Resources and Environment 2018, 8(3): 174-183
DOI: 10.5923/j.re.20180803.09

Modeling of the Dynamics and Prediction of Transparency of a Mesotropic Tropical Lake: Bakré Lake (Abidjan, Côte d'Ivoire)

Yapo Habib Kpidi[1], Mamadou Guy-Richard Koné[2,*], Tiama Guy Ballet[1], Lamoussa Ouattara[2], Ossey Bernard Yapo[1,3], Nahossé Ziao[2]

[1]Laboratoire des Sciences de l'Environnement (LSE), (UFR-SGE), Université Nangui Abrogoua, 02 BP 801 Abidjan 02 - Côte-d'Ivoire
[2]Laboratoire de Thermodynamique et Physico-chimie du Milieu (LTPCM), (UFR-SFA), Université Nangui Abrogoua, 02 BP 801 Abidjan 02 - Côte-d'Ivoire
[3]Laboratoire Central de l'Environnement, Centre Ivoirien Anti-pollution (LCE-CIAPOL)

Abstract In this study, thirty samples were collected bimonthly from August 2015 to December 2016 at six locations of *Bakré* Lake and analyzed. Lake Bakré, located in the sub-equatorial Attiean climate, between the Atlantic Ocean in the South and the Ebrié Lagoon in the North, is an endorheic lake with a surface outfall, Abidjan (Côte d'Ivoire). A modeling study was conducted in order to determine a quantitative and qualitative relationship between transparency and five (5) physico-chemical descriptors (temperature, turbidity, nitrite ((NO_2^-), nitrates (NO_3^-) and ammonium (NH_4^+)). These descriptors were used as the explanatory and predictive parameters of the transparency of the selected samples taken ~~at~~ from Lake M'koa in a previous study. This study was conducted by using the Principal Component Analysis (PCA), the Ascending Hierarchical Classification (AHC). And then multilinear regression (MLR) and partial least squares (PLS) quantitative and qualitative models were proposed. These models were accredited to have good statistical indicators and so have been validated according to the principles established by the Organization for Economic Co-operation and Development (OECD). The statistical indicators of the MLR reveal more efficient predictions with the coefficient of determination $\mathbf{R^2}$ = 0,920, the standard error **RMCE** = 0.169, correlation coefficient of cross-validation $\mathbf{Q^2cv}$ = 0.920, and Fisher test **F** = 285.896. This model is acceptable with $R^2 - Q^2_{CV}$ = 0.000 < 0.3. The values of the Transp theo / Transp exp report of the validation set tend (is closed) to unity, and the obtained results suggest that the combination of these 5 descriptors could be useful for predicting the property of Transparency. In addition, temperature is the priority (useful) descriptor for the prediction of transparency on the *Bakré* Lake stations.

Keywords Eutrophication, Transparency, Quantitative Structure-Property Relationship (QSPR), Physico-Chemical Descriptors

1. Introduction

Water is a useful and indispensable resource for the life of any living organism. Maintaining its quality healthy is a major problem for societies that they must face according to the growing water needs [1, 2]. The deterioration of the quality of these waters seems to come from strong anthropic pressures (agricultural activities, industrial and / or domestic effluents discharged into the receiving environment without being treated before). The quality of water is therefore the main cause of most public health problems in developing countries [3, 4]. The assessment of the trophic state of the aquatic environment consists in carrying out *physico-chemical* and biological measurements of a group of environmental variables those seem to be dependent on several constraints. Nowadays, there is growing interest in developing alternative methods for monitoring the quality of the aquatic environment and the rational management of its resources in order to remove these constraints. Eutrophication is defined as a natural aging process of aquatic environments. It can be accelerated by anthropogenic activities on their watershed [5]. In addition, the functioning of aquatic ecosystems is governed by dynamic equilibria. However, eutrophication is an operating imbalance triggered off by a change in the quantities, relative proportions or forms of nitrogen and phosphorus entering in these aquatic systems. The nature and intensity of the responses therefore depend on environmental factors favoring eutrophication. However, the preservation of these remarkable ecosystems

* Corresponding author:
guyrichardkone@gmail.com (Mamadou Guy-Richard Koné)
Published online at http://journal.sapub.org/re

for their undeniable ecological and socio-economical roles and the sustainability of good environmental health status are therefore an urgent and useful necessity for coming generations (sustainable development). Thus, the control of the factors that can disturb the equilibrium of these lacustrine ecosystems seem to be a major interest [6]. To stop the eutrophication of these aquatic systems, a continuous monitoring of their physico-chemical and biological qualities and other eutrophication factors must be done. This indispensable monitoring seems to be expensive in terms of equipment and reagents. To overcome the considerable demands and investments in terms of logistics and chemical products necessary for taking (sporadic) punctual or continuous measurements [7]. The development of alternative methods for the rational management of water resources is an opportunity. Although mathematical models of eutrophic ecosystems have been developed to understand and represent ecological dynamics, some of them have been used to estimate eutrophication risks and to evaluate the necessary reduction of nutrient supplies. It is an approach which seems to be based on models those are qualified to be statistical. They seek (try) to predict one or more eutrophication descriptors according to causal variables which were measured in the working area [8, 9]. Thus, transparency is a variable which can be modeled and seems to reflect the influence of dissolved organic compound and particle those are released into water. These last ones color water and make it cloudy. Thus, a Quantitative Structure Property Study (QSPR) is used to elaborate models by using *physico-chemical* descriptors. This statistical approach can combine methods such as linear regression, partial least squares regression, and so on. Generally all the descriptors of the equation of the QSPR model must be a function of at least 1/5 of the initial data, here we have thirty (30) samples. The main aim of this study is to design statistical models by multiple regressions (linear and partial least squares) which are capable to predict the transparency property of the waters of Lake *Bakré* by using environmental variables. Specifically, the question is to identify meaningful and expressive explanatory variables of transparency in order to develop simulation tools that will be useful for integrating and sustaining the management of water resources.

2. Materials and Methods

2.1. Presentation of the Study Area

Located in the sub-equatorial Attiean climate, between the Atlantic Ocean in the South and the Ebrié Lagoon in the North, Lake Bakré, Abidjan (Côte d'Ivoire) is an endorheic lake with a surface emissary, the Vri (**Figure 1**). It is located between 4°02'57 "and 4°07'49" west longitude and 5°14'54"and 5°15'48" north latitude. A shallow natural lake with an average depth of 7.07 m, approximately 9 km long with an area of approximately 5.636 km^2, *Bakré* Lake is infiltrated and comprises an East-West main channel and four North-South secondary channels [10].

