Chemosphere 152 (2016) 190–195

Contents lists available at ScienceDirect

Chemosphere

journal homepage: www.elsevier.com/locate/chemosphere

Quantitative modeling of bioconcentration factors of carbonyl herbicides using multivariate image analysis

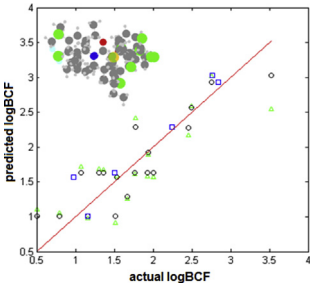
Mirlaine R. Freitas, Stephen J. Barigye, Joyce K. Daré, Matheus P. Freitas*

Department of Chemistry, Federal University of Lavras, P.O. Box 3037, 37200-000, Lavras, MG, Brazil

HIGHLIGHTS

- Image-based descriptors describe the bioconcentration factors of carbonyl herbicides.
- A three-parameter linear equation yielded a predictive and validated QSPR model.
- Interpretation of built model provided information on structural features affecting BCF.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 21 October 2015
Received in revised form
3 March 2016
Accepted 4 March 2016
Available online xxx

Handling Editor: I. Cousins

Keywords:
Bioconcentration
Carbonyl herbicides
QSPR
Image analysis

ABSTRACT

The bioconcentration factor (BCF) is an important parameter used to estimate the propensity of chemicals to accumulate in aquatic organisms from the ambient environment. While simple regressions for estimating the BCF of chemical compounds from water solubility or the *n*-octanol/water partition coefficient have been proposed in the literature, these models do not always yield good correlations and more descriptive variables are required for better modeling of BCF data for a given series of organic pollutants, such as some herbicides. Thus, the logBCF values for a set of carbonyl herbicides comprising amide, urea, carbamate and thiocarbamate groups were quantitatively modeled using multivariate image analysis (MIA) descriptors, derived from colored image representations for chemical structures. The logBCF model was calibrated and vigorously validated ($r^2 = 0.79$, $q^2 = 0.70$ and $r^2_{test} = 0.81$), providing a comprehensive three-parameter linear equation after variable selection (logBCF = 5.682 − 0.00233 × **X9774** − 0.00070 × **X813** − 0.00273 × **X5144**); the variables represent pixel coordinates in the multivariate image. Finally, chemical interpretation of the obtained models in terms of the structural characteristics responsible for the enhanced or reduced logBCF values was performed, providing key leads in the prospective development of more eco-friendly synthetic herbicides.

© 2016 Published by Elsevier Ltd.

1. Introduction

The excessive use of herbicides in agriculture can be an important cause of health risks and hazards in animal and human life (Corsonlini et al., 2005). Some herbicides are persistent organic pollutants and, therefore, they can be strongly sorbed in soil and accumulate in the trophic food chain. On the other hand, herbicides

Abbreviations: BCF, Bioconcentration factor; MLR, Multiple Linear Regression; LOOCV, Leave-one-out Cross-validation; MIA, Multivariate Image Analysis; RGB, Red-Green-Blue; RMSE, Root Mean Square Error; QSPR, Quantitative Structure-Property Relationships.

* Corresponding author.
E-mail address: matheus@dqi.ufla.br (M.P. Freitas).

http://dx.doi.org/10.1016/j.chemosphere.2016.03.011
0045-6535/© 2016 Published by Elsevier Ltd.

can also reach water resources, and thus affecting the aquatic life, due to their propensity to be absorbed through the aquatic organisms' respiratory and dermal surfaces (Gobas and Morrison, 2000). This process is denominated as bioconcentration and is quantified in terms of the equilibrium ratio of the concentration of the chemical substance in the organism to that in the surrounding aquatic environment, known as the bioconcentration factor (BCF) (Mackay, 1982; Chiou et al., 1977; Arnot and Gobas, 2006). Due to the enormous costs incurred in performing standard bioconcentration tests as established by OECD 305 (OECD, 1996), many efforts have been invested in developing *in silico* QSPR models for estimating the BCFs of chemical compounds. Several *in silico* BCF models have been proposed in the literature ranging from simple regression equations using the physicochemical parameters solubility or *n*-octanol/water partition coefficient to more complex models based on topological molecular descriptors and advanced statistical and machine learning methods (Devillers et al., 1998). However, each approach comes along with inherent advantages and limitations. While simple and interpretable models are desirable, they usually possess lower predictive power and robustness and vice versa. The present manuscript is an attempt to strike a tradeoff between the two extremes, using the Multivariate Image Analysis applied to quantitative structure-property relationships (MIA-QSPR) method approach, with the ultimate goal of providing key leads on the structural features necessary for developing environmentally friendly herbicides with a reduced tendency to bioconcentrate in aquatic organisms.

