Computational toxicity evaluation of organophosphorus pesticides

[1]Alina-Maria Petrescu, [1,2]Gheorghe Ilia

[1] West University of Timisoara, Faculty of Chemistry, Biology, Geography, Dept. of Biology-Chemistry, Timisoara, Romania; [2] Institute of Chemistry Timisoara of Romanian Academy, Timisoara, Romania. Corresponding author: A. M. Petrescu, alinapetrescu79@gmail.com

Abstract. Pesticides are broad-spectrum compounds used against insect, arthropod, other pest and fungi that are highly toxic to humans by different routes of exposures, such as dermal absorption, ingestion or inhalation. Fungicides are a group of pesticides used in agriculture to control plant diseases caused by fungal attack. As a result of their relative persistence, fungicides residues present in water, food, animals etc., represent a risk for human health. Due to their widespread presence and toxicity, an environmental monitoring is required to protect the public from possible organic toxins released into the air, soil, and water. All these aspect have originated the need to develop toxicity methodologies, structured-toxicity relationships and biological activity.
Key Words: pesticide, insect, agriculture, phosphates.

Background. Organophosphorus compounds play a central role in the living organism; it can be mentioned photosynthesis, metabolism and involvement in coenzyme systems; and are widely used in the agriculture around the world as pesticides, insecticides, acaricides, and fungicides (Evidente et al 2014). Thiophosphates (or phosphorothioates) are a very toxic, dangerous class of organophosphorus including a family of compounds and anions with the general chemical formula $PS_{4-x}O_x^{3-}$ (x = 0, 1, 2, or 3) (De Costa & Bezerra 2009).

Thiophosphates are a very dangerous class of organophosphorus compounds, because they are very toxic and have such a wide range of properties and applications that they have stimulated an enormous amount of study for several decades (Upadhyay et al 2011). An efficient synthesis of antiviral dendrite prodrug candidates, water-soluble, polyanionic conjugates of 1st and 2nd generation thiophosphate dendrimers with acyclovir. Acyclovir was successfully grafted on the surface of thiophosphate dendrimers via thio- and phosphodiester linkages, providing water-soluble prod rug candidates (Kwak et al 2012).

The present study attempts to evaluate the toxicity of organophosphorus fungicide using a modern computational structural-toxicity method, respectively Hyper Chem. Professional by QSTR system, semi-empirical optimization, using Polak-Ribiere algorithm, with RMS gradient 0.1 kcal/(Åmol), in vacuum. Computational measured toxicity values, LD_{50} (mg kg^{-1})= -A=log(1/LD_{50}) were considered as dependent variables and were related to structural features obtained by molecular and quantum mechanics calculations by MLR approaches (Barnes & Denz 1951).

The study was performed with a series of 20 thiophosphates. The obtained models showed that thiophosphates toxicity and antifungal activity was influenced by molecular geometry, as well as by compound polarity (Greenwood & Earnshaw 1984).

The aim of this study was to evaluate and correlate the toxicity of organophosphorus fungicide with molecular descriptors using QSTR-quantitative structure toxicity relationships model.

Material and Method. The study was performed on a series of 20 thiophosphates (Table 1).

