

Article

Retention Modelling of Phenoxy Acid Herbicides in Reversed-Phase HPLC under Gradient Elution

Alessandra Biancolillo ¹, Maria Anna Maggi ², Sebastian Bassi ³, Federico Marini ³ and Angelo Antonio D'Archivio ^{1,*}

¹ Dipartimento di Scienze Fisiche e Chimiche, Università degli Studi dell'Aquila, Via Vetoio, Coppito, 67100 L'Aquila (AQ), Italy; alessandra.biancolillo@univaq.it

² Hortus Novus srl, Via Campo Sportivo 2, Canistro, 67100 L'Aquila, Italy; maria.maggi79@gmail.com

³ Dipartimento di Chimica, Università degli Studi di Roma “La Sapienza”, Piazzale Aldo Moro 5, 00185 Roma, Italy; sebastian.bassi@hotmail.it (S.B.); federico.marini@uniroma1.it (F.M.)

* Correspondence: angeloantonio.darchivio@univaq.it

Received: 26 January 2020; Accepted: 11 March 2020; Published: 11 March 2020



Abstract: Phenoxy acid herbicides are used worldwide and are potential contaminants of drinking water. Reversed phase high-performance liquid chromatography (RP-HPLC) is commonly used to monitor phenoxy acid herbicides in water samples. RP-HPLC retention of phenoxy acids is affected by both mobile phase composition and pH, but the synergic effect of these two factors, which is also dependent on the structure and pKa of solutes, cannot be easily predicted. In this paper, to support the setup of RP-HPLC analysis of phenoxy acids under application of linear mobile phase gradients we modelled the simultaneous effect of the molecular structure and the elution conditions (pH, initial acetonitrile content in the eluent and gradient slope) on the retention of the solutes. In particular, the chromatographic conditions and the molecular descriptors collected on the analyzed compounds were used to estimate the retention factor k by Partial Least Squares (PLS) regression. Eventually, a variable selection approach, Genetic Algorithms, was used to reduce the model complexity and allow an easier interpretation. The PLS model calibrated on the retention data of 15 solutes and successively tested on three external analytes provided satisfying and reliable results.

Keywords: phenoxy acid herbicides; HPLC; gradient elution; molecular descriptors; retention prediction; PLS regression

1. Introduction

Phenoxy acid herbicides are synthetic analogues of the plant regulator auxins and are extensively used to control broad-leaved weeds in many crops, such as rice, wheat and soybeans [1]. The treatment by phenoxy acids induces “auxin overdose” into the unwanted vegetal species that undergoing disorganized and uncontrolled growth eventually die. 2,4-dichlorophenoxyacetic acid (2,4-D), although first marketed in 1944, is still one of the most widely used herbicides in the world, together with the parent products 2-(4-chloro-2-methylphenoxy) propanoic acid (mecoprop), (4-chloro-2-methylphenoxy)acetic acid (MCPA) and 2-(2,4-dichlorophenoxy)propanoic acid (dichloroprop) [2,3]. Phenoxy acids are highly soluble in the aqueous media and therefore, can easily migrate through the agricultural environment finally leading to groundwater contamination [4,5], which can compromise drinking water quality and affect human health because of chronic toxicity, endocrine-disrupting action and potential carcinogenic effects [6–8]. Phenoxy acids in water or food can be detected and quantified by reversed-phase high-performance liquid chromatography (RP-HPLC) coupled to spectrophotometric or mass-spectrometry detection [3,9–11]. Multi-residue determination in real samples by RP-HPLC is usually preceded by a procedure for concentration and isolation of

the analytes, such as liquid-phase micro-extraction or solid-phase extraction [12], but differently from gas-chromatography [13,14], preliminary derivatization of the phenoxy acids is not needed.

Routine monitoring of trace pollutants in water requests chromatographic methods able to simultaneously detect and quantify many structural congeners or multi-class mixtures. In this context, the empirical trial-and-error strategies still used to develop chromatographic methods may be inefficient or excessively time-consuming. To overcome such limitation various multivariate statistical approaches have been proposed [15–18] with the aim of providing predictive tools for the chromatographic retention. A widely used approach is quantitative structure-retention relationship (QSRR) methodology [19,20] which is aimed at establishing a relationship between the solute structure, encoded by a set of molecular descriptors, and the retention time or a related parameter. Approaches complementary to QSRR, on the other hand, are applied to model the effects of the separation conditions on the chromatographic behavior of a target solute or a mixture of analytes. To this end, multivariate design of experiments combined with response surface methodology can be employed to optimize the chromatographic response(s) of interest [21–25], such as retention time, pair or global resolution and overall analysis time. Apart from the above two kinds of predictive approaches, which can be considered orthogonal according to the complementary nature of the described variability, QSRR-based models combining both molecular descriptors of solutes and the features of the eluent, the column or both have been proposed [26–30]. These comprehensive models, once calibrated on a pool of known solutes and a well-designed set of separation conditions, allow to predict the behavior of unseen molecules in the whole experimental domain covered in calibration or to transfer retention data among different columns [31–34].

In RP-HPLC the composition of the eluent, mostly a binary aqueous mixture, is the experimental parameter commonly tuned to find the optimal separation conditions. Moreover, to overcome the typical disadvantages of isocratic elution, such as poor resolution of early peaks, broadening of late peaks, band tailing, and long separation times, mobile-phase gradients are required [35,36], especially when dealing with complex mixtures or structural congeners. The concentration of organic solvent in the aqueous mobile phase is usually increased during the chromatographic run, which determines a progressive increase of the elution power of the eluent and a consequent decrease in solute retention. Apart from composition, the retention behavior of weak acids is also influenced by the eluent pH, that must be fixed or properly modulated during the chromatographic run [37]. An increase in pH of the mobile phase produces a progressive ionization of the solute and a consequent decrease in its retention time. It must be remarked that the organic modifier of the mobile phase affects the acidity of both medium and solutes, which does not allow to easily deduce the synergic effect of the eluent pH and composition on the retention behavior of the target analytes, especially when gradient elution is applied.