The multivariate image analysis applied to quantitative structure-property relationships (MIA-QSPR) method has proved to be a successful technique in the prediction of the soil sorption of herbicides when this parameter does not adequately correlate with the octanol/water partition coefficient (Freitas et al., 2014, 2015). This is an alignment-based method centered on the notion that 2D images of chemical structures contain important topostructural and topochemical information, which may be employed in explaining the variance in physicochemical, chemical and biological properties of molecules (Freitas et al., 2005, 2008). From a digital image processing perspective, images are simply an array (or grid) of pixels and, therefore, in order to extract relevant chemical information from the molecular structure images, these pixels (*i.e.* their numeric values and coordinate positions) are considered as descriptors for the corresponding chemical structure.

The pioneering MIA-QSPR framework considered binary (black and white) wire-frame graphs of chemical structures. However, despite the reasonably good correlations obtained in modeling experiments with this framework, it exhibited redundancy due to the inability to discriminate heteroatoms/atom types or to use atomic property weighting schemes. Consequently, the **aug-MIA-QSPR** (acronym of **aug**mented **MIA-QSPR**) approach was introduced, distinguished by the incorporation of information on atom types and groups, atomic size, as well as different atomic properties (such as van der Waals radii and electronegativity), by means of different color schemes carefully defined according to the RGB system of colors. This method, which was recently implemented in QSAR and QSPR experiments (aug-MIA-QSAR/QSPR), yields more robust and, most importantly, interpretable models (Nunes and Freitas, 2013; Duarte et al., 2015a, 2015b). For a detailed review on this method, see Barigye and Freitas (2016).

Carbonyl herbicides, such as amides, ureas, carbamates and thiocarbamates are extensively used in the agricultural industry (Grover and Cessna., 1990). For instance, thiocarbamates have been shown to form reactive sulfoxide and sulfone intermediates, which may be involved in toxic reactions through covalent modification of cysteine and serine sites of enzymes (Zimmerman et al., 2004). Thus, it is important to comprehend the structural characteristics responsible for increased or decreased logBCF values. In this sense, the aim of the present report is to build aug-MIA-QSAR based models for the logBCF of a series of carbonyl herbicides to provide insight on structural moieties responsible for the observed BCF profiles within the series of compounds.

2. Materials and methods

The experimental logBCF values of 25 carbonyl herbicides pertaining to the amide, urea, carbamate and thiocarbamate classes were obtained from the literature (see Table 1 for identity and Supplementary Material for the corresponding structures) (Higgins et al., 2010; Grisoni et al., 2016; Mackay et al., 1997; Bermúdez-Saldaña et al., 2005; Jackson et al., 2009; Asano et al., 1989; Vogs et al., 2015).

The MIA-QSPR methodology is a simple and computationally cheap procedure briefly defined as follows: for a given dataset of chemical compound images, these are aligned with respect to a common basic scaffold, and then each image (grid of pixels) is unfolded into a single row pixel values, and thus the entire dataset yields a bi-dimensional matrix of descriptors which is then employed to explain the variance of the modeled property (Freitas et al., 2014, 2015, 2005, 2008; Nunes and Freitas, 2013; Duarte et al., 2015a, 2015b; Barigye and Freitas, 2016). In this sense, to generate the aug-MIA descriptors for carbonyl herbicides employed in the present study, firstly, the chemical structures of the herbicides were drawn using the GaussView program (Dennington et al., 2008) and saved separately as bitmap files, keeping the common substructure (the carbonyl group) aligned. Fig. 1 shows the 25 chemical images superposed to give insight on the structural variance that explains the variance in the logBCF block. The chemical structures were drawn by considering atoms as spheres with sizes proportional to the corresponding van der Waals radii. In addition, the atom types were colored differently to distinguish them, since different numbers are assigned to each color pixel, consistent with the RGB system of colors. According to the RGB model, the entire color spectrum is obtained from the contribution of red (255), green (255) and blue (255) components, thus varying from 0 (black, absence of color) to 765 (white, the sum of all three original components). In this sense, the images in this series of compounds comprised the following pixel values: 765 (blank space, white), 688 (fluorine, light blue), 615 (chemical bond, light grey), 612 (hydrogen, light grey), 493 (sulfur, yellow), 426 (carbon, grey), 289 (chlorine, green), 279 (nitrogen, blue) and 229 (oxygen, red). Since images are simply 2D grids of pixels, each image was expressed as a matrix of pixel values and unfolded to form a row vector, therefore, yielding a 25 row dataset matrix. Fig. 2 summarizes the procedure followed in generating the descriptors according to the aug-MIA-QSAR strategy. In general, the MIA-QSPR approach yields thousands of descriptors. However, many of these descriptors are in fact redundant given that they do not usually vary for a given series of compounds (particularly due to blank spaces or identical substructures), are strongly correlated, or simply do not explain the variance in the Y-block, and are thus not exactly useful in modeling tasks. In this sense, the obtained data matrix was pruned by removing columns with zero variance (representing blank spaces and identical substructures) and using a step-wise variable selection procedure, based on supervised and unsupervised filters (Urias et al., 2015).