Table 1

Thiophosphates used in this study

CASs Number	*IUPAC NAME*	*CANONICAL SMILES*
001_ CID21803	Dimethoxy-sulfanylidene-(3,5,6-trichloropyridin-2-yl)oxy-$l^{5}-phosphane	COP(=S)(OC)OC1=NC(=C(C=C1Cl)Cl)Cl
002_ CID19657	2-methoxy-2-sulfanylidene-4H-1,3,2$l^{5}-benzodioxaphosphinine	COP1(=S)OCC2=CC=CC=C2O1
003_ CID 91664	2,6-dichloro-4-methylphenoxy)-dimethoxy-sulfanylidene-$l^{5}-phosphane	CC1=CC(=C(C(=C1)Cl)OP(=S)(OC)OC)Cl
004_ CID2730	diethoxy-sulfanylidene-(3,5,6-trichloropyridin-2-yl)oxy-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=NC(=C(C=C1Cl)Cl)Cl
005_ CID16213399	Bis(1,1,2,2,2-pentadeuterioethoxy)-sulfanylidene-(3,5,6-trichloropyridin-2-yl)oxy-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=NC(=C(C=C1Cl)Cl)Cl
006_ CID165070	dimethoxy-(4-methylsulfanylphenoxy)-sulfanylidene-$l^{5}-phosphane	COP(=S)(OC)OC1=CC=C(C=C1)SC
007_ CID991	Diethoxy-(4-nitrophenoxy)-sulfanylidene-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=CC=C(C=C1)[N+](=O)[O-]
008_ CID3346	Dimethoxy-(3-methyl-4-methylsulfanylphenoxy)-sulfanylidene-$l^{5}-phosphane	CC1=C(C=CC(=C1)OP(=S)(OC)OC)SC
009_ CID19577	Dimethoxy-(3-methyl-4-methylsulfinylphenoxy)-sulfanylidene-$l^{5}-phosphane	CC1=C(C=CC(=C1)OP(=S)(OC)OC)S(=O)C
010_ CID197971	Dimethoxy-[(5-phenyl-1,2-oxazol-3-yl)oxy]-sulfanylidene-$l^{5}-phosphane	COP(=S)(OC)OC1=NOC(=C1)C2=CC=CC=C2
011_ CID12901	2-(dimethoxyphosphinothioylsulfanylmethyl)isoindole-1,3-dione	COP(=S)(OC)SCN1C(=O)C2=CC=CC=C2C1=O
012_ CID77323	2-(dimethoxyphosphorylsulfanylmethyl)isoindole-1,3-dione	COP(=O)(OC)SCN1C(=O)C2=CC=CC=C2C1=O
013_ CID25834	Diethoxyphosphorylsulfanylmethylbenzene	CCOP(=O)(OCC)SCC1=CC=CC=C1
014_ CID8292	Diethoxy-(4-methylsulfinylphenoxy)-sulfanylidene-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=CC=C(C=C1)S(=O)C
015_ CID74390	Diethoxy-(3-methyl-4-methylsulfanylphenoxy)-sulfanylidene-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=CC(=C(C=C1)SC)C
016_ CID18609	Dipropoxyphosphinothioyloxy-dipropoxy-sulfanylidene-$l^{5}-phosphane	CCCOP(=S)(OCCC)OP(=S)(OCCC)OCCC
017_ CID 29307	Diethoxy-[(5-phenyl-1,2-oxazol-3-yl)oxy]-sulfanylidene-$l^{5}-phosphane	CCOP(=S)(OCC)OC1=NOC(=C1)C2=CC=CC=C2
018_ CID33294	Di(propan-2-yloxy)phosphorylsulfanylmethylbenzene	CC(C)OP(=O)(OC(C)C)SCC1=CC=CC=C1
019_ CID16421	Ethoxy-(4-nitrophenoxy)-phenyl-sulfanylidene-$l^{5}-phosphane	CCOP(=S)(C1=CC=CC=C1)OC2=CC=C(C=C2)[N+](=O)[O-]
020_ CID28292	[Ethoxy(phenylsulfanyl)phosphoryl]sulfanylbenzene	CCOP(=O)(SC1=CC=CC=C1)SC2=CC=CC=C2

QSTR structure modeling. Al parametres were computed by the Hyper Chem Professional, software, version 7.0 using semi-empirical methods, MM^{+} (molecular mechanic) method for generating initial structures. To obtain minimum energy structures, second geometry optimization was performed with and PM_3 semi-empirical calculations. Geometrically optimized structures were used for the calculations of molecular descriptors -Hansch, and QSAR parameters.

Toxicity data. Legislation has been implemented since the early 1970s to ensure that harmful effects of environmental toxicants are minimized for all species. For this problem, this study also, tried to provide a better life with reducing the toxicity, with don't use the pollution substance. The more we can know about the toxicity, the more is much simple to have protection.

Analysis of data proposed for the model. PLS method (Partial least squares regression) is a statistical method related with some relation to principal components regression; instead of finding hyper planes of minimum variance between the response and independent variables, it finds a linear regression model by projecting the predicted variables and the observable variables to a new space. Because both the *X* and *Y* data are projected to new spaces, the PLS family of methods are known as bilinear factor models, and indicate the correlation between values of experimental biological activity - toxicity, parameters Hansch and QSAR parameters, like a simple equation with one identified parameter (ECOSAR, software version v.1.11).

Model validity. Correlation coefficient Pearson (R2, or R^2), and (Q2,or Q^2) is the percent of variation of the training set, between calculated and experimental biological activity-toxicity, are used to evaluate the accuracy of the proposed QSTR model by PLS model.