In a previous work [38], we modelled the combined effect of pH and composition of water-acetonitrile eluent on the retention of phenoxy acids and structurally related carboxylic acids under isocratic elution conditions. The molecular structure of the phenoxy acids was described by a-priori selected quantum-chemical properties (dipole moment, mean polarizability, anisotropy of polarizability, water/n-heptane partition coefficient and a hydrogen bond descriptor) of both neutral and ionized forms together with the experimental pKa, while pH and acetonitrile volume fraction were the variables associated with the eluent. A three-layer artificial neural network was used to generate the retention model. In the present work, we developed a comprehensive model to predict the retention behavior of phenoxy acid herbicides and related carboxylic acids under the application of linear composition gradients at different pH levels. Moreover, the molecular structure of solutes is encoded by non-quantum-chemical theoretical descriptors provided by popular software *Dragon* (Dragon 6, Talete srl, Milan) [39] that, as compared to quantum-chemical molecular properties, are quickly accessible to users with no specific expertise in theoretical chemistry. The regression problem was solved by Partial Least Square [40,41], as it is one of the most commonly applied regression

methods in this context [20,42,43]. Eventually, Genetic Algorithms (GA) [44] were applied in order to reduce the model complexity and make its interpretation easier.

2. Results

The set of solutes investigated in this work (reported in Table A1) consists of eight phenoxy acid herbicides (2,4-D, 2,4,5-T, 2,4,5-TP, MCPA, dichlorprop, mecoprop, clopyralid and trichlopyr) and ten derivatives of benzoic acid, phenylacetic acid and phenoxyacetic acid, pKa of these compounds ranging between 2.29 and 4.31. The RP-HPLC retention data of the 18 analytes were collected at three different values of eluent pH, namely 2, 3 and 4. For each pH level the starting acetonitrile volume fraction φ_i was set to 30, 40, 50, 60 and 70%. While pH of the mobile phase was kept fixed during elution, the acetonitrile content was fixed at the starting value (isocratic elution) or linearly increased to 100% in 15, 20 or 25 min. The descriptors selected to identify the various elution modes were the eluent pH, φ_i and the gradient slope ϕ defined as $(100 - \varphi_i)/t_g$, where t_g was the application time of the linear gradient, and $\phi = 0$ for the isocratic condition. The chromatographic conditions are resumed in Table 1. Nevertheless, in the chromatograms collected under application of the 60 elution modes obtained by the combination of the above designed levels for the three eluent descriptors, we did not consider those peaks given by both poorly or strongly retained compounds. 753 retention time (t_r) values were finally collected and the (base 10) logarithm of the related retention factor k [45–47], defined as $k = (t_r - t_0)/t_0$, t_0 being the column's dead time, was assumed as the model response. The use of the logarithm of the retention factor has been preferred to the retention factor itself because it can be directly related to some physico-chemical parameters governing the retention process. In particular, Ciura et al. [45] claim that the use of $\log k$ as dependent variable in QSRR presents various advantages, one of the most relevant of which being that it can be used as a hydrophobicity index. Similar considerations are reported by Kalisz and collaborators [46] who showed how $\log k$ could be linearly related to the fraction of the organic modifier in the mobile phase.

Table 1. Chromatographic conditions utilized to collect the retention data.

Column	Kinetex C18 (Phenomenex)
Eluent	water-acetonitrile, flux:1 mLmin ⁻¹
Elution mode	Starting acetonitrile volume fraction (φ_i): 30, 40, 50, 60, 70% Eluent pH: 2,3,4 Application time of linear composition gradient (from φ_i to 100%): none, 15, 20, 25 min

2.1. Prediction of the Retention Factor k

Prior to the creation of any prediction model, the three chromatographic parameters (pH, φ_i and ϕ) and the 329 descriptors were organized in a unique data matrix X. Then, in order to externally validate the regression model, data collected on the 18 compounds were divided into a training set (for building the calibration model), and a test set (for validation); in particular, the latter one was constituted by data collected on 2,4-dichlorofenossiacetic acid, trichlopyr and 2-iodobenzoic acid. As mentioned above, a PLS regression model was created (exploiting only the training set) in order to fit the logarithm of the retention factor k to the autoscaled data matrix X. The calibration model, calculated extracting 11 latent variables (LVs) (defined into a cross-validation procedure), provided a coefficient of determination (R_{cv}^2) of 0.887 and a Root Mean Square Error in Cross-Validation (RMSECV) of 0.142. The application of this model to the test set provided a R_{test}^2 of 0.872 and a Root Mean Square Error in Prediction (RMSEP) of 0.156; a graphical representation of these results is displayed in Figure 1 showing the agreement between computed/predicted and experimental responses. The black solid line in the plot depicts the ideal fitting (i.e., RMSEP = 0), while the red one represents the actual one; as it can be seen from the good overlap between the two straight lines, and in accordance with the R_{cv}^2 and R_{test}^2 values shown, the model properly fits the data under study.

