QSPR MODELLING OF STABILITY CONSTANTS OF METAL-THIOSEMICARBAZONE COMPLEXES USING MULTIVARIATE REGRESSION METHODS AND ARTIFICIAL NEURAL NETWORK

NGUYEN MINH QUANG[1,2], TRAN NGUYEN MINH AN[1], NGUYEN HOANG MINH[1], TRAN XUAN MAU[2], PHAM VAN TAT[3]

[1]*Faculty of Chemical Engineering, Industrial University of Ho Chi Minh City*

[2]*Department of Chemistry, University of Sciences – Hue University*

[3]*Faculty of Science and Technology, Hoa Sen University*

nguyenminhquang@iuh.edu.vn

Abstract: In this study, the stability constants of metal-thiosemicarbazone complexes, $\log\beta_{11}$ were determined by using the quantitative structure property relationship (QSPR) models. The molecular descriptors, physicochemical and quantum descriptors of complexes were generated from molecular geometric structure and semi-empirical quantum calculation PM7 and PM7/sparkle. The QSPR models were built by using the ordinary least square regression ($QSPR_{OLS}$), partial least square regression ($QSPR_{PLS}$), primary component regression ($QSPR_{PCR}$) and artificial neural network ($QSPR_{ANN}$). The best linear model $QSPR_{OLS}$ (with *k* of 9) involves descriptors C5, xp9, electric energy, cosmo volume, N4, SsssN, cosmo area, xp10 and core-core repulsion. The $QSPR_{PLS}$, QSPR $_{PCR}$ and $QSPR_{ANN}$ models were developed basing on 9 varibles of the $QSPR_{OLS}$ model. The quality of the QSPR models were validated by the statistical values; The $QSPR_{OLS}$: R^2_{train} = 0.944, Q^2_{LOO} = 0.903 and MSE = 1.035; The $QSPR_{PLS}$: R^2_{train} = 0.929, R^2_{CV} = 0.938 and MSE = 1.115; The $QSPR_{PCR}$: R^2_{train} = 0.934, R^2_{CV} = 0.9485 and MSE = 1.147. The neural network model $QSPR_{ANN}$ with architecture I(9)-HL(12)-O(1) was presented also with the statistical values: R^2_{train} = 0.9723, and R^2_{CV} = 0.9731. The QSPR models also were evaluated externally and got good performance results with those from the experimental literature.

Keywords: QSPR, stability constants $\log\beta_{11}$, ordinary least square regression, partial least square, primary component regression, artificial neural network, thiosemicarbazone.

1. INTRODUCTION

Thiosemicarbazone compounds and its metal complexes were widely researched in the world because of its diversified application areas in fact. In the field of chemistry, thiosemicarbazones are used as analytical reagents [1,2], they are also used as a catalyst in chemical reactions [3,4]. Besides, they also have application in biology [5], environment [6] and medicine [7,8].

For complexes, the stability constant of complexes is an important factor. This is hold to identify the complex stability in solutions with different solvents. The stability constant of complexes is the hinge parameter to explain phenomenon such as the mechanism of reaction and distinct properties of the biological systems. Augmentation, it is also a measure of the power of the interaction between the metal ions and the ligand to form complexes. We can calculate the equilibrium concentration of substances in a solution upon the stability constant. The changes of the complex structure in solutions can be forecasted by using the initial concentration of the metal ion and the ligand.

In recent years, the stability constant of the complexes has been researched by incorporating the UV/VIS spectrophotometric method and the computational chemistry [9]. Furthermore, the in silico methods that QSAR/QSPR methods are also used for predicting properties/activities of complexes based on the relationships between the structural descriptors and the properties/activities [9]. Here, a few complex descriptors between the metal ions and thiosemicarbazone were determined by quantum mechanics methods [10–12].

On the other hand, computer science has evolved dramatically, it has been becoming a helpful tool to develop computational chemistry such as material simulation and data mining [13–16]. The molecular design by means of a computer is also a way to accelerate the discovery process for resulting knowledge of material properties. This is also a tendency to reduce the classical trial-and-error approach [17]. In this case, the development of molecular models such as the quantitative structure and property relationship (QSPR) and conformational search methodologies has also contributed greatly to the discovery and development of new molecules [18,19]. In this way, the multivariate analysis methods have been becoming a convenient and an easy tool for supporting empirical and theoretical models. The multivariable linear relationships can be used to assess the different characteristics of the systems.

In this work, we successfully constructed of the quantitative structure and properties relationships (QSPRs) using the 2D and 3D-descriptors, structural descriptors and stability constant of complexes between the metal ions and thiosemicarbazone. The structural descriptors are calculated by using the semi-empirical quantum chemistry method with new version PM7 and PM7/sparkle [20], molecular mechanics, and connectivity calculation. Three multivariate regression models are established $QSPR_{OLS}$, $QSPR_{PLS}$ and $QSPR_{PCR}$ models by using the ordinary least square regression, partial least square regression and primary component regression methods. In addition, the artificial neural network model $QSPR_{ANN}$ is constructed by the error back-propagation method using multilayer perceptron algorithm with the input layer that includes variables of the best selected $QSPR_{OLS}$ model. The stability constant $\log\beta_{11}$ of the metal-thiosemicarbazone in the test set resulting from the QSPR models is validated and compared with those from experimental data in the published scientific works.

2. COMPUTATIONAL METHODS

In order to develop a QSPR model, there are several steps must be considered [21] which are described in detail in the following subsections.

2.1. Stability constant of complex and data selection

In an aqueous solution, the formation of a complex between a metal ion (M) and a thiosemicarbazone ligand (L) is the general equilibrium reaction [14]

$$p\,\mathrm{M} + q\,\mathrm{L} \rightleftharpoons \mathrm{M}_p\mathrm{L}_q \tag{1}$$

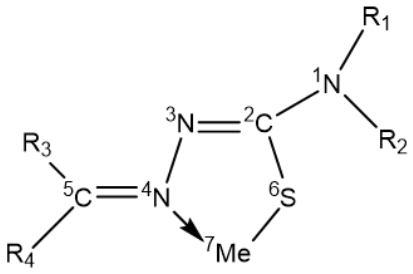
The stability constant, given the symbol β, is the constant for the formation of the complex from the reagents. The stability constant for the formation of M_pL_q is given by

$$\beta_{pq} = \frac{[\mathrm{M_pL_q}]}{[\mathrm{M}]^{\mathrm{p}}[\mathrm{L}]^{\mathrm{q}}} \tag{2}$$

In one step with $p = 1$ and $q = 1$, the stability constant, given the symbol β_{11}, is the stability constant for the formation of ML, it is given by

$$\beta_{11} = \frac{[\mathrm{ML}]}{[\mathrm{M}][\mathrm{L}]} \tag{3}$$

a)



b)

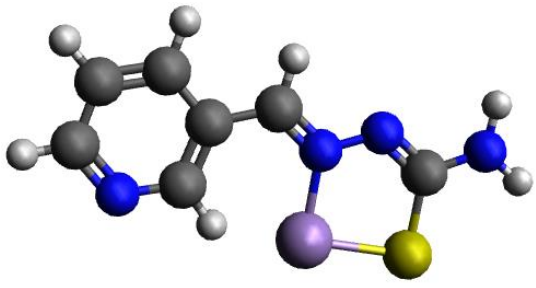


Figure 1. Structure of the metal-thiosemicarbazone complex: a) General complex structure; b) Complex between Mn^{2+}/Ni^{2+} and 3-formylpyridine thiosemicarbazone [22]

A data set of the values logβ_{11} of complexes between metal ions and the ligand thiosemicarbazone were taken from the literature on Table 1.

