E-ISSN: 2664-6773
P-ISSN: 2664-6765
Impact Factor: RJIF 5.6
IJCBS 2025; 7(1): 57-62
www.chemicaljournal.org
Received: 10-01-2025
Accepted: 15-02-2025

Heberth Juliano Vieira
Postgraduate Program in Environmental Science and Technology, Faculty of Exact Sciences and Technology, Federal University of Grande Dourados/UFGD, Dourados/MS, Brazil

Ediléia Amancio da Silva
Postgraduate Program in Environmental Science and Technology, Faculty of Exact Sciences and Technology, Federal University of Grande Dourados/UFGD, Dourados/MS, Brazil

Corresponding Author:
Heberth Juliano Vieira
Postgraduate Program in Environmental Science and Technology, Faculty of Exact Sciences and Technology, Federal University of Grande Dourados/UFGD, Dourados/MS, Brazil

2D QSAR model based on drugs phytoxicity on lactuca sativa seeds

Heberth Juliano Vieira and Ediléia Amancio Da Silva

DOI: https://www.doi.org/10.33545/26646765.2025.v7.i1a.128

Abstract
2D QSAR model based on drugs phytoxicity on lactuca sativa seeds. The concern with monitoring the environmental impact of anthropogenic activities is current and constant. Acute toxicity tests of chemical substances on organisms