The chemical dataset was split into training (*ca.* 75%, 19 compounds) and test sets (*ca.* 25%, 6 compounds) using random sampling, and subsequently a search for the best 3 variable model for the logBCF was performed using the Multiple Linear Regression method coupled with the Genetic Algorithm (MLR-GA). The built QSPR model was validated using leave-one-out cross-validation

Table 1
Carbonyl herbicides and respective logBCF values used in the aug-MIA-QSPR modeling.

Compound no.	Herbicide	logBCF	Compound no.	Herbicide	logBCF
1	Chlortoluron	1.156[b]	**14**	Butachlor	2.430[f]
2	Diuron	1.400[g]	**15**	Metolachlor	1.839[e]
3	Fenuron	0.500[b]	**16**	Propachlor[a]	1.570[c]
4	Fluometuron	1.610[c]	**17**	Propanil[c]	1.940[g]
5	Isoproturon	1.510[g]	**18**	Diphenamid[a]	1.420[d]
6	Monuron	1.260[d]	**19**	Pronamide	2.380[d]
7	Neburon	2.457[g]	**20**	Butylate	2.610[c]
8	Monolinuron	1.301[c]	**21**	Diallate	3.180[d]
9	Linuron	1.362[c]	**22**	Triallate[a]	3.150[c]
10	Barban	1.924[b]	**23**	EPTC	1.900[g]
11	Chlorpropham[a]	2.160[c]	**24**	Pebulate[a]	2.680[d]
12	Propham	1.302[g]	**25**	Molinate[a]	1.415[g]
13	Alachlor	2.450[d]			

[a] Test set compounds chosen by random sampling.
[b] Estimated using the BCFWIN model according to Meylan et al. (1999) as experimental BCF values were unavailable in the literature.
[c] Experimental values obtained from dataset of wet-weight BCF compiled by Grisoni et al. by merging three datasets, *i.e.* VEGA CAESAR, VEGA Meylan and VEGA Read-across (Grisoni et al., 2015).
[d] Data retrieved from the Toxics Release Inventory built according to Vesar and ACQUIRE (Versar Inc., 1990; U.S. Environmental Protection Agency (EPA), 1995).
[e] *Lepomis macrochirus* employed as model (Higgins et al., 2010; Grisoni et al., 2016; Mackay et al., 1997; Bermúdez-Saldaña et al., 2005; Jackson et al., 2009; Asano et al., 1989; Vogs et al., 2015).
[f] *Cyprinus carpio* employed as model (Higgins et al., 2010; Grisoni et al., 2016; Mackay et al., 1997; Bermúdez-Saldaña et al., 2005; Jackson et al., 2009; Asano et al., 1989; Vogs et al., 2015).
[g] Fish models employed not referred (Higgins et al., 2010; Grisoni et al., 2016; Mackay et al., 1997; Bermúdez-Saldaña et al., 2005; Jackson et al., 2009; Asano et al., 1989; Vogs et al., 2015).

