Green Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/greenchem

Aquatic Ecotoxicity of Personal Care Products: QSAR models and ranking for prioritization and safer alternatives' design

Paola Gramatica*, Stefano Cassani and Alessandro Sangion

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Theoretical and Applied Sciences, University of Insubria, Varese, Italy
E-mail: paola.gramatica@uninsubria.it; http://www.qsar.it; Tel: +39-0332-421573

Personal Care Products (PCPs) ingredients, widely used all over the world, during last years became chemicals of increasing environmental concern, mainly because they are detected in water and may harm wildlife. Due to their high structural heterogeneity, to the big number of end-points and the huge lack of experimental data it is very important to have tools able to quickly highlight the most hazardous and toxic compounds, focusing the experiments on the prioritized chemicals. *In silico* tools, like QSAR models based on structural molecular descriptors, can predict missing data for activities and properties necessary to prioritize existing or even not yet synthesized chemicals for their potential hazard. In the present study, new externally validated QSAR models, specific to predict acute PCPs' toxicity in three key organism of aquatic trophic level, i.e. algae, crustacean and fish, were developed according to the OECD principles for the validation of QSARs, using the QSARINS software. These OLS models are based on theoretical molecular descriptors calculated by free PaDEL-Descriptor, selected by Genetic Algorithm: are statistically robust, externally predictive and characterized by a wide structural applicability domain. They were applied to predict acute toxicity for over 500 PCPs without experimental data; a trend of acute aquatic toxicity was highlighted by PCA allowing the ranking of inherently more toxic compounds, using also a MCDM approach for prioritization purposes. Additionally, a QSAR model for the prediction of this aquatic toxicity index (ATI) was proposed to be applicable in QSARINS for the *a priori* chemical design of not environmentally hazardous PCPs.

Keywords
Personal Care Products, QSARINS, QSAR Ecotoxicity models, Aquatic Toxicity Index

Introduction

Personal care products (PCPs) and their ingredients include a broad number of compounds that are used in daily lives, such as soaps, perfumes, detergents, cleaning agents, disinfectants, sunscreens, deodorants, sprays, etc. During the last years, an increasing attention was paid to the concerns related to the environmental occurrence and possible harmful impact of PCPs, which became now a widely and well recognized class of emerging environmental pollutants.[1]

PCPs are detected with a greater frequency mainly in the aquatic environments, where they are continually released by either being rinsed from human bodies and washed down drains and sewer systems. In fact, they are among the most frequently detected chemicals in surface waters,[2] but they have been also found in drinking and ground water, seawater, sewage and wastewater treatment plants (STP, WWTP).[3–7] In addition, they are detected also in soil and sediments.[3–6,8,9] The measured concentrations of PCPs in surface waters around the world mainly range from µg/L to tens/thousands of ng/L, similar to those measured in sediments.[4] High concentrations of synthetic musk fragrances were found in WWTPs worldwide.[10]

Even though the environmental concentrations of these contaminants could be considered relatively low, their regular input to aquatic ecosystems imparts them with 'pseudo-persistence', leading to the long-term exposure of aquatic communities,[11] raising concerns for their potential adverse effects for wildlife and environmental health. The occurrence of UV filters and stabilizers is reported in fish, mussels, crustaceans and marine mammals like dolphins, with concentrations that range from tens to thousands of ng/g.[9,12–18]

Recently, the number of (eco)toxicity studies on PCPs, which have assessed both acute and sub- lethal effects, has significantly grown. The primary issues of concern for PCPs are their ability to bioaccumulate as well as the propensity to cause estrogenic and endocrine effects.[8] In fact, some of these PCPs ingredients, like UV filters and polycyclic musks, have a high potential for bioaccumulation, due to their lipophilic characteristics. Bioconcentration factors (BCF) of some UV filters are greater than 5000 in fish,[8] leading to consider them as B or vB substances, also following REACH criteria for bioaccumulative compounds. Recently, in our screening work on PCPs for PBT (Persistence Bioaccumulation and Toxicity) behavior some UV sun-screeners were identified as potential PBTs.[19] Some UV filters have shown decreased reproduction and increased mortality rates of benthic organisms,[20] as well as a high acute and chronic toxicity towards green algae and *Daphnia magna.*[21]