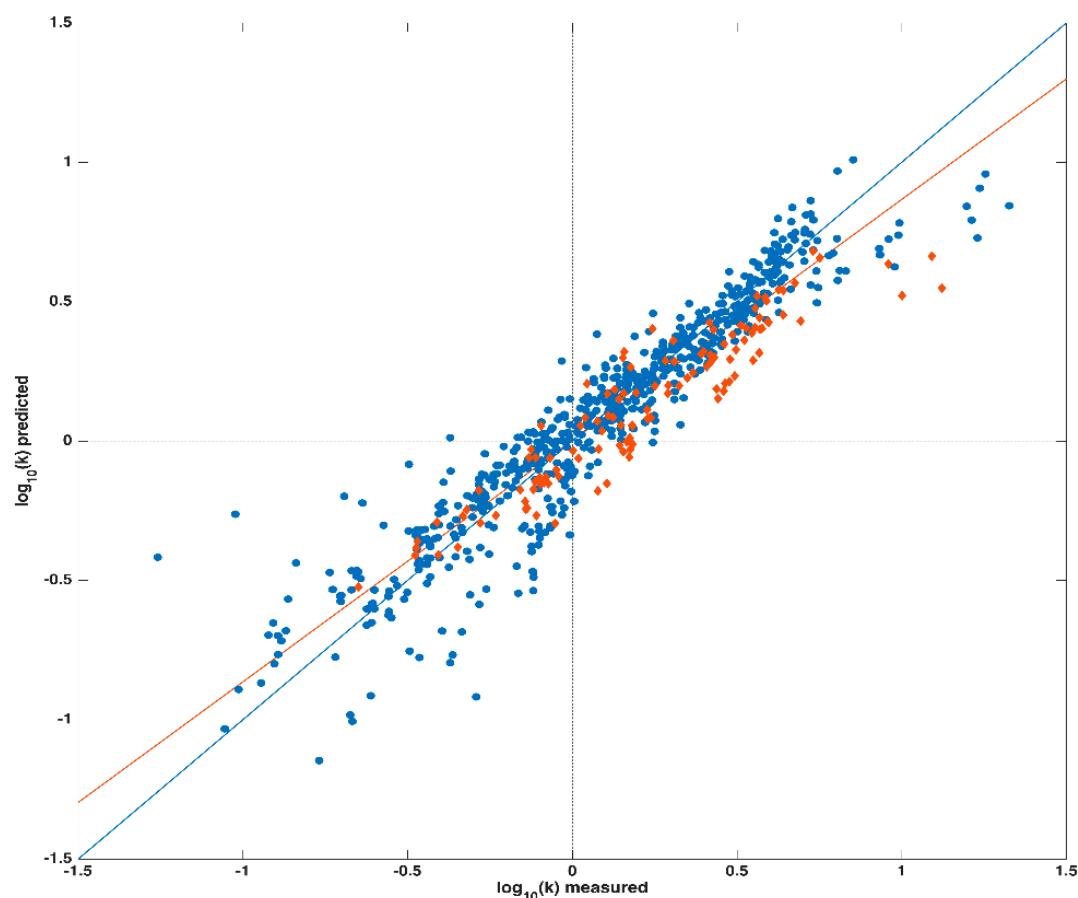


Figure 1. PLS model. Legend: Calibration set: blue dots; Test set: orange diamonds.

Eventually, despite the results provided by the PLS model were quite satisfying, a variable selection approach, Leardi's Genetic Algorithm (GA) [44], was tested aiming at two final goals: evaluating whether variable selection could improve the predictive accuracy and simplifying the interpretation of the regression model.

The genetic algorithm was applied to the training set (with the parameters reported in Table A2 in Appendix A) and the R_{cv}^2 (5 cancelation groups) was used as fitness.

The predictors were sorted according to the frequency of selection in the 100 GA runs and PLS models with an increasing number of variables were built according to a forward stepwise scheme: the model leading to the highest R_{cv}^2 (5 cancelation groups), which included 29 predictors (see Table 2), was selected as the final one. This model was built on 4 latent variables and resulted in a RMSEC = 0.143 and a RMSECV = 0.148 (corresponding to $R^2 = 0.883$ and $R_{cv}^2 = 0.876$, respectively). The model was then applied to the test molecules for the final (external) validation and resulted in a RMSEP of 0.167, corresponding to a R^2_{pred} of 0.847. This means that, in the present case, the use of a variable selection approach mainly helps the interpretation of the regression model in chemical terms, rather than leading to an improvement in the prediction accuracy. Nevertheless, it has to be emphasized that the results obtained are comparable with those of similar works on QSRR published in the literature (for instance [48] where models (built on $\log k$) have in general SEP values between 0.16 and 0.22, and R^2_{pred} around 0.75–0.88), and, framed into a wider scenario, are in line with studies on experimentally measuring retention factor, considering the uncertainty associated to the estimation of this entity [49].

Table 2. List of the 29 predictors selected by the GA in order of decreasing frequency.

Variable	Description	Block
φ_i	starting acetonitrile volume fraction in the eluent	-
ϕ	gradient slope	-
pH	eluent pH	-
MATS2i	Moran autocorrelation of lag 2 weighted by ionization potential	2D autocorrelations
nCb-	number of substituted benzene C(sp ²)	Functional group counts
Mor21u	signal 21/unweighted	3D-MoRSE descriptors
MATS4p	Moran autocorrelation of lag 4 weighted by polarizability	2D autocorrelations
Eta_beta_A	eta average VEM count	ETA indices
TDB05s	3D Topological distance-based descriptors - lag 5 weighted by I-state	3D autocorrelations
RDF070u	Radial Distribution Function - 070/unweighted	RDF descriptors
RDF025s	Radial Distribution Function - 025/weighted by I-state	RDF descriptors
Mor06s	signal 06/weighted by I-state	3D-MoRSE descriptors
G2i	2nd component symmetry directional WHIM index/weighted by ionization potential	WHIM descriptors
MW	molecular weight	Constitutional indices
nCsp2	number of sp ² hybridized Carbon atoms	Constitutional indices
MATS5m	Moran autocorrelation of lag 5 weighted by mass	2D autocorrelations
Mor28v	signal 28/weighted by van der Waals volume	3D-MoRSE descriptors
R7e+	R maximal autocorrelation of lag 7/weighted by Sanderson electronegativity	GETAWAY descriptors
nN	number of Nitrogen atoms	Constitutional indices
piPC06	molecular multiple path count of order 6	Walk and path counts
RDF080u	Radial Distribution Function - 080/unweighted	RDF descriptors
RDF090u	Radial Distribution Function - 090/unweighted	RDF descriptors
RDF070s	Radial Distribution Function - 070/weighted by I-state	RDF descriptors
Mor04m	signal 04/weighted by mass	3D-MoRSE descriptors
Mor21p	signal 21/weighted by polarizability	3D-MoRSE descriptors
G2e	2nd component symmetry directional WHIM index/weighted by Sanderson electronegativity	WHIM descriptors
E2s	2nd component accessibility directional WHIM index/weighted by I-state	WHIM descriptors
R5s	R autocorrelation of lag 5/weighted by I-state	GETAWAY descriptors
CATS2D_03_LL	Lipophilic-Lipophilic at lag 03	CATS2D

3. Discussion

Interpretation of the Model

The PLS model discussed in Section 2.1 provided satisfying and reliable results, indicating that pH, φ_i , ϕ , and the inspected molecular descriptors represent a reasonable set of predictors for the estimation of log(k) for the investigated phenoxy acid herbicides and related carboxylic acids. The regression coefficients from this model are reported in Table A3.

