Short Research Article

DETERMINATION OF DESCRIPTORS WHICH INFLUENCE THE TOXICITY OF ORGANOCHLORINE COMPOUNDS USING QSAR METHOD

SUMMARY

Organochlorine Pesticides (OCP) are organic compounds obtained by the chlorination of various unsaturated hydrocarbons. They are very toxic and therefore belong to the family of persistent organic pollutants. If formerly these pesticides were used ~~to improve the productivity and~~ to fight against certain vectors of diseases and thus improve the productivity of the host, today they are considered as "enemy" of the environment. To understand the origin of the toxicity of organochlorine compounds, we used 73 molecules (test set: 50 and validation: 23) containing at least one chlorine atom and for which the toxicity ($LogLC_{50}$) against *Poecilia reticulata* is known to establish QSAR models. Firstly, we used principal component analysis (PCA) to identify the best descriptors. Then, the different models ~~are~~ were established using the method of multiple linear regression (MLR). Models established with quantum and physicochemical descriptors only showed satisfactory results. But the best model ~~is~~ was determined with the combination of both quantum and physicochemical descriptors. The criteria of this model are as follows:

$R^2 = 0.939$; $R^2_{ajusted} = 0.932$; $P_{value} < 0.0001$; $\alpha = 0.05$

$R^2_{CV} = 0.935$; $R^2 - R^2_{CV} = 0.004$; $MCE = 0.073$; F =134.701

These criteria show that the toxicity of organochlorine compounds is well described by the combination of quantum and physicochemical descriptors namely lipophilia (LogP), polarizability (pol), entropy (S), zero-point energy (ZPE) and the number of chlorine atoms (NCl).

Key words: organochlorine compounds, toxicity, QSAR, quantum descriptors and physicochemical descriptors

1- INTRODUCTION

Organochlorine pesticides use began in the 1940s with the advent of dichlorodiphenyl- trichloroethane (DDT) which was synthesized for the first time by Othmar Zeidler in 1873 [1] and whose insecticidal properties were discovered by Paul Müller in 1939 [2]. These pesticides have contributed to improved and increased agricultural yields and have led to progress in the control of food resources [3]. The use of organochlorine pesticides has not been limited to increasing agricultural yields, ~~,~~ it has also spread to other sectors. Indeed, organochlorine pesticides have long time been used to fight against certain diseases vectors [4-7]. DDT is one of the insecticides recommended by WHO for indoor residual spraying for malaria control [8].

After two decades of intense use, research [7, 9] ~~has~~ begun to show the dangerousness of these chemical compounds to the environment.

Over time, several studies [10-13] ~~will~~ proved their presence in all ecosystems and their effects on elements which live there. The presence of organochlorine pesticides in all ecosystems is assumed to be due to both their persistence their volatility.

Formatted: Subscript

Comment [U1]: Write in full the first time and abbreviate thereafter

The persistence of organochlorine compounds in the environment is largely due to the stability of the carbon-chlorine bond that is resistant to degradation. But their presence throughout the food chain is the consequence of their lipo-solubility. This property gives them the ability to cross the phospholipid structure of biological membranes thereby reaching the adipose tissue in which they accumulate. The rate of these compounds increases in the body of the species that live in these environments ~~when we~~with progress in the food chain in which humans are at the top [14-16].

As for volatility, it is generally due to the organochlorines' relatively high vapor pressure. This property allows them to travel great distances. Thus, according to some studies [17-19], organochlorine pesticides have been found in environments in which these pesticides have never been used and sometimes even far from the places where they were used. In view of all the above-mentioned, the use of organochlorine pesticides ~~will~~ was therefore r~~be r~~egressed from the 1970s and even several ~~will be~~pesticides containing organochlorines were banned in some countries. But it was until the year 2001 that the first Conference on Persistent Organic Pollutants (POPs) was held. Indeed, in Sweden in 2001, the Stockholm Convention marked the first convention on POPs. This convention has been signed by 151countries. This high number of signatories of the convention shows its importance. Since then, it was decided to reduce or even to eliminate the production and the exploitation of persistent organic pollutants. Among the 12 organic pollutants covered by this convention, 9 represent organochlorine pesticides.

The main objective of this study is to perform a QSAR study of organochlorine compounds to determine the descriptors that influence their toxicity.

Crum-Brown and Frazer [20] were considered to be the precursors of the QSAR methodology. Indeed, already in 1868, they postulated that the biological activity of a molecule is a function of its chemical constitution. In 1893, Richet [21] discovered that the toxicity of organic compounds is inversely proportional to their solubility in water. But the era of QSAR really begins in the 1960s with, on the one hand the publication of Hansch and Fujita [22] and on the other hand, the publication of Free and Wilson [23]. And from then, the number of publications containing the word "QSAR" continues to increase due its ability to predict the properties of chemical compounds.

2- MATERIALS AND METHODS

2.1- Materials

a- Selection of data set

We investigated about 73 molecules as displayed in Table 2. These molecules were taken from Alan Katritzky et al.'s article [24]. For each molecule, $LogLC_{50}$ is calculated where LC_{50} stands for the concentration that causes the death of 50% of the population of the test organism. Besides, the most toxic compound is assumed to display the smallest value of LogLC50. Furthermore, it should be noted that lethal dose (LD) and lethal concentration (LC) are identical. But the difference is related to the mode of penetration of the substance into the body. When the administered substance enters the body by inhalation, the notion of dose is replaced by that of concentration. Thus, the lethal dose 50 becomes the lethal concentration 50 [25]. According to Hodge and Sterner [26] chemicals products can be classified into 6 groups according to their toxicities.