are essential to help assess the impact on various environments. The information on the toxicity of pollutants to organisms is essential data, as well as the knowledge of the substance's effect on them. Data on the phytotoxicity of pollutants such as pharmaceuticals are essential in this context. In this work, the development of a QSAR-2D model for predicting the phytotoxicity of drugs on Lactuca sativa was proposed. The model, based on available data on the phytotoxicity of drugs on seed radicle elongation, uses the EC50 of these drugs and their molecular descriptors. The QSAR model, obtained using the BuildQSAR software and descriptors AATS0e, MATS8i, and minHssNH, but also showcases robustness, enhancing trust in its predictive reliability.

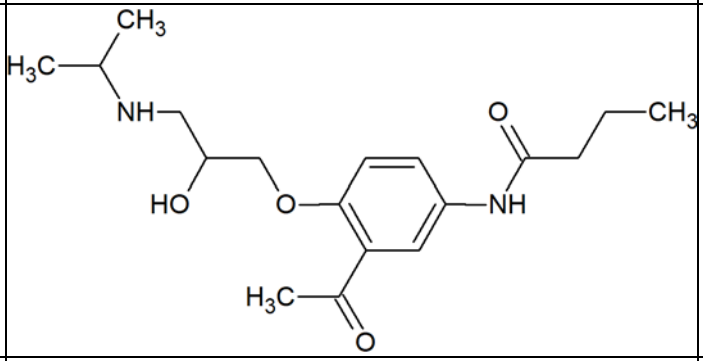
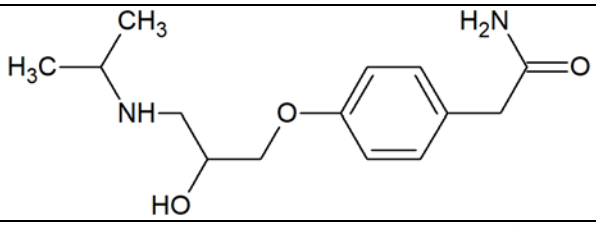
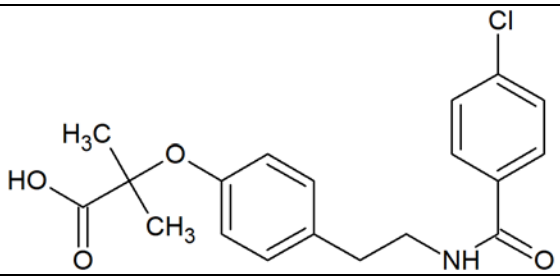
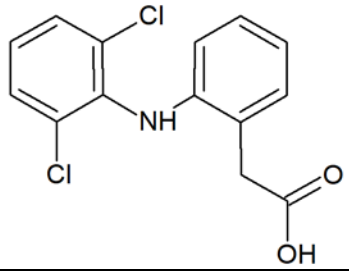
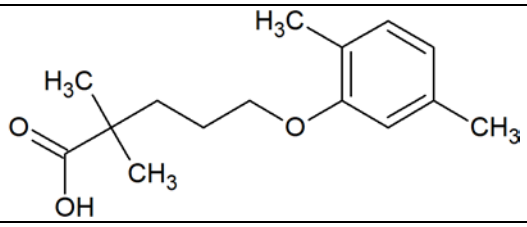
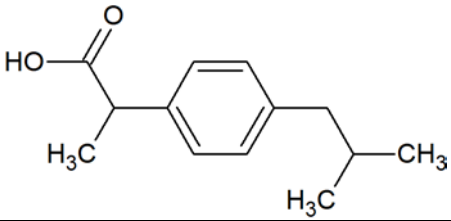
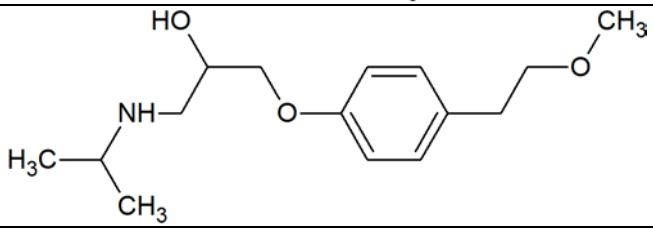
Keywords: Pharmaceuticals, molecular descriptors, biological activity, prediction model, multiple linear regression

