	0.039
9	-10509.41	-4.85	-2.23	2.62	2.41	4.11	301.45	0.043
10	-10513.53	-5.79	0.14	5.93	1.38	4.38	282.99	0.053
11	-11583.83	-5.81	0.15	5.96	1.75	4.58	270.48	0.052
12	-11583.90	-5.85	0.13	5.98	1.58	4.08	303.97	0.060
13	-11583.89	-5.76	0.06	5.82	1.34	3.65	339.71	0.035
14	-12654.38	-5.86	0.09	5.95	1.60	3.29	376.55	0.031
15	-12654.31	-5.75	0.07	5.83	1.37	2.32	534.92	0.051
16	-14663.98	-5.81	-0.69	5.12	1.74	4.97	249.60	0.103
17	-11584.10	-5.63	0.27	5.90	1.69	4.38	283.18	0.049
18	-11584.05	-5.57	0.32	5.88	1.44	5.97	207.83	0.066
19	-11584.15	-5.53	0.30	5.83	1.39	5.78	214.60	0.216
20	-13724.80	-5.56	0.25	5.82	1.17	2.88	429.86	0.012
21	-14790.88	-4.96	-2.48	2.48	6.28	3.01	412.20	0.071
22	-20955.58	-5.74	-0.70	5.04	1.67	1.90	652.23	0.122
23	-18000.09	-4.81	-2.70	2.11	1.74	1.88	658.90	0.005

Principal component analysis

The set of descriptors encoding the 23 phenol derivatives, electronic and energetic parameters are submitted to use for PCA analysis [15]. The first three principal axes are sufficient to describe the information provided by the data

matrix. Indeed, the percentages of variance are 49.97%; 21.91% and 14.15% for the axes F1, F2 and F3, respectively. The total information is estimated to a percentage of 86.03%.

The principal component analysis (PCA) [16] was conducted to identify the link between the different variables. Bold values are different from 0 at a significance level of p= 0.05. Correlations between the eight descriptors are shown in table 3 as a correlation matrix and in Figure 1 these descriptors are represented in a correlation circle. The Pearson correlation coefficients are summarized in the following Table 3. The obtained matrix provides information on the negative or positive correlation between variables.

Table 3: Correlation matrix (Pearson (n)) between different obtained descriptors

Variables	pIC_{50}	E_T	E_{HOMO}	E_{LUMO}	ΔE	μ	Ea	λ_{max}	$f_{(SO)}$
pIC_{50}	**1**								
E_T	**-0.88**	**1**							
E_{HOMO}	0.30	-0.31	**1**						
E_{LUMO}	-0.07	0.25	**-0.82**	**1**					
ΔE	-0.13	0.28	**-0.90**	**0.99**	**1**				
μ	0.05	-0.13	**0.55**	**-0.67**	**-0.66**	**1**			
Ea	-0.40	**0.63**	-0.39	0.36	0.38	-0.23	**1**		
λ_{max}	**0.56**	**-0.77**	0.36	-0.35	-0.37	0.14	**-0.94**	**1**	
$f_{(SO)}$	0.40	-0.19	0.15	-0.09	-0.11	0.10	0.28	-0.16	**1**

Bold values are different from 0 at a level significant for p < 0.05
At a very significant for p < 0.01
At a very significant for p < 0.001

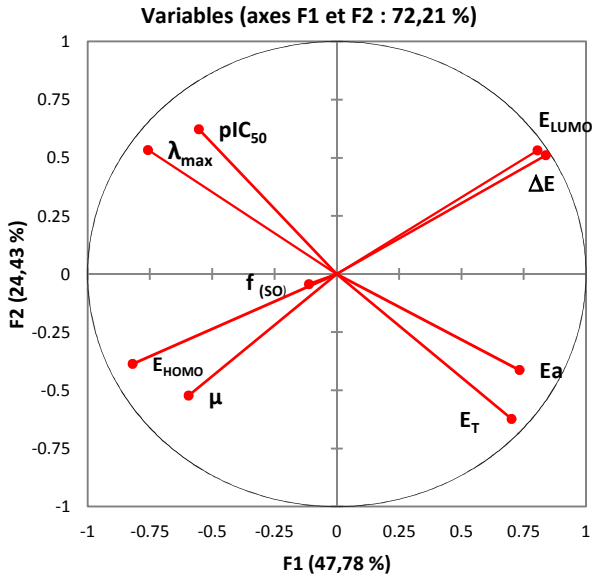

Figure 1: Correlation circle

The analysis of projections according to the plan F1-F2 (72.21% of the total variance) of the studied molecules is shown in Figure 2:

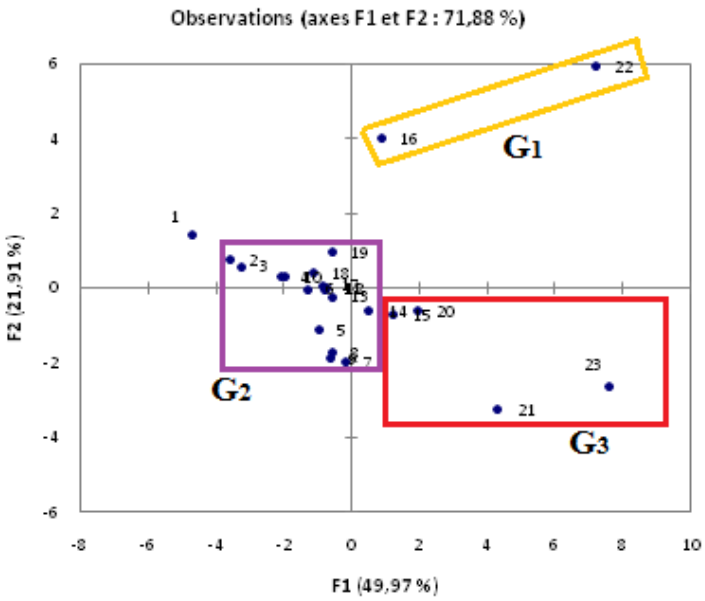

Figure 2: Cartesian diagram according to F1 and F2: Separation between three regions

The obtained matrix provides information on the negative or positive correlation between variables.

*The energy gap **ΔE** is positively correlated with the lowest unoccupied molecular orbital energy E_{LUMO} (r=0.954 and p<0.05) at a significant level.

The principal component analysis revealed from the correlation circle (Figure 1) shows that the F1 axis (47.78% of the variance) is clearly connected to the energy gap ΔE, while the axis F2 (24.43% of the variance) is located by the other parameters of energy.

Analysis of projections according to the plan F1-F2 (71.88% of the total variance) of the studied molecules (Figure 2) shows that the molecules are dispersed, according to the of radicals of phenol derivatives, in three classes of compounds belonging to three groups: the group 1 (G1) containing the phenol substituted by phenyl, the group 2 (G2) containing the phenol substituted *t*-butyl, and the group 3 (G3) containing the phenol substituted by methyl or ethyl. In this representation, compound 1 is an exception because it is a phenol non substituted

Linear Multiple Regressions

To establish quantitative relationships between toxicity pIC_{50} and selected descriptors, our array data are subject to linear multiple and multiple nonlinear regressions. Only variables whose coefficients are statistically significant are retained.

Linear Multiple regression of the variable toxicity (MLR)

Linear multiple regression is carried out to develop a relationship with the indicator variable of toxicity pIC_{50}, but the best relationship obtained by this method is only one corresponding to the linear combination of two descriptors selected the total energy E_T, the activation energy E_a, Refractive energy E_{HOMO}, energy E_{LUMO} and the absorption maximum λ_{max}.

The resulting equation is:

$$\mathbf{pIC_{50}} = -2.502-2.552.10^{-4}\times\mathbf{E_T}+0.371\times\mathbf{E_{HOMO}}+0.159\times\mathbf{E_{LUMO}}+0.303\times\mathbf{Ea}+3.006.10^{-3}\times\boldsymbol{\lambda_{max}} \quad (1)$$

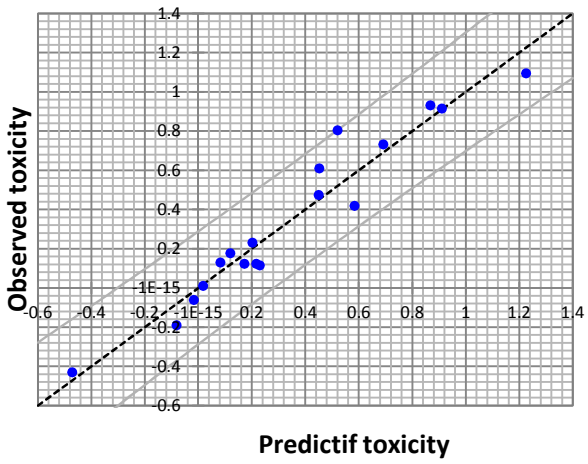

Figure 3: Graphical representation of calculated and observed toxicity by MLR

For our 18 compounds, the correlation between experimental toxicity and calculated one based on this model is quite significant (Figure 3) as indicated by statistical values:

N = 18 R = 0.97 R^2 = 0.94 RMSE = 0.13

With the optimal MLR model, the values of toxicity calculated from equation 1 and the observed values are given in table 5. The correlations of predicted and observed pIC_{50} are illustrated in Figure 3. The descriptors proposed in equation 1 by MLR were, therefore, used as the input parameters in the Multiples nonlinear regression (MNLR).

Nonlinear multiple regression of the variable toxicity (MNLR)

We have also used the technique of nonlinear regression model to improve the structure toxicity relationship in a quantitative way, taking into account several parameters. This is the most common tool for the study of multidimensional data. We have applied it to Table 2 containing 18 molecules associated with eight variables. We used a pre-programmed function of XLSTAT as follows:

$$\mathbf{Y = a + (b\,X_1 + c\,X_2 + d\,X_3 + e\,X_4 \ldots) + (f\,X_1^2 + g\,X_2^2 + h\,X_3^2 + i\,X_4^2 \ldots)}$$

Where a, b, c, d… represent the parameters and X_1, X_2, X_3, X_4…: represent the variables.

The resulting equation is:

$$\mathbf{pLC_{50}} = -3.866 - 1.10010^{-03} \times \mathbf{E_T} + 1.957 \times E_{\mathbf{HOMO}} + 0.179 \times E_{\mathbf{LUMO}} + 0.215 \times \mathbf{Ea} + 2.25010^{-03} \times \boldsymbol{\lambda}_{\mathbf{max}} - 3.48210^{-08} \times \mathbf{E^2_T} + 0.174 \times E_{\mathbf{HOMO}}^2 + 5.99610^{-02} \times E_{\mathbf{LUMO}}^2 + 9.35110^{-03} \times \mathbf{Ea}^2 + 5.03810^{-07} \times \boldsymbol{\lambda}^2_{\mathbf{max}} \quad (2)$$

The obtained parameters describing the electronic aspects of the studied molecules are:

N = 18 R = 0.98 R^2 = 0.96 RMSE = 0.14

The toxicity value pIC_{50} predicted by this model is somewhat similar to that observed. The figure 4 shows a very regular distribution of toxicity values based on the observed values. The values of toxicity calculated from equation 2 and the observed values are given in table 5.

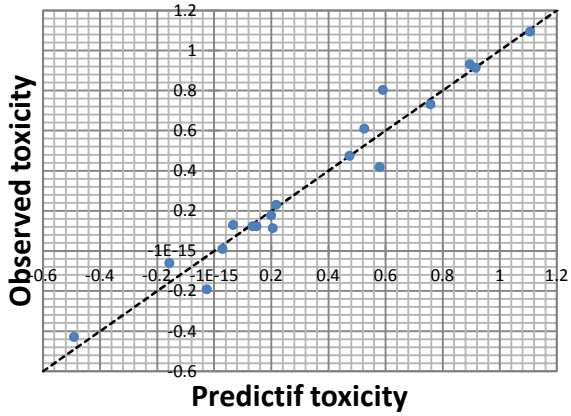

Figure 4: Graphical representation of calculated and observed toxicity by MNLR

The obtained coefficient of determination in equation (2) is quite very interesting (**0.98**).

The proper predictive power of a QSAR model is to test their ability to predict accurately the toxicity of compounds from an external test set (compounds which were not used for the model development), the toxicity of the remained set of 5 compounds are deduced from the quantitative model proposed with the 18 molecules (training set) by MLR and MNLR. The observed and calculated toxicity values are given in tables 4.

The comparison of the values of toxicity pIC_{50}-test to pIC_{50}-obs shows that a good prediction has been obtained for the 5 compounds:

Multiple linear regression: N=5 $r_{test} = 0.80$ $r^2_{test} = 0.64$

Multiple non-linear regression: N=5 $r_{test} = 0.78$ $r^2_{test} = 0.61$

Table 4: Observed values and calculated values of pIC_{50} according to MLR and MNLR for the 5 tested compounds (test set).