Table 1. Complexes of metal ions and thiosemicarbazone and stability constant

Ord	Thiosemicarbazone				Metal ions	logβ_{11}	Ref.
	R_1	R_2	R_3	R_4			
1	H	H	H	$-C_7H_7O_3$	Cu(II)	5.000	[23]
2	H	H	H	$-C_{13}H_{16}NO_3$	Cu(II)	17.540	[24]
3	H	H	H	$-C_{13}H_{16}NO_3$	Fe(III)	19.480	[24]
4	H	H	H	$-C_8H_9O_3$	Cd(II)	5.544	[25]
5	H	H	H	$-C_6H_3OHOCH_3$	Mo(VI)	6.5514	[26]
6	CH_3	$-CH_3$	$-C_5H_4N$	$-C_5H_4N$	Fe(III)	7.060	[27]
7	H	H	H	$-C_{14}H_{12}N$	Cd(II)	5.860	[28]
8	H	H	H	$-C_4H_3O$	Cu(II)	14.670	[29]
9	H	$-C_6H_5$	H	$-C_9H_6NO$	Cu(II)	15.650	[29]
10	H	H	H	$-C_5H_4N$	Zn(II)	7.300	[29]
11	H	$-CH_3$	$-CH_3$	H	Ag(I)	14.500	[30]
12	H	H	H	$-C_7H_7O_3$	Ag(I)	15.700	[30]
13	H	H	H	$-C_5H_4N$	Cu(II)	17.200	[31,32]
14	H	H	H	$-C_6H_3OHOCH_3$	Cd(II)	7.340	[33]
15	H	H	H	$-C_6H_4OH$	Zn(II)	7.470	[33]
16	H	H	H	$-CCH_3NOH$	Mn(II)	5.000	[34]
17	H	H	$-C_6H_5$	$-C_7H_6NO$	Cu(II)	5.7482	[35]
18	H	H	H	$-C_6H_3OHOCH_3$	Cu(II)	11.610	[36]
19	H	H	H	$-C_6H_4NO_2$	La(III)	10.840	[37]
20	H	H	H	$-C_6H_4NO_2$	Pr(III)	11.040	[37]
21	H	H	H	$-C_6H_4NO_2$	Nd(III)	9.090	[37]
22	H	H	$-CH_3$	$-C_6H_4OH$	Cd(II)	10.630	[38]
23	H	H	$-CH_3$	$-C_6H_4OH$	Al(III)	11.240	[38]
24	H	H	-	$-C_9H_8NO$	Cu(II)	5.491	[39]
25	H	H	H	$C_6H_4NH_2$	Cu(II)	5.924	[39]

2.2. Descriptors calculation

Molecular descriptors can be defined as basic numerical characteristics related to chemical structures. So the complexes of metal-thiosemicarbazone were built structure molecular by BIOVIA Draw 2017 R2 [40] and optimized by means of quantum mechanics on the MoPac2016 system [41]. The two and three-dimensional of the molecular in the database were calculated by using the QSARIS system [15,42]. The quantum descriptors were calculated by using the semi-empirical quantum method with new version PM7 and PM7/sparkle for lanthanides [20].

After computation, the proceeding of removing non-conforming variables for resulting receives a set of databases that includes observations with the logβ_{11} values and the variables as the calculated structural parameters. And we use this database to develop regression models and neural networks.

2.3. Multivariate regression model development

The three regression methods were used in this study, which are the ordinary least square regression, primary component regression and partial least square regression. It has the common characteristic of generating models that involve linear combines of explanatory variables. The difference between the three method lies on the way the correlation structures between the variables are handled.

The ordinary least square regression (OLS) is used to model and predict the values of one or more dependent quantitative or qualitative variables by means of a linear combination of one or more explanatory quantitative and/or qualitative variables, without facing the constraints of ordinary least square regression on the number of variables versus the number of observations.

In this case, the regression model with *k* explanatory variables writes

$$Y = \beta_0 + \sum_{j=1}^{k} \beta_j \cdot X_j + \varepsilon \quad (3)$$

where Y is the dependent variable, β_0, is the intercept of the model, X_j corresponds to the j^{th} explanatory variable (with j = 1 to k), and ε is the random error with expectation 0 and variance σ^2.

In the case of k observations, the estimation of the predicted value of the dependent variable Y is given by expression (4)

$$\hat{Y} = \beta_0 + \sum_{j=1}^{k} \beta_j \cdot X_j \quad (4)$$

The principal components regression (PCR) can be divided into three steps: firstly, it calculates a principal components analysis (PCA) on the table of the explanatory variables, secondly, it calculates an OLS regression on the selected components, then it computes the parameters of the model that correspond to the input variables.

PCA allows to transform an X table with n observations described by variables into an S table with n scores described by q components, where q is lower or equal to p and such that (S'S) is invertible. An additional selection can be applied on the components so that only the r components that are the most correlated with the Y variable are kept for the OLS regression step. We then obtain the R table.

The partial least square regression method is quick, efficient and optimal for a criterion based on covariance. It is recommended in cases where the number of variables is high, and where it is likely that the explanatory variables are correlated.

The idea of PLS regression is created, starting from a table with n observations described by p variables, a set of h components with h < p. The method used to build the components differs from PCA, and presents the advantage of handling missing data. The determination of the number of components to keep is usually based on a criterion that involves a cross-validation. The equation of the PLS regression model writes

$$Y = T_h C'_h + E_h = X W^*_h C'_h + E_h = X W_h \left(P'_h W_h\right)^{-1} C'_h + E_h \quad (5)$$

where Y is the matrix of the dependent variables, X is the matrix of the explanatory variables. T_h, C_h, W^*_h, W_h and P_h, are the matrices generated by the PLS algorithm, and E_h is the matrix of the residuals.

The matrix B of the regression coefficients of Y on X, with h components generated by the PLS regression algorithm is given by

$$B = W_h \left(P'_h W_h\right)^{-1} C'_h \quad (6)$$

The three methods give the same results if the number of components obtained from the PCA (in PCR) or from the PLS regression is equal to the number of explanatory variables. The components obtained from the PLS regression are built so that they explain as well as possible Y, while the components of the PCR are built to describe X as well as possible.

The models were screened by using the values R^2_{train} and Q^2_{LOO}. These were assessed by the same formula (6)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} \quad (7)$$

where Y_i, $\hat{Y}_i$, and $\bar{Y}$ are the experimental, calculated and average value, respectively.