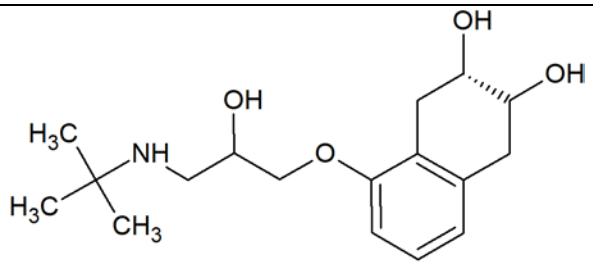
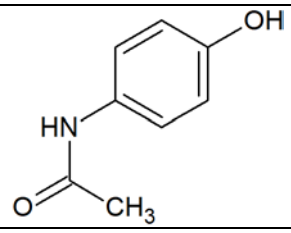
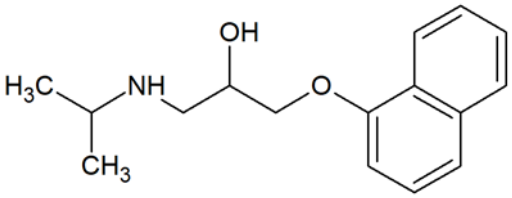
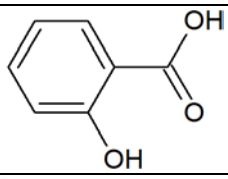
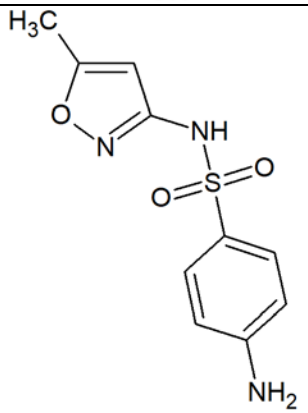
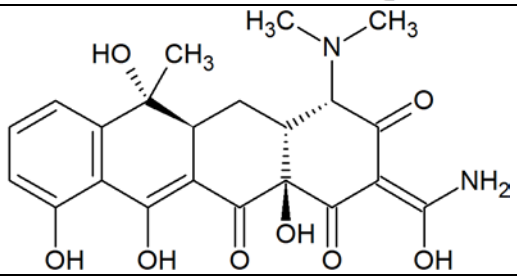
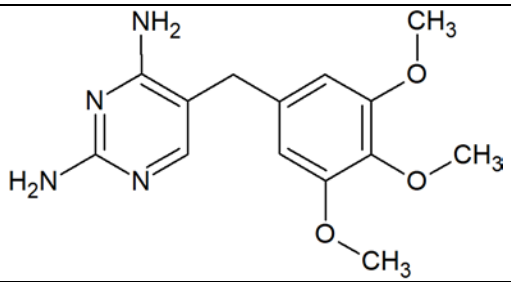
Introduction
With the constant concern of monitoring industrial and anthropogenic environmental impacts, the use of substance toxicity on various organisms as an analytical parameter has been widely utilized. Assessing the toxicity of water bodies, soils, effluents, and waste/effluents on organisms shows a strong indication of the impact caused by pollutant component substances in these environments [1, 2]. The monitoring of phytotoxicity, which assesses the effect of chemical substances on the germination of *Lactuca sativa* seeds, has been a key biological test in evaluating the efficiency of toxicity attenuation in remediation technologies for various effluents and waste. Published works have shown a strong correlation between different pharmaceuticals in impacted water samples and the germination of *Lactuca sativa* seeds, underscoring the practical application of this research. The results of different remediation processes in various samples were evaluated. Works involving the degradation of phenacetin and atenolol [3], and metformin [4] were described, demonstrating the applicability of the biological test using the germination of *Lactuca sativa*. Ozone-assisted degradation processes for substances have been described for the degradation of tannery wastewater [5], pentachlorophenol [6], and carbamazepine [7]. The germination of Lactuca sativa seeds was applied in the evaluation of toxicity attenuation in photolytic-assisted processes of drugs such as metformin [4], sulfamethoxazole [8], triclosan [9], and diclofenac [10]. QSAR modeling, or Quantitative Structure-Activity Relationship, is a computational method that relates the molecular structure of a given chemical substance to its biological properties in a specific organism. Based on data from the acute toxicity of substances, several methods have been described, obtaining the molecular descriptors of these substances with the available computational methods [11, 12]. Some studies have developed QSAR models for predicting the acute toxicity of chemical substances using Danio rerio. [13, 17] and Daphnia magna [17, 18]. Understanding a substance's toxicity on a given organism is crucial for assessing its environmental impact. The main objective of this work is to develop a QSAR model for predicting the acute phytotoxicity of drugs in the germination of *Lactuca sativa* seeds using 2D molecular descriptors. The successful development of this model will significantly contribute to our ability to assess and mitigate the environmental impact of pharmaceuticals.