In regard to synthetic musks, they have been shown to bioaccumulate to a great extent in aquatic wildlife, including marine mussels, different fish species and mammals.[10,17,22] Concentrations of galaxolide and tonalide, two widely used synthetic fragrances frequently found in aquatic compartments, induce oxidative and genetic damage in the zebra mussel, suggesting the involvement of oxidative stress in the mechanism of action of these aquatic pollutants to this freshwater bivalve.[10]

Additionally, it has been demonstrated that, among PCPs, UV filters, polycyclic musks, parabens and phthalates showed potential endocrine disruption.[17,23]

However, information about ecotoxicity is not so extensive if compared with other pollutants, especially regarding primary producers (algae) and consumers (aquatic invertebrates),[1] and in general little is known about the ecotoxicological characteristics of these substances. QSAR (Quantitative Structure Activity Relationship) modelling is an important and useful structural tool for discovering the potential inherent hazard of chemicals. QSAR models can predict missing experimental data, finding out the relationship between chemical structure and biological activity,

only on the existing experimental data. The aims of the present study are: a) to fill the data gaps of ecotoxicity on aquatic organisms by QSAR models, specifically developed for PCPs; b) to compare our results with the predictions obtained by a widely applied online modelling tool such as ECOSAR;[24] c) to apply our models for the ranking of a big set of more than 500 PCPs, collected in our previous study;[19] d) to highlight and prioritize the most hazardous PCPs, ranking them by tools able to combine toxicity data from multiple sources, such as Principal Component Analysis (PCA) and Multi-Criteria Decision Making (MCDM); and e) to propose a final QSAR model of the obtained cumulative aquatic toxicity trend, defined as an aquatic toxicity index (ATI). All the proposed QSAR models will be applicable by using the QSARINS-Chem module in the software QSARINS.[25,26]

The whole goal of this work is to propose a comprehensive framework for identifying *a priori* the potentially more toxic PCPs, an approach useful also for avoiding the synthesis, and subsequent introduction to the market and into the environment, of harmful compounds, as supposed "safer alternatives". This can be done by combining QSAR models and chemometric methods, exploiting the fundamental information inherent in the chemical structure, in the benign by design approach of the Green Chemistry.

Materials and Methods

Dataset and data curation

Experimental data for the acute toxicity of the whole set of 534 PCPs,[19] in the three organisms studied (i.e. algae, *Daphnia*, fish), were mainly harvested from the ECOTOX database.[27] Great attention was devoted in collecting homogeneous data, because this is the first care that QSAR modellers must apply in preparing the input datasets for their modelling. Only a careful data curation for obtaining homogeneous data sets can guarantee reliable QSAR models.[28] Experimental data were specifically filtered for the species (only data obtained on a single specific species for each trophic level were selected), defined time of exposure (only data obtained at the same exposure time were collected), endpoint and measured effect , trying to ensure the highest degree of homogeneity in experimental measures, that derive from different sources. If different and multiple values were found for a specific chemical, the minimum LC/EC_{50} value was taken and modelled, considering the "worst case scenario" (i.e. the most toxic value).

Once these experimental values were selected and filtered, the data were additionally carefully checked, removing duplicates and measures reported as "nominal concentration". All the values, selected as input responses, were converted to mol/l and then expressed as negative logarithm of the concentration "-log(mol/l)" (or pEC_{50}/LC_{50}).

Out of the 534 PCPs ingredients in our study, experimental ecotoxicological data, which satisfied the filters for homogeneity previously reported, were found only for 107 molecules (20%) and only 11 chemicals (2%) possess experimental data for all the three studied trophic levels,

For *Pseudokirchneriella subcapitata* only 20 PCPs with consistent data for growth rate inhibition at 96h (EC_{50}96h), according to the OECD test 201,[29]were found .

For *Daphnia magna* 54 data immobilization at 48h (EC_{50}48h) were collected from ECOTOX database, while the remaining 18 data, for the same end point,were taken from different sources, including literature[30 28] and safety data sheets, verifying their consistency with the studied endpoint, for a final data set of 72 PCPs with homogeneous data according the OECD test 202.[31]

For *Pimephales promelas* homogeneus data for mortality at 96h (LC_{50}96h), according to the OECD test 203[32], for 67 PCPs were taken from the ECOTOX database.