Table 1: Class of toxicity according to the scale Hodge and Sterner [27]

Index or class of toxicity	Term commonly used	Toxicological parameter (DL_{50})
1	Extremely toxic	$DL_{50} \leq 1$ mg/kg
2	Highly toxic	1 mg/kg $\leq DL_{50} \leq$ 50 mg/kg
3	Moderately toxic	50 mg/kg $\leq DL_{50} \leq$ 500mg/kg
4	Slightly toxic	500 mg/kg $\leq DL_{50} \leq$ 5000mg/kg
5	almost not toxic	5000 mg/kg $\leq DL_{50} \leq$ 15000 mg/kg
6	relatively harmless	$DL_{50} \geq$ 15000 mg/kg

The set of all 73 molecules that were used in this study are ~~registered~~ shown in Table 2.

Table 2 *:* Names and LogLC_{50} of 73 organochlorine molecules

N°	Compounds	Log LC_{50}			
Test set					
1	3-chloroaniline	2.02	38	2-chlorophenol	1.94
2	3-chlorophenol	1.70	39	3,4-dichloroaniline	1.61
3	4,5-dichloro-2-methoxyphenol	1.40	40	3,4-dichlorotoluène	1.40
4	4-chloro-3,5-dimethylphenol	1.34	41	3,5-dichloroaniline	1.38
5	4-chloroaniline	2.33	42	3-chlorotoluène	2.16
6	4-chlorophenol	1.82	43	chlorobenzene	2.23
7	4-chlorotoluène	1.67	44	lindane	-0.69
8	Chloroform	2.93	45	trichloroethene	2.58
9	dichloromethane	3.54	46	1,2,3-trichlorobenzene	1.11
10	hexachlorobutadiene	-0.20	47	1,3-dichlorobenzene	1.72
11	pentachlorobenzene	-0.15	48	2,4-dichlorophenol	1.41
12	pentachloroethane	1.74	49	2,5-dichlorophenol	1.42
13	pentachlorophenol	0.22	50	3,4,5,6-tetrachloro-2-hydroxyphenol	1.00
14	tetrachloroethene	1.98	**Validation set**		
15	tetrachloromethane	2.64	51	1,1-dichloroéthane	3.31
16	α,α-dichloro-m-xylyene	-0.16	52	1,2,3,4-tetrachlorobenzene	0.65
17	1,1,1-trichloroéthane	3.00	53	1,2,3,5-tetrachlorobenzene	0.57
18	1,1,2-trichloroéthane	2.82	54	1,2,4,5-tetrachlorobenzene	0.15
19	1,2-dichlorobenzene	1.60	55	1,2,4-trichlorobenzene	1.17
20	2,3,4,5-tetrachloroaniline	0.19	56	1,2-dichloroethane	3.06
21	2,3,5-trichlorophenol	1.08	57	1,2-dichloropropane	3.01
22	2,4-dichloroacetophenone	1.80	58	1,3-dichloropropane	2.87
23	2,4-dichloroaniline	1.59	59	1,4-dichlorobenzene	1.44
24	2-chloro-4-methylphenol	2.40	60	2,2,2-trichloroethanol	3.31
25	Dieldrin	-1.78	61	2,3,4,6-tétrachlorophenol	0.67
26	hexachloroethane	0.81	62	2,3,6-trichloroaniline	1.27
27	1,1,2,2-tetrachloroethane	2.23	63	2,4,5-trichlorophenol	0.80
28	1,2,3-trichloropropane	2.45	64	2,4,6-trichlorophenol	1.06
29	1,3,5-trichlorobenzene	1.26	65	2,4,α-trichlorotoluène	0.08
30	1-chlorobutane	3.02	66	2,4-dichlorotoluène	1.46
31	2,3,4,5-tetrachlorophenol	0.48	67	2,6-dichlorophenol	1.68
32	2,3,5,6-tetrachloroaniline	0.07	68	2-chloroaniline	1.69
33	2,3,5,6-tetrachlorophenol	0.74	69	3,4,5-trichloro-2,6-dimethoxyphenol	1.12
34	2,3,6-trichlorophenol	1.44	70	3,4,5-trichloro-2-methoxyphenol	1.03
35	2,4,5-trichloroaniline	1.08	71	3,4,5-trichlorophenol	0.92
36	2,4,5-trichlorotoluène	0.94	72	3,5-dichlorophenol	1.22
37	2,5-dichloroaniline	1.01	73	4-chloro-3-methylphenol	1.67

Formatted: Subscript

Formatted: Subscript

b- Molecular descriptors

We have calculated several quantum and physicochemical descriptors to carry out the QSAR model.

➢ Quantum descriptors

For quantum descriptors we determined highest occupied molecular orbital energy (E_{HOMO}), lowest unoccupied molecular orbital energy (E_{LUMO}), total energy (E_T), dipole moment (DM), constant volume heat capacity (C_v), entropy (s), thermal energy (E_{th}), highest Mulliken electronic charge (CAE), lowest Mulliken electronic charge (CAF), ionization potential (PI) ; electronic affinity (AE), energy gap between E_{HOMO} and E_{LUMO} (ΔE), absolute electronegativity (χ), Chemical potential (μ), absolute hardness (η), mollesse ($\mathcal{S}$) and reactivity index (ω). ~~W~~We also considered the number of chlorine (N_{Cl}) atoms as a descriptor. These descriptors were generated or calculated by Gaussian 03 **[27]** **technique**. The calculations were performed, thanks to DFT method with B3LYP as the functional and 6-311++g (d,p) as basis set. Moreover, they were determined on reference to equations below.

Comment [U2]: Write in full first then abbreviate thereafter

Comment [U3]: As U2

$$PI = -E_{HOMO}\,; \qquad AE = -E_{LUMO}\,; \qquad \Delta E = E_{LUMO} - E_{HOMO}$$

$$\chi = \frac{E_{LUMO}+E_{HOMO}}{2} = \frac{AE+PI}{2} = -\mu\,; \qquad \eta = \frac{E_{LUMO}-E_{HOMO}}{2} = \frac{PI-AE}{2}\,; \qquad \mathcal{S} = \frac{1}{\eta}\,; \qquad \omega = \frac{\mu^2}{2\eta}.$$

➢ **Physicochemical descriptors**

For physico-chemical descriptors, we used ChemSketch **[28]** to determine lipophilicity (LogP), formula weight (M), molar refractivity (Rm), molar volume (Vm), parachlor (Pc), index of refraction (Ir), surface tension (ɣ), density (d) and polarizability (Pol).