For a PLS model, UMETRICS software: SIMCA V.14.0., 1998-2015, was used, which displays the following coefficient plot. The coefficients (B) refer to the PLS model being rewritten as a regression model:

$$Y = Y_{avg} + XB + F \quad (1)$$

The constant Y_{avg} represent the average of Y, where Y is not displayed in the coefficient plot, but is found in the list of the coefficients. SIMCA is versatile multivariable data analysis software appreciated for its accessibility and ease of use. With SIMCA 14 Umetrics further complements the toolbox for extensive data analysis as well as reducing work effort to analyze data.

Results and Discussion. The values of toxicity data, LD_{50} (mg kg^{-1}), were collected from literature, and were converted: - $A=\log(1/LD_{50})$ (Table 2).

Table 2

Toxicity values from literature

LD_{50} (mg kg^{-1})	$-A=\log(1/LD_{50})$
2032	3.30
88	1.94
3500	3.54
60	1.77
0	0
10	1
5	0.69
180	2.25
220	2.34
593	2.77
26	1.41
50	1.69
230	2.36
2	0.3
30	1.47
8	0.9
40	1.6
435	2.63
12.2	1.08
1443	3.159

The values of the toxicity from literature were compared with the values of the predicted toxicity (Table 3). This study used ECOSAR v.1.11, software, for calculating the predicted toxicity to fish, for the compounds proposed in this study (HyperChem Professional software, version 7.0).

Table 3

Toxicity values from calculated data

LD_{50} observed	LD_{50} predicted
3.3	0.205
1.94	0.753
3.54	0.080
1.77	0.02
0	0.02
1	0.377
0.69	0.317
2.25	0.189
2.34	3.774
2.77	352.08
1.41	9.725
1.69	18.40
2.36	3.132
0.3	2.190
1.47	0.055
0.9	0.02
1.6	0.271
2.63	1.768
1.08	0.128
3.159	1.839

LD_{50} observed = - A=log(1/LD_{50}), biological activity-toxicity, measured mg kg^{-1} to the mice, to monothiophosphates proposed in this study. LD50 predicted by ECOSAR v.1.11, software = LC50, measured mg L^{-1} (ppm) to the fish about 96-hr to Esters, monothiophosphates proposed in this study.

All the structures were optimized with HyperChem Professional software , using the Semi Empirical method, MM+, and selected PM_3, when Hansch and QSAR parameters are obtained (Table 4).

Table 4

The obtained Hansch and QSAR Parameters

Entry	LD_{50} (mg kg^{-1})	Hansch Parameters				HOMO/ LUMO (kcal mol^{-1})	QSAR Parameters			
		TE (a.u)	logP	POL ($Å_3$)	HE (kcal mol^{-1})		SAA ($Å_2$)	VOL ($Å_3$)	MASS amu	REF kcal mol^1
1	2032	-132.4728	4.2	25.21	-128.065	-2209.625	472.94	741.35	322.53	69.5
2	88	-83.64160	2.61	19.36	-127.943	-2284.627	375.74	581.3	216.19	54.42
3	3500	-122.5912	4.39	25.82	-146.439	-2636.197	455.35	748.85	301.12	70.52
4	60	-143.9265	4.88	28.88	-141.821	-2773.569	523.66	855.22	350.58	79
5	0	-143.9279	4.88	28.88	-142.678	-2774.426	504.1	830.79	350.58	79
6	10	-103.2760	2.98	24.97	-132.509	-2734.891	460.19	736.75	264.29	68.77
7	5	-122.6380	-0.34	25.64	-139.764	-3138.471	496.23	801.21	291.26	71.68
8	180	-102.6010	3.44	26.8	-135.004	-3012.480	477.24	778.69	278.32	73.81
9	220	-113.3703	2.31	24.37	-160.853	-3097.888	487.71	796.11	294.32	75.19
10	593	-115.0526	2.86	26.24	-90.6436	-3088.863	484.44	783.54	285.25	73.11
11	26	-121.9598	2.73	28.75	-128.039	-3192.658	506.17	827.48	317.31	78.59
12	50	-125.9607	2.08	25.59	-203.64	-3261.419	496.16	801.85	301.25	70.6
13	230	-101.2677	3.36	24.84	-152.207	-3238.378	505.22	797.83	260.29	68.65
14	2	-118.8655	2.52	26.2	-166.084	-3378.213	539.38	872.08	308.35	79.64
15	30	-113.5892	4.13	30.47	-144.076	-3571.741	542.79	890.65	306.37	83.31
16	8	-144.8677	5.54	34.21	-329.818	-4421.088	644.03	1089.7	378.42	96.52
17	40	-126.0407	3.54	29.91	-99.6502	-3648.058	549.44	896.71	313.31	82.61
18	435	-112.2541	4.19	28.51	-160.175	-3796.533	520.65	861.43	288.34	77.49
19	12.2	-129.2011	0.42	30.99	-54.0759	-3669.169	530.09	877.37	323.3	85.13
20	143	-109.1594	4.95	31.36	-38.7182	-3540.196	517.74	862.08	310.37	85.29