No	Obs.	MLR	MNLR
		Pred-test	Pred-test
19	1.695	0.844	0.850
20	1.297	1.139	1.005
21	1.245	1.182	1.052
22	2.113	3.136	3.455
23	1.788	2.423	1.381

The results obtained by MLR and MNLR are very sufficient to conclude the performance of the model; it's confirmed by the test done with the 5 compounds. Even if it is possible that this good prediction is found by chance we can claim that it is a positive result. So, this model could be applied to all phenol derivatives accordingly to table 1.

A comparison of the quality of MLR and MNLR models shows that the approach 2 is better predictive capability because it gives better results. MLR and MNLR were able to establish a satisfactory relationship between the molecular descriptors and the toxicity pIC_{50} of the studied compounds.

The developed equations can be used for the designing of new phenol derivatives with improved the Activity.

To optimize the error standard deviation and a better finish to building our model, we involve in the next part artificial neural networks (ANN).

Artificial neural networks ANN

In order to increase the probability of good characterization of studied compounds, neural networks (ANN) can be used to generate predictive models of quantitative structure–activity relationships (QSAR) between a set of molecular descriptors obtained from the MLR and observed activity. The ANN calculated toxicity model develop using the properties of several studied compounds. The correlation between ANN calculated and experimental toxicity values are very significant as illustrated in figure 8 and as indicated by R and R^2 values.

N = 23 R = 0.95 R^2 = 0.90 RMSE = 0.05

The R^2 = 0.90 value confirms that the results of the artificial neural network were the best for building quantitative structure–activity models. In order to validate the generated ANN model 'Leave-one-out' method is use to check their predictivity and robustness, test sets of new compounds, not included in the model development set, must be used. The 'Leave-one-out' is an approach particularly well adapted to the estimation of that ability. In this procedure, one compound is removed from the data set, the network is train with the remaining compounds and used to predict the discarded compound. The process is repeated in turn for each compound in the data set. In this paper the 'leave-one-out' procedure is use to evaluate the predictive ability of the ANN.

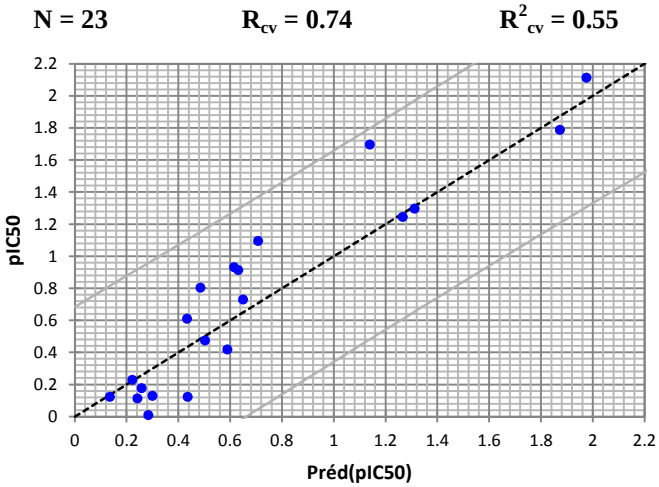

Figure 5: Correlations of observed and predicted activities calculated using ANN

The correlation between the calculated and experimental artificial neural network values were highly significant, as illustrated in Fig. 5 and as indicated by the R and R^2 values. The predicted activities calculated with the artificial neural network and the observed values are given in Table 5.

The obtained squared correlation coefficient (R^2) value is 0.96 for this data set of halogenated phenols. It confirms that the multiple non-linear regression results were the best to build the quantitative structure activity relationship models. In this study, we investigated the best linear QSAR regression equations established in this study. Based on this result, a comparison of the quality of the CPA, MLR, MNLR and ANN models shows that the MNLR model has substantially better predictive capability because the MNLR approach gives better results than MLR and ANN. MNLR is able to establish a satisfactory relationship between the molecular descriptors and the activity of the studied compounds.