Adjusted R^2 (R^2_{adj}) is the adjusted determination coefficient for the model. The value R^2_{adj} can be negative if the R^2 is near to zero. This coefficient is only calculated if the constant of the model has not been fixed by the user. R^2_{adj} is defined by

$$R^2_{adj} = R^2 - \frac{k-1}{N-1}\cdot\left(1-R^2\right) \tag{8}$$

The R^2_{adj} is a correction to R^2, which takes into account the number of variables used in the model. The mean squared error (*MSE*) is defined by

$$MSE = \frac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{N-k-1} \tag{9}$$

The root mean square of the errors (*RMSE*) and the standard errors (*SE*) is the square root of the *MSE*.

2.4. ANN model development

Artificial neural network (ANN) is computing systems dubiously inspired by the biological neural networks that create animal brains. An ANN is based on a collection of connected units or nodes called artificial neurons which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transfer a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it [43].

In common ANN implementations, the signal at a connection between artificial neurons are real number, and the output of each artificial neuron is computed by some non-linear function of the sum of its inputs. The connections between artificial neurons are called 'edges'. Artificial neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that the signal is only sent if the aggregate signal crosses that threshold. Typically, artificial neurons are aggregated into layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first layer (the input layer), to the last layer (the output layer), possibly after traversing the layers multiple times. [44,45].

Neural network models can be viewed as simple mathematical models defining a function f: X → Y or a distribution over X or both X and Y. The functions applied at the nodes of the hidden layers are called activation functions. The activation function is a transformation of a linear combination of the *X* variables. The function applied at the response is a linear combination of continuous responses, or a logistic transformation for nominal or ordinal responses [44,45]. There are three transfer functions, namely sigmoid, hyperbolic tangent, and Gaussian transfers function.

The main advantage of the neural network model is that it can model efficiently different response surfaces. Neural networks are very flexible models and have a tendency to overfit data. The main disadvantage of a neural network model is that the results are not easily interpretable, since there are intermediate layers rather than a direct path from the *X* variables to the *Y* variables, as in the case of regular regression [44,45].

In this work, we used a typical feed-forward neural network with an error back-propagation learning algorithm to train it. This neural network style propagates information in the feed-forward direction using equation (10) [46]

$$b_j = f\left(\sum_{i=0}^{N} w_{i,j}\cdot a_i - T_j\right) \tag{10}$$

where a_i is the input factor, b_j is the output factor, w_{ij} is the weight factor between two nodes, T_j is the internal threshold, and f is the transfer function. There are many transfer functions that are used in neural networks where hyperbolic tangent function is used in this study, a hyperbolic tangent learning algorithm is based on a generalized delta-rule accelerated by a momentum term. To increase the efficiency of the neural network, both the weight factors and the internal threshold values were adjusted using equations (11) and (12) [46]

$$W^{new}_{i,j} = w^{old}_{i,j} + \eta.\sum_k \delta_{k,j}.O_{k,i} + \alpha.\Delta W^{old}_{i,j} \tag{11}$$

$$T_j^{new} = T_j^{old} + \eta.\sum_k \delta_{k,j} + \alpha.\Delta T_j^{old} \quad (12)$$

where η is the learning rate; α is the momentum coefficient; Δ*W* is the previous weight factor change; Δ*T* is the previous threshold value change; *O* is the output – the gradient-descent correction term; and *k* stands for the pattern.

The performance of the trained network was verified by determining the error between the predicted value and the real value. All the data of the patterns were normalized to be less than 1 before training the neural network; the initial weight factors were randomly generated from –0.2 to 0.2, and the initial internal threshold values were set to zero [46,47].

3. RESULTS AND DISCUSSION

3.1. Constructing models $QSPR_{OLS}$, $QSPR_{PCR}$ and $QSPR_{PLS}$

The construction of $QSPR_{OLS}$ model was performed using back-elimination and the forward regression technique on the Regress system [48] and MS-EXCEL [13,15,49]. The construction of $QSPR_{PLS}$ and $QSPR_{PCR}$ models were effectuated using XLSTAT2016 [50] and MS-EXCEL [13,15,49]. The predictability of QSPR models was cross-validated by means of the leave-one-out method (LOO) using the statistic Q^2_{LOO}.

The multivariate regression models were constructed based on the training set and the test set, in which the portion of the test set is 20 %. The quality of models were evaluated by means of statistical values R^2_{train}, R^2_{adj}, Q^2_{LOO} and F_{stat} (Fischer's value). The $QSPR_{OLS}$ models and the statistical values are shown in Table 2.

Table 2. Selected model $QSPR_{OLS}$ (*k* of 2 to 10) and statistical values

k	Variables	*SE*	R^2_{train}	R^2_{adj}	Q^2_{LOO}	F_{stat}
2	x_1/x_2	3.149	0.394	0.368	0.274	15.28537
3	$x_1/x_2/x_3$	2.716	0.559	0.530	0.429	19.42606
4	$x_1/x_2/x_3/x_4$	2.586	0.609	0.574	0.486	17.52034
5	$x_1/x_2/x_3/x_4/x_5$	2.346	0.685	0.650	0.554	19.16658
6	$x_1/x_2/x_3/x_4/x_5/x_6$	2.089	0.756	0.722	0.622	22.20887
7	$x_1/x_2/x_3/x_4/x_5/x_6/x_7$	1.875	0.808	0.776	0.685	25.27557
8	$x_1/x_2/x_3/x_4/x_5/x_6/x_7/x_8$	1.586	0.866	0.840	0.782	33.12386
9	$\mathbf{x_1/x_2/x_3/x_4/x_5/x_6/x_7/x_8/x_9}$	**1.035**	**0.944**	**0.932**	**0.903**	**75.28873**
10	$x_1/x_2/x_3/x_4/x_5/x_6/x_7/x_8/x_9/x_{10}$	0.940	0.955	0.944	0.880	83.25919

Notation of molecular descriptors			
C5	x_1	SsssN	x_6
xp9	x_2	cosmo area	x_7
electric energy	x_3	xp10	x_8
cosmo volume	x_4	core-core repulsion	x_9
N4	x_5	Hmax	x_{10}

The best linear models $QSPR_{OLS}$ were selected with the critical value α = 0.05; the important descriptors selected were based on the changes of the statistical parameters: standard error – *SE*, R^2_{train}, R^2_{adj}, Q^2_{LOO}, and F_{stat}. The number of descriptors *k* was selected in range 2 to 10. The change of the amount of structural parameter leads to the change of the values *SE*, R^2_{train} and Q^2_{LOO} (Figure 2a).

The selected variables included in the $QSPR_{OLS}$ models (Table 2), showed that the R^2_{train}, Q^2_{LOO} and F_{stat} values change and increase with *k* variables. When k values increase from 9 to 10, the corresponding statistical values add up negligibly and tend to decrease as Q^2_{LOO} values, so choosing the *k* of 9 was

appropriated for the change trend. The variables from x_1 to x_9 were examined for the internal correlation between two or more variables based on the Pearson correlation coefficient matrix, which determines the significant correlation for logβ_{11}. The correlation matrix is given in Table 3.