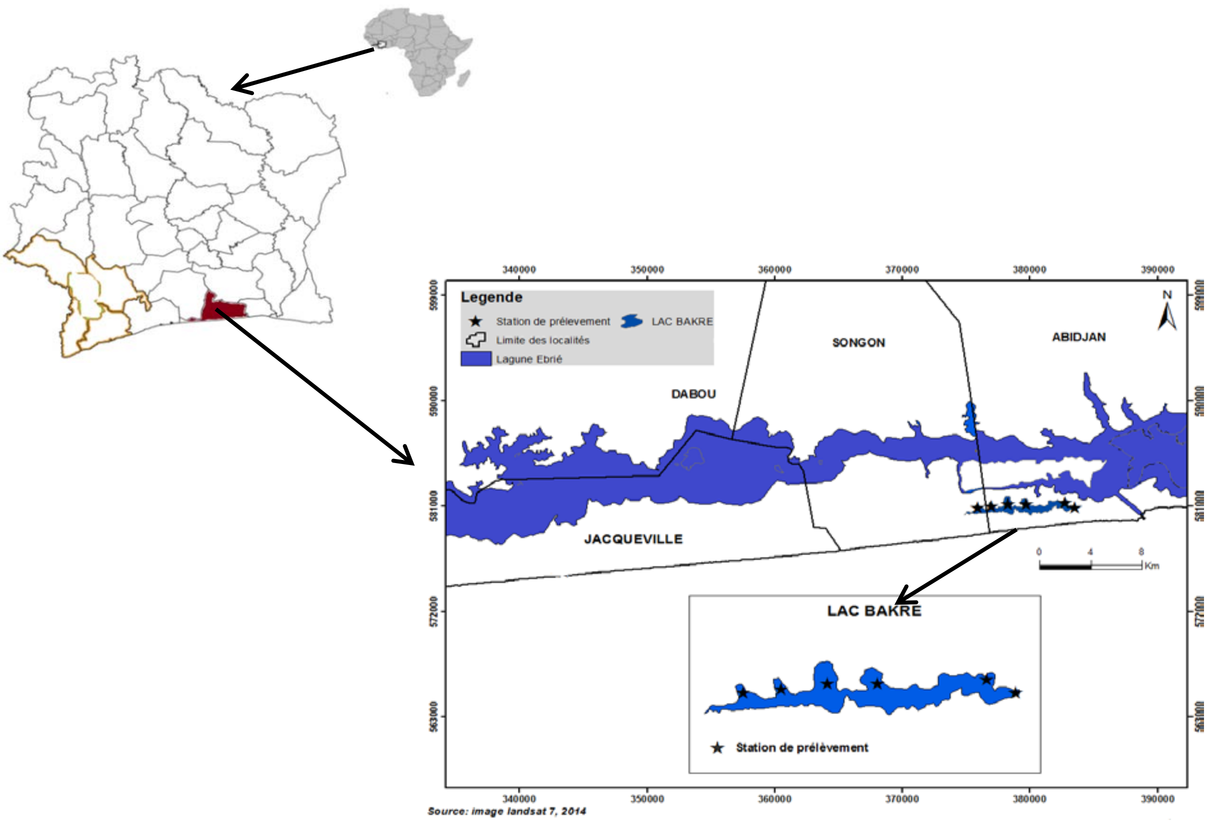

Figure 1. Mapping of the study area

2.2. Physico-Chemical Descriptors

Temperature, turbidity, nitrate (NO_3^-), nitrite (NO_2^-) and ammonium (NH_4^+) are the *physico-chemical* descriptors used for transparency modeling. The transparency property was measured by the Secchi disk. Indeed, immersed until complete disappearance, the Secchi disk came slowly back and the depth at which the disc becomes visible then determines the transparency of the area or the depth according to Rodier *et al.* [11].

Water transparency depends on the amount of particles in the water. These particles can be algae or sediment from erosion, the more particles the less water transparency. In other words, when the water is murky or cloudy and contains a lot of particles, the light cannot penetrate as deeply into the water column. These materials may be of mineral origin (silt, clay) or organic (plant and animal debris, microorganisms, algae, chemical compounds). This property is inversely proportional to the turbidity which allows to identify the visual information on the water. Turbidity reflects the presence of suspended particles in the water (organic debris, clays, microscopic organisms, etc.). It reduces both photosynthesis and the dissolved oxygen content because of the presence of biodegradable colloids [11]. Water quality monitoring using modeling is essential to map its turbidity [12]. Water turbidity is a good indicator in predicting the state of health of the wetland in order to maintain its ecosystem [13]. Temperature is the most important kinetic factor for all chemical and biological reactions in aquatic environments. It plays a fundamental role in the kinetics of *physico-chemical* and biological reactions and the value of equilibrium constants. In addition, a temperature above 15°C promotes the development of microorganisms, intensifies the biodegradation of organic matter [14]. Dissolved oxygen is too important as parameter in assessing the health status of a lake. It depends mainly on the (breathing) respiration of the planktonic populations' photosynthesis and the mineralization of biomass [15]. Nitrate (NO_3^-), nitrite (NO_2^-) and ammonium (NH_4^+) are the inorganic nitrogen forms and current pollutants in both surface water and groundwater causing enormous health problems in humans and animals [16, 17]. These different *physico-chemical* descriptors were determined by kpidi *et al.* [18]. The modeling was done using the multilinear and least square regression method implemented in Excel spreadsheets [19] and XLSTAT version 2014 [20].