Experimental Procedures

Activity data: The biological activity data used in this work consists of 14 pharmaceuticals tested in the germination of the *Lactuca sativa* seed, shown in *Table 1*. The biological activity used in this work is the EC50, the concentration that causes a 50% decrease in the length of the radicle elongation. [20, 21].

Table 1: Pharmaceuticals used to determine pEC50 in germination of *Lactuca sativa*

	pharmaceutical	structure	pEC50[(a)]	logP[(b)]
1	Acebutolol		1.85	1.71
2	Atenolol		2.42	0.16
3	Bezafribate		2.56	3.97(*)
4	Diclofenac		2.53	4.2
5	Gemfibrozil		2.44	3.89
6	Ibuprofen		2.31	3.97
7	Metoprolol		1.95	1.88

8	Nadolol		2.49	0.81
9	Paracetamol		1.73	0.46
10	Propranolol		3.07	3.48
11	Salicylic Acid		2.91	2.26
12	Sulfamethoxazole		2.40	0.89
13	Tetracycline		2.99	-1.37
14	Trimethoprim		3.46	0.91

(a) Data obtained from reference [20];

(b) Partition coefficient values were obtained from https://pubchem.ncbi.nlm.nih.gov/. When more than one logP value was available, the median was considered.

(*) The logP for bezafibrate was obtained using the ALOGPS2.1 software found on the VCCLAB platform (https://vcclab.org/lab/alogps/).

Molecular descriptors and QSAr model

All the 3D pharmaceutical structures were obtained in https://pubchem.ncbi.nlm.nih.gov/ and incorporated into PaDEL Descriptor Computation software for the obtaining of 2D structural descriptors [22]. A total number of 1,038 2D descriptors were obtained with PaDEL. The next step removed molecular descriptors with constant and highly intercorrelated values, taking the variance and correlation coefficient. Before selecting variables, the descriptor data set was centered on the mean. The reduced descriptor data set has been used to select molecular descriptors with the software BuildQSAr [23]. At this stage, variables were selected using a genetic algorithm, aiming for three variables per model. After selecting the variables, the QSAR model was obtained and validated using R software. The values of the obtained descriptors were mean-centered for the construction of the multivariate calibration model.

Discussion

Some hypotheses about the basis of the toxicity of these drugs in Lactuca sativa seeds can be found [24]:

a) Interference with essential metabolic processes: Some drugs, such as diclofenac and propranolol, can interfere with crucial metabolic processes in seeds, such as cellular respiration and protein synthesis. These pharmaceuticals can alter the enzymatic activity or generate reactive oxygen species (ROS), leading to oxidative stress.
b) Alteration in cell membrane permeability: Drugs such as bezafibrate and metoprolol can affect the integrity of seed cell membranes. This can result in the loss of essential ions and nutrients, compromising seedling germination and initial growth.
c) Accumulation of toxic compounds in the root environment: Drugs in the growing medium can lead to the accumulation of toxic compounds around the seeds. These compounds can inhibit water absorption and nutrients, impairing root development and plant growth.

The biological activity of chemical substances in organisms occurs due to a series of structural factors. Different chemical structures lead to different chemical properties, which translocate the chemical substance through different cellular barriers in organisms [25]. In this transport, substances can alter the structures of cellular barriers, leading to metabolic changes, or they can act as enzyme inhibitors or modifiers of cellular chemical composition. It is challenging to determine which chemical structure causes the acute toxicity of the substances, but molecular descriptors can be employed to explain each substance's action. In this work, after the initial selection of descriptors, the following molecular descriptors were found to be significant:

- AATS0e: Average Broto-Moreau autocorrelation - lag 0 / weighted by Sanderson electronegativities;
- MATS8i: Moran autocorrelation - lag 8 / weighted by first ionization potential;
- minHssNH: Minimum atom-type H E-State: -NH-.

The AATS0e descriptor considers the electronic property of atomic partial charge. It evaluates the partial charges of the atoms in the molecule, that is, the distribution of electric charge within each atom, based on factors such as electronegativity and chemical connectivity calculated based on a topological distance of zero. The MATS8i descriptor uses atomic polarizability, which reflects the ability of an atom to deform in response to an electric field, which is relevant for understanding molecular interactions. In the case of MATS8i, this property is evaluated over a topological distance of 8 bonds in the molecule. The molecular descriptor minHssNH refers to the minor hydrogen charge on a nitrogen atom connected to hydrogen within a molecule. In simpler terms, it identifies the hydrogen associated with a nitrogen atom with the lowest charge density (or partial charge) in a molecule.

The model obtained after variable selection can be described by the equation below

pEC50 = − 0.2698 (±0.1353) ATS0e - 0.1978 (±0.1407) MATS8i 4 - 0.1656 (±0.1372) minHssNH + 2.5091 (± 0.1257).

