This article was published in Environmental Technology, 35 (8), 945-955, 2014
http://dx.doi.org/10.1080/09593330.2013.856956

Multiple linear and principal component regressions for modelling ecotoxicity bioassay response

Ana I. Gomes [a], José C.M. Pires [b], Sónia A. Figueiredo [c], Rui A.R. Boaventura [a,*]

[a] *LSRE* ,[b] *LEPAE, Departamento de Engenharia Química, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal*

[c] *REQUIMTE, Instituto Superior de Engenharia do Porto, Rua Dr. António Bernardino de Almeida 431, 4200-072 Porto, Portugal*

[*] Corresponding author: Tel.: +351225081683; Fax.: +351225081674

e-mail address: bventura@fe.up.pt (Rui A.R. Boaventura)

Abstract

The ecotoxicological response of the living organisms in an aquatic system depends on the physical, chemical and bacteriological variables, as well as the interactions between them. An important challenge to scientists is to understand the interaction and behaviour of factors involved in a multidimensional process such as the ecotoxicological response. With this aim, multiple linear regression (MLR) and principal component regression (PCR) were applied to the ecotoxicity bioassay response of *Chlorella vulgaris* and *Vibrio fischeri* in water collected at seven sites of Leça river during five monitoring campaigns (February, May, June, August and September of 2006). The river water characterization included the analysis of 22 physicochemical and 3 microbiological parameters. The model that best fitted the data was MLR, which shows: (i) a negative correlation with dissolved organic carbon (DOC), zinc and manganese, and a positive one with turbidity and arsenic, regarding *Chlorella vulgaris* toxic response; (ii) a negative correlation with conductivity and turbidity and a positive one with phosphorus, hardness, iron, mercury, arsenic and fecal coliforms, concerning *Vibrio fischeri* toxic response. This integrated assessment may allow the evaluation of the effect of future pollution abatement measures over the water quality of Leça River.

Keywords: *Chlorella vulgaris*; Ecotoxicological assessment; Multiple linear regression; Principal component regression; Surface water quality; *Vibrio fischeri*.

1. Introduction

Pollution of surface water with toxic chemicals and excess of nutrients, resulting from storm water runoff, mains leakage leaching, and groundwater discharges, has been an issue of worldwide environmental concern [1]. The water quality assessment must comprise an ecotoxicological characterization, which allows properly evaluating the potential risks of effluent discharges, especially when they are complex [2]. The ecotoxicity evaluation by means of acute bioassays may bring quick and valuable information [3, 4]. However, most of the ecotoxicity test methods were established to measure the toxicity of pure single chemicals, and not to be applied to unknown environmental water samples with complex components. Since chemicals are present in environmental water as a complex mixture, their potential ecotoxicological effects are much complicated due to their interactions [**5-9**]. In addition, even if the toxicity of an environmental sample is tested, there is no guidance on how to evaluate the water quality in terms of protection of aquatic living organisms [6]. It is difficult to extrapolate the potential damage on the aquatic ecosystem from the test results with specific species, particularly because not all species respond identically to the same pollution stresses [**10**]. It is also quite difficult to evaluate the actual exposure levels and ecotoxicological effects of all coexisting chemicals on aquatic organisms by measuring concentrations of individual chemicals (United States Environmental Protection Agency - USEPA [**11, 12**]). It must also be kept in mind that there is an uncertainty factor when laboratory results are extrapolated to field conditions because of the simultaneous influence of a number of environmental and biological factors (bioavailability, toxicokinetics, sensitivity of organisms, etc.) [4]. However, direct toxicity test of environmental water sample can provide an integral view on ecotoxicological effects of

all chemicals coexisting in water as a mixture and has been widely used in safety assessment of water quality [**6, 13, 14**].

The study of ecological properties of different organisation levels may reveal changes of potential ecological signification that cannot be detected by other analyses [1]. The bacterium *Vibrio fischeri* (decomposer) and the alga *Chlorella vulgaris* (1st producer) were selected for this study because they belong to different trophic levels and are widely used in ecotoxicity tests [1, 2]. One of the advantages of these tests is the fast assessment of ecotoxicity.

The ecotoxicological response of the living organisms in an aquatic system depends on several variables, such as nutrient quantitative and qualitative profiles, temperature, physicochemical properties of the water and grazing pressure [**15**]. An important challenge for scientists is to develop analytical tools that could be used to understand the interaction and behaviour of factors involved in a multidimensional process [**16**] such as the ecotoxicological response, and to provide the necessary tools for monitoring and management of resources. Modelling is regarded as an important analytical tool for biological and ecological studies [**17, 18**].

Multivariate statistical techniques are useful for evaluation and interpretation of large and complex water quality data sets [**19**]. Multiple linear regression (MLR) is one of the most widely used methodologies for expressing the dependence of a response variable on several explanatory (predictor) variables [**16, 20-22**]. Principal component analysis (PCA) is useful in pre-processing methodology for mitigating the problem of multicollinearity (when the explanatory variables are correlated with each other) and for exploring the relations among the input variables, particularly if it is not obvious which of the variables should be the predictors. PCA creates new variables, the principal components (PCs), by linear combination of the original variables. PCs are uncorrelated

to each other, removing the multicollinearity problem. They are interpreted by the association with original variables through the corresponding factor loadings. Principal component regression is the linear model that relates the dependent variable with these PCs. Both MLR [**20, 23**] and principal component regression (PCR) [**16**] approaches have been applied in studies of water quality.

The present study aims to model *Chlorella vulgaris* and *Vibrio fischeri* bioassays toxic response in concern to the Leça river water characterization by MLR and PCR. The achieved models lead to infer possible influences of physicochemical and microbiological variables of river water in bioassay results.

2. Materials and methods

2.1 Area description – sampling sites

The Leça river flows through a highly populated and industrialized area in the north of Portugal and receives a complex mixture of pollutants from poorly treated or untreated domestic, agricultural and industrial effluents, and other contaminated waters both from point and diffuse sources.

Figure 1 presents the location of Leça river in the north of Portugal. It rises in the Mountain of Santa Luzia at Santo Tirso and flows for approximately 48 km until the Atlantic Ocean. Water samples were collected in seven sampling sites along the river: site 1 is located in the upstream part of the river in a mainly rural area; sites 2 and 4 are both located downstream from wastewater treatment plants in a highly populated area; sites 3 and 5 are situated in a strongly populated and industrialized area; site 6 is in a revitalized area with a recreational park; and site 7 is some meters upstream from the river mouth, before a waterfall, and therefore it does not receive any marine influence. Water samples were collected in five different periods - February, May, June, August

and September of 2006, one day in each month (not always the same). Most of the samples were collected from bridges, in order to obtain samples from running water which were representative of the river water. Grab samples were manually collected by immersion of plastic bottles into the river.