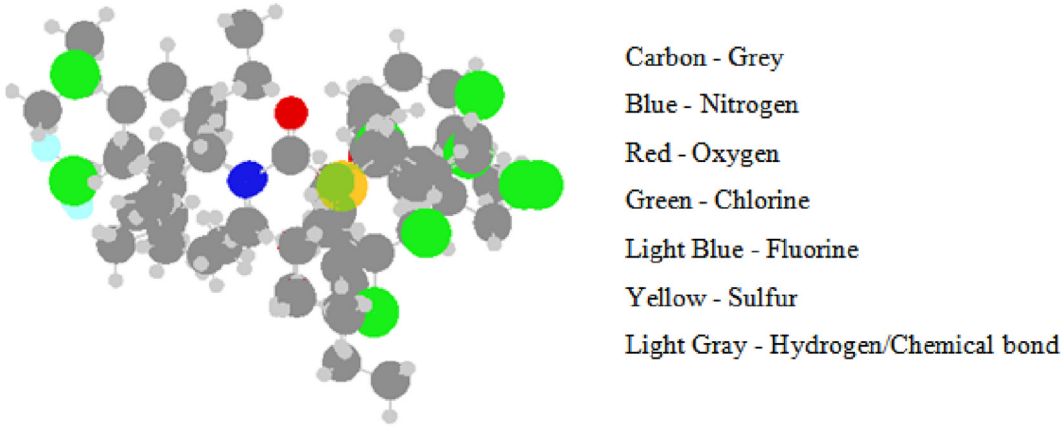


Fig. 1. Superposed chemical images used to generate the aug-MIA descriptors.

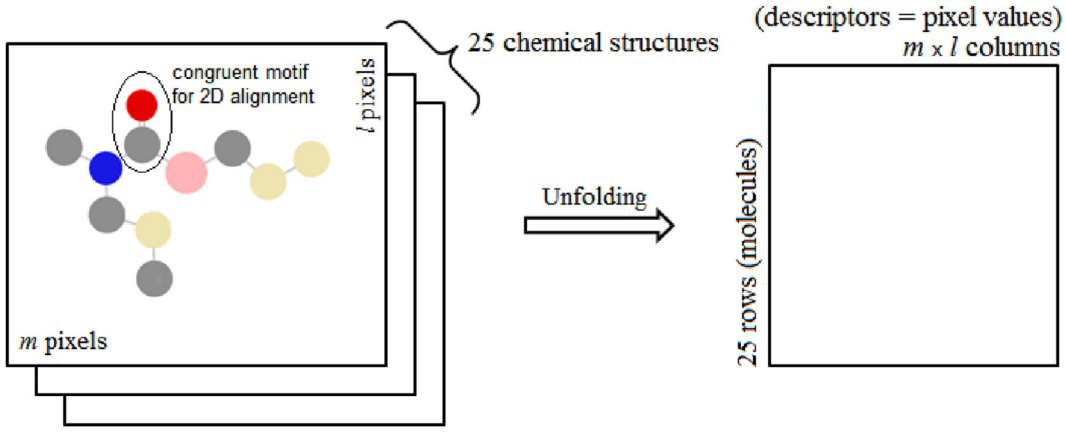


Fig. 2. Illustration of workflow followed in the generation of aug-MIA descriptors.

(LOOCV) and external validation procedures. Other measures considered in the assessment of the quality of the built model include: the determination coefficient between actual and predicted logBCF (q^2 and r^2_{test}), root mean square error of prediction (RMSECV and RMSEP) and the modified r^2_{test} (r^2_m) parameter, according to the criteria established in the literature (Golbraikh and Tropsha, 2002; Roy et al., 2013). In addition, the reliability of the model was attested using the y-randomization test [analyzed in terms of the corrected penalized r^2 ($^cr^2_p$)] (Mitra et al., 2010), in which the y-block is shuffled and regression performed to verify the inexistence of chance correlation. The image treatment and statistical analysis were performed using the Chemoface program (Nunes et al., 2012).

On the other hand, in order to compare the performance of the aug-MIA descriptors with the physicochemical parameters hydrophobicity (logP) and solubility (at pH = 6.5), which are classically used in simple regression models to estimate the BCF of chemical compounds, these were calculated for the dataset employed in this study using the Percepta module of the ACD/Labs program (ACD/Percepta, 2013) and correlated with the logBCF values.

3. Results and discussion

Although the logBCF of chemical compounds is often estimated by simple linear regressions with water solubility (S or logS) or octanol/water partition coefficient (logP), such a relationship was not successful for the class of carbonyl herbicides employed in the present study (see Supplementary Material). The linear regression between the logBCF data of Table 1 with calculated S (in mg/mL), logS and logP generally yielded unsatisfactory correlations (*i.e.* r^2 = 0.33, 0.22 and 0.66, respectively). Additionally, these models do not allow for the interpretation and deep comprehension of the structural characteristics responsible for increasing or decreasing the logBCF values. In this sense, molecular descriptors encoding the precise positions and chemical groups affecting the logBCF would be useful in the development of herbicide candidates with lower ecotoxicity. Thus, the aug-MIA-QSAR method was applied to model the logBCF for the series of 25 carbonyl herbicides. Eq. (1) shows the best 3-variable model obtained with the aug-MIA-QSPR descriptors using the MLR-GA procedure, as well as the corresponding statistical parameters.