2.2- Methods

a- Descriptive analysis

XLSTAT **[29]** software and the principal component analysis (PCA) method were used to realize the matrix of correlation. Principal component analysis (PCA) permits to examine descriptors set and to select the good ones that give the best model at prediction **[30].** It allows to identify the descriptors which correlate well with the biological activity Log (LC_{50}).

Formatted: Font: 10 pt, Subscript

b- Statistical analysis

The establishment of QSAR model consists of making out mathematical relationship between biological activity and chemical descriptors. Thus, QSAR model is considered satisfied when the following conditions are satisfied. The choice of the best statistical model has to satisfy the following criteria that encompass the highest coefficient of determination (R^2), the highest of adjusted determination coefficient ($R^2_{ajusted}$), the highest coefficient of Fischer (F), the highest coefficient of cross validation (R^2_{CV}), the lowest values of Mean Square Error (MSE) and the difference $R^2 - R^2_{CV} < 0.3$.

The calculation of these parameters requires a statistical analysis. The most available methods allowing that calculation are Simple Linear Regression (SLR), Multiple Linear Regression (MLR), neurons networks and Partial Least Squares (PLS). In this article, MLR method of XLSTAT ~~is~~ was used to perform the prediction **[29].**

3- RESULTS AND DISCUSSIONS

3.1- Matrix of correlation

The principal component analysis permits to perform the matrix of correlation. Any descriptor having a partial correlation coefficient with the biological activity Log (LC_{50}) less than 0.5 is removed from the descriptors set. For 2 descriptors having partial correlation coefficients with biological activity greater than 0.5 and whose partial correlation coefficient between both descriptors is greater than 0.95 then the one with the smallest partial correlation coefficient with the biological activity is also removed from the descriptors set. And we obtain the matrix of correlation presented ~~by~~ in Tables 3 and 4. These tables show that the selected descriptors correlate well with toxicity, ~~.~~ b~~B~~ecause all of correlation's coefficients are higher than 0.5.

Table 3: Matrix of correlation (Pearson (n)) of quantum descriptors used for model 1

Variables	Log LC_{50}	E_{LUMO}	S	ZPE	N_{Cl}	ΔE
Log LC_{50}	1					
E_{LUMO}	0,522	1				
S	-0,848	-0,425	1			
ZPE	0,52	0,668	-0,542	1		
N_{Cl}	-0,569	-0,772	0,536	-0,9	1	
ΔE	0,68	0,336	-0,581	0,182	-0,091	1

Table 4 : Matrix of correlation (Pearson (n)) of Physico-chemical descriptors used for model 2

Variables	Log LC_{50}	LogP	M	Pc	Ir	ɣ	D	Pol
Log LC_{50}	1							
LogP	-0,847	1						
M	-0,871	0,826	1					
Pc	-0,777	0,668	0,779	1				
Ir	-0,728	0,622	0,628	0,576	1			
ɣ	-0,679	0,59	0,682	0,55	0,941	1		
D	-0,574	0,603	0,797	0,413	0,527	0,656	1	
Pol	-0,925	0,832	0,912	0,853	0,774	0,749	0,552	1

3.2- Quantum descriptors model

- Equation of model 1

$$\boldsymbol{Log\ LC_{50} = 3.11 - 12.66{*}E_{LUMO} - 0.049{*}S - 4.92E\text{-}04{*}ZPE - 0.51{*}NCl + 13.05{*}\Delta E}$$

$N = 50$; $R^2 = 0.88$; $R^2_{adjusted} = 0.86$; $R^2_{CV} = 0.86$; $R^2 - R^2_{CV} = 0.02$;

$MCE = 0.15$; $F = 63.40$; $P_{value} < 0.0001$; $\alpha = 0.05$

Here, $\boldsymbol{N}$ is the number of molecule, ***MCE*** is Mean Square Error, $\boldsymbol{R^2}$ is the coefficient of determination, $\boldsymbol{R^2_{adjusted}}$ is the adjusted coefficient of determination, $\boldsymbol{R^2_{CV}}$ is the coefficient of cross validation and $\boldsymbol{F}$ is the coefficient of Fischer.

Furthermore, the small value of P_{value} means that the selected variables do bring a significant amount of information. The high values of R^2 (0.88) and $R^2_{adjusted}$ (0.86) and the low value of MSE (0.15) including $0.5 < R^2_{CV} < 0.9$ and $R^2 - R^2_{CV} < 0.3$ indicate that the proposed model is satisfactory and can predict our biological activity (LC_{50}).

$R^2_{adjusted} = 0.86$ indicates that the toxicity of 86% of our compounds are described with reliability by the selected descriptors. The high value of the coefficient of Fischer ($F = 63.40$) shows the strong relation which exists between the toxicity and the selected descriptors.

We can notice from the analysis of the equation of model 1 that Log (LC_{50}) increases when energy of the lowest unoccupied molecular orbital (E_{LUMO}), the entropy (s), the zero-point energy (ZPE) and the number of chlorine atoms (N_{Cl}) decrease while the gap energy (ΔE) increases. Thus, the toxicity of organochlorine compounds evolves in the same direction as ΔE and in the opposite direction as E_{LUMO}, S, ZPE and N_{Cl}.

The regression line between the experimental and theoretical LogLC_{50} of the test and the validation set is illustrated in Figure 1. The high value of determination coefficient ($R^2 = 0.88$) and the low value of mean square error $MCE = 0.15$ prove a good similarity between the predicted and experimental values. This good similarity is highlighted also ~~in~~to Figure 2.