Chemical structures and molecular descriptors

The chemical structures for all the studied PCPs ingredients were carefully verified and drawn in HyperChem software.[33] The structures, saved as .hin files, were then converted by OpenBabel (ver. 2.3.2)[34] into MDL MOL (.mol), which is the recommended input format in PaDEL-Descriptor software.[35] Various molecular descriptors (only mono- and bi-dimensional to avoid the complexity of 3D- conformation), that encode for the different structural features of the chemicals, including various calculated logP such as XlogP, MlogP, AlogP and CrippenLogP, were then calculated using PaDEL-Descriptor (ver. 2.21), implemented in QSARINS ver. 2.2.1.[25,26] In order to minimize redundant and not useful information, constant and semi-constant (80%) values and descriptors found to be pair-wise correlated more than 0.98 were excluded in QSARINS.

A total of 453, 623, and 602 descriptors, respectively for algae, *Daphnia* and fish toxicity datasets, were used as input variables for QSAR modelling. Finally, a total of 481 descriptors were used as a structural input for the development of the QSAR model related to the Aquatic Toxicity Index (ATI) of PCPs.

QSAR Modelling and ranking

QSAR models were developed by Multiple Linear Regression (MLR) using the Ordinary Least Squares (OLS) method and the Genetic Algorithm-Variable Subset Selection (GA-VSS), included in QSARINS,[25,26] was applied for the selection of modelling descriptors. Following the OECD principles for QSAR validation,[36] several statistical parameters were used to verify the internal stability and the external predictivity of the developed QSAR models.[37,38] The coefficient of determination R^2 was used as a measure of the goodness-of-fit, while internal robustness was verified by the cross-validation coefficient Q^2_{LOO} (leave-one-out, used also as fitness function in GA) and Q^2_{LMO} (leave-many-out, 30 %). In order to exclude chance correlation between the selected descriptors and the modelled endpoint, the Y-Scrambling method was applied.[37,39] The external validation of the models was performed by calculating different parameters, i.e. Q^2ext-F1,[40] Q^2ext-F2,[41] Q^2ext-F3,[42] and external CCC (CCCext, Concordance Correlation Coefficient).[43–45] In addition, the Root Mean Squared of Errors (RMSE), that summarizes the overall error, was used to measure and compare the accuracy of the reproduced data in the training set ($RMSE_{TR}$) and predicted data in the prediction set ($RMSE_P$). In order to verify the actual predictive capability of the selected models, the datasets were split *a priori*, before model development, into a training set (~70% of compounds) used for model development, and a prediction set (~30% of compounds) used for external validation. The chemicals included in the prediction sets were never used in the model development step to select the modelling descriptors.[37,38] Three different splitting techniques, all available in QSARINS,[25,26] were applied: a) randomly b) by ordered response, and c) by structural similarity.[39]

In the first splitting, compounds were randomly divided into training and prediction set, using the randomization algorithm included in QSARINS,[25,26] always setting 30% of chemicals for the inclusion in the prediction set. In the split by ordered response, molecules were ordered according to their increasing toxicity value (pEC_{50} or pLC_{50}), and one out of every three chemicals was put in the prediction set (including always the most and the least toxic compounds in the training sets, to cover the response range).

The splitting of the data set, based on structural similarity, is realized by Principal Component Analysis (PCA) on the available molecular descriptors. The chemicals were sorted by PC1 score that explain the majority of the total structural variance; then, again, one out of every three chemicals was put in the prediction set. In this way, the selection of a structurally meaningful training set and an equally representative prediction set was realized.

Finally, the same set of structural descriptors, which had demonstrated their modelling ability for the prediction of chemicals not used in model development (i.e. compounds in the prediction sets) in all three independent splitting procedures, was applied to derive a full model from the whole data set, in order to exploit, at the end, all the available information and ensuring a wider applicability domain of this full model.

Principal Component Analysis of the three modelled end-points (i.e. pEC_{50} in algae and *Daphnia* and pLC_{50} in fish) was finally performed on the available experimental data implemented with the toxicity values predicted by our specific QSAR models, in order to rank the studied PCPs according to their cumulative acute aquatic toxicity.

The Multi-Criteria Decision Making (MCDM) method included in QSARINS-Chem[26] was used to highlight the most toxic chemicals within the studied simplified aquatic scenario. MCDM is a technique that summarizes the performances of a certain number of criteria simultaneously, as a single number (score) between 0 and 1. This is done associating to every criteria, in our case different predictions for the studied endpoints, a desirability function which values range from 0 to 1 (where 0 represents the less toxic compound and 1 the most toxic), and giving different weights to the selected criteria. The sum of the weights of the criteria must be 1, and in our case, we used the same weight for each criteria: 0.333, which is 1/3 (total/number of criteria). The geometric average of all the values obtained from the desirability functions gives the MCDM value (i.e. the ranking).