The application of GA leads to a deeper insight into the system under study; in fact, it provides information about the contributions of the 29 selected variables.

The sample scores along the first two latent variables of the model are reported in Figure 2, where their relation to the values of the response is also highlighted. By inspecting the scores in Figure 2, it is apparent how LV1 accounts mostly for the differences among substances while the mobile phase conditions are mainly responsible for the variance along LV2. In general, log k increases both along LV1 and along LV2.

Investigation of the corresponding loadings (Figure 3) confirms what already evidenced by the scores plot and provides details for a comprehensive interpretation. Indeed, the three predictors related to the mobile phase conditions have practically zero loadings along LV1, thus confirming that the latent variable is affected only by the molecular descriptors characterizing the analytes. In particular, MATS2i, nCb-, MATS4p, MW, nCsp2 and CATS2D_03_LL have the highest positive loadings, while Mor21u, Eta_beta_A, TDB05s, G2i, nN and Mor21p are those giving the largest negative contribution. This means that the value of log k is higher for molecules having higher values of the former descriptors and lower values of the latter, and vice versa.

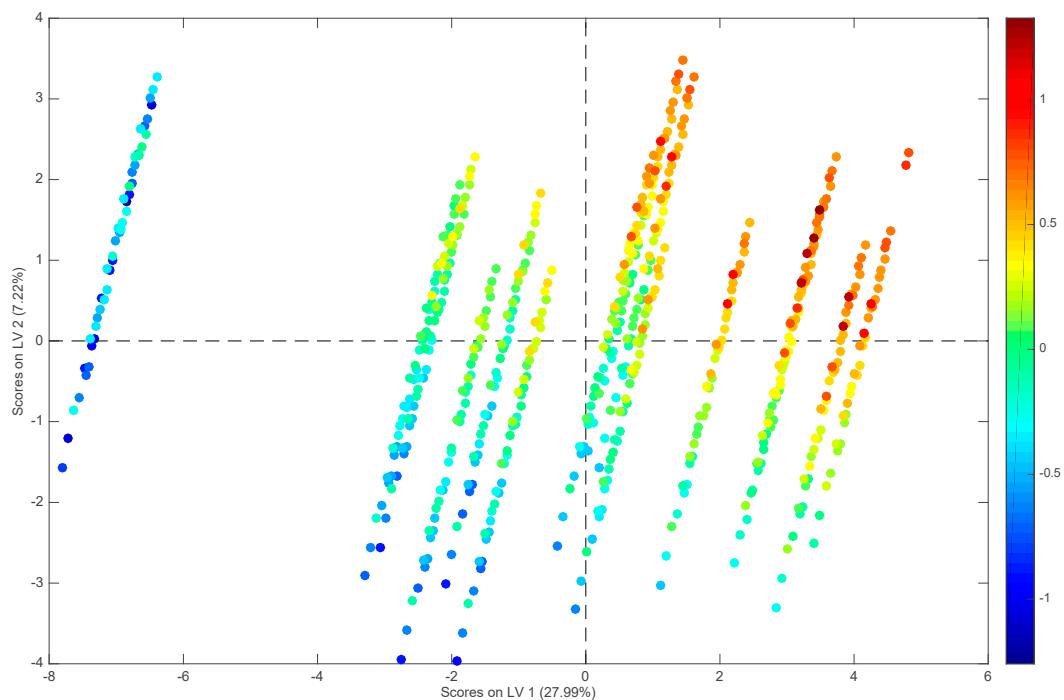


Figure 2. PLS model built on the reduced set of 29 predictors selected by GA: scores plot of the training samples. The points are colored according to their values of $\log k$.

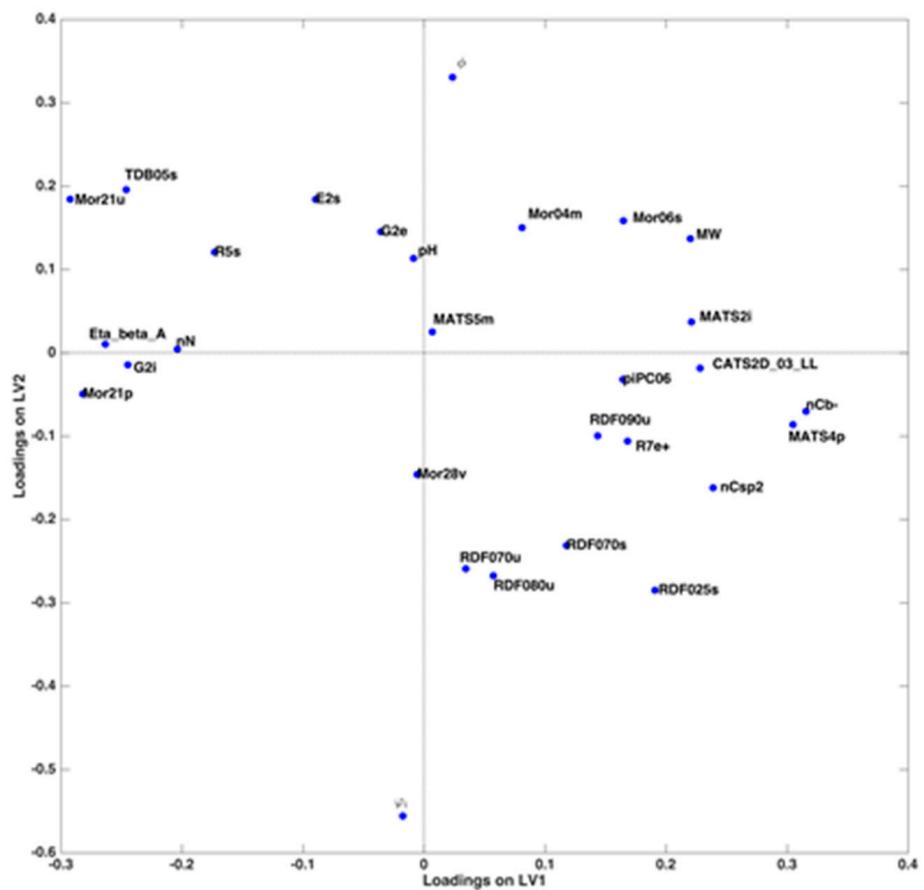


Figure 3. Variables loadings projected on the first two latent variables of the PLS model.