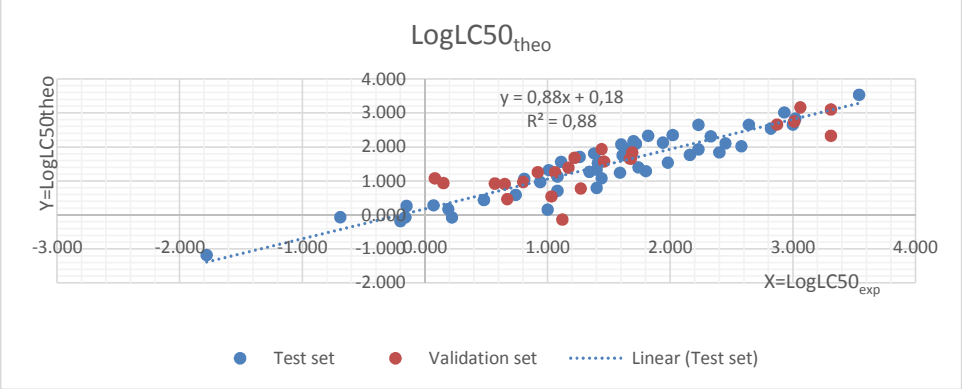

Figure 1: Regression line of the test and validation sets

The regression line of the test set indicates that: $LogLC_{50theo} = 0.88 * LogLC_{50exp} + 0.18$

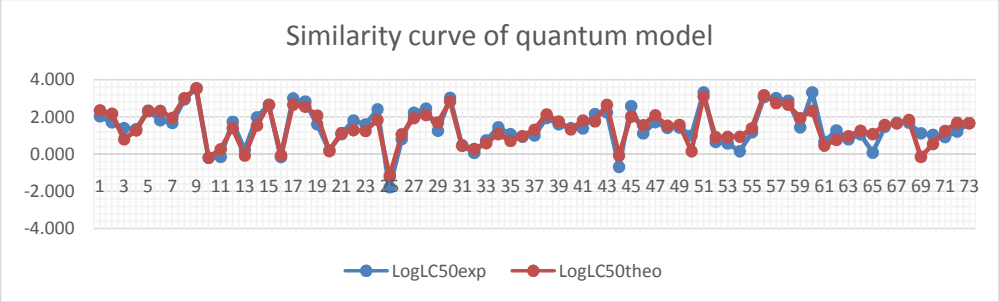

Figure 2: Similarity curve of the experimental and predicted values of the quantum descriptors model

Formatted: Font: 10 pt, Subscript

Formatted: Subscript

Formatted: Font: 10 pt, Subscript

➢ Contribution of descriptors

The contributions of the quantum descriptors in the prediction of the organochlorine molecules toxicity were illustrated ~~by~~ in the Figure 3. The classification of the contribution of the descriptors in the model is as follows:

$$N_{Cl}(-0.74) > S(-0.57) > ZPE(-0.37) > \Delta E(0.35) > ELUMO(-0.19)$$

According to the contribution of these descriptors, the number of chlorine molecules (N_{Cl}) is the most important descriptor and lowest unoccupied molecular orbital energy (E_{LUMO}) is the least important descriptor for the prediction of toxicity of organochlorine molecules.

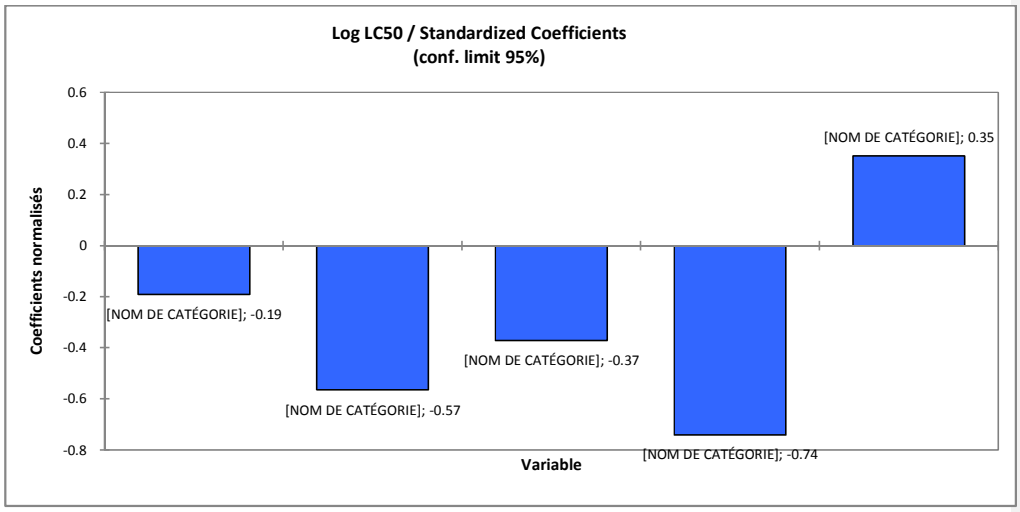

Figure 3: Contribution of descriptors in model 1

3.3- Physico-chemical descriptors model

➢ Equation of model 2

$$\textit{Log } LC_{50} = 9{,}95 - 0{,}11*\textbf{LogP} - 3{,}31*\textbf{Ir} + 0{,}03*\gamma - 0{,}73*\textbf{d} - 0{,}22*\textbf{pol}\ ;\quad N = 50;\ R^2 = 0.90;\ MCE = 0.12;$$

$$F = 80.29;\ P_{value} < 0.0001;\ \alpha = 0.05\quad R^2_{ajusted} = 0.89;\ R^2_{CV} = 0.89;\ R^2 - R^2_{CV} = 0.01$$

The various calculated parameters prove that the model established with the physicochemical descriptors is predictive and reliable. In this model 89% of our compounds are described with reliability by the selected descriptors. Therefore, this model is ~~assumed~~ shown to be satisfactory.

The toxicity of organochlorine compounds increases when lipophilicity (LogP), density (d), polarisability (pol) and index of refraction decrease and surface tension (γ) increases.