The statistical evaluation of the obtained QSAr model was carried out in sequence. In evaluating the obtained QSAr model, the relative errors of the calculated pEC50 of drugs 11, 12, 13, and 14 were 8.1, 5.3, 1.7, and 2.0%, respectively. The correlation between the predicted and observed values of the phytotoxicity endpoint is shown in Figure 1. The QSAr model presented a coefficient of determination of R2 = 0.853, indicating that the model explains 85% of the data variability. The analysis of variance of the obtained QSAr model was statistically significant at the 95% level since $F(2,10)_{calc}$ = 19.37 ($F(2,10)_{tab}$ = 4.10), indicating that the model is reliable. The multiple regression analysis showed adequate residuals for determining the pEC50 of drugs on the germination of *Lactuca sativa* seeds. The residues are randomly dispersed over the determination range evaluated by the model. The drugs atenolol (02) and metoprolol (06) were those that presented a relative error of −13.5 and 12%.

The descriptors used to obtain the QSAr model did not show a significant correlation. The following correlation coefficients were obtained between the variables: ATS0e and MATS8i (R=0.259), ATS0e and minHssNH (R=0.137), and MATS8i and minHssNH (R=0.303). The individual correlations between the molecular descriptors and acute toxicity for ATS0e, MATS8i, and minHssNH were −0.712, −0.658, and −0.544, respectively.

Cross-validation was used to assess the degree of predictability of the QSAr model. The regression presented a Spress cross-validation standard deviation of 0.702 and a cross-validation correlation value of Q2=0.769, indicating good predictive capacity. The correlation between pEC50 predicted and observed pEC50 obtained in the germination of *Lactuca sativa* is showed in figure 1. The QSAr model allows the prediction capacity of 77% of the biological activity of drugs not included in the model. The RMSEP calculated by comparing the predictions made by the model with those of a validation set was 0.304. The RMSE of the proposed QSAr model was calculated as 0.305, indicating the overall accuracy of the model. The QSAr model is good at predicting the biological activity of drugs on the germination of Lactuca sativa seeds.

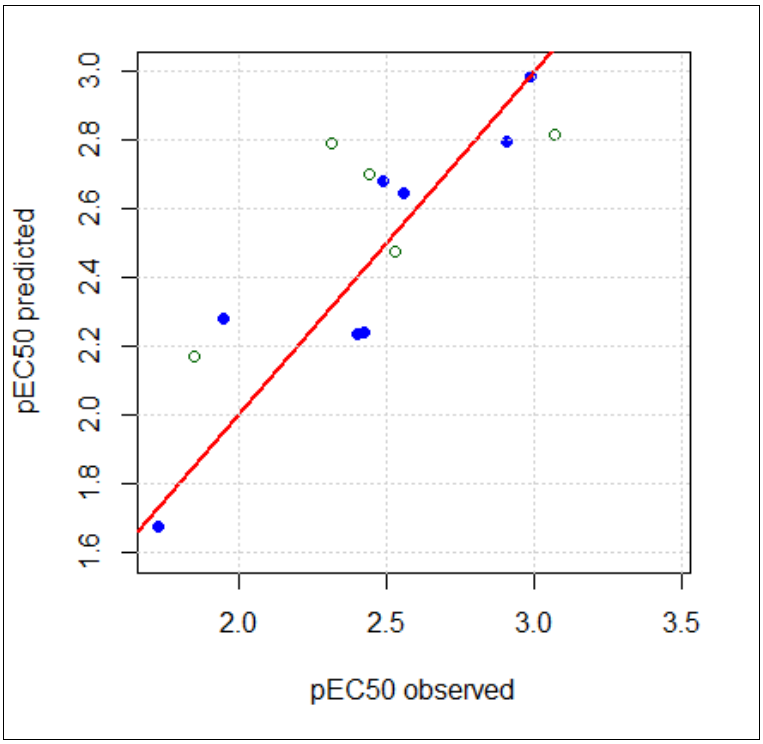


Fig 1: Correlation between pEC50 predicted and observed pEC50 obtained in the germination of *Lactuca sativa*. white circle: training data; blue circle: test data.