On the other hand, LV2 is characterized by a very negative contribution of the % of acetonitrile and a highly positive loading of the slope of the gradient; pH of the mobile phase also has a positive, although lower, loading on this component. For each substance, the value of $\log k$ increases when a lower initial amount of acetonitrile is used, together with a steeper slope of the gradient; a more basic pH also contributes to increase the value of the retention factor.

In summary, the PLS regression model discussed in Section 2.1 provided satisfying and reliable results, indicating that pH, φ_i , ϕ , and the inspected molecular descriptors represent a reasonable set of predictors for the estimation of $\log k$ for the investigated herbicides. The use of a variable selection approach in this specific case did not lead to an improvement in the prediction accuracy but was useful for the interpretation of the regression model in chemical terms.

4. Materials and Methods

4.1. Chemicals

Pestanal standards (purity greater than 99%) of the phenoxy acid herbicides 2-(4-chloro-2-methylphenoxy) propanoic acid (mecoprop), 2,4-dichlorophenoxyacetic acid (2,4-D), (4-chloro-2-methylphenoxy)acetic acid (MCPA), (2,4,5-trichlorophenoxy) acetic acid (2,4,5-T), 2-(2,4,5-trichlorophenoxy)propanoic acid (2,4,5-TP), 2,6-dichloro-pyridine-2-carboxylic acid (clopyralid), 2-(2,4-dichlorophenoxy)propanoic acid (dichloroprop) and [(3,5,6-trichloro-2-pyridinyl)oxy]acetic acid (triclopyr) were obtained from Labor Dr. Ehrenstorfer-Schäfers (Augsburg, Germany). Chemical standards of phenoxyacetic acid, benzoic acid and phenylacetic acid derivatives with purity greater than 95% were purchased from Sigma-Aldrich (St. Louis, MO, USA). Stock solutions (1.00 g L^{-1}) of individual analytes were prepared by dissolving accurately weighed 10 mg of each solute in 10 mL of HPLC-grade acetonitrile (Carlo Erba Reagenti, Milano, Italy) and stored at 4°C . HPLC-grade acetonitrile was acquired from Sigma-Aldrich (St. Louis, MO, USA). Double deionized water was obtained from a Milli-Q filtration/purification system (Millipore, Bedford, MA, USA).

4.2. HPLC Apparatus

Retention data were collected with a Waters chromatographic system (Waters, Milford, MA, USA) consisting of a Model 600 pump, a 600 Pump Controller Module, an Acqua 717 plus auto-sampler and a 996-photodiode array detector. Chromatographic data management was automated using the Empower data acquisition system (Waters). Eluent degassing was performed by an Agilent 1200 system (Agilent Technologies, Waldbronn, Germany). The analyses were carried on a $250 \times 4.6 \text{ mm}$ Kinetex C18 (Phenomenex, Torrance, CA, USA) analytical column preceded by a $4 \times 3\text{mm}$ UHPLC guard cartridge (Phenomenex) both packed with octadecyl-silica having $5 \mu\text{m}$ particle size.

4.3. Collection of Retention Data

The HPLC analyses were carried out at room temperature using water-acetonitrile mobile phases with a flow-rate of 1mLmin^{-1} . The retention time of the analytes was collected at fixed values of the eluent pH obtained by mixing the aqueous phase buffered at pH 2, 3 or 4. To control the acidity of the mobile phase a phosphate buffer at 0.1 M concentration was used. pH of water, before mixing with acetonitrile, was adjusted to the desired value by addition of HCl. pH measurements were carried out by means of an Orion 420 A (Beverly, MA, USA) pH meter equipped with an Orion 9107 electrode. Five levels for the initial acetonitrile volume fraction φ_i were considered, namely 30, 40, 50, 60 and 70%. For each combination of the eluent pH and φ_i pair, the acetonitrile content in the mobile phase was kept constant for the whole chromatographic run (isocratic elution) or was linearly increased until 100% in 15, 20, or 25 min. After gradient elution the column was re-equilibrated by fluxing the mobile phase at the initial composition for ten minutes before successive analysis. Retention data were collected by injecting $10 \mu\text{L}$ of acetonitrile solutions of single compounds (5mgL^{-1}) or mixtures providing well-resolved chromatograms. The detection wavelength (220 nm for most analytes) was the

one providing the maximum peak height. The dead time of the column was determined by injection of an aqueous solution of sodium nitrate.

4.4. Molecular Descriptors

Starting geometries of phenoxy acid herbicides and the other carboxylic acids were drawn by means of the MacroModel 7.1 molecular modelling program package [50]. The global energy minimum of each molecule was searched using the MM2 force field. Software Dragon 6 [39] was used to compute the molecular descriptors from the optimised geometries. The version utilised in this work provides 4885 descriptors classified as zero- (0D), one- (1D), two- (2D) and three-dimensional (3D) descriptors depending on the fact they are computed from the chemical formula, the substructure list representation, the molecular graph or the geometrical representation of the molecule, respectively. After removal of constant and highly correlated variables ($r > 0.85$), 329 molecular descriptors were retained for further analyses.