2.2 Analysis of the water samples

The analytical procedures used to characterize the water samples are presented in Table 1. All used reagents were analytic grade.

Temperature, pH and oxidation-reduction potential, dissolved oxygen and conductivity were measured *in situ*. Water samples were stored at 4 ºC (no chemical preservatives were added) and analyzed in duplicate within 24 hours. For dissolved organic carbon and metals a filtration by 0.45 μm pore diameter membrane filter was performed. Bioassays were performed within (the maximum) 48 hours after sampling.

The bioluminescent inhibition toxicity tests (ISO 11348) **were performed** using the bacteria *Vibrio fischeri* (NRRL B 11177). Tested concentrations were 5.6%, 11.3%, 22.5% and 45% (v/v). The values of EC_{50} (effective concentration of the sample that causes 50 inhibition to the test-organisms) and the corresponding 95% confidence intervals were determined for 5 and 15 minutes of bacterial exposure.

The green algae inhibition growth tests were performed with the microalgae *Chlorella vulgaris* according to USEPA Guideline (2002). Three replicates of each sample were tested for five different concentrations (10%, 20%, 40% 60% and 80%). The test solutions were incubated for 72 hours, under continuous cool white fluorescent

light. Agitation was performed manually twice per day. Initial and final absorbance were measured at 440 nm [24] in order to evaluate the growth of the algal population. A calibration curve was used to convert the absorbance in cell concentration. The acceptability criterion considered was variability less than 20% among replicates. Shapiro-Wilk's Normality Test and Bartlett's Test for Homogeneity of Variance were performed to validate data, and Dunnett's procedure was followed (USEPA 2002). Since these assumptions were met, EC_{50} was calculated by linear interpolation.

The reference toxicants used to validate tests were phenol and potassium dichromate, respectively for *V. fisheri* and *C. vulgaris* bioassays.

The toxic response was evaluated through the calculation of EC_{50}, effective concentration that causes 50% of inhibition to test-organism. For regression models purpose EC_{50} was converted in toxicity units, TU_{50} ($TU_{50} = 100/EC_{50}$), as suggested by Wisconsin Department of Natural Resources [25]. Because EC_{50} was expressed in percentage, the sample is considered "not toxic" when $TU_{50} = 1$ and biostimulated when $TU_{50} < 1$.

2.3 Regression models

The data considered for this analysis were the mean of replicates. Before the determination of the models, the data were Z standardized to have zero mean and unit standard deviation. MLR attempts to model the relationship between two or more explanatory variables and a response variable, by fitting a linear equation to the observed data [26, 27]. The dependent variable (y) is given by:

$$y = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i + \varepsilon \quad (1)$$

were x_i (i = 1, …, k) are the explanatory variables, $\hat{\beta}_i$ (i = 0, …, k) are the regression coefficients, and ε is the error associated with the regression and assumed to be normally distributed with both expectation value zero and constant variance [28].

The predicted value given by the regression model ($\hat{y}$) is calculated by:

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i \quad (2)$$

To estimate the regression coefficients $\hat{\beta}_i$ the minimization of the sum of squared errors (SSE) method is used, as follows:

$$\hat{\beta}i = arg\ min \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \quad (3)$$

PCR is a method that combines linear regression and PCA [27]. Essentially, PCA maximizes the correlation between the original variables to form new variables, the principal components (PCs) that are orthogonal and uncorrelated. These variables are linear combinations of the original variables. The PCs are ordered in such a way that the first component has the largest fraction of the original data variability [16, 29]. To evaluate the influence of each variable in the PCs, varimax rotation is generally used to obtain the rotated factor loadings that represent the contribution of each variable in a specific PC. PCR establishes a relationship between the output variable (y) and the selected PC obtained from the explanatory variables (x_i) [27].

The significance of the regression coefficients in the MLR and PCR models was evaluated through the calculation of their confidence intervals [27, 30]. The regression coefficient $\hat{\beta}_i$ is statistically significant if:

$$|\hat{\beta}i| > \frac{t^{\alpha/2}_{n-k-1}\hat{\sigma}}{\sqrt{Sxx_i}} \quad (4)$$

where t is the Student t distribution, n is the number of points, k is the number of parameters, α is the significance level, $\hat{\sigma}$ is the standard deviation given by $\sqrt{\mathrm{SSE}/(n-k-1)}$ and Sxx_i is the sum of the squares related to x_i given by $\sum_{j=1}^{n}(x_{i,j}-\bar{x}_i)^2$.

Hence, several MLR and PCR models were determined by testing all combinations of the explanatory variables, selecting the ones that presented the lowest SSE and all statistically significant regression coefficients [27].

The PCs were calculated using Matlab, while MLR and PCR models were evaluated by developed subroutines in Microsoft Visual Basic for Applications (Microsoft Excel).

2.4 Performance indexes

The performances of MLR and PCR models in the prediction of *Chlorella vulgaris* and *Vibrio fischeri* toxic response were evaluated through calculation of the coefficient of determination (R^2), mean absolute error (MAE), root mean squared error (RMSE) and index of agreement (d_2) [31, 32]. The MAE and the RMSE measures residual errors which gives a global idea of the difference between the observed and modelled values. The values d_2 indicate the degree of which the predictions are error free, because it compares the difference between the mean, the predicted and the observed concentrations.

3. Results

The physicochemical, bacteriological and ecotoxicological results were presented in a previous study [33]. The models were determined to model *Chlorella vulgaris* and

Vibrio fischeri toxic response using physicochemical and bacteriological variables as predictors. Regarding *Vibrio fischeri* results, only the 15 min-toxic responses were used in the regression models. From the 25 monitored variables, only 15 were applied for models development. Variables that were measured *in situ* and that presented always values below the detection limit were not considered. Both MLR and PCR models were determined by statistically significant regression coefficients with a significance level of 0.05.