Applicability Domain assessment

The Applicability Domain (AD) assessment of the new proposed models for PCPs was performed, at the model development step, following the leverage approach.[37,39] The leverage method is based on the calculation of the *hat* matrix: the diagonal values of this matrix (the leverage values: *h*) are used to verify the presence of structural outliers, i.e. those compounds with *h* greater than the cut off values h*. The h* value is here calculated as 3p'/n, where p' is the number of the model variables plus one, and n is the number of training compounds. Response outliers (i.e. the chemicals badly predicted by the models) are identified as those compounds with cross-validated standardized residuals greater than 2.5 standard deviation units. Both types of outliers have been detected using the Williams plot, i.e. the plot of the diagonal values of the hat matrix (*h, diagonal values of the hat matrix*) which represent the similarity of a given compound to the training set versus the differences between predicted and observed values(standardized residuals).

When the developed models were applied to the whole set of 534 PCPs, to define more rigorously the degree of extrapolation of the obtained predictions, three different and complementary approaches for the study of AD were performed, based on: leverage (*hat*), range of modelling descriptors and PCA bounding box.

In leverage approach, again, chemicals with *h* greater than the cut off value h* were considered out of the AD, as can be observed in the pertinent Insubria Graphs.[39] In this graph the h values for each chemical (the leverages) together with the h* cut off value (vertical line) are reported on the abscissa axis against the predictions, reported on the axis of ordinates. The training set values are delimited by horizontal lines.

For the range of descriptors, if a compound shows a descriptor value out of the range of the modelling descriptors, different in any model, it is considered out of the model AD for this approach. Within the PCA approach, a sub-structural box space (based on the PC1 and PC2 scores of the modelling descriptors) was defined, delimited by the training chemicals; if a predicted chemical fell into this chemical space box was considered inside the AD.

The chemicals "view" inside the AD for all the three approaches altogether were defined as "In AD" (into the AD), with more reliable predictions due to their high interpolation degree; the chemicals "view" outside the AD for all the approaches considered, were defined as "Out AD" (out the AD), with predictions that could be also corrected but less reliable due to their

extrapolation degree; the chemicals view inside the AD for at least one method, were considered "Bor AD" (borderline the AD).

Results

Development of QSAR models for PCP acute aquatic toxicity

Before to present and discuss the results of the QSAR modelling performed in this study, it is important to highlight that the quality of QSAR models cannot be better that the input data on which they are developed on. For this reason, the first important aspect that QSAR modellers must deal with is the preparation of a reliable data set on which to develop QSAR models for finding the relationship between chemical structure and biological activity . Therefore, the preliminary careful data curation for the preparation of input data, to be used as training set, is a very important point, highlighted also in the "OECD guidance document on the validation of (Quantitative) Structure Activity Relationships [(Q)SAR] models"[46] and in literature.[28,39] The chemical structure must be carefully verified and experimental data must be selected specifically to prepare an homogenous data set. The biological end points to be modelled by QSAR must have been obtained on the same species measuring the same toxicological response, applying the same experimental conditions, preferably following the OECD protocols. Otherwise QSARs will model mainly the variability present in the biological data, due to the different experiments, instead to model the variance of the studied biological activity which is related to the chemical variance, i.e. the differences in the molecular structure which influence the specific end point of interest. This is the only variance that must be modelled by QSARs.

For this reason, careful selection and filter of experimental data must be a preliminary step and the derived availability of experimental data, useful for reliable QSAR modelling, is generally limited for some species, as in our study for algae.

Pseudokirchneriella subcapitata. The dataset for acute toxicity of PCPs ingredients on *P. subcapitata* was composed of 20 chemicals, for which homogeneous experimental data satisfying the condition of reliable QSAR modelling are available. Thus, a QSAR model based only on two theoretical molecular descriptors is here selected as the best predictive for this endpoint, among the GA population of models. Even if the available homogeneous experimental data are limited, the input data set was split before the model development to verify the model external predictivity. Three different training sets of 12 chemicals, obtained by the splitting techniques (explained in Materials and Methods paragraph) were used for this purpose, and the developed models were externally validated on the respective three prediction sets (8 chemicals). The finally proposed QSAR model (full), here developed for the prediction of acute toxicity of PCPs ingredients in *P. subcapitata*, based on two molecular descriptors, was the best combination of variables selected by the GA-Variable Subset-Selection (VSS) procedure in the three different and independent populations of split models. In Table 1, the statistical parameters related to internal and external validation of the split models are reported, as well as the equation of the final full model, recalibrated using all the 20 experimental data, once the descriptors selected in the split models had guarantee external predictivity.