Hansch Parameters: TE-total energy (a.u.), Log P-coefficient partition octanol-water, POL-polarizability, He-hydration energy, HOMO-the highest orbital energy, LUMO-the smallest orbital energy; QSAR parameters: SAA-molar surface grid, VOL-molar volume, MASS-molar mass, REF-refractivity (HyperChem Professional software, version 7.0).

Description of the dataset. The project was started with an dataset who was created 12.03.2015 22:26:08 having 19 variables, 18 beeng X variables and 1 Y variable - LD_{50}(mg kg^{-1}), and 20 observations. Missing values: 0 (0%). All variables are scaled to Unit Variance.

PLS model Summary Plot. Then follows the model type, N (number of observations) and K (total number of X variables and Y variables), and the number of expanded terms in the X matrix. The columns for the PLS components are: Component index; R2X - Fraction of X variation modeled in the component; R2X(cum) - Cumulative R2X up to the specified component; Eigenvalue - The number of X variables times R2X; R2Y - Fraction of Y variation modeled in the component; R2Y(cum) – Cumulative; R2Y up to the specified component; Q2 - Overall cross-validated R2 for the component; Limit - Critical value of Q2 under which the component is insignificant according to CV rule 1; Q2(cum) - Cumulative Q2 up to the specified component. Note that unlike R2(cum), Q2(cum) is not additive; Significance - CV insignificant (NS) or significant according to rule R_z; Iterations - The number of NIPALS iterations until convergence.

Variables. From the PLS model, the software computes PLS regression coefficients for each Y variable, and the coefficients for the first Y variable are displayed above. By default, the regression coefficients relate to the centered and scaled data and are computed from all extracted components. These coefficients are not independent (as in an MLR model), as the X variables are not independent. The coefficients are displayed with a confidence interval computed by jack-knifing, like in the Figure 1.

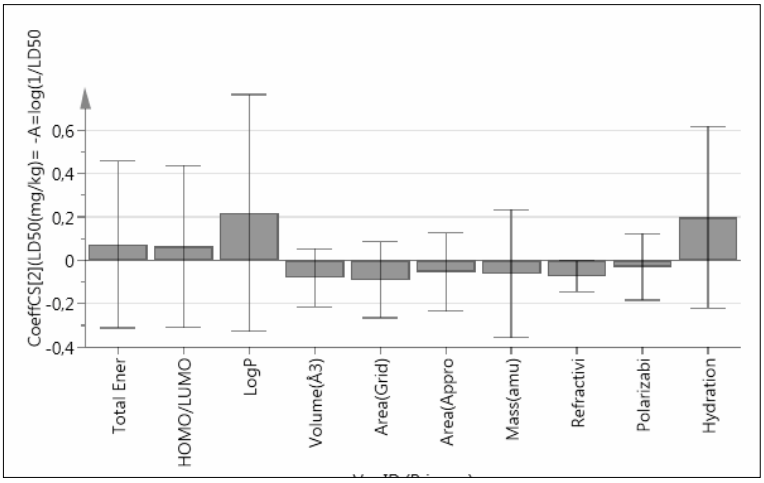

Figure 1.Corelation of variables.