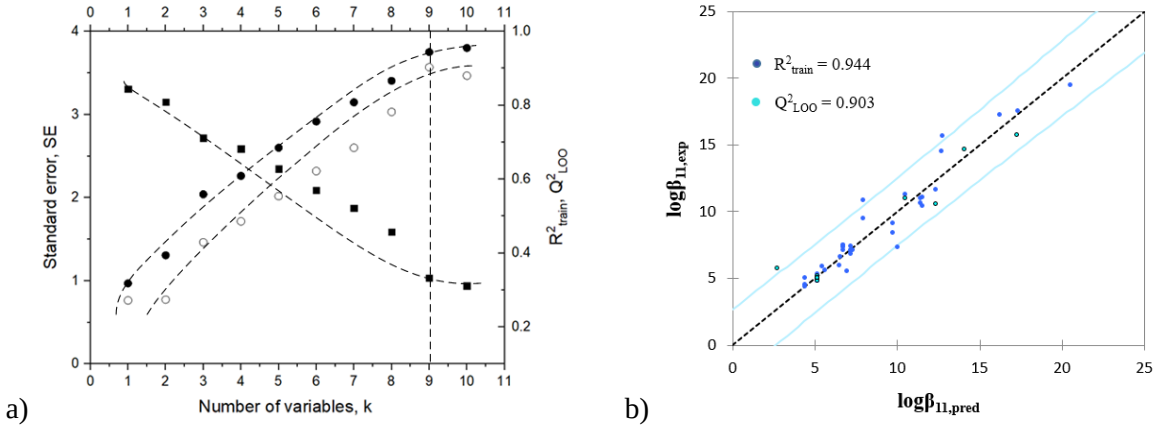


Figure 2. a) Change trend line of values *SE*, R^2_{train} and Q^2_{LOO} according to *k* descriptors; b) Correlation of experimental versus predicted values logβ_{11} of the test compounds using the QSPR$_{OLS}$ model (with k = 9)

Table 3. Pearson correlation matrix of variables in the QSPR$_{OLS}$ model with *k* of 9

Variables	logβ_{11}	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
logβ_{11}	**1**	-0.517	0.251	-0.451	0.420	0.288	0.347	0.440	0.444	0.305
x_1	-0.517	**1**	0.041	-0.046	-0.233	-0.381	-0.274	-0.273	0.046	-0.076
x_2	0.251	0.041	**1**	-0.798	0.682	-0.133	0.640	0.704	0.799	0.989
x_3	-0.451	-0.046	-0.798	**1**	-0.868	0.132	-0.634	-0.853	-1.000	-0.792
x_4	0.420	-0.233	0.682	-0.868	**1**	0.095	0.550	0.994	0.876	0.723
x_5	0.288	-0.381	-0.133	0.132	0.095	**1**	0.159	0.076	-0.119	-0.087
x_6	0.347	-0.274	0.640	-0.634	0.550	0.159	**1**	0.557	0.635	0.638
x_7	0.440	-0.273	0.704	-0.853	0.994	0.076	0.557	**1**	0.861	0.752
x_8	0.444	0.046	0.799	-1.000	0.876	-0.119	0.635	0.861	**1**	0.794
x_9	0.305	-0.076	0.989	-0.792	0.723	-0.087	0.638	0.752	0.794	**1**

Based on the results of Table 3, the correlation coefficients of 9 independent variables and a dependent variable logβ_{11} showed that the selected variables in the QSPR$_{OLS}$ model with *k* of 9 were consistent and statistically acceptance and correlated t-student characterized the variables. The linear regression equation of the QSPR$_{OLS}$ model with the statistical values follows

$$\log\beta_{11} = -64.63 - 24.58 \cdot x_1 + 26.71 \cdot x_2 - 0.02334 \cdot x_3 - 0.355 \cdot x_4 + 25.47 \cdot x_5 - 2.143 \cdot x_6 + 0.531 \cdot x_7 - 38.16 \cdot x_8 - 0.02505 \cdot x_9 \quad (13)$$

$$n = 50;\ R^2_{train} = 0.944;\ Q^2_{LOO} = 0.903;\ MSE = 1.035$$

Thus, the training dataset used to build the QSPR$_{OLS}$ model satisfies the statistical requirements and good prediction. The predictability of the QSPR$_{OLS}$ model is well suited to the group of complexes. The selected parameters in the model have no correlation between the selected variables. This modeling data will be used to develop the QSPR$_{PCR}$ and QSPR$_{PLS}$ models.

Using a matrix of data with independent variables (k = 9) and a dependent variable logβ_{11}, the QSPR$_{PCR}$ model was constructed from the results of the primary components analysis PCA, which showed that 9 major components were statistically significant. The regression equation of the QSPR$_{PCR}$ model with the statistical values follows

$$\log\beta_{11} = -64.064 - 23.655 \cdot x_1 + 24.918 \cdot x_2 - 0.022 \cdot x_3 - 0.400 \cdot x_4 + 26.040 \cdot x_5 - 1.840 \cdot x_6 + 0.574 \cdot x_7 - 36.476 \cdot x_8 - 0.024 \cdot x_9 \quad (14)$$

n = 50; R^2_{train} = 0.934; R^2_{CV} = 0.9485; *MSE* = 1.147; *RMSE* = 1.071

Similarly from the results of the $QSPR_{PCR}$ modeling, proceed to construct a $QSPR_{PLS}$ model based on a data matrix with 9 independent variables. The quality of the $QSPR_{PLS}$ model was assessed based on statistical indicators with cumulative statistical values Q^2_{cum} = 0.177; R^2_{Ycum} = 0.934 and R^2_{Xcum} = 0.999. In addition, based on the Variable Importance for the Projection (VIP) of the variables X affects $\log\beta_{11}$ in the $QSPR_{PLS}$ model and the deviation value of the variables, from which the model variables are selected. So the $QSPR_{PLS}$ model gives the following results

$$\log\beta_{11} = -55.976 - 26.729 \cdot x_1 + 25.082 \cdot x_2 - 0.020 \cdot x_3 - 0.353 \cdot x_4 + 24.146 \cdot x_5 - 2.277 \cdot x_6 + 0.504 \cdot x_7 - 36.044 \cdot x_8 - 0.021 \cdot x_9 \quad (15)$$

n = 50; R^2_{train} = 0.934; R^2_{CV} = 0.9658; *MSE* = 0.982; *RMSE* = 0.991

In the QSPR models, the R^2_{train} value is the coefficient of multiplication correlation that multiplied by 100 times with variance will explain the stability constant $\log\beta_{11}$. The predictability of QSPR models is evaluated by R^2_{CV} and Q^2_{LOO}. The F_{stat} values reflect the variance ratio explained by the model and the variance from the regression error. The high F_{stat} value indicates that the model is statistically significant. The low *MSE* and *RMSE* values also indicate that the model is statistically significant. The predictive power of the model is shown by the value of the Q^2_{test} for the non-original compounds group.