The MLR led to the following results: (i) *Chlorella vulgaris* toxic response was negatively affected by DOC, Zn and Mn, and positively affected by turbidity and As; and (ii) *Vibrio fischeri* toxic response was negatively affected by conductivity and turbidity, and positively affected by phosphorus, hardness, Fe, Hg, As and fecal coliforms. The regression models obtained by MLR were as follows:

$$C.\ vulgaris = 2.719 - 2.193\ (\text{DOC}) - 1.399\ (\text{Zn}) - 0.782\ (\text{Mn}) + 1.651\ (\text{turbidity}) + 3.643\ (\text{As}) \quad (5)$$

$$V.\ fischeri = 1.849 - 5.845\ (\text{conductivity}) - 0.860\ (\text{turbidity}) + 0.971\ (\text{phosphorus}) + 2.951\ (\text{hardness}) + 0.551\ (\text{Fe}) + 1.624\ (\text{Hg}) + 0.595\ (\text{As}) + 0.657\ (\text{fecal coliforms}) \quad (6)$$

PCA was performed to obtain in the PCs all variance contained in the original data. Thus, fifteen PCs were determined. Table 2 presents the results from PCA showing the rotated factor loadings for all fifteen PCs. Values in bold correspond to the greatest contributions of the original variables on the PCs. PC1 had important contributions from conductivity, DOC, total nitrogen, total phosphorus, hardness and Hg. PC3 was heavily loaded by all bacteriological parameters. PC2, PC4, PC5, PC6, PC7 and PC8 had important contributions from Mn, Zn, turbidity, As, Fe and colour,

respectively. PC9 to PC15 did not present any significant contribution of the original variables; however, they were used in PCR to analyse if these minor contributions are statistically significant in the ecotoxicological response of living organisms. The regression models using PCs as input variables (PCR) were the following:

$$\textit{C. vulgaris} = 2.719 + 0.683\,(PC3) - 1.899\,(PC6) - 1.677\,(PC8) + 2.841\,(PC9) \quad (7)$$

$$\textit{V. fischeri} = 1.849 - 0.442\,(PC4) - 1.304\,(PC8) + 1.087\,(PC9) + 8.596\,(PC15) \quad (8)$$

Table 3 presents the matrix that multiplied by the original variables matrix gives the values of PCs. These values show how a PC was influenced by each original variable. For instance, negative values showed that the original value and the PC are negatively correlated. Taking values in Table 3 corresponding to high factor loadings (in Table 2) and the regression coefficients for each PC, it is possible to infer the relationship between the original variables and the output variable. If both values have the same signal, the influence is positive; otherwise, the influence is negative. According to this transformation and the regression coefficients given by the models, PCR showed that: (i) *Chlorella vulgaris* toxic response was negatively influenced by colour and DOC, and positively by As, Hg and all bacteriological parameters, especially fecal coliforms; and (ii) *Vibrio fischeri* toxic response was negatively correlated with colour and DOC, and positively with Zn and fecal coliforms.

Figures 2 and 3 present the comparison between toxicity experimental and calculated values (TU_{50}) from MLR and PCR, respectively. Table 4 shows the performance indexes for MLR and PCR. MLR is the regression model that best fit the

Chlorella vulgaris and *Vibrio fischeri* toxic response in respect to the Leça river water characterization.

4. Discussion

4.1 Multiple linear regression

The MLR results for *Chlorella vulgaris* showed a negative correlation between the toxic response and the DOC, Zn and Mn parameters. DOC is extremely important in the transport of metals in aquatic systems, forming strong complexes with metals, enhancing metal solubility while also reducing metal bioavailability. Studies using multispecies laboratory bioassays proved *Chlorella vulgaris* resistance to toxicants like Zn [**34, 35**].

Turbidity is considered an important variable relative to transport and bioavailability of contaminants in natural waters [**36**]. In addition, turbidity affects the results of tests based on photometric measurements, produces light losses and leads to toxicity overestimation [**37**]. In the present study turbidity was positively related to *Chlorella vulgaris* toxic response results due to the scattering of incident light by colloidal and particulate matter in water.

The *Vibrio fischeri* toxic response, according to MLR, presented a negative relation with conductivity and turbidity. Conductivity is related to ionic concentrations and pH. The Microtox® test procedure, based on the inhibition of *Vibrio fischeri* marine bacteria, involves the addition of sodium chloride, therefore possibly changing sample ionic concentration and, consequently, metals toxic potential. This effect may be due to competition between toxic ions and chloride ions in the cellular membrane [**38**]. Some

studies showed silver toxicity diminishing with the raise of salinity up to 25‰, however for salinity above 25‰ it was observed an increase in the metal toxicity, which was attributed to osmotic imbalance caused by chloride ions [**39-41**].

The hardness, the metals Fe, Hg and As and the fecal coliforms presented a positive correlation with the toxic response of *Vibrio fischeri*. Concerning the effect of hardness on metals toxicity, it is known that the presence of calcium and magnesium carbonates in water can cause the precipitation of metals, making them insoluble and therefore not available to penetrate in the membranes of living organisms. This effect was observed for manganese chronic toxicity in aquatic species *Salmo trutta*, and also for other metals, such as copper, zinc and cadmium [**42-44**]. The hardness values obtained for Leça river were normal for surface water, and therefore the metals Fe, Hg and As contributed to global toxic effect. Nevertheless Microtox® test is especially sensitive to several metals, such as Hg, Pb, Zn and Cu [**45, 46**], the toxicity of heavy metals is highly influenced by matrix effects, conditions and concentration [**47, 48**]. The fecal coliforms in Leça river presented extremely high concentrations showing positive correlation with the *Vibrio fischeri* toxic response, probably due to competition between the bacteria, both Gram negative, heterotrophic and facultative anaerobes. This competition may be for oxygen, which would influence the luminescence produced once its mechanism is intrinsically connected to the respiratory metabolism [**49**].

4.2 Principal component regression

The PCR results for *Chlorella vulgaris* toxic response showed a negative correlation with colour and DOC parameters. In the specific case of surface water samples in the natural environment, the colour is related to high concentrations of DOC, which could explain the inclusion in the same PC (PC8). As algae absorb light energy

for photosynthesis, in coloured samples the light provided during the toxicity bioassay may be partially absorbed by the coloured compounds of the surface waters [**50**]. Arsenic, mercury and all bacteriological parameters (especially fecal coliforms) showed a positive correlation with *Chlorella vulgaris* toxic response. Algae generally are hyper-accumulators of heavy metals [1, **51-54**]. However, some studies showed that arsenic is toxic to algae but highly variable data have been reported due to different experimental conditions [**48**]. As concerns the bacteriological parameters, it would be expected a negative instead of a positive correlation once bacteria respiration releases carbon dioxide, essential for algae photosynthesis.