$$\textbf{logBCF} = 5.682 - 0.00233 \times \textbf{X9774} - 0.00070 \times \textbf{X813} - 0.00273 \times \textbf{X5144} \quad (1)$$

RMSEc = 0.287, r^2 = 0.793, RMSE y-rand = 0.569, r^2*y-rand* = 0.174, $^cr^2p$ = 0.701,

RMSEcv = 0.349, q^2 = 0.700, RMSEp = 0.379, r^2*test* = 0.805, r^2_m = 0.704

The resulting model was found to be predictive and reliable, since the calibration parameter r^2 was about 0.8, while q^2, r^2_{test}, $^cr^2_p$ and r^2_m were all above 0.5, as recommended in the literature for validation purposes. The calibration and validation performance is illustrated in Fig. 3 (predictions for each compound in dataset are available as Supplementary Material). The applicability domain (AD) of the built aug-MIA-QSPR model was also examined using a simple standardization approach proposed by Roy et al. (2015). This analysis showed that there existed no X-outlier in the training set and that all test set compounds lied within the model's AD. In addition to the appreciable predictability of the model, useful in the estimation of the logBCF of prospective herbicide candidates, a major advantage of the three-parameter aug-MIA-QSPR model over logP or S-based approaches, is that it offers chemical information

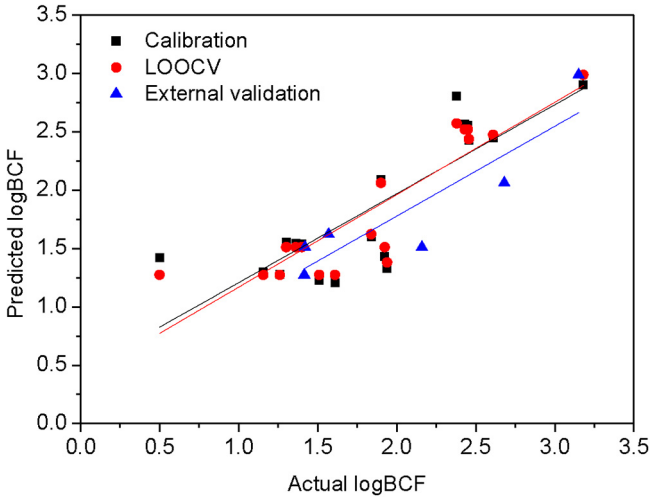


Fig. 3. Plot of actual *vs.* predicted logBCF obtained by the aug-MIA-QSPR model (Eq. (1)).

on the structural characteristics that favor lower ecotoxicity and thus offering key leads in the synthesis of novel environmentally friendly herbicides. The three variables in the model (*i.e.* **X9774**, **X813** and **X5144**) represent pixels in specific coordinates in the workspace where images were drawn, which after a painstaking analysis were identified and are qualitatively shown in Fig. 4.

Because of the negative coefficients in the MLR Eq. (1), high pixel values for all three variables **X9774**, **X5144** and **X813** lead to low logBCF values. The variables **X9774** and **X5144** have comparable contribution (in terms of the coefficient magnitudes), while **X813** demonstrates a much lower contribution. An analysis of these variables reveals the following trends: low pixel values for **X9774**, such as 229 corresponding to the oxygen in compounds **13** and **14**, play an important role in increasing logBCF, while short carbon chains or hydrogen [represented by high pixel values, *i.e.* 765 (blank space in analyzed pixel coordinate due to short carbon chain) and 612] bonded to the nitrogen in same coordinate relative to the common scaffold tend to decrease the bioconcentration factor. The variables **X5144** and **X813** follow a similar trend, *i.e.* long carbon chains (encoded by 426 pixel at the **X5144** coordinate), and