Conclusion

The QSAR model developed from acute phytotoxicity data of 14 drugs reported in the literature showed high accuracy. The approach employed demonstrated effectiveness in obtaining results, evidenced by the adequate adjustment of the regression and the satisfactory predictive capacity. The proposed model, built using open-source software, proved suitable for predicting the acute phytotoxicity of a specific class of drugs on Lactuca sativa seeds.

Acknowledgments

The authors would like to thank CAPES for the scholarship provided to Ediléia Amancio da Silva.
The authors have no conflict of interest.

References

1. Khan AH, Aziz HA, Palaniandy P, Naushad M, Cevik E, Zahmatkesh S. Pharmaceutical residues in the ecosystem: Antibiotic resistance, health impacts, and removal techniques. Chemosphere. 2023; 339:139647.
2. Shukla A, Ross R, Bhattacharya B, Stumpf A. Autonomous water sampling and quality monitoring in remote locations: A novel approach using a remote-controlled boat. HardwareX. 2025; e00634.
3. Cardoso R, da Silva TF, Cavalheri PS, Machado BS, Nazario CED, Machulek Junior A, *et al.* Effective degradation of phenacetin in wastewater by (photo) electro-Fenton processes: Investigation of variables, acute toxicity, and intermediates. J Environ Chem Eng. 2024 Jun; 12:112704.
4. Carbuloni CF, Savoia JE, Santos JSP, Pereira CAA, Marques RG, Ribeiro VAS, *et al.* Degradation of metformin in water by TiO2-ZrO2 photocatalysis. J Environ Manage. 2020 May; 262:110347.
5. Aguilar-Ascon E, Marrufo-Saldana L, Barra-Hinojosa JA. Toxicity Assessment of Tanning Effluents Treated via Electrocoagulation and Ozonation Using a Bioassay with Lactuca sativa L. J Ecol Eng. 2024; 25:316-327.
6. Quispe C, Valdes C, Delgadillo A, Villasenor J, Cheel J, Pecchi G. Toxicity Studies During The Degradation Of Pentachlorophenol By Ozonation In The Presence Of Mno2/Tio2. J Chil Chem Soc. 2018; 63:4090-4097.
7. Aguilar CM, Vazquez-Arenas J, Castillo-Araiza OO, Rodriguez JL, Chairez I, Salinas E, *et al.* Improving ozonation to remove carbamazepine through ozone-assisted catalysis using different NiO concentrations. Environ Sci Pollut Res. 2020 Jun; 27:22184-22194.
8. Al-Maqdi KA, Hisaindee S, Rauf MA, Ashraf SS. Detoxification and degradation of sulfamethoxazole by soybean peroxidase and UV + H2O2 remediation approaches. Chem Eng J. 2018 Nov; 352:450-458.
9. Kosera VS, Lumbaque EC, Dallegrave A, Gomes MF, de Paula VCS, de Freitas AM, *et al.* A Comparison of the Photolytic and Photocatalytic Degradation of Triclosan: Identification of Transformation Products and Ecotoxicity Evaluation. J Braz Chem Soc. 2021 Aug; 32:1531-1540.
10. Ardila PLK, da Silva BF, Spadoto M, Rispoli BCM, Azevedo EB. Which route to take for diclofenac removal from water: Hydroxylation or direct photolysis? J Photochem Photobiol A Chem. 2019 Sep; 382:111879.
11. OECD. (Q)SAR Assessment Framework: Guidance for the regulatory assessment of (Quantitative) Structure-Activity Relationship models, predictions, and results based on multiple predictions. Paris: Organisation for Economic Co-operation and Development; 2023.
12. Barreiro EJ, Rodrigues CR, Albuquerque MG, Sant'Anna CMRd, Alencastro RBd. Modelagem Molecular: Uma Ferramenta para o Planejamento Racional de Fármacos em Química Medicinal. Quim Nova. 1997 May; 20:300-310.
13. Lovric M, Malev O, Klobucar G, Kern R, Liu JJ, Lucic B. Predictive Capability of QSAR Models Based on the CompTox Zebrafish Embryo Assays: An Imbalanced Classification Problem. Molecules. 2021 Mar; 26:1617.
14. Castanha RF, Pereira AdES, Villarreal GPU, Vallim JH,