According to PCR, the *Vibrio fischeri* toxic response presented a negative correlation with colour and DOC. A coloured sample may potentially absorb a portion of the light produced by the *Vibrio fischeri* before it reaches the photomultiplier, and the sample may appear more toxic than it really is [**55**]. In this way, colour should present a positive and not a negative correlation. The DOC biodegradable fraction consists of organic molecules that can be used by heterotrophic bacteria, such as *Vibrio fischeri*, as a source of energy and carbon, thus contributing to bacterial metabolism. Zn and fecal coliforms presented positive correlation with *Vibrio fischeri* toxic response, which agrees with the result obtained by MLR, confirming the idea of competition between *Vibrio fischeri* and coliforms.

5. Conclusions

In order to better understand the interaction of physical, chemical and bacteriological factors involved in a multidimensional process such as the ecotoxicological response, multiple linear regression (MLR) and principal component regression (PCR) were applied to the results of *Chlorella vulgaris* and *Vibrio fischeri*

toxic response to the Leça river water characterization, both physicochemical and microbiological. In a general way, and supported by the performance indexes, the MLR seems to be the most appropriate model to the Leça river data, presenting: (i) a negative correlation with DOC, Zn and Mn, and a positive one with turbidity and As for *Chlorella vulgaris* toxic response; and (ii) a negative correlation with conductivity and turbidity, and a positive one with phosphorus, hardness, Fe, Hg, As and fecal coliforms for *Vibrio fischeri* toxic response. The results obtained may be useful in the future to evaluate the effect of pollution abatement measures over the water quality of Leça River.

Acknowledgements

J.C.M. Pires thanks the Foundation for Science and Technology, POPH-QREN and FSE for the Post-Doctoral fellowship SFRH/BPD/66721/2009.

References

[1] H.F. Olguin, A. Puig, C.R. Loez, A. Salibian, M.L. Topalian, P.M. Castane, and M.G. Rovedatti, *An integration of water physicochemistry, algal bioassays, phytoplankton, and zooplankton for ecotoxicological assessment in a highly polluted lowland river*, Water Air Soil Poll. 155 (2004), pp. 355-381.

[2] A. Picado, E. Mendonca, L. Silva, S.M. Paixao, F. Brito, M.A. Cunha, S. Leitao, I. Moura, and R. Hernan, *Ecotoxicological assessment of industrial wastewaters in Trancao River Basin (Portugal)*, Environ. Toxicol. 23 (2008), pp. 466-472.

[3] L.A. Athey, J.M. Thomas, and J.R. Skalski, *Role of acute toxicity bioassays in the remedial action process at hazardous waste sites, Office of Water-USEPA*, 1987.

[4] L. Ferrari, F.R. de la Torre, S.O. Demichelis, M.E. Garcia, and A. Salibian, *Ecotoxicological assessment for receiving waters with the premetamorphic tadpoles acute assay*, Chemosphere 59 (2005), pp. 567-575.

[5] H.I. Zeliger, *Toxic effects of chemical mixtures*, Arch. Environ. Health 58 (2003), pp. 23-29.

[6] D.B. Wei, A. Kisuno, T. Kameya, and K. Urano, *A new method for evaluating biological safety of environmental water with algae, daphnia and fish toxicity ranks*, Sci. Total Environ. 371 (2006), pp. 383-390.

[7] A.A. Meharg, *Industrial-Accidents Involving Release of Chemicals into the Environment - Ecotoxicology*, Environ. Technol. 15 (1994), pp. 1041-1050.

[8] A. Saddoud, S. Abdelkafi, and S. Sayadi, *Effects of domestic wastewater toxicity on anaerobic membrane-bioreactor (MBR) performances*, Environ. Technol. 30 (2009), pp. 1361-1369.

[9] A. Gurung, S.H.A. Hassan, and S.E. Oh, *Assessing acute toxicityof effluent from a textile industry and nearby river waters using sulfur-oxidizing bacteria in continuous mode*, Environ. Technol. 32 (2011), pp. 1597-1604.

[10] N.M. Van Straalen, and C.A.J. Denneman, *Ecotoxicological Evaluation of Soil Quality Criteria*, Ecotox. Environ. Safe. 18 (1989), pp. 241-251.

[11] USEPA, *Permit writer's guide to water quality-based permitting for toxic pollutants EPA 440/4-87-005*, US Environmental Protection Agency, Washington D. C., 1987.

[12] USEPA, *Technical support document for water quality-based control EPA-505/2-90-001*, US Environmental Protection Agency, Washington D. C., 1991.

[13] H. Dizer, E. Wittekindt, B. Fischer, and P.D. Hansen, *The cytotoxic and genotoxic potential of surface water and wastewater effluents as determined by bioluminescence, umu-assays and selected biomarkers*, Chemosphere 46 (2002), pp. 225-233.

[14] USEPA, *Methods for measuring the acute toxicity of effluents and receiving waters to freshwater and marine organisms (Fifth Edition) EPA-821-R-02-012*, US Environmental Protection Agency, Washington D. C., 2002.

[15] L. Hakanson, J.M. Malmaeus, U. Bodemer, and V. Gerhardt, *Coefficients of variation for chlorophyll, green algae, diatoms, cryptophytes and blue-greens in rivers as a basis for predictive modelling and aquatic management*, Ecol. Model. 169 (2003), pp. 179-196.

[16] H. Çamdevýren, N. Demyr, A. Kanik, and S. Keskyn, *Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs*, Ecol. Model. 181 (2005), pp. 581-589.

[17] O.J. Cacho, *Systems modelling and bioeconomic modelling in aquaculture*, Aquaculture Economical Management 1 (1997), pp. 45-64.

[18] D.L. Tufford, and H.N. McKellar, *Spatial and temporal hydrodynamic and water quality modeling analysis of a large reservoir on the South Carolina (USA) coastal plain*, Ecol. Model. 114 (1999), pp. 137-173.

[19] K.P. Singh, A. Malik, and S. Sinha, *Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques - a case study*, Anal. Chim. Acta 538 (2005), pp. 355-374.

[20] K.H. Cho, J.H. Kang, S.J. Ki, Y. Park, S.M. Cha, and J.H. Kim, *Determination of the optimal parameters in regression models for the prediction of chlorophyll-a: A case study of the Yeongsan Reservoir, Korea*, Sci. Total Environ. 407 (2009), pp. 2536-2545.

[21] S.Y. Li, S. Gu, X. Tan, and Q.F. Zhang, *Water quality in the upper Han River basin, China: The impacts of land use/land cover in riparian buffer zone*, J. Hazard. Mater. 165 (2009), pp. 317-324.

[22] S.L. Su, D. Li, Q. Zhang, R. Xiao, F. Huang, and J.P. Wu, *Temporal trend and source apportionment of water pollution in different functional zones of Qiantang River, China*, Water Res. 45 (2011), pp. 1781-1795.