The regression line between the experimental and theoretical $LogLC_{50}$ of the test and the validation set is illustrated in Figure 4. Here, the high value of determination coefficient ($R^2 = 0.90$) and the low value of mean

Formatted: Subscript

Formatted: Font: 10 pt, Subscript

square error $MCE = 0.12$ prove a good similarity between the predicted and experimental values. This good similarity is emphasized also into Figure 5.

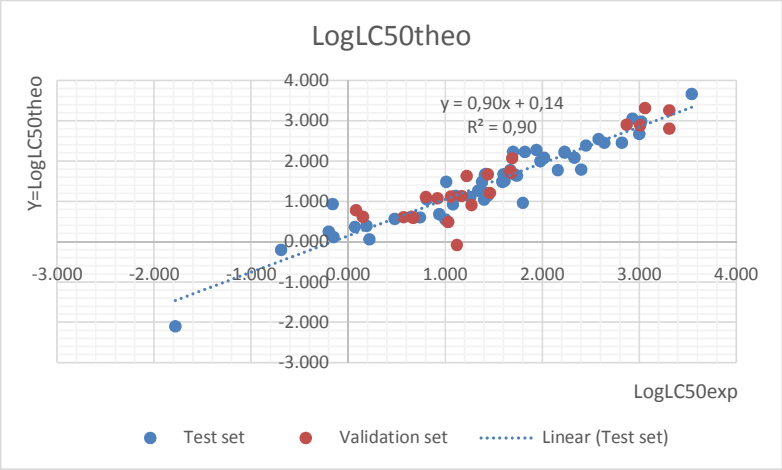

Figure 4: Regression line of the test and validation set

The regression line of the test set indicates that: $LogLC_{50theo} = 0.90 * LogLC_{50exp} + 0.14$

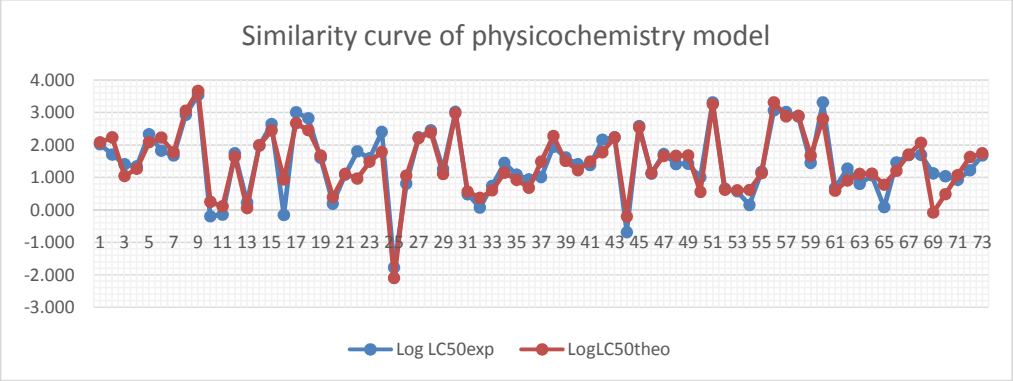

Figure 5: Similarity curve of the experimental and predicted values of the physicochemical descriptors model

- Contribution of descriptors

The contributions of the physicochemical descriptors in the prediction of the organochlorine molecules toxicity were illustrated ~~by~~ in the Figure 6. The classification of the contribution of the descriptors in the model is as follows:

$$Pol(-0.86) > \gamma(0.31) > Ir(-0.20) > d(-0.16) > LogP(-0.10)$$

According to this classification, polarizability is the most important descriptor and lipophilicity is the least important descriptor to prediction of toxicity of organochlorine molecules.

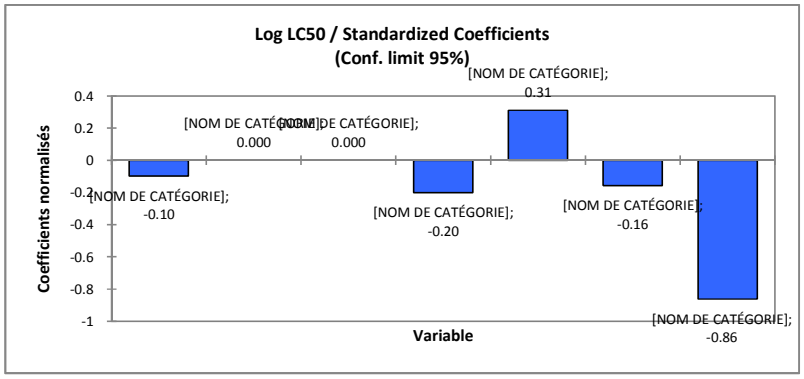

Figure 6: Contribution of descriptors in model 2

3.4- Quantum and Physicochemical descriptors model

➢ Equation of model

Log LC_{50} = 3,72 - 0,14***LogP** - 0,29***pol** + 3,17E-02***S** - 5,70E-04***ZPE** - 0,36*N_{Cl}

$R^2 = 0.94$; $R^2_{ajusted} = 0.93$; $R^2_{CV} = 0.93$; $R^2 - R^2_{CV} = 0.004$; $MCE = 0.07$; F =134.70; $P_{value} < 0.0001$; $\alpha = 0.05$

The greatest values of R^2 ; $R^2_{adjusted}$ and F and the lowest values of MCE and P_{value} shows the strong relation which exists between the toxicity and the selected descriptors. In this model 93% of our compounds are described with reliability by the selected descriptors. Then $R^2_{CV}(= 0.93) > 0.9$ *and* $R^2 - R^2_{CV}(= 0.004) < 0.3$ prove that this model is excellent. This equation shows that LogLC50 of organochlorine molecules increases when LogP, Pol, ZPE and NCl decrease and S increases.

The regression line between the experimental and theoretical LogLC50 of the test and the validation set is illustrated in Figure 7. The high value of determination coefficient ($R^2 = 0.94$) and the low value of mean square error $MCE = 0.073$ prove a good similarity between the predicted and experimental values. This good similarity is demonstrated also through Figure 8.