4.5. Development and Validation of Retention Models

The data matrix constituted of the three chromatographic parameters and the 329 molecular descriptors were used as predictors to estimate $\log k$ by Partial Least Squares (PLS) [40,41]. The regression model was built on a restricted sub-set of data (the training set) in order to pursue validation on the test set. The optimal complexity of the calibration model (i.e., the number of LVs to be extracted) were defined into a cross-validation procedure (not involving the test set). Once the PLS model parameters were defined (by inspection of the lowest RMSECV) it was validated by predicting the external test set. The goodness of the model was estimated by inspecting R_{cv}^2 and R_{test}^2 .

In order to test whether feature reduction could improve the model from the prediction and the interpretation standpoint, Genetic Algorithms (GA) [44] were applied for variable selection. Concisely, this iterative approach starts picking random sets of variables, which, by means of genetic operators, lead to the generation of *offspring*s. Calculations are then re-run on the subsets of features, and the variables providing the most performant solution are retained. The procedure is repeated until the optimal solution is reached. For more details, the reader is addressed to [51].

All the calculations were run in Matlab 2015b (The Mathworks Inc., Natick, MA) using in-house routines or by the PLS toolbox Version 8.1 (Eigenvector Research, Inc. Manson, WA).

Author Contributions: Conceptualization, A.A.D. and A.B.; methodology and data curation, F.M.; formal analysis, A.B.; investigation, M.A.M. and S.B.; writing—original draft preparation, review and editing, A.A.D. and A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Molecular structure and pKa value (from literature [38]) of the investigated acids.

Name	Structure	pKa
2,4-D		2.73
2,4,5-T		2.83
2,4,5-TP		2.84
Clopyralid		2.29
Dichlorprop		3.10
MCPA		3.13
Mecoprop		3.10
Tryclopir		3.97
Benzoic acid		4.19
Salicylic acid		2.97

Table A1. *Cont.*

Name	Structure	pKa
2-Iodobenzoic acid		2.93
3-Iodobenzoic acid		3.85
4-Iodobenzoic acid		4.00
Phenylacetic acid		4.31
4-Chlorophenylacetic acid		2.76
4-Nitrophenylacetic acid		3.85
Phenoxyacetic acid		3.17
4-Chlorophenoxyacetic acid		3.56

Table A2. Parameters of the genetic algorithm.

Parameters	Value
Number of chromosomes	30
Probability of selection in the original population	0.015
Maximum number of variables per chromosome	30
Probability of mutation	0.01
Probability of cross-over	0.50
Backward stepwise selection every	100 iterations
Number of runs	100
Number of evaluations per run	200

Table A3. Regression coefficients extracted from the PLS model based on the 29 descriptors selected by GA.

Descriptor	Coefficient
φ_i	-0.686
ϕ	-0.091
pH	0.212
MATS2i	0.061
nCb-	0.097
Mor21u	-0.067
MATS4p	0.080
Eta_beta_A	-0.073
TDB05s	-0.042
RDF070u	-0.025
RDF025s	0.0366
Mor06s	0.0900
G2i	-0.0400
MW	0.093
nCsp2	0.039
MATS5m	0.008
Mor28v	0.001
R7e+	0.034
nN	-0.048
piPC06	0.0658
RDF080u	-0.015
RDF090u	0.0188
RDF070s	0.016
Mor04m	0.025
Mor21p	-0.081
G2e	-0.007
E2s	0.025
R5s	-0.037
CATS2D_03_LL	0.0300

References

1. Tu, M. General Properties of Herbicides. In *Weed Control Methods Handbook*; The Nature Conservancy: Arlington, VA, USA, 2003.
2. Peterson, M.A.; McMaster, S.A.; Riechers, D.E.; Skelton, J.; Stahlman, P.W. 2,4-D Past, Present, and Future: A Review. *Weed Technol.* **2016**, *30*, 303–345. [[CrossRef](#)]
3. Liu, J.F.; Toräng, L.; Mayer, P.; Jönsson, J.Å. Passive extraction and clean-up of phenoxy acid herbicides in samples from a groundwater plume using hollow fiber supported liquid membranes. *J. Chromatogr. A* **2007**, *1160*, 56–63. [[CrossRef](#)] [[PubMed](#)]
4. Jacobsen, C.S.; van der Keur, P.; Iversen, B.V.; Rosenberg, P.; Barlebo, H.C.; Torp, S.; Vosgerau, H.; Juhler, R.K.; Ernstsæn, V.; Rasmussen, J.; et al. Variation of MCPA, metribuzine, methyltriazine-amine and glyphosate degradation, sorption, mineralization and leaching in different soil horizons. *Environ. Pollut.* **2008**, *156*, 794–802. [[CrossRef](#)] [[PubMed](#)]
5. Spliid, N.H.; Helweg, A.; Heinrichson, K. Leaching and degradation of 21 pesticides in a full-scale model biobed. *Chemosphere* **2006**, *65*, 2223–2232. [[CrossRef](#)] [[PubMed](#)]
6. Jayakody, N.; Harris, E.C.; Coggon, D. Phenoxy herbicides, soft-tissue sarcoma and non-Hodgkin lymphoma: A systematic review of evidence from cohort and case-control studies. *Br. Med. Bull.* **2015**, *114*, 75–94. [[CrossRef](#)] [[PubMed](#)]
7. Bukowska, B.; Hutnik, K. 2,4-D and MCPA and their derivatives: Effect on the activity of membrane erythrocytes acetylcholinesterase (in vitro). *Pestic. Biochem. Physiol.* **2006**, *85*, 174–180. [[CrossRef](#)]
8. Murschell, T.; Farmer, D.K. Real-time measurement of herbicides in the atmosphere: A case study of MCPA and 2,4-D during field application. *Toxics* **2019**, *7*, 40. [[CrossRef](#)]