2.3. Estimation of the Predictive Capacity of a Model

The transparency of the thirty (30) samples of our study shows a variation ranging from 1.75 m to 3.50 m. Thus, this range of reduced variation in concentrations allows to define a better quantitative relationship between the transparency and the physico-chemical descriptors of taken samples. The quality of a model is determined by basing ourselves on various statistical analysis indicators such as the coefficient of determination $\mathbf{R^2}$, the standard error (**RMCE**), the correlation coefficients of the cross-validation $\mathbf{Q^2cv}$ and Fischer's one which is **F**. $\mathbf{R^2}$, **RMCE** and **F** relate to the adjustment of calculated and experimental values. They describe the predictive power within the limits of the model and allow to estimate the accuracy of the calculated values on the test set [21-23]. Concerning the cross-validation coefficient $\mathbf{Q^2cv}$, it gives information on the predictive power of the model, which can be "internal" because it is calculated from the structures which are used to build this model. The coefficient of determination R^2 gives an evaluation of the dispersion of the calculated values around the experimental ones. The quality of the modeling is better when the points evaluated by the coefficient $\mathbf{R^2}$ are closer to the adjustment line [24, 25]. $\mathbf{R^2}$ is expressed as follows:

$$\boldsymbol{R}^2 = 1 - \frac{\Sigma\left(y_{i,exp} - \hat{y}_{i,theo}\right)^2}{\Sigma\left(y_{i,exp} - \bar{y}_{i,exp}\right)^2} \tag{3}$$

Where:

$y_{i,exp}$: Experimental value of transparency; $\hat{y}_{i,theo}$: Calculated value of transparency and

$\bar{y}_{i,exp}$: Experimental mean value of transparency

The closer the value of R^2 is to 1, the better the computed and experimental values are correlated. Moreover, the variance $\boldsymbol{\sigma}^2$ is determined by the relation 4 below:

$$\boldsymbol{\sigma}^2 = (\mathbf{RMCE})^2 = \frac{\Sigma\left(y_{i,exp} - y_{i,theo}\right)^2}{n-k-1} \tag{4}$$

Where **k** is the number of independent variables (descriptors), **n** is the observation number of the test or learning set and **n-k-1** is the degree of freedom.

The standard error or standard deviation **RMCE** is another used statistical indicator. It allows to evaluate the reliability and the accuracy of a model:

$$\mathbf{RMCE} = \sqrt{\frac{\Sigma\left(y_{i,exp} - y_{i,theo}\right)^2}{n-k-1}} \tag{5}$$

The Fisher **F** test is also used to measure the level of statistical significance of the model, that is to say, the quality of the choice of descriptors constituting the model.

$$\mathbf{F} = \frac{\Sigma\left(y_{i,theo} - y_{i,exp}\right)^2}{\Sigma\left(y_{i,exp} - y_{i,theo}\right)^2} * \frac{n-k-1}{k} \tag{6}$$

The coefficient of determination of the cross-validation $\mathbf{Q^2_{CV}}$ which permits to evaluate the accuracy of the prediction on the test set is defined by the following relation:

$$\boldsymbol{Q}^2_{cv} = \frac{\Sigma\left(y_{i,theo} - \bar{y}_{i,exp}\right)^2 - \Sigma\left(y_{i,theo} - y_{i,exp}\right)^2}{\Sigma\left(y_{i,theo} - \bar{y}_{i,exp}\right)^2} \tag{7}$$

2.4. Statistical Analyzes

Principal Component Analysis (PCA) is a data analysis tool that allows to explain the structure of correlations or covariances using linear combinations of original data. Its use makes it possible to interpret the data in a reduced space [26, 27]. It was used to appreciate, certainly, the relations between the different measured variables, but especially to access their structure in order to be able to group them by zone. Grouping by zone thus meets the checked target of this approach, which is to correlate the classes of physico-chemical descriptors obtained at the sampling stations. The Ascending Hierarchical Classification (AHC) aims to partition a set of samples into homogeneous classes [28, 29]. It organizes samples according to a number of variables and modalities. It groups them too hierarchically on a dendrogram. It aggregates samples that have a similarity between them by using the measurement of the distance between samples in order to form classes. It is realized by using the data of individuals and variables. AHC allowed to establish a typology of the samples according to the temperature, turbidity, NO_2^-, NO_3^- and NH_4^+. The Multiple Linear Regression Statistical Technique (MLR) is used to study the relationship between a dependent variable (Property) and several independent variables (descriptors). This statistical method minimizes the differences between the actual and predicted values. It also allowed to select descriptors used as input parameters in Partial Least Squares (PLS) regression. The Partial Least Squares (PLS) regression analysis also improves the structure-property relationship in order to evaluate quantitatively the property. It is the most current tool for studying multidimensional data.

MLR and PLS were generated from XLSTAT version 2014 [20] to predict transparency. The equations of the different models were evaluated by the coefficient R^2, the mean squared error (RMCE), the Fischer test (F) and the cross correlation coefficient (Q^2_{CV}) [30, 31, 32].

2.5. Criterion of Acceptance of a Model

The performance of a mathematical model, according to Eriksson et al. [33, 34], is characterized by a value of $Q^2_{cv} > 0{,}5$ for a satisfactory model and an excellent one is characterized by, $Q^2_{cv} > 0{,}9$. According to these authors, a test game, will be qualified to be an efficient model if the following acceptance criterion R^2-Q^2_{CV} < 0,3 is respected.

3. Results and Discussion

The set of descriptor values for twenty (20) samples of the test set and ten (10) other samples of the validation set are presented in **Table 1**.

Table 1. Experimental *physico-chemical* and transparency descriptors of test and validation test

Observations	Temperature (°C)	Turbidity (NTU)	NO_2^- (mg.L^{-1} N-NO_2^-)	NO_3^- (mg.L^{-1} N-NO_3^-)	NH_4^+ (mg.L^{-1} N-NH_4^+)	Transp Exp
			Test Set			
1	28.50	3.61	0.002	0.400	0.010	3.350
2	27.60	3.49	0.003	0.300	0.040	3.400
3	31.50	2.51	0.003	0.300	0.020	2.750
4	30.90	3.05	0.002	0.300	0.020	2.400
5	30.70	3.53	0.001	0.500	0.000	2.350
6	31.20	3.39	0.002	0.300	0.010	2.050
7	31.25	2.96	0.003	0.400	0.115	2.700
8	31.80	2.43	0.004	0.400	0.075	2.575
9	31.60	2.64	0.003	0.350	0.090	2.225
10	31.40	3.10	0.002	0.400	0.135	2.150
11	31.65	3.37	0.003	0.300	0.055	1.975
12	31.90	3.07	0.003	0.300	0.190	2.050
13	32.30	2.23	0.003	0.300	0.160	2.050
14	32.10	2.67	0.002	0.400	0.270	1.950
15	32.10	3.35	0.003	0.300	0.100	1.900
16	30.60	3.38	0.004	0.400	0.030	3.100
17	30.20	2.87	0.004	0.400	0.020	3.500
18	30.90	3.64	0.004	0.400	0.040	2.900
19	30.80	3.97	0.005	0.300	0.030	2.300
20	30.60	3.94	0.005	0.300	0.010	2.500
			Validation Set			
21	30.60	3.38	0.004	0.400	0.030	3.100
22	30.20	2.87	0.004	0.400	0.020	3.500
23	30.90	3.64	0.004	0.400	0.040	2.900
24	30.80	3.97	0.005	0.300	0.030	2.300
25	30.60	3.94	0.005	0.300	0.010	2.500
26	31.20	2.25	0.004	0.500	0.000	2.400
27	31.60	2.27	0.004	0.300	0.000	2.680
28	31.00	2.43	0.006	0.300	0.000	2.150
29	33.60	3.00	0.003	0.300	0.000	1.750
30	33.00	3.57	0.004	1.100	0.000	1.920