Table 5: Observed values and calculated values of pIC_{50} according to different methods

	pIC_{50} (obs.)	pIC_{50} (calc.)			
N°		MLR	NMLR	ANN	CV
1	-0.431	-0.471	-0.491	-0.560	-0.017
2	-0.192	-0.081	-0.025	-0.035	0.234
3	-0.062	-0.016	-0.157	0.013	0.131
4	0.229	0.203	0.218	0.198	0.258
5	0.176	0.121	0.200	0.199	-0.007
6	0.122	0.217	0.149	0.203	0.316
7	0.128	0.083	0.067	0.032	0.222
8	0.009	0.019	0.029	-0.045	0.222
9	0.122	0.173	0.135	0.204	0.222
10	0.113	0.231	0.205	0.230	0.184
11	0.803	0.521	0.591	0.584	0.465
12	0.609	0.453	0.527	0.457	0.391
13	0.473	0.450	0.475	0.459	0.448
14	0.730	0.692	0.759	0.725	0.641
15	0.913	0.911	0.917	1.113	0.584
16	1.094	1.226	1.107	1.171	1.728
17	0.418	0.584	0.580	0.674	0.581
18	0.930	0.869	0.897	1.786	0.706

CONCLUSION

Multiple linear and non-linear regression analysis and artificial neural networks are used to construct a relationship between several descriptors and inhibition values pIC_{50} of phenol derivatives. The multiple nonlinear regression is substantially better predictive capability than the other two models, with greater predictive power. We establish a relationship between several descriptors and inhibition values pIC_{50} of phenol derivatives, with external validation. The results show that the model proposed in this paper can predict activity accurately and that the selected descriptors are pertinent.

The accuracy and predictability of the proposed models are illustrated by comparison of the key statistical terms R or R^2 for the different models (table 6) and the predictive powers of the equations are validated by an external test set (Table4). The proposed methods will reduce the time and cost of synthesis and determination of the toxicity pIC_{50} of phenol derivatives.

Furthermore, the descriptors are sufficiently rich in chemical and electronic information to encode structural features that could be used with other descriptors in the development of predictive QSAR models.

Acknowledgment
We are grateful to the “Association Marocaine des Chimistes Théoriciens” (AMCT) for its pertinent help concerning the programs.

REFERENCES

[1] Y. B. Zang, *Chinese Agricultural Science Bulletin* **2012**, 28, 282-285.
[2] P. R. Zhan, H. T. Wang, Z. X. Chen, *Journal of Agro-Environment Science* **2008**, 27, 801-804.
[3] J. W. Chen, W. J. Peijnenburg, X. Quan, The Use of PLS, *Chemosphere* **2000**, 40, 1319-1326.
[4] M. J. Zhu, F. Ge, R. L. Zhu, *Chemosphere* **2010**, 80, 46-52.
[5] F. A. Pasha, H. K. Srivastava, P. P. Singh, *Bioorganic & Medicinal Chemistry*, **2005**, 13, 6823–6829.
[6] M. Parac, S. Grimme, All calculations were done by GAUSSIAN 03 W software, *J. Phys. Chem*., **2003**, 106, 6844-6850.
[7] M. Larif, A. Adad, R. Hmamouchi, A. I. Taghki, A. Soulaymani, A. Elmidaoui, M. Bouachrine, T. Lakhlifi, Article in press in *Arabian Journal of Chemistry*, **2013**.
[8] XLSTAT **2013**software (XLSTAT Company). http://www.xlstat.com.
[9] A. Ousaa, B. Elidrissi, M. Ghamali, S. Chtita, M .Bouachrine ,T. Lakhlifi, *Journal of Computational Methods in Molecular Design*, **2014**, 4 (3), 10-18.
[10] J. Zupan, J. Gasteiger, Neural Networks for Chemists an Introduction, VCH Publishers Weinheim, **1993**.
[11] D. Cherquaoui, D. Villemin, *J. Chem. Soc. Faraday. Trans*., **1994**, 90, 97-102.
[12] J. A. Freeman, D. M. Skapura, *J. Operat. Res. So*., **1991**, 43(11), 1106.
[13] M. Ghamali, S. Chtita, A. Adad, R. Hmamouchi, M. Bouachrine, T. Lakhlifi, *IJARCSSE*, **2014**, 4, 536-546.
[14] G. Hea, L. Fenga, H. Chena, International Symposium on Safety Science and Engineering in China, Procedia Engineering, **2012**, 43, 204 – 209.
[15] STATITCF Software, Technical Institute of cereals and fodder, Paris, France **1987**.
[16] N. Jonathan, H. Nobuyasu, S. Yuso, K. Ahsan, H. Jae-Jak, N. N. Jong-Sik, L. L. Q. Xiang, L. L. Jun, Z. Gan, M. Shigeki, *Chemosphere*, **2012**, 86, 718–726.