9. Ji, Z.; Cheng, J.; Song, C.; Hu, N.; Zhou, W.; Suo, Y.; Sun, Z.; You, J. A highly sensitive and selective method for determination of phenoxy carboxylic acids from environmental water samples by dispersive solid-phase extraction coupled with ultra high performance liquid chromatography-tandem mass spectrometry. *Talanta* **2019**, *191*, 313–323. [[CrossRef](#)]
10. Wu, J.; Kim, H.E.; Hian, K.L. Automated dynamic liquid-liquid-liquid microextraction followed by high-performance liquid chromatography-ultraviolet detection for the determination of phenoxy acid herbicides in environmental waters. *J. Chromatogr. A* **2005**, *1082*, 121–127. [[CrossRef](#)]
11. McManus, S.L.; Moloney, M.; Richards, K.G.; Coxon, C.E.; Danaher, M. Determination and occurrence of phenoxyacetic acid herbicides and their transformation products in groundwater using ultra high performance liquid chromatography coupled to tandem mass spectrometry. *Molecules* **2014**, *19*, 20627–20649. [[CrossRef](#)]
12. Mei, X.Y.; Hong, Y.Q.; Chen, G.H. Review on Analysis Methodology of Phenoxy Acid Herbicide Residues. *Food Anal. Methods* **2016**, *9*, 1532–1561. [[CrossRef](#)]
13. Crespo-Corral, E.; Santos-Delgado, M.J.; Polo-Díez, L.M.; Soria, A.C. Determination of carbamate, phenylurea and phenoxy acid herbicide residues by gas chromatography after potassium tert-butoxide/dimethyl sulphoxide/ethyl iodide derivatization reaction. *J. Chromatogr. A* **2008**, *1209*, 22–28. [[CrossRef](#)] [[PubMed](#)]
14. Hua, K.; Xiao-Gang, C.; Yu-Xia, H.; Chuan-Lai, X. Simultaneous determination of 13 phenoxy acid herbicide residues in soybean by GC-ECD. *Anal. Lett.* **2006**, *39*, 2617–2627. [[CrossRef](#)]
15. Poole, C.F.; Lenca, N. Applications of the solvation parameter model in reversed-phase liquid chromatography. *J. Chromatogr. A* **2017**, *1486*, 2–19. [[CrossRef](#)] [[PubMed](#)]
16. Vitha, M.; Carr, P.W. The chemical interpretation and practice of linear solvation energy relationships in chromatography. *J. Chromatogr. A* **2006**, *1126*, 143–194. [[CrossRef](#)]
17. Cela, R.; Ordoñez, E.Y.; Quintana, J.B.; Rodil, R. Chemometric-assisted method development in reversed-phase liquid chromatography. *J. Chromatogr. A* **2013**, *1287*, 2–22. [[CrossRef](#)]
18. Torres-Lapasió, J.R.; García-Alvarez-Coque, M.C.; Rosés, M.; Bosch, E.; Zissimos, A.M.; Abraham, M.H. Analysis of a solute polarity parameter in reversed-phase liquid chromatography on a linear solvation relationship basis. *Anal. Chim. Acta* **2004**, *515*, 209–227. [[CrossRef](#)]
19. Héberger, K. Quantitative structure-(chromatographic) retention relationships. *J. Chromatogr. A* **2007**, *1158*, 273–305. [[CrossRef](#)]
20. Kaliszan, R. QSRR: Quantitative structure-(chromatographic) retention relationships. *Chem. Rev.* **2007**, *107*, 3212–3246. [[CrossRef](#)]
21. D'Archivio, A.A.; Donato, F.D.; Foschi, M.; Maggi, M.A.; Ruggieri, F. Uhplc analysis of saffron (*crocus sativus* l.): Optimization of separation using chemometrics and detection of minor crocetin esters. *Molecules* **2018**, *23*, 1851. [[CrossRef](#)]
22. Sadoughi, N.; Schmidtke, L.M.; Antalick, G.; Blackman, J.W.; Steel, C.C. Gas chromatography-mass spectrometry method optimized using response surface modeling for the quantitation of fungal off-flavors in grapes and wine. *J. Agric. Food Chem.* **2015**, *63*, 2877–2885. [[CrossRef](#)] [[PubMed](#)]
23. Skartland, L.K.; Mjøs, S.A.; Grung, B. Experimental designs for modeling retention patterns and separation efficiency in analysis of fatty acid methyl esters by gas chromatography-mass spectrometry. *J. Chromatogr. A* **2011**, *1218*, 6823–6831. [[CrossRef](#)] [[PubMed](#)]
24. Missaoui, I.; Sayedi, L.; Jamoussi, B.; Hassine, B. Ben Response surface optimization for determination of volatile organic compounds in water samples by headspace-gas chromatography-mass spectrometry method. *J. Chromatogr. Sci.* **2009**, *47*, 257–262. [[CrossRef](#)]
25. D'Archivio, A.A.; Maggi, M.A.; Marinelli, C.; Ruggieri, F.; Stecca, F. Optimisation of temperature-programmed gas chromatographic separation of organochloride pesticides by response surface methodology. *J. Chromatogr. A* **2015**, *1423*, 149–157. [[CrossRef](#)] [[PubMed](#)]
26. D'Archivio, A.A.; Maggi, M.A.; Mazzeo, P.; Ruggieri, F. Quantitative structure-retention relationships of pesticides in reversed-phase high-performance liquid chromatography based on WHIM and GETAWAY molecular descriptors. *Anal. Chim. Acta* **2008**, *628*, 162–172. [[CrossRef](#)]
27. D'Archivio, A.A.; Maggi, M.A.; Ruggieri, F. Prediction of the retention of s-triazines in reversed-phase high-performance liquid chromatography under linear gradient-elution conditions. *J. Sep. Sci.* **2014**, *37*, 1930–1936. [[CrossRef](#)]