Table 2. Correlation matrix (Pearson (n)) between the different *physico-chemical* descriptors

Variables	Temperature	Turbidity	NO_2^-	NO_3^-	NH_4^+	Transparency
Temperature	**1.000**					
Turbidity	-0.341	**1.000**				
NO_2^-	-0.061	0.149	**1.000**			
NO_3^-	0.225	0.083	-0.013	**1.000**		
NH_4^+	0.290	-0.270	**-0.372**	-0.136	**1.000**	
Transparency	**-0.771**	0.094	0.172	-0.061	-0.329	**1.000**

Bold values are different from 0 to a significant level for p <0.05. Very significant for p <0.01. Very highly significant for p <0.001.

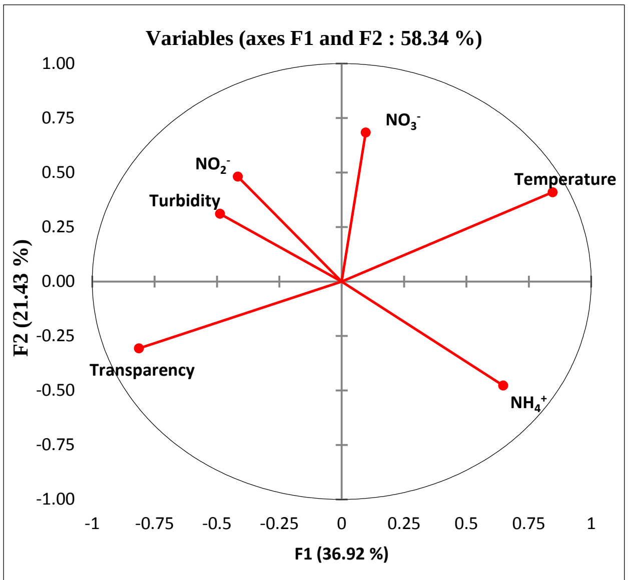

Figure 2. Correlation circle of physico-chemical descriptors according to F1 x F2

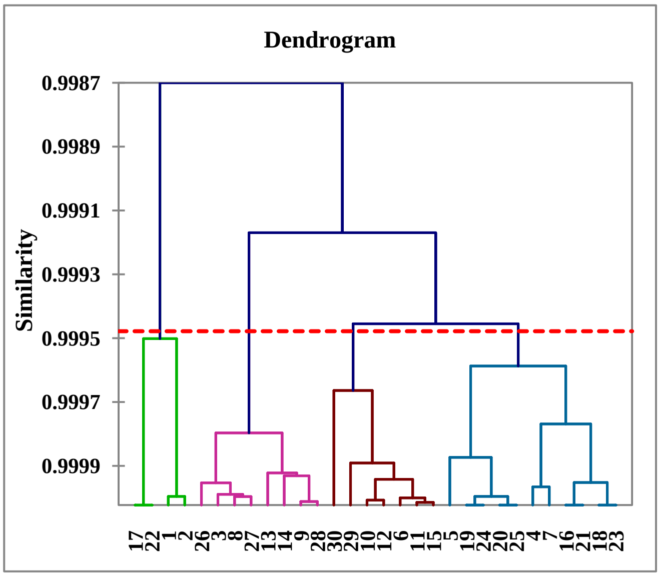

Figure 4. Dendrograms of the samples

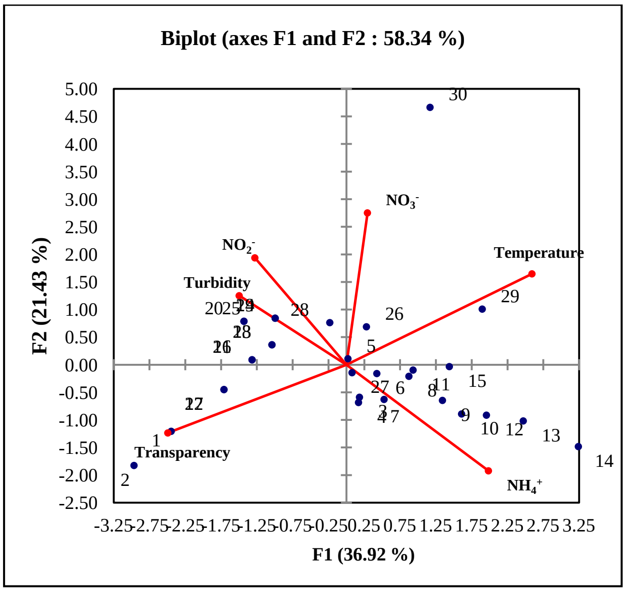

Figure 3. Cartesian diagram according to F1 and F2: correlation between used descriptors and transparency

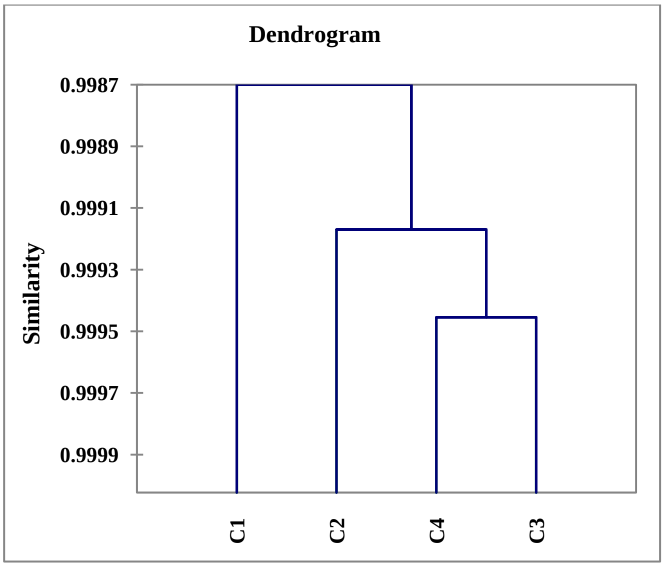

Figure 5. Dendrograms of the groups

3.1. Typology of the Waters of Lake Bakré

The Pearson correlation matrix (n) (**Table 2**), the correlation circle (**Figure 2**), the Cartesian diagrams according to F1 and F2 (**Figures 3 and 4**) and the dendrogram of the stations are shown above.