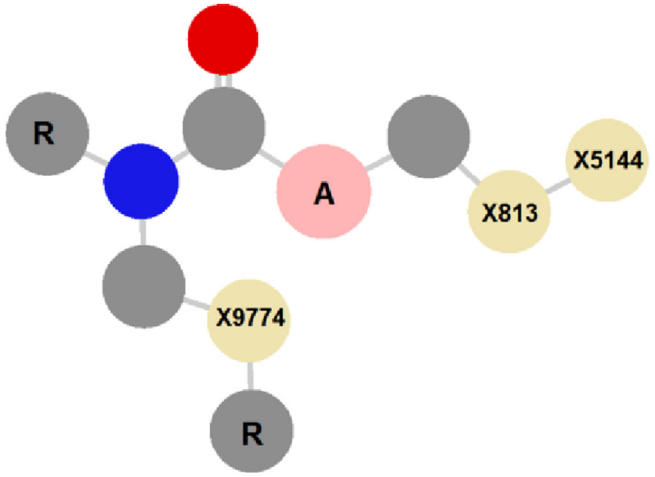
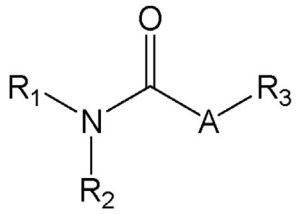


Fig. 4. Basic scaffold of the series of herbicide compounds (**A** = CR_2 for amides, NR for ureas, O for carbamates and S for thiocarbamates; **R** = H and alkyl chains, if applicable) indicating the approximate position of pixels (variables) **X9774**, **X813** and **X5144**.



A = CR_2, NR, O or S
R_1 = any (better if it probably improves bioactivity, *e.g.* 4-Cl-Ph)
R_2 = not -CH_2-**O**-R
R_3 = not long carbon chain

Fig. 5. Framework prospectively useful for the design of active and environmentally friendly herbicides.

chlorine (289 pixel) or carbon atoms (426 pixel, also describing long carbon chains) at position **X813**, are characteristic of herbicides with higher logBCF. Thus, the absence of electronegative and hydrophobic substituents, such as halogens and long carbon chains, at these positions would yield compounds with low logBCF values, which would in turn cause less environmental pollution.

A quick inspection of **6** (Monuron) and **7** (Neburon) confirms this hypothesis, since these molecules differ from one another only by a chlorine at position 3 of the aromatic ring and by the long carbon chain corresponding to variables **X5144** and **X813**; this small structural change strongly affects logBCF from 1.26 in **6** to 2.46 in **7**. The effect of changing **X9774** is also evident to a lesser extent when Metolachlor (**15**, logBCF = 1.84) and Alachlor (**13**, logBCF = 2.45) are compared. Based on these chemical insights, promising candidates may be proposed as herbicides with reduced bioconcentration, for instance with the basic scaffold in Fig. 5 (which may be an amide, urea, carbamate or thiocarbamate nucleus), where the R groups are managed to improve the bioactivity (generally chlorinated moieties) without deteriorating the environment in bioconcentration terms.

Despite the fact that in some cases hydrophobicity and solubility correlate with the bioconcentration factor, these do not offer mechanistic insight on the chemical sorption by an aquatic organism. Some efforts have been made to build mechanistically interpretable models for this endpoint. In particular, a previous study based on a careful analysis of the Dragon descriptors revealed that five molecular descriptors appropriately describe the BCF profile of 238 non-ionic compounds (Gramatica and Para, 2005). This study emphasized the role of descriptors encoding molecular branching and cyclicity, descriptors highlighting the relevance of polarity properties (in agreement with Lu et al., 2000), hydrogen bond acceptors (since compounds containing these groups tend to interact better with water than with aquatic organisms, and thus lowering the BCF) and two topological descriptors related with molecular geometry and atomic polarizability, in describing the BCF of chemicals. The selected aug-MIA descriptors are in fact consistent with some of the general properties suggested (in the Dragon descriptors analysis) as relevant in the description of the BCF, *e.g.* long carbon chains [relative to both the IDDM (encodes molecular branching) (Bonchev, 1983) and **X813** aug-MIA descriptors] and the atomic polarizability [for instance the H6p Dragon descriptor (representative of a topological distance of 6 from an atom weighted by polarizability) and the **X5144** aug-MIA descriptor, which corresponds to chlorine (characterized by high polarizability) *e.g.* compound **19**]. However, it is important to highlight that while the analysis of Dragon descriptors computed over the structurally diverse dataset provided valuable information on the molecular properties affecting the BCF in a general sense (advantageous in itself due to the broad applicability domain achieved), the inherent drawback of such an analysis lies in the vague notion on the group types and/or molecular positions that most affect the BCF. In this sense, the contribution of the aug-MIA approach is appreciated, as the ensuing descriptors allow for the identification of the particular structural features and/or positions responsible for enhanced or attenuated bioconcentration of the studied herbicides in aquatic organisms.