28. Fatemi, M.H.; Abraham, M.H.; Poole, C.F. Combination of artificial neural network technique and linear free energy relationship parameters in the prediction of gradient retention times in liquid chromatography. *J. Chromatogr. A* **2008**, *1190*, 241–252. [[CrossRef](#)]
29. Golubović, J.; Protić, A.; Zečević, M.; Otašević, B.; Mikić, M.; Živanović, L. Quantitative structure-retention relationships of azole antifungal agents in reversed-phase high performance liquid chromatography. *Talanta* **2012**, *100*, 329–337. [[CrossRef](#)]
30. Tham, S.Y.; Agatonovic-Kustrin, S. Application of the artificial neural network in quantitative structure-gradient elution retention relationship of phenylthiocarbamyl amino acids derivatives. *J. Pharm. Biomed. Anal.* **2002**, *28*, 581–590. [[CrossRef](#)]
31. D'Archivio, A.A.; Incani, A.; Ruggieri, F. Cross-column prediction of gas-chromatographic retention of polychlorinated biphenyls by artificial neural networks. *J. Chromatogr. A* **2011**, *1218*, 8679–8690. [[CrossRef](#)]
32. D'Archivio, A.A.; Giannitto, A.; Maggi, M.A.; Ruggieri, F. Cross-column retention prediction in reversed-phase high-performance liquid chromatography by artificial neural network modelling. *Anal. Chim. Acta* **2012**, *717*, 52–60. [[CrossRef](#)] [[PubMed](#)]
33. D'Archivio, A.A.; Maggi, M.A.; Ruggieri, F. Cross-column prediction of gas-chromatographic retention indices of saturated esters. *J. Chromatogr. A* **2014**, *1355*, 269–277. [[CrossRef](#)] [[PubMed](#)]
34. D'Archivio, A.A.; Maggi, M.A.; Ruggieri, F. Multiple-column RP-HPLC retention modelling based on solvatochromic or theoretical solute descriptors. *J. Sep. Sci.* **2010**, *33*, 155–166. [[CrossRef](#)] [[PubMed](#)]
35. Fanali, S.; Haddad, P.R.; Poole, C.F.; Schoenmakers, P.; Lloyd, D. *Liquid Chromatography: Fundamentals and Instrumentation*; Elsevier: Amsterdam, The Netherlands, 2013; ISBN 9780124158078.
36. Jandera, P.; Churáček, J. Gradient elution in liquid chromatography. II. Retention characteristics (retention volume, band width, resolution, plate number) in solvent-programmed chromatography - theoretical considerations. *J. Chromatogr. A* **1974**, *91*, 223–235. [[CrossRef](#)]
37. Kaliszan, R.; Wiczling, P.; Markuszewski, M.J. pH Gradient Reversed-Phase HPLC. *Anal. Chem.* **2004**, *76*, 749–760. [[CrossRef](#)] [[PubMed](#)]
38. Aschi, M.; D'Archivio, A.A.; Mazzeo, P.; Pierabella, M.; Ruggieri, F. Modelling of the effect of solute structure and mobile phase pH and composition on the retention of phenoxy acid herbicides in reversed-phase high-performance liquid chromatography. *Anal. Chim. Acta* **2008**, *616*, 123–137. [[CrossRef](#)]
39. Todeschini, R.; Consonni, V. *Molecular Descriptors for Chemoinformatics*, 2nd ed.; Wiley-VCH: Weinheim, Germany, 2009; pp. 1–1257.
40. Geladi, P.; Kowalski, B.R. Partial least squares regression: A tutorial. *Anal. Chim. Acta.* **1986**, *185*, 1–17. [[CrossRef](#)]
41. Martens, H.; Naes, T. *Multivariate calibration*; John Wiley & Sons: New York, NY, USA, 1991.
42. Nasab, S.G.; Semnani, A.; Marini, F.; Biancolillo, A. Prediction of viscosity index and pour point in ester lubricants using quantitative structure-property relationship (QSPR). *Chemometr. Intell. Lab. Syst.* **2018**, *183*, 59–78. [[CrossRef](#)]
43. Wen, Y.; Amos, R.I.J.; Talebi, M.; Szucs, R.; Dolan, J.W.; Pohl, C.A.; Haddad, P.R. Retention prediction using quantitative structure-retention relationships combined with the hydrophobic subtraction model in reversed-phase liquid chromatography. *Electrophoresis* **2019**, *40*, 2415–2419. [[CrossRef](#)]
44. Leardi, R.; Boggia, R.; Terrile, M. Genetic algorithms as a strategy for feature selection. *J. Chemometrics.* **1992**, *6*, 267–281. [[CrossRef](#)]
45. Ciura, K.; Kawczak, P.; Nowkowska, J.; Bczek, T. Introduction to Quantitative Structure-Retention Relationships. In *Chemometrics for Chromatography*; Komsta, L., Hyden, Y., Sherma, J., Eds.; CRC Press: Boca Raton, FL, USA, 2016; pp. 303–318.
46. Kaliszan, R.; Markuszewski, M. In Quantitative Structure-Retention Relationship in Studies of Monolithic Materials. In *Monolithic silicas in separation science*; Unger, K., Tanaka, U., Machtejevas, E., Eds.; Wiley-VHC: Weinheim, Germany, 2011; pp. 159–171.
47. Poole, C.F. Influence of Solvent Effects on Retention of Small Molecules in Reversed-Phase Liquid Chromatography. *Chromatographia* **2019**, *82*, 49–64. [[CrossRef](#)]
48. Nekrasova, N.A.; Kurbatova, S.V. Quantitative structure-chromatographic retention correlations of quinoline derivatives. *J. Chromatogr. A* **2017**, *1492*, 55–60. [[CrossRef](#)] [[PubMed](#)]

49. Escuder-Gilabert, L.; Bermúdez-Saldaña, J.M.; Villanueva-Camañas, R.M.; Medina-Hernández, M.J.; Sagrado, S. Reliability of the retention factor estimations in liquid chromatography. *J. Chromatography A* **2004**, *1033*, 247–255. [[CrossRef](#)] [[PubMed](#)]
50. Mohamadi, F.; Richards, N.G.J.; Guida, W.C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W.C. Macromodel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990**, *11*, 440–467. [[CrossRef](#)]
51. Leardi, R. Genetic algorithm-PLS as a tool for wavelength selection in spectral data sets. In *Nature Inspired Methods in Chemometrics: Genetic Algorithms and Artificial Neural Networks*; Leardi, R., Ed.; Elsevier: Amsterdam, The Netherlands, 2003; pp. 169–196.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).