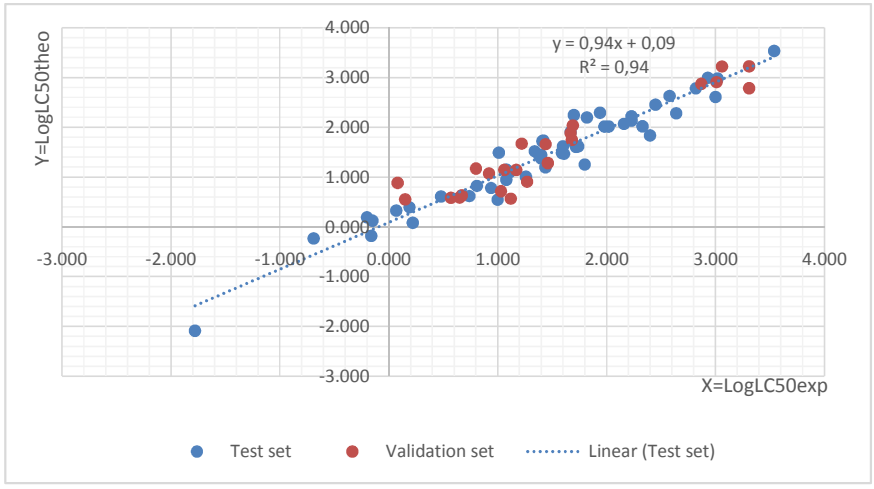

Figure 7: Regression line of the test and validation sets

Formatted: Subscript

The regression line of the test set indicates that: $LogLC_{50theo} = 0.94 * LogLC_{50exp} + 0.09$

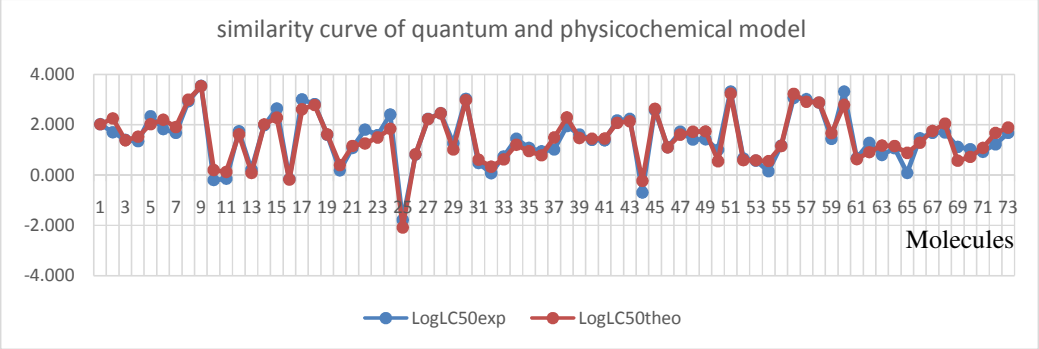

Figure 8: Similarity curve of the experimental and predicted values of the quantum and physicochemical descriptors model

- Contribution of descriptors

The contributions of the five descriptors in the prediction of the organochlorine molecules toxicity are illustrated by the Figure 9. The classification of the contribution of the descriptors in the model is as follows:

$$Pol(1.14) > N_{Cl}(0.53) > ZPE(0.43) > S(0.37) > LogP(0.13)$$

According to the contribution of these descriptors, polarizability is the most important descriptor and lipophilicity is the least important descriptor for prediction of toxicity of organochlorine molecules.

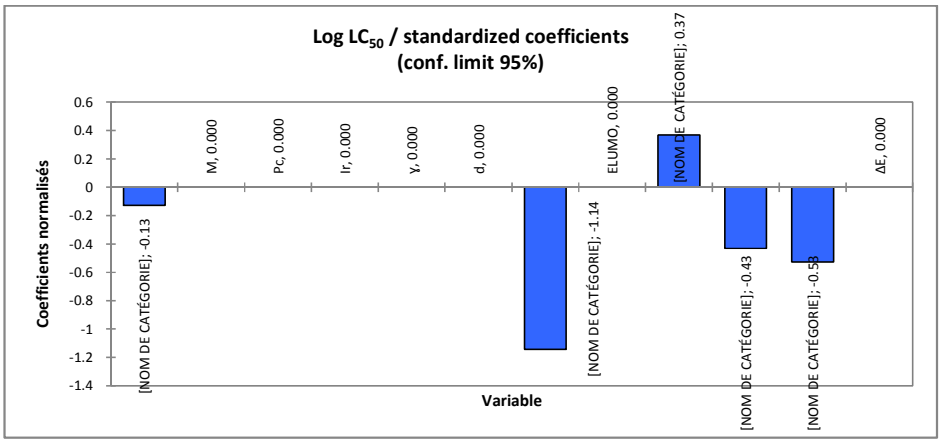

Figure 9: Contribution of descriptors in model

3.5- Comparison of different models

The comparison of the criteria of validation enables us to choose the best model for the toxicity of the organochlorine molecules prediction. These criteria are summarized in the Table 5.

Table 5: Table of comparison of different models

Nature of descriptors	Quantum	Physicochemical	Quantum and Physicochemical
Criteria			
N	50	50	50
R^2	0.88	0.90	0.94
$R^2_{ajusted}$	0.86	0.89	0.93
MCE	0.15	0.12	0.07
F	63.40	80.29	134.70
R^2_{CV}	0.86	0.89	0.93
$R^2 - R^2_{CV}$	0.02	0.01	0.004
P_{value}	< 0.0001	< 0.0001	< 0.0001
α	0.05	0.05	0.05

After comparing the validation criteria contained in Table 5, we notice that the best model is that obtained with the union of both quantum and physicochemical descriptors. This is proven by the highest value of $R^2, R^2_{ajusted}, F\ and\ R^2_{CV}$ and the lowest values of $MCE\ and\ R^2 - R^2_{CV}$. The values of R^2_{CV} superior to 0,9 and $R^2 - R^2_{CV}$ lower than 0,3 show that the established model is excellent [31]. Besides, The table below summarizes the different molecules, the experimental values of toxicity (Log (LC_{50}))$_{exp}$, the theoretical values of toxicity (Log (LC_{50}))$_{theo}$ and the ratio (Log (LC_{50}))$_{exp}$ / (Log (LC_{50}))$_{theo}$.