Prediction model

Observed vs Predicted. The Observed vs. Predicted plot displays the observed values vs. the fitted values for the first Y-variable using the PLS model. By default, regression coefficients related to scaled and centered X-variables are displayed. This scaling of the data makes the coefficients comparable. The size of the coefficient represents the change in the Y-variable when the X-variable varies from 0 to 1, in coded units, (one standard deviation when the data are scaled to unit variance UV) while the other variables are kept at their averages, and shows in the Figure 2. Thus, these coefficients express how strongly Y is correlated to the systematic part of each of the X-variables. Figure 2 illustrates how many variables are correlated in dataset, with fraction of X variation modeled in the component: R2x[1]=0.626, for the first plot, and R2x[2]=0.202, for the second plot. Eigenvalue =3.52; Iterations=30.

Result is: R2(cum)=R2X[1]+R2X[2]=0.626+0.202=0.828.

The t1 vs t2 plot of the X scores can be interpreted as a window into the X space. This plot shows how the modeled observations are situated in the X space. Thus, it shows the possible presence of outliers, groups, time trends, and other patterns in the data.

Observations that lie close to each other are more similar than observations that lie relatively distant from each other. Two observations are more distanced in the plot, the 16 and 2 observations number, like in the Figure 2, that's means not so well validity. For this reason, 16 and 2 observations are excluded, and next results are obtained: R2X[1] =1 and a probability of validity this model 0.95=95%, as in the Figure 3, with RMSEE (Root Mean Square Error Estimation) is 0.845906, RMSEcv (Root Mean Square Error of Cross-Validation) is 1.00055, as in the Figure 4.

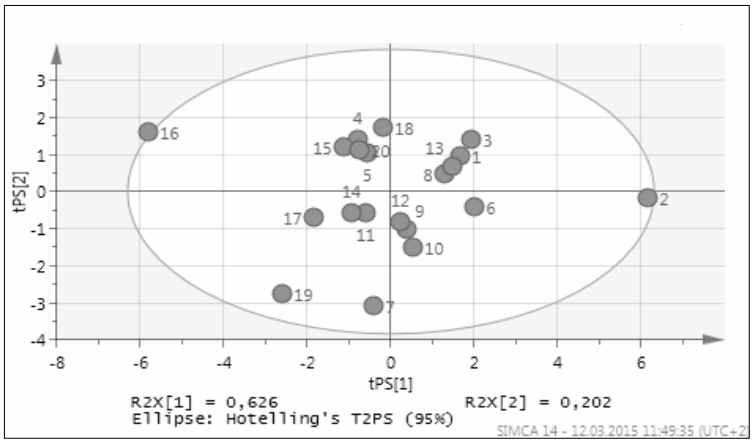

Figure 2. Predicted model of variables.

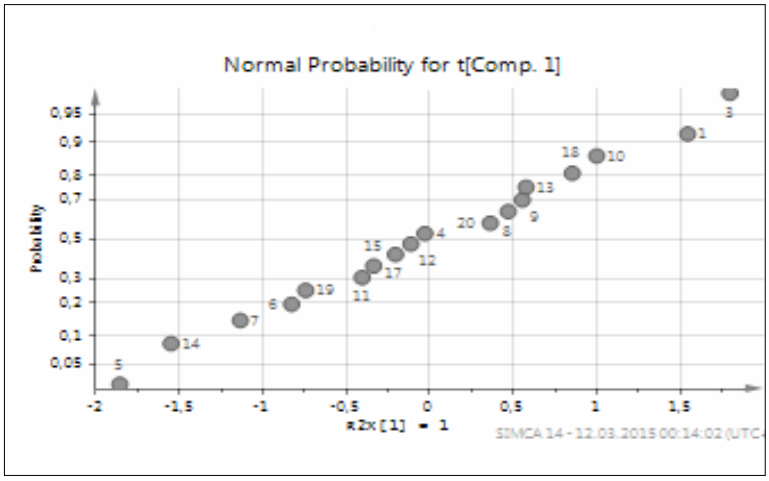

Figure 3. Validation of the proposed model.