3.2. Constructing model $QSPR_{ANN}$

In addition to regression models, the $QSPR_{ANN}$ model is also developed with the neural network technique on the Visual Gene Developer system [46] upon 9 variables of model $QSPR_{OLS}$. The architecture of the neural network consist of three layers I(9)-HL(12)-O(1) (Fig. 3); the input layer I(9) includes 9 neurons that are C5, xp9, electric energy, cosmo volume, N4, SsssN, cosmo area, xp10 and core-core repulsion; the output layer O(1) includes 1 neuron that is the $\log\beta_{11}$; the hidden layer includes 12 neurons.

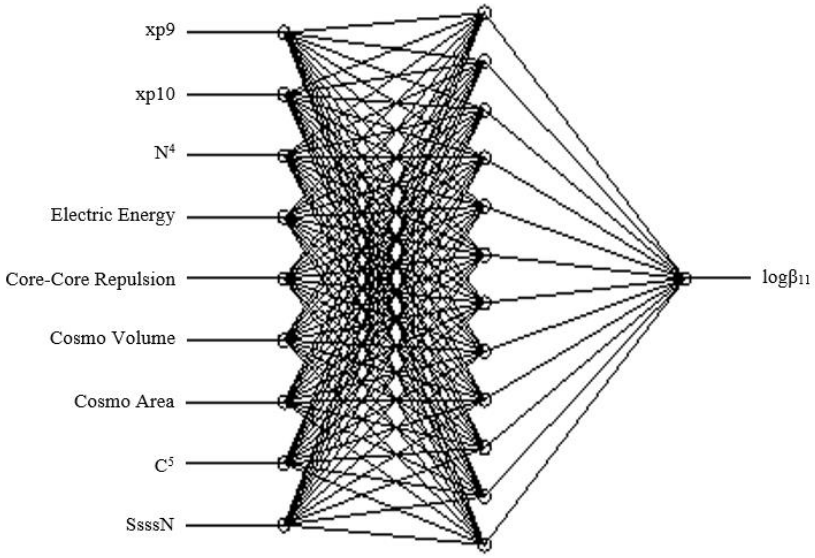


Figure 3. Architecture of neural network I(9)-HL(12)-O(1)

The error back-propagation algorithm is used to train the network. The hyperbolic tangent transfer function sets on each node of the layer neural network; the training network parameters include the learning rate of 0.01; the momentum coefficient of 0.1. The results got the sum of error 0.000021 with 1,500,000 loops and the regression coefficients of the training process are given in Table 4.

Table 4. Training quality of neural network $QSPR_{ANN}$ I(9)-HL(12)-O(1)

Data set	Regression coefficient	Slope	*y*-intercept
Training	0.9723	0.9659	0.187
Validation	0.9731	0.9938	–0.1134

As observation of eq. 13-15 and table 4, the neural model $QSPR_{ANN}$ based on the architecture of neural network I(9)-HL(12)-O(1) adapts better than the built QSPR models. In fact, neural model $QSPR_{ANN}$ exhibits a better fit and correlation between the predicted values and the experimental values than the $QSPR_{OLS,}$ $QSPR_{PLS}$ and $QSPR_{PCR}$ models through Q^2_{test} values (Table 5b and Fig. 4).

3.3. Predictability of QSPR models

The predictability of the QSPR models was carefully evaluated by means of the phasing-each-case technique. The predicted results received for 10 randomly chosen substances with the experimental values are described in Table 5a and 5b.

The average absolute values of the relative error *MARE (%)* used to assess the overall error of the QSPR models are calculated according to formula (16)

$$MARE,\% = \frac{\sum_{i=1}^{n} ARE_i,\%}{n} \quad \text{where} \quad ARE,\% = \frac{\left|\log\beta_{11,\text{exp}} - \log\beta_{11,\text{cal}}\right|}{\log\beta_{11,\text{exp}}}100 \tag{16}$$

n is the number of test substances; $\beta_{11,\text{exp}}$ and $\beta_{11,\text{cal}}$ are the experimental and calculated stability constants.

Table 5a. Stability constant of 10 test substances for validated externally

Ord	Thiosemicarbazone				Metal Ions	$\log\beta_{11,\text{exp}}$	Ref.
	R_1	R_2	R_3	R_4			
1	H	$-C_6H_5$	$-CH_3$	$-C_2H_3NOH$	V(V)	5.3222	[51]
2	$-CH_3$	$-CH_3$	$-C_5H_4N$	$-C_5H_4N$	Co(II)	11.970	[52]
3	H	H	H	$-C_{13}H_{16}NO_3$	Co(II)	5.360	[53]
4	H	H	H	$-CH=CHC_6H_5$	Co(II)	5.099	[54]
5	H	H	CH_3	$-CH=N-NHC_6H_5$	Co(II)	9.900	[55]
6	H	H	CH_3	$-CH=N-NHC_6H_5$	Mn(II)	9.600	[55]
7	H	H	H	$-C_6H_3OHOCH_3$	Cu(II)	11.980	[55]
8	H	$-C_2H_5$	H	$-C_9H_5NOH$	Cu(II)	19.100	[31,32]
9	H	H	-	$-C_9H_8NO$	Zn(II)	7.654	[56]
10	H	H	-	$-C_9H_8NO$	Cd(II)	6.611	[56]

Table 5b. Stability constant of 10 test substances resulting from the QSPR models

Ord	$\log\beta_{11,\text{exp}}$	$QSPR_{OLS}$		$QSPR_{PLS}$		$QSPR_{PCR}$		$QSPR_{ANN}$	
		$\log\beta_{11,\text{cal}}$	*ARE*, %	$\log\beta_{11,\text{cal}}$	*ARE*, %	$\log\beta_{11,\text{cal}}$	*ARE*, %	$\log\beta_{11,\text{cal}}$	*ARE*, %
1	5.3222	4.322	18.798	4.718	11.352	3.807	28.473	5.296	0.497
2	11.970	13.537	13.090	13.217	10.416	13.309	11.185	12.110	1.166
3	5.360	3.808	28.954	4.226	21.156	3.999	25.393	4.831	9.867
4	5.099	4.559	10.581	5.026	1.427	4.699	7.845	5.489	7.647
5	9.900	8.836	10.744	8.642	12.710	9.301	6.054	10.801	9.101

6	9.600	9.779	1.866	9.374	2.358	10.211	6.368	8.003	16.637
7	11.980	10.628	11.284	10.438	12.875	11.039	7.854	11.897	0.689
8	19.100	14.591	23.607	14.742	22.814	15.482	18.942	15.958	16.451
9	7.654	6.136	19.837	6.911	9.712	6.397	16.417	7.696	0.546
10	6.611	5.066	23.363	5.643	14.635	5.209	21.213	5.242	20.706
		MARE,%	16.212	*MARE*,%	11.945	*MARE*,%	14.975	*MARE*,%	8.331

The single factor ANOVA method was used to evaluate the difference between the experimental and predictive logβ_{11} values from the QSPR models. Consequently, the differences between the experimental and calculated values of stability constants logβ_{11} resulting from the QSPR models are insignificant ($F = 0.043509 < F_{0.05} = 2.866266$). Hence, the predictability of all QSPR models turns out to be in a good agreement with the experimental data.