The PCA data matrix gathers the mean values of five (5) variables representing *physico-chemical* descriptors and thirty (30) individuals representing the samples of the stations. The resulting matrix provides information about the negative or positive correlation between the variables. The temperature is negatively correlated with transparency (r = -0.771 and p <0.05) at a significant level.

The examination of the community circle in **Figure 2** associated with the analysis of the PCA factor structure (**Table 1**), indicate two (02) principal components represented by F1 which corresponds to 36.92% of the explained variance and F2 with 21.43% of the explained variance. The two factors F1 and F2 totalize 58.34% of the total variance are sufficient to interpret the entire PCA data. Each variable is associated with its factorial weight. Thus, temperature (0.845) and NH_4^+ (0.647) are positively correlated with factor 1. Conversely, transparency (-0.813) is negatively and strongly correlated. Factor 1 reflects a pollution gradient. As for factor 2, NO_3^- (0.684), which participates in the oxidation of organic matter, is positively correlated with it. Factor 2 seems to indicate an oxidation gradient.

The Cartesian diagram of **figure 3** shows that samples **1, 2, 17** and **22** negatively correlated with factor 1 with a small value of transparency are inversely proportional to sample **29** positively correlated with factor 1 with a high value of temperature. However, the light (temperature) entering water side depends on many factors such as the sum of the composition and size of the dissolved and suspended substances. Thus, the transparency decreases with the presence of molecules and particles that can absorb or diffuse the light so creating an increasingly turbid environment [11]. In addition, samples **16, 18, 19, 20, 21, 23, 24** and **25** are turbid and negatively correlated with factor 1 and are inversely proportional to NH_4^+ and grouping together samples **9, 10, 12, 13** and **14**.

Figure 4 illustrates a distribution of samples in four classes **(C1, C2, C3 and C4**) characterizing the *physico-chemical* descriptors. Thus, the class **C1** (**1, 2, 17** and **22**) which is expressed by the transparency is inversely proportional to the temperature. Indeed, the increase in temperature induces a reduction in transparency. Classes **C2** (**3, 8, 9, 13, 14, 26, 27** and **28**) and **C4** (**6, 10, 11, 12, 15, 29,** and **30**) are correlated with the inorganic nitrogen compound (NH_4^+). Class **C3** (**4, 5, 7, 16, 18, 19, 20, 21, 23, 24** and **25**) of the classification is influenced by organic and particulate pollution expressed by low turbidity and high content of NO_2^-. This classification is corroborated by the Ascending Hierarchical Classification which also highlights these groups of samples.

3.2. Multilinear Regression (MLR) and Partial Least Squares (PLS) Models

It should be noted that the negative or positive sign of the coefficient of a descriptor of the model reflects the effect of proportionality between the evolution of transparency and this physico-chemical parameter of the regression equation. Thus, the negative sign indicates that when the value of the descriptor is high, the transparency decreases while the positive sign reflects the opposite effect. **Table 3** presents the best MLR and PLS models obtained for the different experimental values of transparency as well as the statistical indicators. **Figures 6** and **7** show the adjustment line of the experimental and theoretical transparency data of the test sets (blue dots) and validation sets (red dots) of the model. It should be underlined that these models were established using the same test and validation sets of **Table 1**.

Model MLR:

$$\boldsymbol{Transp_i^{theo}} = \mathbf{14.18348 - 0.37521 * Temperature - 0.48276 * Turbidity + 202.76075} * \boldsymbol{NO_2^-} + \mathbf{2.77602} * \boldsymbol{NO_3^-} - \mathbf{1.43687} * \boldsymbol{NH_4^+}$$

Model PLS:

$$\boldsymbol{Transp_i^{theo}} = \mathbf{12.71554 - 0.34118 * Temperature - 0.41065 * Turbidity + 158.87859} * \boldsymbol{NO_2^-} + \mathbf{3.71937} * \boldsymbol{NO_3^-} - \mathbf{1.68015} * \boldsymbol{NH_4^+}$$

Table 3. Statistical analysis report on the transparency of Lake *Bakré*

Statistical indicators of Multilinear Regression	Model MLR	Model PLS
Number of observations **N**	20	20
Squared regression correlation coefficient $\mathbf{R^2}$	0.920	0.893
Standard error of the regression **RMCE**	0.169	0.172
Statistical significance of regression, Fisher **F**-test **F**	285.896	149.664
Cross-validation correlation coefficient Q_{CV}^2	0.920	0,893
$\mathbf{R^2}$ - Q_{CV}^2	0.000	0.000
Level of Statistical Significance **α**	> 95%	

The negative correlation coefficient between turbidity and transparency indicates that these two variables are inversely proportional, that is to say the increase in turbidity favors the reduction of transparency in both models. The significance of these models is shown by the high values of the Fischer coefficient **F** which are respectively 285.896 and 149.664 for the MLR and PLS models. In addition, the robustness of these models is reflected by the correlation coefficient of the cross validation $\mathbf{Q^2cv}$ which are also 0.920 and 0.893. These MLR and PLS models are all acceptable because $R^2 - Q_{CV}^2 = 0.000 < 0.3$.. The regression lines established by the experimental and theoretical test and validation set transparency data for the MLR and PLS models are shown in **Figures 6 and 7.**