Table 6 : Ratio (Log (LC50))$_{exp}$ / (Log (LC50))$_{theo}$ of mixed model

N°	Compounds	Log(LC_{50exp})	Log(LC_{50theo})	Log(LC_{50exp})/ Log(LC_{50theo})
Validation set				
51	1,1-dichloroéthane	3.31	3.225	1.026
52	1,2,3,4-tetrachlorobenzene	0.65	0.590	1.101
53	1,2,3,5-tetrachlorobenzene	0.57	0.583	0.978
54	1,2,4,5-tetrachlorobenzene	0.15	0.552	0.272
55	1,2,4-trichlorobenzene	1.17	1.145	1.022
56	1,2-dichloroethane	3.06	3.221	0.950
57	1,2-dichloropropane	3.01	2.908	1.035
58	1,3-dichloropropane	2.87	2.875	0.998
59	1,4-dichlorobenzene	1.44	1.660	0.867
60	2,2,2-trichloroethanol	3.31	2.783	1.189
61	2,3,4,6-tétrachlorophenol	0.67	0.632	1.060
62	2,3,6-trichloroaniline	1.27	0.905	1.403
63	2,4,5-trichlorophenol	0.8	1.170	0.684
64	2,4,6-trichlorophenol	1.06	1.145	0.926
65	2,4,α-trichlorotoluène	0.08	0.882	0.091
66	2,4-dichlorotoluène	1.46	1.282	1.139
67	2,6-dichlorophenol	1.68	1.746	0.962
68	2-chloroaniline	1.69	2.038	0.829
69	3,4,5-trichloro-2,6-dimethoxyphenol	1.12	0.565	1.982

Formatted: Subscript
Formatted: Subscript
Formatted: Subscript
Formatted: Subscript
Formatted: Subscript
Formatted: Subscript
Formatted: Subscript
Formatted: Subscript

70	3,4,5-trichloro-2-methoxyphenol	1.03	0.717	1.437
71	3,4,5-trichlorophenol	0.92	1.073	0.858
72	3,5-dichlorophenol	1.22	1.670	0.730
73	4-chloro-3-methylphenol	1.67	1.884	0.887

The ratio $(\text{Log } (LC50))_{exp}$ / $(\text{Log } (LC50))_{theo}$ varies around 1 for most compounds, which justifies that our model is excellent.

4- CONCLUSION

The organochlorine pesticides constitute a subgroup of the organochlorine compounds. These compounds are well-known for their toxicity. What led us to determine, by QSAR method, the theoretical descriptors which could better explain this toxicity. Firstly, we determined on the one hand the quantum descriptors and on the other hand the physicochemical descriptors. Then, a Principal component analysis enabled us to select the best descriptors. Finally, three QSAR models were established. If the models established with the quantum descriptors only and the physico-chemical descriptors only gave good results with respectively $R^2 = 0.88$; $R^2_{ajusted} = 0.86$; $R^2_{CV} = 0{,}86$; $R^2 - R^2_{CV} = 0.02$; $MCE = 0.15$; $F = 63.40$ and $R^2 = 0.90$; $R^2_{ajusted} = 0.89$; $R^2_{CV} = 0.89$; $R^2 - R^2_{CV} = 0.01$; $MCE = 0.12$; F =80.29, the best model was obtained with the combination of the two types of descriptors whose equation and criteria of validation are:

Log LC_{50} = 3,72 **-** 0,14***LogP* -** 0,29***pol* +** 3,17E-02***S* -** 5,70E-04***ZPE* -** 0,36***N_{Cl}*

$R^2 = 0.94$; $R^2_{ajusted} = 0.93$; $R^2_{CV} = 0.94$; $R^2 - R^2_{CV} = 0.004$; $MCE = 0.07$; $F = 134.70$; $P_{value} < 0.0001$; $\alpha = 0.05$.

In perspective, we ~~will~~ determined the descriptors that influence the half-life time and the bioaccumulation factor of organochlorine compounds and then propose less toxic, less bio-accumulative and less persistent organochlorine pesticides.

Formatted: Subscript

REFERENCES

[1] Zeidler, O., Beitrag zur Kentniss der Verbindungen zwisschen Aldehyden und aromatischer Kohlenwasserstoffen. Inaugural Dissertation der Philosophen facultat der Universitat-Strasbourg, Wien, 1873.

[2] Smith, A.G. Chlorinated hydrocarbon insecticides. In : Handbook of Pesticides Toxicology. San Diego/New York: Academic Press Inc., 1991, p.731-915.

[3] M. L. BOUGUERRA et B. PHILOGENE, Éditions scientifiques, techniques et médicales (ESTEM). Paris, 1994, 35-66.

[4] Hayes WJ, Laws ER, Handbook of pesticide toxicology, volume I and II, 1991.

[5] OMS, Aldrine et Dieldrine. Critère d'hygiène et environnementale 1989, 91.

[6] OMS, Questions fréquemment posées à propos de l'utilisation du DDT pour la lutte antivectorielle, (2004). https://apps.who.int/iris/handle/10665/68629 (consulted 01/04/2019).

[7] Jean Mouchet Le DDT en santé publique, cahiers santé ; 1994, 4, 257-262

http://www.documentation.ird.fr/hor/fdi:41553 (consulté le 01/04/2019).

[8] WHO, The use of DDT in malaria vector control. WHO position statement; Geneva: World Health Organization, 2011. https://www.who.int/malaria/publications/atoz/who_htm_gmp_2011/en/ (consulté le 01/04/2019).

[9] Carson R , Silent Spring, (eds.) Hall GK & Company, Boston, 1962.