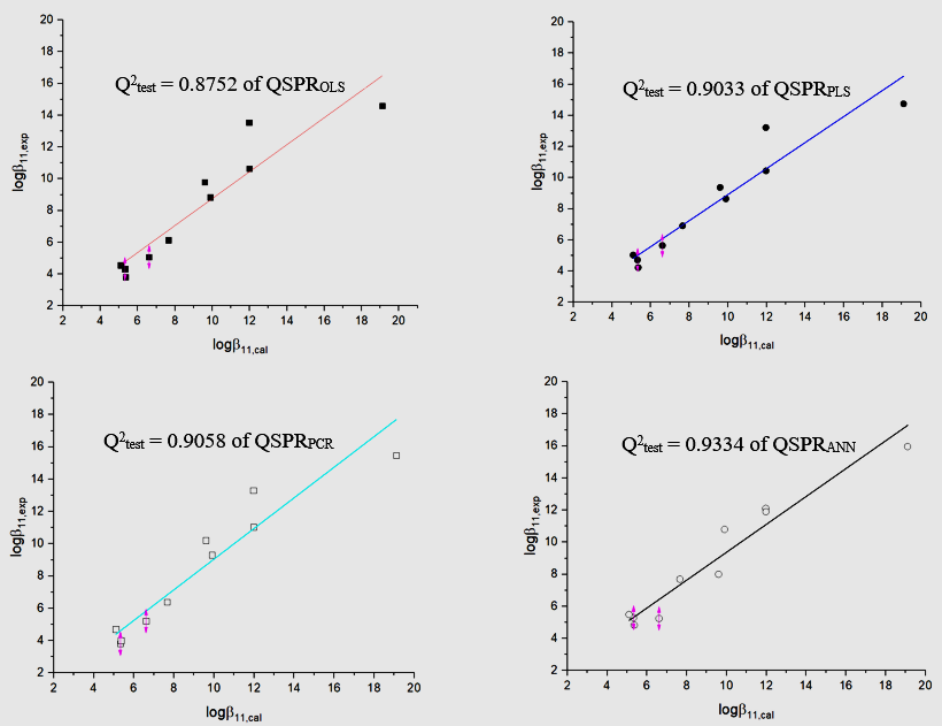


Figure 4. Correlation of experimental vs. predicted values of test set from the QSPR models

As Table 5b, the *MARE* values of models $QSPR_{OLS}$, $QSPR_{PCR}$, $QSPR_{PLS}$ and $QSPR_{ANN}$ I(9)-HL(12)-O(1) are 16.212%, 14.975%, 11.945% and 8.331%, respectively, indicating that model $QSPR_{ANN}$ displays highest predictability next model $QSPR_{PLS}$, $QSPR_{PCR}$ and $QSPR_{OLS}$. The logβ_{11} values resulting from model $QSPR_{ANN}$ are closer to the experimental values.

The results of analysis data in Table 5b are presented Fig. 4, it can show that the predictability of the models is very good. Whereby, neural model $QSPR_{ANN}$ exhibits a best fit and correlation between the predicted values and the experimental values, next $QSPR_{PLS}$ and $QSPR_{PCR}$ models and the last $QSPR_{OLS}$ models with Q^2_{test} of 0.9334, 0.9033, 0.9058 and 0.8752, respectively.

4. CONCLUSION

This work has successfully built the quantitative structure and property relationship (QSPR) incorporating ordinary least square regression ($QSPR_{OLS}$), partial least square ($QSPR_{PLS}$), primary component regression ($QSPR_{PCR}$) and artificial neural network ($QSPR_{ANN}$). The QSPR models were constructed by using the dataset of structural descriptors resulting from the semi-empirical quantum

calculation and molecular mechanics. The QSPR models were cross-validated carefully using the leave-one-out method upon statistical values R^2_{train}, Q^2_{LOO}, *MARE*, %, and one-way ANOVA method. The $QSPR_{ANN}$ model I(9)-HL(12)-O(1) turns out to be satisfactory for actual applicability. The results from this study are in the service of designing new thiosemicarbazone derivatives that are helpful to find new complexes in the many fields such as analytical chemistry, pharmacy, and environment.

ACKNOWLEDGMENT

The authors thank the financial support from Industrial University of Ho Chi Minh City for conducting this study (Project code: **184.HH09**).

REFERENCES

1. B. H. Patel, J. R. Shah, and R. P. Patel, Stability constants of complexes of 2-hydroxy-5-methylacetophenone-thiosemicarbazone with Cu(II), Ni(II), Co(II), Zn(II) and Mn(II), J. *Ind. Chem. Soc*, vol. 53, pp. 9-10, 1976.
2. R. B. Singh, B. S. Garg, and R. P. Singh, Analytical applications of thiosemicarbazones and semicarbazones: A review, *Talanta*, vol 25, no. 11–12, pp. 619–632, 1978.
3. M. Rajendran, A. Panneerselvam, V. Periasamy, and M. J. Grzegorz, Palladium(II) pyridoxal thiosemicarbazone complexes as efficient and recyclable catalyst for the synthesis of propargylamines by a three-component coupling reactions in ionic liquids, *Polyhedron*, vol. 119, pp. 300–306, 2016.
4. R. Ramachandran, G. Prakash, P. Vijayan, P. Viswanathamurthi, and J. G. Malecki, Synthesis of Heteroleptic Copper(I) Complexes with Phosphine-Functionalized Thiosemicarbazones: Efficient Catalyst for Regioselective N-Alkylation Reactions, *Inor. Chim. Acta*, vol. 464, pp. 88-93, 2017.
5. E. B. Seena, R. Bessy, M. R. Prathapachandra Kurup, and I. E. Suresh, A crystallographic study of 2-hydroxyacetophenone *N (4)* cyclohexyl thiosemicarbazone, *J. Chem. Crystallogr*, vol. 36, pp. 189, 2006.
6. K. Pyrzynska, Determination of molybdenum in environmental samples, *Anal. Chim. Acta*, vol. 590, pp. 40–48, 2007.
7. Ezhilarasi et al, Synthesis Characterization and Application of Salicylaldehyde Thiosemicarbazone and Its Metal Complexes, *Int. J. Res. Chem. Environ*, vol. 2, no. 4, pp. 130–148, 2012.
8. A. Nagajothi, A. Kiruthika, S. Chitra, and K. Parameswari, Fe(III) Complexes with Schiff base Ligands: Synthesis, Characterization, Antimicrobial Studies, *Res. J. chem. Sci*, vol. 3, no. 2, pp. 35–43, 2013.
9. R. Chaudhary and Shelly, Synthesis, Spectral and Pharmacological Study of Cu (II), Ni (II) and Co (II) Coordination Complexes, *Res. J. chem. Sci*, vol. 1, no. 5, pp. 1–5, 2011.
10. M. Ante and N. Raos, Estimation of Stability Constants of Mixed Copper(II) Chelates Using Valence Connectivity Index of the 3rd Order Derived from Two Molecular Graph Representations, *Acta. Chim. Slov*, vol. 56, pp. 373–378, 2009.
11. M. Ante and N. Raos, Estimation of Stability Constants of Copper(II) Bis-chelates by the Overlapping Spheres Method, *Croatica Chemica Acta*, vol. 79, no. 2, pp. 281–290, 2006.
12. S. Nikolic and N. Raos, Estimation of Stablity Constantsof Mixed Amino Acid Complexes with Copper(II) from Topological Indices, *Croatica Chemica Acta*, vol. 74, no. 3, pp. 621–631, 2001
13. E. J. Billo, *Excel For Scientists And Engineers: Numerical Methods*, John Wiley and Sons, Inc, Hoboken, NJ, USA, 2007.
14. D. Harvey, *Modern analytical Chemistry*, Mc.Graw Hill, Boston, Toronto, 2000.
15. Pham Van Tat, *Development of QSAR and QSPR*, Publisher of Natural sciences and Technique, Ha Noi, 2009.
16. K. Roy, S. Kar and R.N. Das, *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*, Academic Press, Amsterdam, 2015.
17. A. Speck-Planche, V. V. Kleandrova, L. Feng, M. Natália, and D. S. Cordeiro, Rational drug design for anti-cancer chemotherapy: multi-target QSAR models for the in silico discovery of anti-colorectal cancer agents, *Bioorg. Med. Chem*, vol. 20, no. 15, pp. 4848–4855, 2012.
18. A. Speck-Planche, V. V. Kleandrova, L. Feng, M. Natália, and D. S. Cordeiro, Chemoinformatics in anti-cancer chemotherapy: Multi-target QSAR model for the in silico discovery of anti-breast cancer agents, *Eur. J. Pharm. Sci*, vol. 47, no. 1, pp. 273–27, 2012.
19. R. Sabet, M. mohammadpour, A. Sadeghi, and A. Fassihi, QSAR study of isatin analogues as in vitro anti-cancer agent., *Eur. J. Med. Chem*, vol. 45, no. 3, pp. 1113–1118, 2010.
20. James J. P. Stewart, Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters, *J. Mol. Model*, vol. 19, pp. 1–32, 2013.