Table 1*: statistical parameters, related to internal and external validation, of the QSAR model for algae toxicity of PCPs. The equation of the full model is reported.*

Model	N_{TR}	N_{PRED}	R^2	Q^2_{LOO}	Q^2_{LMO}	R^2Y_{scr}	CCC_{EXT}	Q^2_{EXT} Fn	RMSEtr	RMSEp	AD to 534 PCPs
Split by random	12	8	0.93	0.89	0.88	0.19	0.94	0.88-0.90	0.38	0.50	-
Split by ordered response	12	8	0.95	0.92	0.91	0.18	0.95	0.89-0.89	0.34	0.50	-
Split by structural similarity	12	8	0.94	0.89	0.88	0.18	0.94	0.87-0.90	0.40	0.50	-

pEC50 *P.subcapitata* = -10.580 + 13.045 **GGI8** + 21.881 **Mp**	**20**	**-**	**0.93**	**0.91**	**0.90**	**0.11**	**0.95cv**	**-**	**0.40**	**0.45cv**	**98%**

All the models demonstrated high statistical performances: they were internally stable and robust (high Q^2_{LOO} and Q^2_{LMO}), not given by chance correlation (low R^2_{Yscr}), and externally predictive (high values of CCC_{ext} and Q^2_{extFn}) also on chemicals never seen during the model development (i.e. chemicals in different prediction sets)., The good predictive ability of these two, simple descriptors in each split model is a proof of the validity of the encoded structural information, regardless of the different composition of the three different training sets. The errors in predicting chemicals in training (Root Mean Squared Error RMSEtr) are similar to those in predicting external chemicals, not used in model development.

It is particularly interesting to note that the finally proposed full model, even if developed on the limited amount of experimentally available data, has a large structural applicability domain (AD) to the big heterogeneous set of 534 PCPs ; in fact, the coverage of the full model AD, verified with the methods explained in the Methods paragraph and with the results commented in the specific following section, is near to 100%. This could be considered as a guarantee of the reliability and generalizability of this QSAR model.

To help the evaluation of the proposed QSAR model, in Figure 1 the graph of experimental vs. predicted values and the Williams Plot for model AD, related to the split model by ordered response, are reported. The equations and the remaining plots, related to other splitting schemes and the full model, are reported in ESI (Figures ES1-ES3).

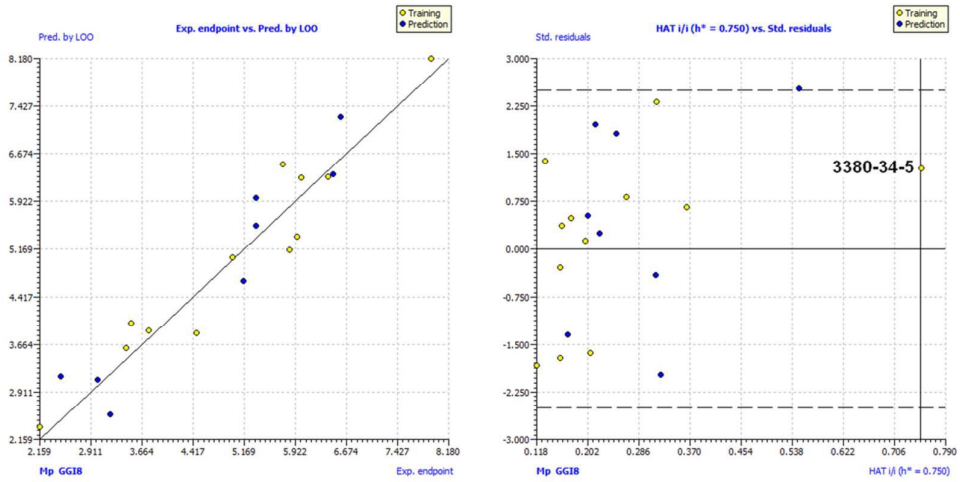


Fig. 1*. Graph of experimental vs. predicted values (on the left) and Williams Plot (on the right) for the P.subcapitata toxicity model (splitting by ordered response).*