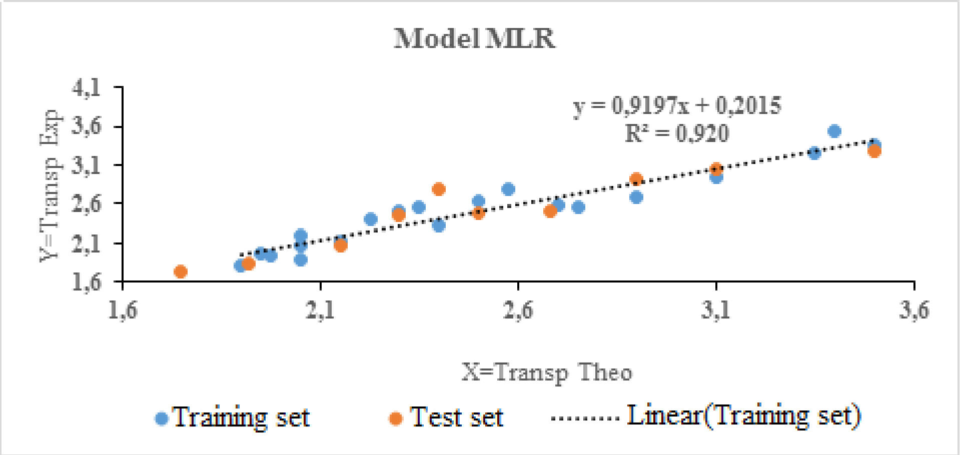

Figure 6. Regression line of the MLR test and validation phases

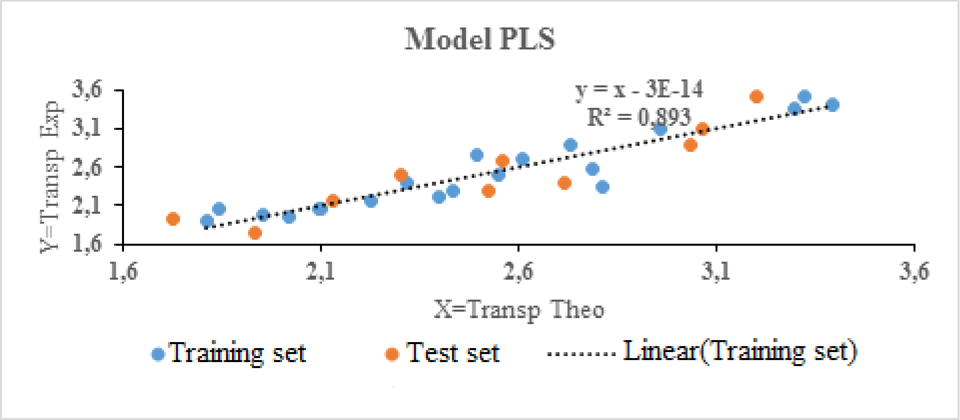

Figure 7. Regression line of the PLS test and validation phases

Table 4. Values of the relationship between theoretical and experimental transparency of the validation set

	Validation set				
	Model MLR Transp theo	**Model PLS Transp theo**	**Transp exp**	**Model MLR Transp theo/Transp exp**	**Model PLS Transp theo/Transp exp**
21	3.057	3.066	3.100	0.99	0.99
22	3.289	3.201	3.500	0.94	0.91
23	2.913	3.036	2.900	1.00	1.05
24	2.471	2.525	2.300	1.07	1.10
25	2.485	2.305	2.500	0.99	0.92
26	2.792	2.717	2.400	1.16	1.13
27	2.521	2.559	2.680	0.94	0.95
28	2.071	2.129	2.150	0.96	0.99
29	1.749	1.937	1.750	1.00	1.11
30	1.852	1.728	1.920	0.96	0.90

The external validation test was verified by calculating the Transp theo /Transp exp report of the MLR and PLS models. These values are (recorded) confined in **Table 4**.

The values of the report of Transp theo / Transp exp of the validation set model tend to unity (**Table 4**) reflecting the good correlation between the theoretical and experimental transparency of the observations. These models are therefore acceptable for predicting the transparency of Lake *Bakré*.

In addition, the low standard error values (**RMCE**) of both MLR and PLS are respectively 0.169 and 0.172, and indicate good similarity between predicted and experimental values (**Figure 8 and 9**). These curves show a similar evolution of the data of these two models for the prediction of the transparency of Lake Bakré, despite some differences.

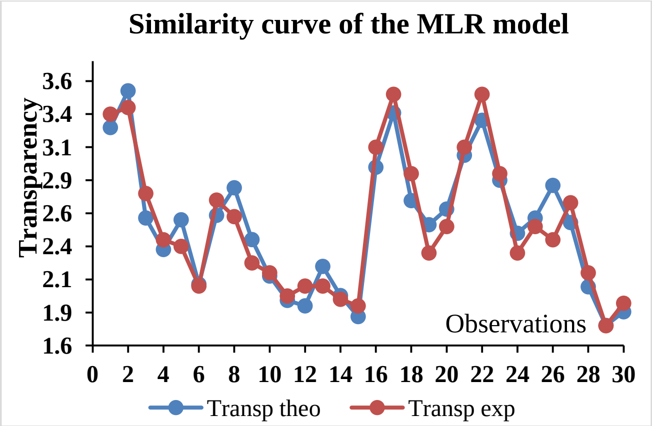

Figure 8. Similarity curve of the experimental and predicted values of the MLR model

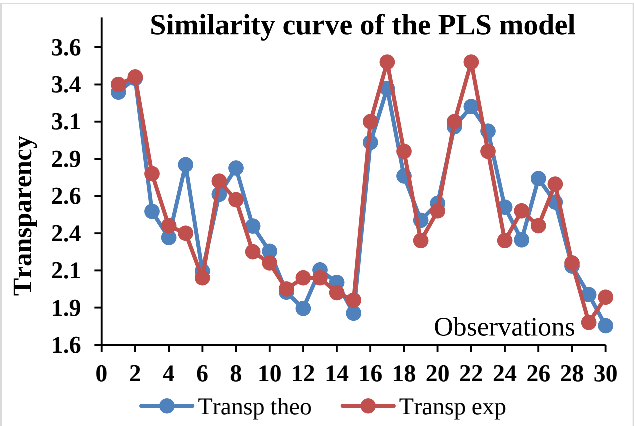

Figure 9. Similarity curve of the experimental and predicted values of the PLS model

The values of the Transp theo/Transp exp report of the validation set that tend towards the unit (**Table 3**) reflect the good correlation between the theoretical and experimental Transparencies of the observations. This model is therefore acceptable for predicting the transparency of Lake *Bakré.*

Among the two obtained models, the PLS statistical one has a much better predictive skill. However, this model is determined by taking into account four *physico-chemical* descriptors, it is essential to determine the contribution of each of them in the prediction of the transparency property. Indeed, the knowledge of this contribution allows to establish the order of priority of different descriptors and define the choice of parameters to optimize for the good prediction and understanding of the transparency of Lake *Bakré.*

3.3. Analysis of the Contribution of the Descriptors

The contributions of the four physico-chemical descriptors in the prediction of the transparency of the sampled waters were illustrated by the normalized coefficients shown in **Figure 10**.