21. Bagheri et al, Simple yet accurate prediction of liquid molar volume via their molecular structure, *Fluid Phase Equilibria*, vol. 337, pp. 183-190, 2013.
22. D. N. Kenie and A. Satyanarayana, Protolitic Equilibria and Stability Constants of Mn (II) and Ni (II) Complexes of 3-formylpyridine Thiosemicarbazone in Sodium Dodecyl Sulphate (SDS)-Water Mixture, *J. Technol. Arts Sci. Res*, vol. 4, no. 1, pp. 74–79, 2015.
23. R. Biswas, D. Brahman, and B. Sinha, Thermodynamics of the complexation between salicylaldehyde thiosemicarbazone with Cu(II) ions in methanol–1,4-dioxane binary solutions, *J. Serb. Chem. Soc*, vol. 79, no. 5, pp. 565–578, 2014.
24. M. N. M. Milunovic, E. A. Enyedy, N. V. Nagy, T. Kiss, R. Trondl, M. A. Jakupec, B. K. Keppler, R. Krachler, G. Novitchi, and V. B. Arion, L- and D-Proline Thiosemicarbazone Conjugates: Coordination Behavior in Solution and the Effect of Copper(II) Coordination on Their Antiproliferative Activity, *Inorg. Chem*, vol. 51, pp. 9309−9321, 2012.
25. D. G. Krishna and C. K. Devi, Determination of cadmium (II) in presence of micellar medium using cinnamaldehyde thiosemicarbazone by spectrophotometry, *Int. J. Green Chem. Biopro*, vol. 5, no. 2, pp. 28-30, 2015.
26. D. G. Krishna and G. V. K. Mohan, A Facile Synthesis, Characterization of Cinnamaldehyde Thiosemicarbazone and Determination of Molybdenum (VI) by Spectrophotometry In Presence of Micellar Medium, *Ind. J. Appl. Res*, vol. 3, no. 8, pp. 7-8, 2013.
27. A. Gaál, G. Orgován, Z. Polgári, A. Réti, V. G. Mihucz, S. Bősze, N. Szoboszlai, and C. Streli, Complex forming competition and in-vitro toxicity studies on the applicability of di-2-pyridylketone-4,4,-dimethyl-3-thiosemicarbazone (Dp44mT) as a metal chelator, *J. Inorg. Biochem*, vol. 130, pp. 52–58, 2014.
28. J. R. Koduru and K. D. Lee, Evaluation of thiosemicarbazone derivative as chelating agent for the simultaneous removal and trace determination of Cd(II) and Pb(II) in food and water samples, *Food Chem*, vol. 150, pp. 1–8, 2014.
29. D. Rogolino, A. Cavazzoni, A. Gatti, M. Tegoni, G. Pelosi, V. Verdolino, C. Fumarola, D. Cretella, P.G. Petronini, and M. Carcelli, Anti-proliferative effects of copper(II) complexes with Hydroxyquinoline-Thiosemicarbazone ligands, *Eu. J. Med. Chem*, vol. 128, pp. 140-153, 2017.
30. M. A. Jiménez, M. D. Luque De Castro, and M. Valcárcel, Potentiometric Study of Silver(l)-Thiosemicarbazonates, *Microchem. J*, vol. 25, pp. 301-308, 1980.
31. M. A. Jiménez, M. D. Luque De Castro, and M. Valcárcel, Titration of Thiosemicarbazones with Cu(ll) and Vice Versa by Use of a Copper Selective Electrode in Acetone-Water Mixture: Determination of the Conditional Formation Constants of the Cupric Thiosemicarbazonates, *Microchem. J*, vol. 32, pp. 166-173, 1985.
32. T. Atalay, and E. Ozkan, Thermodynamic studies of some complexes of 4’-morpholinoacetophenone thiosemicarbazone, *Thermochimica Acta*, vol. 237, pp. 369-374, 1994.
33. B. S. Garg, and V. K. Jain., Determination of thermodynamic parameters and stability constants of complexes of biologically active o-vanillinthiosemicarbazone with bivalent metal ions, *Thermochimica Acta*, vol. 146, pp. 375-379, 1989.
34. B. S. Garg, S. Ghosh, V. K. Jain, and P. K. Singh, Evaluation of thermodynamic parameters of bivalent metal complexes of 2-hydroxyacetophenonethiosemicarbazone (2-HATS), *Thermochimica Acta*, vol. 157, pp. 365-368, 1990.
35. K. H. Reddy and N. B. L. Prasad, Spectrophotometric determination of copper (II) in edible oils and seed using novel oxime-thiosemicarbazones, *India J. Chem*, vol. 43A, pp. 111-114, 2004.
36. S. S. Sawhney and S. K. Chandel, Solution chemistry of Cu(II)-, Co(II)-, Ni(II)-, Mn(II)- and Zn(II)-p-aminobenzaldehyde thiosemicarbazone systems, *Thermochimica Acta*, vol. 71, pp. 209-214, 1983.
37. S. S. Sawhney and S. K. Chandel, Stability and thermodynamics of La(III)-, Pr(III)-, Nd(III)-, Gd(III)- and Eu(III)-p-nitrobenzaldehyde thiosemicarbazone systems, *Thermochimica Acta*, vol. 72, pp. 381-385, 1984.
38. S. S. Sawhney and R. M. Sati, pH-metric studies on Cd(II)-, Pb(II)-, AI(III)-, Cr(III)- AND Fe(III)-p-nitrobenzaldehyde thiosemicarbazone systems, *Thermochimica Acta*, vol. 66, pp. 351-355, 1983.
39. D. Admasu, D. N. Reddy, and K. N. Mekonnen, Spectrophotometric determination of Cu(II) in soil and vegetable samples collected from Abraha Atsbeha, Tigray, Ethiopia using heterocyclic thiosemicarbazone, *SpringerPlus*, vol.5, no. 1169, pp. 1-8, 2016.
40. *BIOVA Draw 2017 R2*, Version: 17.2.NET, Dassault Systèmes, France, 2016.
41. J. J. P. Stewart, *MOPAC2016*, Version: 17.240W, Stewart Computational Chemistry, USA, 2002.
42. *QSARIS 1.1*, Statistical Solutions Ltd, USA, 2001.
43. M. V. Gerven and S. Bohte, *Artificial Neural Networks as Models of Neural Information Processing*, Frontiers in Computational Neuroscience, 2018.
44. J. Gasteiger and J. Zupan, Neural Networks in Chemistry, *Chiw. Inr. Ed. EngI*, vol. 32, pp. 503–521, 1993.