4. Conclusions

The BCFs for a series of carbonyl compounds were quantitatively described by only three descriptors derived from multivariate 2D image analysis. The identification of the location of each selected pixel allowed for chemical interpretation based on structural and atomic properties/coordinates responsible for increased or decreased BCF. Therefore, more ecofriendly herbicides may be synthesized on the basis of the outcomes of the present QSPR modeling.

Acknowledgements

The authors are thankful to FAPEMIG for the financial support as well as to CAPES/Rede Mineira de Química (to M.R.F.) and CNPq (to S.J.B. and M.P.F.) for the fellowships. This work is a collaboration research project of members of the Rede Mineira de Química (RQ-MG) supported by FAPEMIG (Project: CEX - RED-00010-14).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.chemosphere.2016.03.011.

References

ACD/Percepta, 2013. Advanced Chemistry Development, Inc., Toronto.
Arnot, J.A., Gobas, F.A.P.C., 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environ. Rev. 14, 257–297.
Asano, K., Iwamura, H., Nishiuchi, Y., 1989. Bioconcentration of Paddy-herbicides in Carp, Cyprinus carpio. Bulletin of the Agricultural Chemicals Inspection Station, Japan.
Barigye, S.J., Freitas, M.P., 2016. Ten years of the MIA-QSAR strategy: historical development and applications. Int. J. Quant. Struct. Prop. Relat. 1, 64–77.
Bermúdez-Saldaña, J.M., Escuder-Gilabert, L., Medina-Hernández, M.J., Villanueva-Camañas, R.M., Sagrado, S., 2005. Modelling bioconcentration of pesticides in fish using biopartitioning micellar chromatography. J. Chromatogr. A 1063,