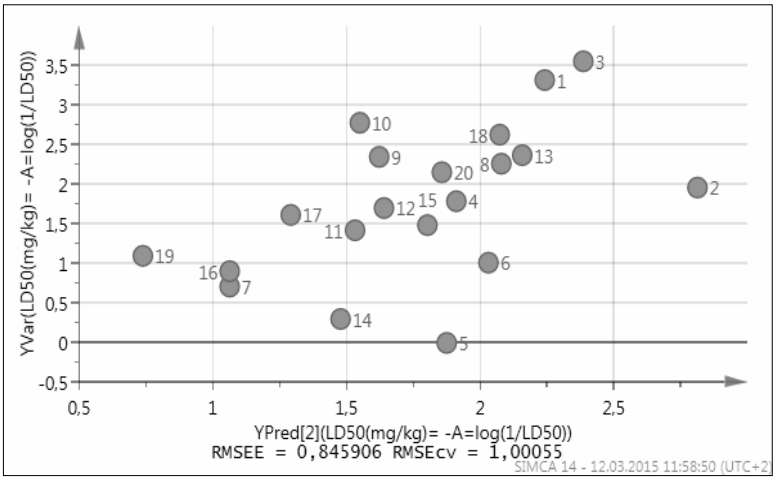

Figure 4. Root Mean Square Error of Cross-Validation.

Conclusions. The training set compounds that have a high toxicity; Class III contains functional groups associated with enhanced toxicity. Substances with chemical structures that permit no strong initial presumption of safety or may even suggest significant toxicity or have reactive functional groups are: Aromatic-H, Benzene, $-CH_2-$ [linear], Methyl [$-CH_3$], Phosphate ester (P=O type), Unsubstituted phenyl group (C_6H_5-).

Acknowledgements. Thanks to the project POSDRU: Invest in people. European Social Fund Operational Program Human Resources Development 2007-2013 Priority Axis 1 "Education and training in support of growth and development of knowledge based society "Key Area of Intervention 1.5." Support for PhD students and postdoctoral researchers'. Title project: Doctoral and Postdoctoral Programs - support for increasing the competitiveness of research in sciences. Contract Code: HRD/159/1.5/S/137750. Client: "Alexandru Ioan Cuza" University Iasi and West University Timisoara. Thank also to Prof. Mircea Mracec for giving access to HyperChem software.

References

Barnes J. M., Denz F. A., 1951 The chronic toxicity of p-nitrophenyl diethyl thiophosphate (E. 605). A long–term feeding experiment with rats. Journal of Hygiene 49(4):430-441.

De Costa P., Bezerra P., 2009 Fungicides: chemistry, environmental impact and health effects. Nova Science Publishers, New York, ISBN: 978-1-60692-631-4.

Evidente A., Kornienko A., Cimmino A., Andolfi A., Lefranc F., Mathieu V., Kiss R., 2014 Fungal metabolites with anticancer activity. Natural Product Reports 31(5):617-627.

Greenwood N. N., Earnshaw A., 1984 Chemistry of the elements. Pergamon Press, Oxford, 1542 pp.

Kwak Y., Kim J. E., Lee I. J., Kim J. G., Rhee I. K., Shin J. H., 2012 Biodegradation of tolclofos-methyl by extracellular secreted organophosphorus hydrolase in recombinant *Escherichia coli*. Journal of the Korean Society for Applied Biological Chemistry 55(3):337-384.

Upadhyay S., Sharma M. K., Rao V. K., Bhattacharya B. K., Sharda D., Vijayaraghavan R., 2011 Organophosphorous compounds - toxicity and detection approach. In: Pesticides - strategies for pesticides analysis. Stoytcheva M. (ed), InTech, Rijeka, pp. 315-332.

*** HyperChem Professional software, version 7.0, http://www.hyper.com/Products/tabid/354. Accessed: 19.02.2015.

*** ECOSAR, software version v.1.11, software. Accesed: 15.02.2015.

*** UMETRICS software: SIMCA V.14.0., 1998-2015. Accessed: 12.03.2015.

Received: 21 July 2015. Accepted: 20 September 2015. Published online: 31 October 2015.
Authors:
Alina-Maria Petrescu, West University of Timisoara, Faculty of Chemistry, Biology, Geography, Dept. of Biology-Chemistry, 16 Pestalozzi Street, 300115 Timisoara, Romania, e-mail: alinapetrescu79@gmail.com
Gheorghe Ilia, West University of Timisoara, Faculty of Chemistry, Biology, Geography, Dept. of Biology-Chemistry, 16 Pestalozzi Street, 300115 Timisoara, Romania; Institute of Chemistry Timisoara of Romanian Academy, 24 Mihai Viteazu Bvd., 300223, Timisoara, Romania

How to cite this article:
Petrescu A. M., Ilia G., 2015 Computational toxicity evaluation of organophosphorus pesticides. Ecoterra 12(3):90-96.