45. R. Rojas, *Neural Networks*, Springer-Verlag, Berlin, 1996.
46. S. K. Jung and K. McDonald, Visual Gene Developer: a fully programmable bioinformatics software for synthetic gene optimization, *BMC Bioinformatics*, vol. 12, no. 1, pp. 340, 2011.
47. J. S. Kyu and L. S. Bok, In Situ Monitoring of Cell Concentration in a Photobioreactor Using Image Analysis: Comparison of Uniform Light Distribution Model and Artificial Neural Networks, *Biotechnology Progress*, vol. 22, no 5, pp. 1443-1450, 2006
48. D. D. Steppan, J. Werner, and P. R. Yeater, *Essential Regression and Experimental Design for Chemists and Engineers*, Germany, 1998.
49. E. J. Billo, *Excel for chemists*, Wiley-VCH, Weinheim, 1997.
50. *XLSTAT Version 2016.02.28451*, Addinsoft, USA, 2016.
51. N. S. R. Reddy and D. V. Reddy, Spectrophotometric determination of vanaditun(V) with salicylaldehyde thiosemicarbazone, *J. Indian. Inst. Sci*, vol. 64(B), pp. 133-136, 1983.
52. D. K. Singh, P.K. Jha, Raman Kant Jha, P. M. Mishra, A. Jha, S. K. Jha, and R. P. Bharti, Equilibrium Studies of Transition Metal Complexes with Tridentate Ligands Containing N, O, S as Donor Atoms, *Asian J. Chem*, vol. 21, no. 7, pp. 5055-5060, 2009.
53. D. N. Kenie and A. Satyanarayana, Solution Equilibrium Study of the Complexation of Co(II) and Zn(II) with Nicotinaldehyde Thiosemicarbazone, *Sci. Technol. Arts Res. J*, vol. 4, no. 3, pp. 145-149, 2015.
54. V. Veeranna, V. S. Rao, V. V. Laxmi, and T. R Varalakshmi, Simultaneous Second Order Derivative Spectrophotometric Determination of Cadmium and Cobalt using Furfuraldehyde Thiosemicarbazone (FFTSC), *Res. J. Pharm. and Tech*, vol. 6, no. 5, pp. 577-584, 2013.
55. A. T. A. El-Karim and A. A. El-Sherif, Potentiometric, equilibrium studies and thermodynamics of novel thiosemicarbazones and their bivalent transition metal(II) complexes, *J. Mol. Liq*, vol. 219, pp. 914–922, 2016.
56. K. Sarkar and B. S. Garg, Determination of thermodynamic parameters and stability constants of the complexes of p-MITSC with transition metal ions, *Thermochimica Acta*, vol. 113, pp. 7-14, 1987.

MÔ HÌNH HÓA QSPR HẰNG SỐ BỀN CỦA PHỨC GIỮA ION KIM LOẠI VÀ THIOSEMICARBAZONE SỬ DỤNG CÁC PHƯƠNG PHÁP HỒI QUY ĐA BIẾN VÀ MẠNG THẦN KINH NHÂN TẠO

Tóm tắt. Trong nghiên cứu này, các mô hình quan hệ định lượng giữa cấu trúc-tính chất (QSPR) của các phức chất giữa ion kim loại và thiosemicarbazone được xây dựng bằng các phương pháp hồi quy đa biến và mạng thần kinh nhân tạo. Bộ mô tả phân tử, tham số hóa lý và các mô tả lượng tử của phức chất được tính toán từ cấu trúc phân tử và lượng tử theo phương pháp bán thực nghiệm PM7 và PM7/spakle. Mô hình $QSPR_{OLS}$ tốt nhất được xây dựng dựa trên phương pháp hồi quy đa biến thường bao gồm 9 biến là C5, xp9, electric energy, cosmo volume, N4, SsssN, cosmo area, xp10 và core-core repulsion. Các mô hình $QSPR_{PLS}$ và $QSPR_{PCR}$ được phát triển tương ứng theo phương pháp bình phương tối thiểu riêng phần và phương pháp hồi quy thành phần chính từ 9 biến của mô hình $QSPR_{OLS}$. Chất lượng các mô hình được đánh giá qua các giá trị thống kê. Mô hình $QSPR_{OLS}$: R^2_{train} = 0,944; Q^2_{LOO} = 0,903; *MSE* = 1,035. Mô hình $QSPR_{PLS}$: R^2_{train} = 0,929; R^2_{CV} = 0,938; *MSE* = 1,115. Mô hình $QSPR_{PCR}$: R^2_{train} = 0,934, R^2_{CV} = 0,9485; *MSE* = 1,147. Mô hình mạng thần kinh $QSPR_{ANN}$ với cấu trúc I(9)-HL(12)-O(1) cũng được xây dựng từ các biến đầu vào của mô hình $QSPR_{OLS}$ với các giá trị thống kê R^2_{train} = 0,9723 và R^2_{CV} = 0,9731. Các mô hình QSPR này cũng được đánh giá ngoại và cho khả năng dự đoán phù hợp với thực nghiệm.

***Từ khóa*:** $QSPR_{OLS}$, $QSPR_{PLS}$, $QSPR_{PCR}$, $QSPR_{ANN}$, hằng số bền logβ_{11}, thiosemicarbazone.

Ngày nhận bài: 27/06/2018
Ngày chấp nhận đăng: 02/01/2019