[23] J.H. Kang, S.W. Lee, K.H. Cho, S.J. Ki, S.M. Cha, and J.H. Kim, *Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin*, Water Res. 44 (2010), pp. 4143-4157.

[24] F. Carvalho, L. Guilhermino, R. Ribeiro, F. Goncalves, and A.M.V.M. Soares, *Metier (Modular Ecotoxicity Tests Incorporating Ecological Relevance) .2.*

Ecotoxicity of Poorly Water-Soluble Compounds - Concentration Versus Dose, Arch. Environ. Con. Tox. 29 (1995), pp. 431-434.

[25] WDNR, *State of Wisconsin Aquatic Life Toxicity Testing Methods Manual*, 2004.

[26] L. Mutihac, and R. Mutihac, *Mining in chemometrics*, Anal. Chim. Acta 612 (2008), pp. 1-18.

[27] J.C.M. Pires, F.G. Martins, S.I.V. Sousa, M.C.M. Alvim-Ferraz, and M.C. Pereira, *Selection and validation of parameters in multiple linear and principal component regressions*, Environ Modell Softw 23 (2008), pp. 50-55.

[28] E. Archives of Environmental HealthEcotoxicology and Environmental SafetyAgirre-Basurko, G. Ibarra-Berastegi, and I. Madariaga, *Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area*, Environ Modell Softw 21 (2006), pp. 430-446.

[29] H.H. Harman, *Modern Factor Analysis, third edition*, University of Chicago Press, 1976.

[30] A.J. Hayter, H.P. Wynn, and W. Liu, *Slope modified confidence bands for a simple linear regression model*, Stat. Method. 3 (2006), pp. 186-192.

[31] S.I.V. Sousa, J.C.M. Pires, F.G. Martins, M.C. Pereira, and M.C.M. Alvim-Ferraz, *Potentialities of quantile regression to predict ozone concentrations*, Environmetrics 20 (2009), pp. 147-158.

[32] R.D. Harmel, and P.K. Smith, *Consideration of measurement uncertainty in the evaluation of goodness-of-fit in hydrologic and water quality modeling*, J. Hydrol. 337 (2007), pp. 326-336.

[33] A.I. Gomes, *Ecotoxicity evaluation in surface waters – Aplication to the water basin of Leça river*, in *Department of Chemical Engineering*, Faculdade de

Engenharia, Universidade do Porto, (Portuguese) http://repositorio-aberto.up.pt/bitstream/10216/12720/2/Texto%20integral.pdf accessed in 17th May 2012, Porto, 2007.

[34] C.R. Loez, A. Salibián, and M.L. Topalián, *Associations phytoplanctoniques indicatrices de la pollution par zinc*, Rev. Science Eau 11 (1998), pp. 315-332.

[35] C.R. Loez, M.L. Topalián, and A. Salibián, *Effects of zinc on the structure and growth dynamics of a natural freshwater phytoplankton assemblage reared in the laboratory*, Environ. Pollut. 88 (1995), pp. 275-281.

[36] P.M. Castane, M.G. Rovedatti, M.L. Topalian, and A. Salibian, *Spatial and temporal trends of physicochemical parameters in the water of the Reconquista River (Buenos Aires, Argentina)*, Environ. Monit. Assess. 117 (2006), pp. 135-144.

[37] M. Farre, E. Martinez, and D. Barcelo, *Validation of interlaboratory studies on toxicity in water samples*, Trac-Trend. Anal. Chem. 26 (2007), pp. 283-292.

[38] G. Gunkel, *Bioindikation in aquatischen Ökosystemen*, Jena, Germany, 1994.

[39] E.A. Ferguson, and C. Hogstrand, *Acute silver toxicity to seawater-acclimated rainbow trout: Influence of salinity on toxicity and silver speciation*, Environ. Toxicol. Chem. 17 (1998), pp. 589-593.

[40] C. Hogstrand, and C.M. Wood, *Toward a better understanding of the bioavailability, physiology and toxicity of silver in fish: Implications for water quality criteria*, Environ. Toxicol. Chem. 17 (1998), pp. 547-561.

[41] J.R. Shaw, C.M. Wood, W.J. Birge, and C. Hogstrand, *Toxicity of silver to the marine teleost (Oligocottus maculosus): Effects of salinity and ammonia*, Environ. Toxicol. Chem. 17 (1998), pp. 594-600.

[42] D. Hart, M. Rinker, M. Dutton, R. Gould, W. Snodgrass, and C. Hogan, *Literature review of environmental toxicity of mercury, cadmium, selenium and antimony in metal mining effluents*, Beak International Incorporated, Brampton, Ontario, 2002.

[43] W.A. Stubblefield, S.E. Brinkman, P.H. Davies, and T.D. Garrison, *Effects of water hardness on the toxicity of manganese to developing brown trout (Salmo trutta)*, Environ. Toxicol. Chem. 16 (1997), pp. 2082-2089.

[44] K.A. Williams, D.W.J. Green, and D. Pascoe, *Studies on the acute toxicity of pollutants to freshwater macroinvertebrates*, Archiv fur Hydrobiologie 102 (1985), pp. 461-471.

[45] P. Bogaerts, J. Bohatier, and F. Bonnemoy, *Use of the cilliated protozoan Tetrahymena pyriformis for the assessment of toxicity and quantitative structure-activity relationships of xenobiotics: comparison with the Microtox test*, Ecotox. Environ. Safe. 49 (2001), pp. 293-301.

[46] C.Y. Hsieh, M.H. Tsai, D.K. Ryan, and O.C. Pancorbo, *Toxicity of the 13 priority pollutant metals to Vibrio fisheri in the Microtox (R) chronic toxicity test*, Sci. Total Environ. 320 (2004), pp. 37-50.

[47] M. Farre, E. Martinez, M.D. Hernando, A. Fernandez-Alba, J. Fritz, E. Unruh, O. Mihail, V. Sakkas, A. Morbey, T. Albanis, F. Brito, P.D. Hansen, and D. Barcelo, *European ring exercise on water toxicity using different bioluminescence inhibition tests based on Vibrio fischeri, in support to the implementation of the water framework directive*, Talanta 69 (2006), pp. 323-333.

[48] T. Tisler, and J. Zagorc-Koncan, *Acute and chronic toxicity of arsenic to some aquatic organisms*, B. Environ. Contam. Tox. 69 (2002), pp. 421-429.

[49] M.T. Rodrigues, *Limistox aplication in the determination of sediment samples toxicity*, Minho University, Braga (Portuguese), 2005.