As can be seen in the right part of Figure 1, chemical with CAS 3380-34-5 (triclosan, a widely used antibacterial agent) is the unique, borderline, structural outlier (on the cut –off value of h*).This is most evident in the full model (see Figure ES3), and thus triclosan can be considered a relatively high leverage compound, influent for the selection of the two modelling variables when it is included in the training set. Regarding outliers for the response, the compound on the upper limit for the response outliers in Figure 1 is not an outlier in the other splitting and in the full model. Concerning the modelling descriptors, it is interesting to note that none of the various LogP descriptors was selected by the GA-VSS procedure as good modelling variable for this end-point, even though they were available as input variables. The two selected descriptors, both positively correlated with the modelled response and therefore with an increasing incidence on this specific toxicity, are the mean atomic polarizability (Mp,

the most important) and a topological index (GGI8, the Topological charge index of order 8,[47] which represents the total charge transfer between atoms). Both are related not only to molecular dimension, but mainly to the electronic distribution in the chemical structure, represented by polarizability and charges, features probably not adequately taken into account by the various calculated LogP parameters.

Daphnia magna. The dataset for acute toxicity of PCPs ingredients on *D. magna* was composed of 72 compounds, and a QSAR model based on three molecular descriptors was chosen as the best for predictivity. The whole set was split into three training sets used for model development and internal validation, and three prediction sets used for external validation of the proposed model. Random splitting gave a training set of 50 PCPs and a prediction set of 22 PCPs, while the remaining both splitting schemes (i.e. ordered response and structural similarity) resulted in training sets of 49 PCPs and prediction sets of 23 PCPs. The proposed QSAR model, here developed for prediction of acute toxicity of PCPs ingredients in *D.magna*, was built on three molecular descriptors, selected as the best combination of variables by the GA-VSS procedure on the three independent populations of split models. In Table 2, the statistical parameters related to the internal and external validation of the split models as well as of the final full model , recalibrated using all the 72 experimental data, once the external predictivity of the three split models were evaluated and verified, (with its equation) are reported.

Table 2*: statistical parameters, related to the internal and external validation, of the proposed model for Daphnia magna toxicity of PCPs. The equation of the full model is reported.*

Model	N_{TR}	N_{PRED}	R^2	Q^2_{LOO}	Q^2_{LMO}	R^2Y_{scr}	CCC_{EXT}	Q^2_{EXT} Fn	RMSEtr	RMSEp	AD to 534 PCPs
Split by random	50	22	0.87	0.84	0.84	0.06	0.95	0.85-0.91	0.51	0.54	-
Split by ordered response	49	23	0.89	0.87	0.86	0.06	0.94	0.87-0.90	0.54	0.50	-
Split by structural similarity	49	23	0.91	0.89	0.89	0.06	0.89	0.80-0.88	0.50	0.57	-
pEC50 *D.magna* = 4.485 + 0.015 **MW** - 3.205 **ATSC0c** - 1.518 **GATS1p**	**72**	**-**	**0.89**	**0.88**	**0.87**	**0.04**	**0.93cv**	**-**	**0.52**	**0.55cv**	**98%**

The statistical parameters show that the models are stable and robust, not given by chance correlation, and also externally predictive when applied to the three, different and independent prediction sets. The very similar values of the Root Mean Squared Error (RMSE) on training and prediction sets guarantee the great ability of the models to predict the chemicals, which are never seen during the modelling step (i.e. in prediction sets) , with similar errors to those obtained in the training sets.. The structural coverage of the full model on the whole set of 534 PCP ingredients is very large, again near to 100%, showing a wide AD also for this *D. magna* full model. In Figure 2 the graph of experimental vs. predicted values and the Williams Plot related to the Structural similarity split model, to study the model AD, are reported. The equations and the remaining plots, related to other splitting schemes and full model, are reported in ESI (Figures ES4-ES6).