[10] Russel RW, Gobas FA, Hafïher GD, Role of chemical and ecological factors in trophic transfer of organic chemicals in aquatic food webs Environmental Toxicology and Chemistry, 1999, 44: 1250.

[11] Boseret J-PH; Pollution des sols: les pesticides; 2000.

[12] Waliszewski SM, Villalobos-Pietrini R, Gomez-Arroyo S, Infanzon RM, Persistent organochlorine pesticides in cow's milk samples from tropical regions of Mexico. Food additives and contaminants, 2003, 20: 270. https://doi.org/10.1080/0265203031000062091.

[13] El Nemr A, Abd-Allah AMA, Organochlorine contamination in some marketable fish in Egypt. Chemosphere; 2004, 54:1401.

[14] Binelli A, Provini A, The PCB pollution of Lake Iseo (N. Italy) and the role of biomagnification in the pelagic food web. Chemosphere, 2003, 53: 143.

[15] Couch JA., Winstead JT and Goodman LR. "Kepone-induced scoliosis and its histological consequences in fish," Science, 1977 vol. 197, 4303, 585–587. https://doi.org/10.1126/science.69318.

[16] Traoré KS., Mamadou K., Dembélé A., Lafrance P, Banton O., Houenou P., Résidus de pesticides organochlorés dans le lait humain d'une zone agricole de Côte d'Ivoire. Journal Ouest Africain de chimie, 2002,13, 99-109.

[17] Gregor DJ, Gummer WD, Evidence of atmospheric transport and deposition of organochlorine pesticides and polychlorinated biphenyls in Canadian arctic snow. Environmental Science & Technology; 1989, 23: 561.

[18] Schomburg CJ., Glotfelty DE., & Seiber JN., Pesticide occurrence and distribution in fog collected near Monterey, California. Environmental Science & Technology, 1991, 25(1), 155–160.

[19] Dietz R, Riget K, Cleemann M, Aarkrog A, Johansen P, Hansen JC, Comparison of contaminants from different trophic levels and ecosystems. The Science of Total Environment; 2000, 245: 221. https://doi.org/10.1016/S0048- 9697(99)00447-7

[20] Alexander Crum Brown and Thomas R. Fraser, On the connection between chemical constitution and physiological action with special reference to the physiological action of the salts of the ammonium bases derived from strychnia, brucia, thebaia, codeia, morphia, and nicotia. Journal of Anatomy; 1868, 2(2): 224–242.

[21] Richet MC., Note sur le rapport entre la toxicité et les propriétés physiques des corps, Compt. Rend. Soc. Biol. (Paris), 1893, 45, 775-776.

[22] Hansch C., Fujita T., *p*-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure; Journal of American Chemical Society, 1964, 86, 1616-1626. https://pubs.acs.org/doi/pdf/10.1021/ja01062a035.

[23] Free SM , Wilson, JW., A Mathematical Contribution to Structure-Activity Studies Journal of Medicinal Chemistry 1964, 7, 395-399 https://pubs.acs.org/doi/10.1021/jm00334a001.

[24] Alan R. Katritzky, Douglas B. Tatham, and Uko Maran, Theoretical Descriptors for the Correlation of Aquatic Toxicity of Environmental Pollutants by Quantitative Structure-Toxicity Relationships; Journal of Chemical Information and Computer Sciences ; 2001, 41, 1162-11760. Doi: 10.1021/ci010011r.

[25] Jean-Claude Amiard, Les risques chimiques environnementaux - Méthodes d'évaluation et impacts sur les organismes 2e Edition revue et augmentée.

[26] Hodge H.C. & Sterner J.H., Determination of substances acute toxicity by LDB50B. Amer. Industrial Hyg. Assoc.; 1943, 10: 93.

[27] Frisch M. J., Trucks G. W., Schlegel H. B., Scuseria G. E., Robb M. A., Cheeseman J. R., Montgomery J. A., Jr., Vreven T., Kudin K. N., Burant J. C., Millam J. M., Iyengar S. S., Tomasi J., Barone V., Mennucci B., Cossi M., Scalmani G., Rega N., Petersson G. A., Nakatsuji H., Hada M., Ehara M., Toyota K., Fukuda R., Hasegawa J., Ishida M., Nakajima T., Honda Y., Kitao O., Nakai H., Klene M., Li X., Knox J. E., Hratchian H. P., Cross J. B., Adamo C., Jaramillo J., Gomperts R., Stratmann R. E., Yazyev O., Austin A. J., Cammi R., J. J., Zakrzewski V. G., Dapprich S., Daniels A. D., Strain M. C., Farkas O., Malick D. K., Rabuck A. D., Raghavachari K., Foresman J. B., Ortiz J. V., Cui Q., Baboul A. G., Clifford S., Cioslowski J., Stefanov B. B., Liu G., Liashenko A., Piskorz P., Komaromi I., Martin R. L., Fox D. J., Keith T., Al-Laham M. A., Peng C. Y., Nanayakkara A., Challacombe M., Gill P. M. W., Johnson B., Chen W., Wong M. W., Gonzalez C., and Pople J. A. (2004) Gaussian 03, Revision B.02. Gaussian, Inc., Pittsburgh PA.

[28] ACDLABS 10, Advanced Chemistry Development Inc., Toronto, ON, Canada, 2015.

[29] XLSTAT Version 2014.5.03 Copyright Addinsoft 1995-2014 (2014) XLSTAT and Addinsoft Are Registered Trademarks of Addinsoft. https://www.xlstat.com.

[30] Brereton, R.G., Chemometrics: Data Analysis for the Laboratory and Chemical Plantl, John Wiley & Sons, Chichester, UK, 2003.

[31] Lennart Eriksson, Joanna Jaworska, Andrew P Worth, Mark T D Cronin, Robert M Mc Dowell, and Paola Gramatica Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs, Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification- and Regression-Based QSARs, Environmental Health Perspectives, 2003, 111(10), 1361-1375.
https://doi.org/10.1289/ehp.5758.