[50] V. Reginatto, *Evaluation of toxicity bioassay with algae Scenedesmus subspicatus for the study of industrial effluents*, Universidade Estadual de Campinas, Campinas, 1998, pp. 138.

[51] G.N. Reddy, and M.N.V. Prasad, *Characterization of Cadmium Binding-Protein from Scenedesmus-Quadricauda and Cd Toxicity Reversal by Phytochelatin Constituting Amino-Acids and Citrate*, J. Plant Physiol. 140 (1992), pp. 156-162.

[52] P.S. Lau, H.Y. Lee, C.C.K. Tsang, N.F.Y. Tam, and Y.S. Wong, *Effect of metal interference, pH and temperature on Cu and Ni biosorption by Chlorella vulgaris and Chlorella miniata*, Environ Technol 20 (1999), pp. 953-961.

[53] A. Toumi, M. Belkoura, S. Benabdallah, M. El Alami, L.L. Idrissi, and A. Nejmeddine, *Effect and bioaccumulation of heavy metals (Zn, Cd) on Micractinium pusillum alga*, Environ Technol 28 (2007), pp. 19-23.

[54] Z. Aksu, and T. Kutsal, *A Comparative-Study for Biosorption Characteristics of Heavy-Metal Ions with C-Vulgaris*, Environ Technol 11 (1990), pp. 979-987.

[55] C. Microbics, *A Microtox Manual - How to reduce microtox test data*, (1988), pp.

[56] A.D. Eaton, L.S. Clesceri, E.W. Rice, and A.B. Greenberg, *Standard Methods for the Examination of Water and Wastewater*, Washington, DC, 2005.

[57] USEPA, *Methods for Chemical Analysis of Water and Wastes - Method 110.2, EPA/600/4-79-020*, US Environmental Protection Agency, Washington D. C., 1983.

[58] ISO 9308-01, *Water Quality - Detection and Enumeration of Escherichia coli and Coliform Bacteria - Part 1: Membrane Filtration Method*, International Organization for Standardization, Geneva, Switzerland, 2000.

[59] ISO 7899-2, *Water Quality - Detection and Enumeration of Intestinal enterococci - Part 2: Membrane Filtration Method*, International Organization for Standardization, Geneva, Switzerland, 2000.

[60] ISO 11348-1, *Water Quality-Determination of the Inhibitory Effect of Water Samples on the Light Emission of Vibrio fischeri (Luminescent Bacteria Test)*, International Organization for Standardization, Geneva, Switzerland, 1998.

Captions of Figures

Figure 1 - Leça river basin showing the geographical location of the sampling sites

Figure 2 – Comparison between experimental values and values given by MLR and PCR models for *Chlorella vulgaris* toxic response

Figure 3 – Comparison between experimental values and values given by MLR and PCR models for *Vibrio fischeri* toxic response

Figure 1

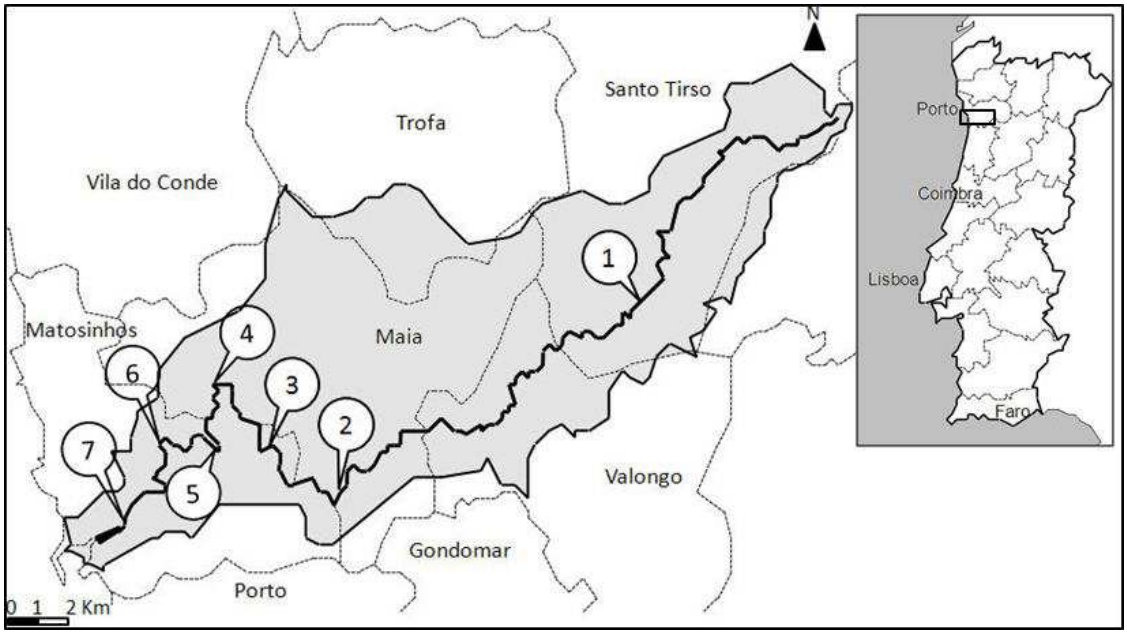
N
Trofa
Santo Tirso
Vila do Conde
Matosinhos
Maia
Valongo
Gondomar
Porto
1
2
3
4
5
6
7
0 1 2 Km
Porto
Coimbra
Lisboa
Faro