The order of priority of the *physico-chemical* descriptors with their respective normalized coefficients is classified according to the following sequence:

Temperature > Turbidity > $N0_2^-$ > $N0_3^-$ > NH_4^+

According to the contribution of these descriptors, the temperature displays the highest normalized coefficient (-0.854) followed by the turbidity with -0.465 and, NH_4^+ holds the lowest coefficient (-0.202) compared to the other descriptors. It should be noted that temperature is the most influential *physico-chemical* descriptor. Thus, to improve the quality of the water of Lake *Bakré*, we must play at maximum temperature to reduce the heat by setting up a dense plant belt around the lake.

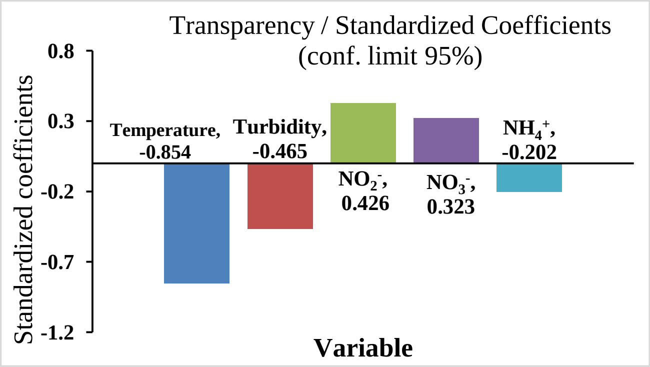

Figure 10. Contribution of *physico-chemical* descriptors in models

4. Conclusions

This study has permitted to highlight relationships between transparency, which is a fundamental property of water to transmit light, and the measured *physico-chemical* descriptors of the samples. The *physico-chemical* descriptors (Temperature, turbidity, NO_3^-, NO_2^-, NH_4^+) allow to explain and predict the property of transparency because there are strong correlations between these calculated and experimental values. Multivariate analysis revealed 4 classes. Class **C1** is characterized by the transparency of water expressing its ability to let the light pass through water. This parameter is inversely proportional to the temperature. Classes **C2** and **C4** are characterized by inorganic nitrogen pollution (NH_4^+) and finally class **C3** is determined by turbidity. The study of the robustness of the two models (MLR and PLS) presented a good stability and an excellent power of prediction. In addition, the MLR model (**R^2**= 0.920, **RMCE** = 0.169, **F**= 285.896) is better and is an effective tool for predicting transparency at *Bakré* Lake. Moreover, the study of the contribution of *physico-chemical* descriptors has shown that temperature is the most important one in terms of priority to predict the transparency of Lake *Bakré.*

REFERENCES

[1] S. Haouchine, "Recherche sur la qualité faunistique et l'écologie des macro-invertébrés benthiques des cours d'eau de Kabylie," *Mém. Magister, Faculté des sciences biologique et sciences agronomiques, Université Mouloud MAMMERI,*

Tizi-Ouzou, Algérie. 116p + annexes, 2011.

[2] M. S. Foto, T. S. H. Zebaze, T. N. L. Nyamsi, G. A. Ajeagah, & T. Njiné, “Evolution spatiale de la diversité des peuplements de macroinvertébrés benthiques dans un cours d’eau anthropisé en milieu tropical (Cameroun),” *European Journal of Scientific Research*, vol. 55, no 2, pp. 291-300, 2011.

[3] M. A. Montgomery, & M. Elimelech, “Water and sanitation in developing countries: Including health in the equation,” *Environmental Science & Technology*, vol. 41, pp. 17-24, 2007.

[4] W. H. Organization, “Water, sanitation and hygiene links to health facts and figures-updated,” November 2004, Geneva, 2010.

[5] K. A. N’Guessan, K. F. Konan, Y. B. Kotchi, O. E. Edia, T. Gnagne, K. S. Traoré, & V. P. Houenou, “Prospects for rehabilitation of man-made lake system of Yamoussoukro (Ivory Coast),” *Procedia Environmental Sciences*, vol. 9, pp. 140-147, 2011.

[6] P. Durand, L. Breuer, P. Johnes, G. Billen, A. Butturini, G. Pinay, H. J. M. van Grinsven, J. Garnier, M. Rivett, D. S. Reay, C. Curtis, J. Siemens, S. Maberly, Ø. Kaste, C. Humborg, R. Loeb, J. De Klein, J. Hejzlar, N. Skoulikidis, P. Kortelainen, A. Lepistö, R. F. Wright, “Nitrogen processes in aquatic ecosystems. In The European Nitrogen Assessment: Sources, Effects and Policy Perspectives,” *Eds. Cambridge University Press, Cambridge*, pp. 126-146, 2011.

[7] C. Neal, “The potential for phosphorus pollution remediation by calcite precipitation in UK freshwaters,” *Hydrology and Earth System Sciences*, vol. 5, pp. 119-131, 2001.

[8] V. Villeneuve, S. Légaré, J. Painchaud & V. Warwick, “Dynamique et modélisation de l’oxygène dissous en rivière,”. *Revue des Sciences de l’Eau*, vol. 19, pp. 259-274, 2006.

[9] A. M. Kouassi, A. Mamadou, K. E. Ahoussi & J. Biémi, “Conception de modèles statistiques à variables hydrochimiques pour la prédiction de la conductivité électrique des eaux souterraines,”. *Larhyss Journal*, vol. 20, pp. 189-207, 2014.

[10] C. A. Aka, K. F. Koffi Bi, N. J-P. Yao, K. L. Adopo, K. S. Yao, N. Aka, S. Monde & K. Aka, “Caractères sédimentologiques d’un environnement lacustre du littoral de Côte d’Ivoire: Cas du lac M’bakré,” *Afrique SCIENCE*, vol. 11, no. 5, pp. 184-198, 2015.

[11] J. Rodier, B. Legube & N. Merlet, “L’Analyse de l’Eau,” *9è Eds, Dunod, Paris*, pp. 1579, 2009.

[12] B. Nechad, K. G. Ruddick & G. Neukermans “Calibration and Validation of a Generic Multisensor Algorithm for Mapping of Turbidity in Coastal Waters,” *SPIE European International Symposium on Remote Sensing*, Berlin. 2009.