Pertrini FS, Jonsson CM, *et al*. Ecotoxicity studies of two atrazine nanoformulations: From the evaluation of stability in media to the effects on aquatic organisms. Environ Pollut. 2023 Oct; 335:122235.
15. De Almeida MC, Machado MR, Costa GG, de Oliveira GAR, Nunes HF, Veloso DFMC, *et al.* Influence of different concentrations of plasticizer diethyl phthalate (DEP) on toxicity of Lactuca sativa seeds, Artemia salina and Zebrafish. Heliyon. 2023 Sep; 9:e18855.
16. Nath A, De P, Roy K. In silico modelling of acute toxicity of 1, 2, 4-triazole antifungal agents towards zebrafish (Danio rerio) embryos: Application of the Small Dataset Modeller tool. Toxicol *In Vitro*. 2021 Sep; 75:105205.
17. Roy J, Roy K. Insights into nanoparticle toxicity against aquatic organisms using multivariate regression, read-across, and ML algorithms: Predictive models for Daphnia magna and Danio rerio. Aquat Toxicol. 2024 Nov; 276:107114.
18. Khan K, Khan PM, Lavado G, Valsecchi C, Pasqualini J, Baderna D, *et al.* QSAR modeling of Daphnia magna and fish toxicities of biocides using 2D descriptors. Chemosphere. 2019 Aug; 229:8-17.
19. Trinh TX, Seo M, Yoon TH, Kim J. Developing random forest based QSAR models for predicting the mixture toxicity of TiO2 based nano-mixtures to Daphnia magna. NanoImpact. 2022; 25:100383.
20. Rosa Pino M, Muniz S, Val J, Navarro E. Phytotoxicity of 15 common pharmaceuticals on the germination of Lactuca sativa and photosynthesis of Chlamydomonas reinhardtii. Environ Sci Pollut Res. 2016 Nov; 23:22530-22541.
21. Sobrero MC, Ronco A. Ensayo de toxicidad aguda con semillas de lechuga Lactuca sativa L. Ensayos toxicológicos y métodos de evaluación de calidad de aguas: estandarización, intercalibración, resultados y aplicaciones. IMTA. 2004:63-70.
22. Yap CW. PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. J Comput Chem. 2011; 32:1466-1474.
23. De Oliveira DB, Gaudio AC. BuildQSAR: A New Computer Program for QSAR Analysis. Quant Struct-Act Relat. 2000; 19:599-601.
24. Geng N, Wu Y, Zhang M, Tsang DCW, Rinklebe J, Xia Y, *et al.* Bioaccumulation of potentially toxic elements by submerged plants and biofilms: A critical review. Environ Int. 2019; 131:105015.
25. Wang Y, Li X, Chen S, Yang J, Fang B, Chen H, *et al.* Structure-Dependent Distribution, Metabolism, and Toxicity Effects of Alkyl Organophosphate Esters in Lettuce (Lactuca sativa L.). Environ Sci Technol. 2024 Sep; 58:17441-17453.
26. Ronco A, Sobrero C, Grassi V, Kaminski L, Massolo L, Mina L. WaterTox bioassay intercalibration network: Results from Argentina. Environ Toxicol. 2000 Sep; 15:287-296.
27. Ronco A, Gagnon P, Diaz-Baez MC, Arkhipchuk V, Castillo G, Castillo LE, *et al.* Overview of results from the WaterTox intercalibration and environmental testing phase II program: Part 1, statistical analysis of blind sample testing. Environ Toxicol. 2002 Jun; 17:232-240.