153–160.
Bonchev, D., 1983. Information Theoretic Indices for Characterization of Chemical Structures. Research Studies Press, Chichester.
Chiou, C.T., Freed, V.H., Schmedding, D.W., Kohnert, R.L., 1977. Partition coefficient and bioaccumulation of selected organic chemicals. Environ. Sci. Technol. 29, 475–478.
Corsonlini, S., Ademollo, N., Romeo, T., Greco, S., Focardi, S., 2005. Persistent organic pollutants in edible fish: a human and environmental health problem. Microchem. J. 79, 115–123.
Dennington, R.D., Keith, T.A., Millam, J.M., 2008. GaussView 5.0. Gaussian, Inc., Wallingford.
Devillers, J., Domine, D., Bintein, S., Karcher, W., 1998. Comparison of fish bioconcentration models. In: Comparative QSAR. Taylor and Francis, Washington, D.C, pp. 1–50.
Duarte, M.H., Barigye, S., Freitas, M.P., 2015. Exploring MIA-QSAR's for antimalarial quinolon-4(1*H*)-imines. Comb. Chem. High. Throughput Screen. 18, 208–216.
Duarte, M.H., Barigye, S., da Mota, E.G., Freitas, M.P., 2015. Computational modelling of the antischistosomal activity for neolignan derivatives based on the MIA-SAR approach. SAR QSAR Environ. Res. 26, 205–216.
Freitas, M.P., Brown, S.D., Martins, J.A., 2005. MIA-QSAR: a simple 2D image-based approach for quantitative structure–activity relationship analysis. J. Mol. Struct. 738, 149–154.
Freitas, M.P., da Cunha, E.F.F., Ramalho, T.C., Goodarzi, M., 2008. Multimode methods applied on MIA descriptors in QSAR. Curr. Comput. Aided Drug Des. 4, 273–282.
Freitas, M.R., Freitas, M.P., Macedo, R.L.G., 2014. Aug-MIA-QSPR modeling of the soil sorption of carboxylic acid herbicides. Bull. Environ. Contam. Toxicol. 93, 489–492.
Freitas, M.R., Barigye, S.J., Freitas, M.P., 2015. Coloured chemical image-based models for the prediction of soil sorption of herbicides. RSC Adv. 5, 7547–7553.
Gobas, F.A.P.C., Morrison, H.A., 2000. Bioconcentration and biomagnification in the aquatic environment. In: Boethling, R.S., Mackay, D. (Eds.), Handbook of Property Estimation Methods for Chemicals: Environmental and Health Sciences. Lewis, Boca Raton, pp. 189–231.
Golbraikh, A., Tropsha, A., 2002. Beware of q^2! J. Mol. Graph. Model. 20, 269–276.
Gramatica, P., Para, E., 2005. An update of the BCF QSAR model based on theoretical molecular descriptors. QSAR Comb. Sci. 24, 953–960.
Grisoni, F., Consonni, V., Villa, S., Vighi, M., Todeschini, R., 2015. QSAR models for bioconcentration: is the increase in the complexity justified by more accurate predictions? Chemosphere 127, 171–179.
Grisoni, F., Consonni, V., Vighi, M., Villa, S., Todeschini, R., 2016. Investigating the mechanisms of bioconcentration through QSAR classification trees. Environ. Int. 88, 198–205.
Grover, R., Cessna, A.J., 1990. Environmental Chemistry of Herbicides. CRC Press, Boca Raton.
Higgins, T.E., Sachdev, J.A., Engleman, S.A., 2010. Toxic Chemicals: Risk Prevention Through Use Reduction. CRC Press, Boca Raton.
Jackson, S.H., Cowan-Ellsberry, C.E., Thomas, G., 2009. Use of quantitative structural analysis to predict fish bioconcentration factors for pesticides. J. Agric. Food Chem. 57, 958–967.
Lu, X., Tao, S., Hu, H., Dawson, R.W., 2000. Estimation of bioconcentration factors of non-ionic organic compounds in fish by molecular connectivity indices and polarity correction factors. Chemosphere 41, 1675–1688.
Mackay, D., 1982. Correlation of bioconcentration factors. Environ. Sci. Technol. 16, 274–278.
Mackay, D., Shiu, W.Y., Ma, K.C., 1997. Illustrated Handbook of Physical-chemical Properties of Environmental Fate for Organic Chemicals. CRC Press, Boca Raton.
Meylan, W.M., Howard, P.H., Boethling, R.S., Aronson, D., Printup, H., Gouchie, S., 1999. Improved method for estimating bioconcentration/bioaccumulation factor from octanol/water partition coefficient. Environ. Toxicol. Chem. 18, 664–672.
Mitra, I., Saha, A., Roy, K., 2010. Exploring quantitative structure–activity relationship studies of antioxidant phenolic compounds obtained from traditional Chinese medicinal plants. Mol. Simul. 36, 1067–1079.
Nunes, C.A., Freitas, M.P., 2013. Introducing new dimensions in MIA-QSAR: a case for chemokine receptor inhibitors. Eur. J. Med. Chem. 62, 297–300.
Nunes, C.A., Freitas, M.P., Pinheiro, A.C.M., Bastos, S.C., 2012. Chemoface: a novel free user-friendly interface for chemometrics. J. Braz. Chem. Soc. 23, 2003–2010.
OECD: Bioaccumulation: Flow-through Fish Test, 1996. Organisation for Economic Cooperation and Development (OECD). OECD Guideline for the Testing of Chemicals No. 305, Paris, France.
Roy, K., Chakraborty, P., Mitra, I., Ojha, P.K., Kar, S., Das, R.N., 2013. Some case studies on application of r^2_m metrics for judging quality of quantitative structure-activity relationship predictions: emphasis on scaling of response data. J. Comput. Chem. 34, 1071–1082.
Roy, K., Kar, S., Ambure, P., 2015. On a simple approach for determining applicability domain of QSAR models. Chemom. Intell. Lab. Syst. 145, 22–29.
Urias, R.W.P., Barigye, S.J., Marrero-Ponce, Y., García-Jacas, C.R., Valdes-Martiní, J.R., Perez-Gimenez, F., 2015. IMMAN: free software for information theory-based chemometric analysis. Mol. Divers. 19, 305–319.
U.S. Environmental Protection Agency (EPA), 1995. Aquatic Information Retrieval (AQUIRE) Database. Environmental Research Laboratory, Duluth.
Versar Inc., 1990. Issue Paper: Bioaccumulation Potential Based on Ambient Water Quality Criteria Documents (VER_BCF). EPA Office of Emergency and Remedial Response, Washington DC.
Vogs, C., Kühnert, A., Hug, C., Küster, E., Altenburger, R., 2015. A toxicokinetic study of specifically acting and reactive organic chemicals for the prediction of internal effect concentrations in Scenedesmus vacuolatus. Environ. Toxicol. Chem. 34, 100–111.
Zimmerman, L.J., Valentine, H.L., Valentine, W.M., 2004. Characterization of *S*-(*N*,*N*-dialkylaminocarbonyl)cysteine adducts and enzyme inhibition produced by thiocarbamate herbicides in the rat. Chem. Res. Toxicol. 17, 258–267.