[13] A. Raihan, S. Mehebub, & S. Haroon, “Preparing turbidity and aquatic vegetation inventory for waterlogged wetlands in Lower Barpani sub-watersheds (Assam), India using geospatial technology,” *The Egyptian Journal of Remote Sensing and Space Sciences*, pp. 1-7, 2016.

[14] N. Gallina, O. Anneville & M. Beniston, “Impacts of extreme air temperatures on cyanobacteria in five deep peri-Alpine lakes,” *J. Limnol.*, vol. 70, pp. 186-196, 2010.

[15] S. Hebert & S. Légre, “Direction du suivi de l’état de l’environnement,” *Ministère de l’Environnement, Gouvernement du Québec*, 2000.

[16] Q. H. Wang, L. J. Yu, Y. Liu, L. Lin, R. Lu, J. p. Zhu, L. He & Z. L. Lu “Methods for the detection and determination of nitrite and nitrate,” A review. *Talanta*, vol. 165, pp. 709-720. 2017.

[17] J. Zhang, C. Yang, X. Wang & X. Yang, “Colorimetric recognition and sensing of nitrite with unmodified gold nanoparticles based on a specific diazo reaction with phenylenediamin,” *Analyst*, vol. 137, pp. 3286-3292. 2012.

[18] Y. H. Kpidi, O. B. Yapo, M. J. Yao-Ohou, & T. G Ballet, “Variabilité journalière de la qualité physico-chimique du lac M’koa de Jacqueville (Côte d’Ivoire),” *Int. J. Biol. Chem. Sci.*, vol. 11, pp. 901-910, 2017.

[19] Microsoft ® Excel ® 2013 (15.0.4420.1017) MSO (15.0.4420.1017) 64 Bits, Partie de Microsoft Office Professionnel Plus, 2013.

[20] XLSTAT Version 2014.5.03 Copyright Addinsoft 1995-2014; XLSTAT and Addinsoft are Registered Trademarks of Addinsoft, https://www.xlstat.com, 2014.

[21] N. J.-B. Kangah, M. G.-R. Koné, C. G. Kodjo, B. R. N’guessan, A. L. C Kablan, S. A. Yéo, & N. Ziao,. “Antibacterial Activity of Schiff Bases Derived from Ortho Diaminocyclohexane, Meta-Phenylenediamine and 1,6-Diaminohexane: Qsar Study with Quantum Descriptors,” *International Journal of Pharmaceutical Science Invention*, vol. 6, no. 3, pp. 38-43, 2017.

[22] G. W. Snedecor, & W. G. Cochran, “Statistical Methods,” *Oxford and IBH, New Delhi, India,* pp. 381, 1967.

[23] M. V. Diudea, “QSPR/QSAR Studies for Molecular Descriptors,” *Nova Science: Huntingdon, New York.* 2000.

[24] J. S. N’dri, M. G.-R. Koné, C. G. Kodjo, S. T. Affi, A. L. C. Kablan, Z. A. Ouattara, & N. Ziao, “Quantitative Structure antifungal Activity Relationship (QSAR) study of a series of Schiff bases derivatives from 4aminobenzenesulphonamide By DFT method,” *IOSR. Journal Of Pharmacy*, vol. 7, no. 4, pp. 27-33. 2017.

[25] E. X. Esposito, A. J. Hopfinger, & J. D. Madura, “Methods for Applying the Quantitative Structure-Activity Relationship Paradigm,” *Methods in Molecular Biology*, vol. 275, pp. 131-213, 2004.

[26] Y. H. Kpidi, O. B Yapo, M. G.-R. Koné, G. A. Gadji, A. E. J. E. Y. Gnagne, J. S. N’dri, N. Ziao, “Monitoring and Modeling of Chlorophyll-a Dynamics in a Eutrophic Lake: M'koa Lake (Jacqueville, Ivory Coast),” *American Journal of Environmental Protection*, vol. 6, no. 11, pp. 1-9, 2018.

[27] D. Soro, L. Ekou, M. G.-R. Koné, T. Ekou, S. T. Affi, L. Ouattara, & N. Ziao “Prediction of the Inhibitory Concentration of Hydroxamic Acids by DFT-QSAR Models on Histone Deacetylase 1,” *International Research Journal of Pure & Applied Chemistry*, vol. 16, no. 2, pp. 1-13, 2018.

[28] J. S. N’dri, M. G-R. Koné, C. G. Kodjo, S. T. Affi, A. C. L Kablan, O. Ouattara, D. Soro, & N. Ziao, “Quantitative Activity Structure Relationship (QSAR) of a Series of Azetidinones Derived from Dapsone by the Method of Density Functional Theory (DFT),” *IRA International Journal of Applied Sciences*, vol. 8, no. 2, pp. 55-62, 2017.

[29] K. R. Clarke, & M. Ainsworth, "A method of linking multivariate community structure to environmental variables," *Marine Ecology*, vol. 92, pp. 205-219, 1993.

[30] J. S. N'dri, M. G.-R. Koné, C. G. Kodjo, A. L. C. Kablan, L. Ouattara, O. Ouattara, N. Ziao, "Combining of DFT and QSAR Results to Predict the Antibacterial Activity of a Series of Azetidinones derived from Dapsone as Inhibitors of Bacillus Subtilis and Pseudomonas Aeruginosa," *Journal of Computational Chemistry & Molecular Modelling*, vol. 2, no. 2, pp. 1-9, 2018.

[31] B. Escofier, & J. Pagès, "Analyses factorielles simples et multiples: Objectifs, méthodes et interprétation," Dunod, Paris, pp. 318, 2008.

[32] N. N. Kouakou, M. G.-R. Koné, K. Bamba, W. P. Ouattara, N. Ziao, "Quantitative Structure Anti-Cancer Activity Relationship (QSAR) of a Series of Ruthenium Complex Azopyridine by the Density Functional Theory (DFT) Method," *Computational Molecular Bioscience*, vol. 7, pp. 19-31, 2017.

[33] C. Rücker & G. J. Rücker, "Y-Randomization is a tool used in validation of QSPR/QSAR models," *Chem. In. Model*," vol. 47, pp. 2345-2357, 2007.

[34] L. Eriksson, J. Jaworska, A. Worth, M. T. D. Cronin, R. M. Mc Dowell & P. Gramatica "Methods for Reliability and Uncertainty Assessment and for Applicability Evaluations of Classification and Regression-Based QSARs," *Environmental Health Perspectives*, vol. 111, pp. 1361-1375, 2003.