Figure 2

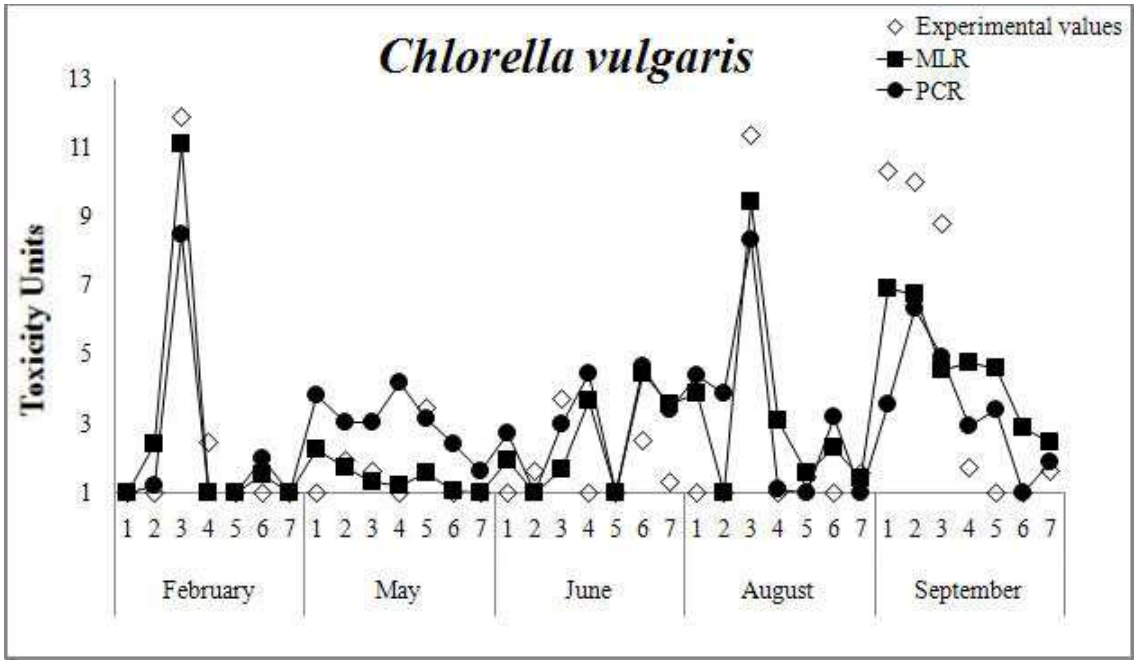
Chlorella vulgaris
Experimental values
MLR
PCR
Toxicity Units
13
11
9
7
5
3
1
1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
February
May
June
August
September

Figure 3

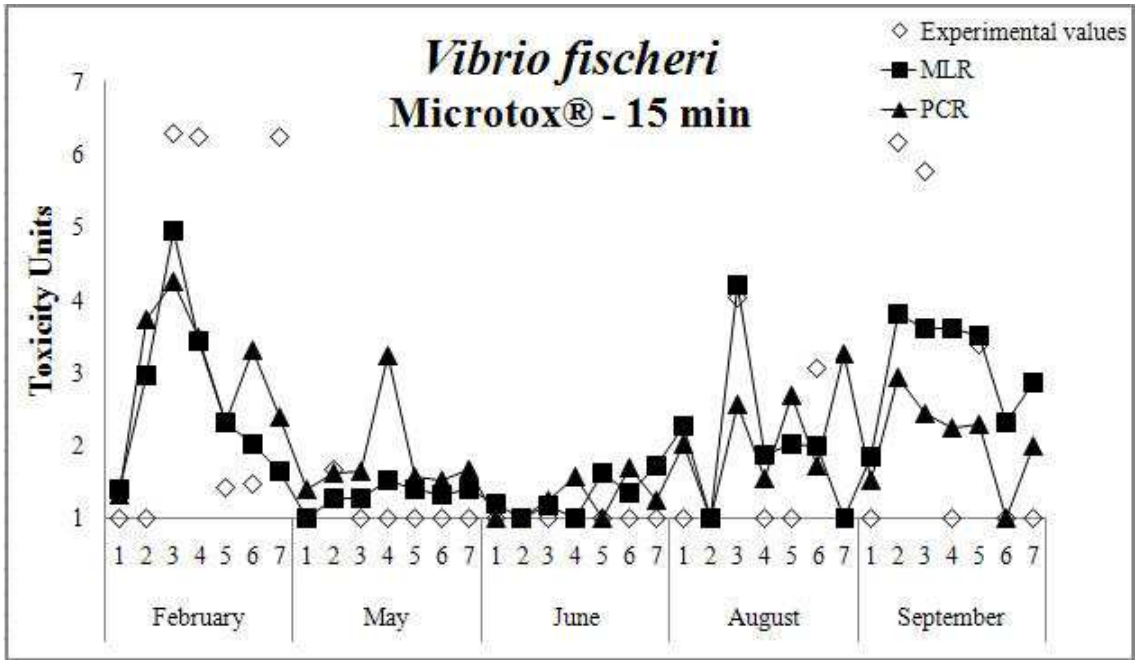
Vibrio fischeri
Microtox® - 15 min
Experimental values
MLR
PCR
Toxicity Units
7
6
5
4
3
2
1
1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7 1 2 3 4 5 6 7
February
May
June
August
September

Table 1 – Analytical procedures

Parameter	Method	Equipment
Physicochemical parameters		
Temperature	Thermometry	Multiparameter analyser HANNA Instruments model 991003
pH	Electrometry	
ORP	Electrometry	
Dissolved oxygen (DO)	Membrane electrode	DO meter HANNA Instruments model 9143
Conductivity	Conductimetry	Conductivity meter WTW model LF 330
Turbidity	Nephelometry Method 2130 B [56]	Turbiquant 3000 IR, Merck -
Colour	Sprectrophotometry (platinum-cobalt) Method 110.2 [57]	UV/Vis Spectrometer PYE Unicam PU 8600
Dissolved organic carbon (DOC)	High-temperature combustion Method 5310 B [56]	Shimadzu analyser 5000 A -
Biochemical oxygen demand (BOD)	5-Day BOD test Method 5210 B [56]	DO meter Crison OXI 45 -
Total nitrogen	Persulfate digestion Method 4500N C [56]	UV/Vis Spectrometer PYE Unicam PU 8600
Total phosphorus	Persulfate digestion + Ascorbic acid Method 4500P E [56]	UV/Vis Spectrometer PYE Unicam PU 8600
Hardness	EDTA titrimetry Method 2340 C [56]	
Dissolved Cd, Cr, Cu, Fe, Mn, Ni, Pb and Zn	Atomic absorption spectrometry - flame Methods 3111 B and D [56]	AAS GBC 932 plus
Dissolved As and Hg	Hydride generation /Cold-vapor atomic absorption spectrometry Methods 3112 B and 3114 C [56]	AAS GBC 932 plus and GBC HG 3000

Table 1 – Analytical procedures (*Continued*)

Parameter	Method	Equipment
Bacteriological parameters		
Total coliforms	Membrane filtration ISO Standard [**58**]	
Fecal coliforms	Membrane filtration ISO Standard [**58**]	
Fecal streptococcus	Membrane filtration ISO Standard [**59**]	
	Ecotoxicological parameters	
Microtox® inhibition	Bioluminiscent inhibition test of bacteria *Vibrio fischeri* (15 min) ISO Standard [**60**]	Microtox Analyzer 2055, Microbics Corporation (at present time, AZUR) Environmental)
Green algae inhibition	Inhibition growth test of microalgae *Chlorella vulgaris* USEPA Guideline [**14**]	Shimadzu UV-Vis spectrometer

Table 2 – Rotated factor loadings for all principal components (PC) of the physical, chemical and bacteriological variables

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
Conductivity (µS/cm)	**-0.896**	0.034	0.223	-0.130	0.213	0.137	-0.088	-0.194	-0.035	-0.072	0.068	0.029	0.003	-0.007	0.085
DOC (mg/L)	**-0.615**	0.052	0.415	-0.217	-0.064	0.299	0.030	-0.501	-0.019	-0.107	0.048	0.030	0.199	-0.018	-0.001
Turbidity (NTU)	0.169	-0.067	-0.011	0.062	**-0.963**	-0.082	0.159	-0.064	0.009	-0.013	-0.003	-0.001	0.001	-0.001	-0.001
Color (Pt-Co)	-0.352	0.148	0.139	-0.109	-0.091	0.218	0.332	**-0.809**	-0.002	0.040	-0.015	-0.005	-0.030	0.004	0.001
Total Nitrogen (mgN/L)	**-0.884**	0.005	0.235	0.188	0.254	0.101	-0.090	-0.075	-0.028	-0.056	0.097	-0.024	-0.029	0.159	-0.007
Total Phoshorus (mgP/L)	**-0.938**	0.007	0.028	-0.126	-0.043	0.101	-0.037	-0.143	-0.001	0.017	-0.257	-0.053	-0.025	-0.017	-0.003
Hardness ($mgCaCO_3/L$)	**-0.940**	0.013	0.178	-0.177	0.058	0.055	0.040	-0.110	0.034	0.075	0.094	0.059	0.032	-0.094	-0.061
Zn (mg/L)	0.131	-0.037	-0.056	**0.980**	-0.059	0.063	0.049	0.084	0.003	-0.009	0.003	-0.003	-0.003	0.004	0.000
Fe (mg/L)	0.146	0.184	-0.111	0.064	-0.186	0.092	**0.917**	-0.204	0.008	0.043	0.001	-0.004	0.002	-0.002	-0.001
Mn (mg/L)	-0.009	**0.978**	-0.025	-0.037	0.064	0.082	0.153	-0.089	-0.001	0.014	0.000	-0.002	0.001	0.000	0.000
Hg (µg/L)	**-0.725**	-0.083	0.181	0.051	-0.070	0.323	-0.259	0.062	0.006	-0.502	0.004	0.012	0.008	0.004	0.002
As (µg/L)	-0.296	0.120	0.206	0.094	0.120	**0.877**	0.119	-0.213	0.000	-0.052	-0.003	0.007	0.002	0.002	0.001
Total Coliforms (C.F.U./100mL)	-0.207	-0.046	**0.884**	-0.055	0.007	0.266	-0.094	-0.045	0.102	-0.067	0.046	0.268	0.014	-0.009	0.001
Fecal Coliforms (C.F.U./100mL)	-0.190	0.009	**0.905**	-0.065	0.144	0.007	-0.127	-0.079	-0.305	0.019	-0.001	-0.068	0.006	0.007	0.002
Fecal Streptococcus (C.F.U./100mL)	-0.161	-0.006	**0.960**	0.016	-0.083	0.044	0.047	-0.080	0.138	-0.021	-0.024	-0.120	-0.008	0.005	-0.001

Values in bold correspond to the greatest contributions of the original variables on the PCs.

Table 3 – Transformation matrix used to calculate the PCs from the physical, chemical and bacteriological variables

Variables	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14	PC15
Conductivity	0.367	0.020	-0.212	0.048	-0.011	0.072	-0.087	0.002	-0.138	0.041	-0.130	0.015	-0.113	-0.215	-0.841
DOC	0.354	-0.177	0.083	0.067	0.132	-0.164	0.052	0.277	-0.231	0.091	-0.117	0.009	0.601	0.519	-0.040
Turbidity	-0.103	-0.233	0.225	-0.274	0.654	0.229	0.379	0.078	0.101	-0.100	-0.285	0.180	-0.187	0.020	-0.105
Color	0.225	-0.484	0.075	0.060	0.094	-0.159	-0.249	0.548	-0.132	0.030	0.225	-0.074	-0.306	-0.323	0.193
Total Nitrogen	0.338	0.094	-0.205	-0.192	-0.128	0.236	-0.211	-0.010	-0.106	0.073	-0.103	0.432	-0.460	0.453	0.224
Total Phoshorus	0.318	-0.067	-0.303	-0.031	0.205	0.178	-0.003	-0.050	0.579	-0.184	0.346	-0.429	-0.010	0.231	0.002
Hardness	0.344	-0.039	-0.222	0.061	0.143	0.199	-0.136	-0.258	-0.117	-0.370	-0.306	0.104	0.320	-0.444	0.358
Zn	-0.098	0.018	0.066	-0.785	-0.273	0.266	-0.199	0.249	-0.057	-0.118	-0.021	-0.202	0.232	-0.075	-0.065
Fe	-0.067	-0.592	0.144	-0.094	0.007	0.043	-0.336	-0.598	-0.052	0.319	-0.050	-0.151	0.003	0.102	-0.045
Mn	0.029	-0.384	0.012	0.246	-0.504	0.518	0.509	0.072	-0.046	-0.040	0.009	-0.018	0.003	0.005	0.016
Hg	0.299	0.153	-0.163	-0.303	0.090	-0.082	0.457	-0.147	-0.204	0.577	0.188	-0.141	0.001	-0.237	0.193
As	0.240	-0.226	0.077	-0.266	-0.343	-0.576	0.240	-0.122	0.401	-0.194	-0.152	0.255	-0.031	-0.067	-0.031
Total Coliforms	0.265	0.169	0.446	-0.008	-0.039	-0.103	0.119	-0.206	-0.380	-0.403	-0.029	-0.460	-0.306	0.142	0.029
Fecal Coliforms	0.244	0.225	0.424	0.156	-0.095	0.172	-0.152	0.153	0.423	0.389	-0.462	-0.184	-0.021	-0.115	0.097
Fecal Streptococcus	0.231	0.112	0.523	-0.005	0.042	0.210	-0.078	-0.142	0.081	-0.030	0.582	0.431	0.184	-0.119	-0.095

Table 4 – Performance indexes for MLR and PCR in the fitting of the *Chlorella vulgaris* and *Vibrio fischeri* toxic responses

	MLR				PCR			
	MAE	RMSE	d_2	R^2	MAE	RMSE	d_2	R^2
Chlorella vulgaris	1.532	1.945	0.884	0.643	1.901	2.364	0.797	0.473
*Vibrio fischeri (*15 min.)	0.613	0.860	0.911	0.711	0.817	1.008	0.864	0.603