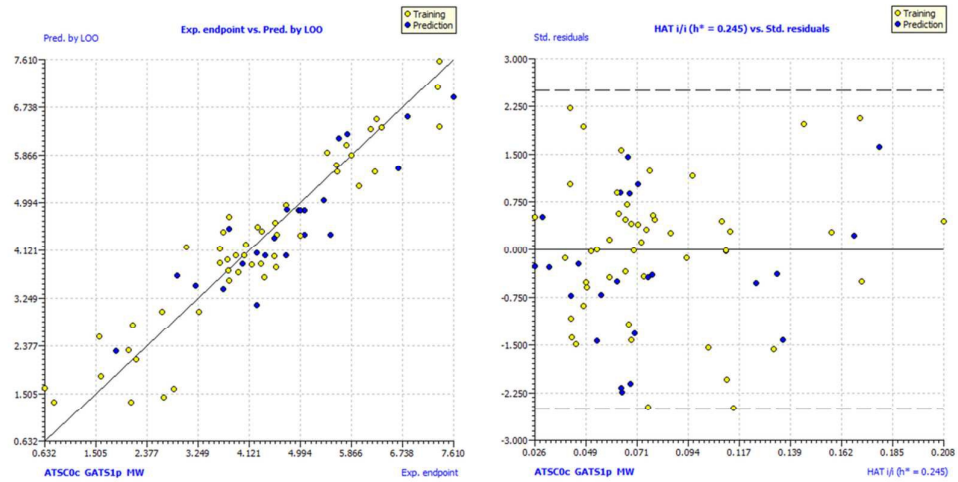


Fig. 2*. Graph of experimental vs. predicted values (on the left) and Williams Plot (on the right) for the D. magna toxicity model (splitting by structural similarity)*

As can be seen in the whole Figure 2, no response and structural outliers were detected in this model.

The most important descriptor in this model equation is MW, molecular weight, which is positively correlated with the modelled response; it means that, in this training sets, the PCPs of higher MW are the most toxic in *D. magna*. The remaining descriptors, both negatively correlated with this toxicity, are ATSC0c (Centered Broto-Moreau autocorrelation lag0 weighted by charges) and GATS1p (Geary autocorrelation lag1 weighted by polarizabilities), belonging to the autocorrelation 2-D descriptors, weighted by electronic features.[47] These descriptors represent the distribution of the electronic properties (polarizability and charges as weights) in the molecule. Again, as for algae toxicity, the various LogP descriptors were not selected by the GA-VSS procedure as good modelling variable for this end-point, probably because in calculated LogPs, which are mainly related to molecular dimension, the features related to the polarity of the chemicals are not sufficiently taken into consideration.

Pimephales promelas*.* Regarding this endpoint, i.e. acute toxicity on *P. promelas*, for which we had a whole dataset composed of 67 chemicals, two different models were selected as the best for predictivity: with LogP (XLogP calculated in PaDEL-Descriptor 2.21)[35] and without logP, "LogPfree" (see Table ES1 and Figure ES7-ES10 for more details of LogPfree model). Every model was developed and internally validated on three different training sets and externally validated on the respective three prediction sets. Random splitting gave a training set of 47 PCPs and a prediction set of 20 PCPs, while the Ordered by Response and the Structural Similarity splitting gave a training set of 46 PCPs and a prediction set of 21 PCPs. The proposed QSAR models, here developed for prediction of acute toxicity of PCPs ingredients in *P. promelas*, were based on different number of modelling molecular descriptors, selected as the best combinations of variables by the GA-VSS procedure on the different populations of split models; the model with LogP (XlogP) was based on three descriptors, while the LogPfree models was based on four descriptors. In Table 3, the statistical parameters related to the internal and external validation of the split models and of the full model with LogP descriptor (with the relative equation) are reported.

Table 3*: statistical parameters, related to the internal and external validation, of the proposed model for P.promelas toxicity of PCPs, with LogP The equation of the full model is also reported.*

LogP Model	N_{TR}	N_{PRED}	R^2	Q^2_{LOO}	Q^2_{LMO}	R^2Y_{scr}	CCC_{EXT}	Q^2_{EXT} Fn	RMSEtr	RMSEp	AD to 534 PCPs
Split by random	47	20	0.81	0.78	0.77	0.07	0.90	0.80-0.85	0.62	0.56	-
Split by ordered response	46	21	0.84	0.81	0.81	0.07	0.85	0.72-0.81	0.60	0.63	-
Split by structural similarity	46	21	0.83	0.81	0.80	0.07	0.86	0.73-0.81	0.60	0.63	-
pLC50 *P.promelas* = 1.701 + 0.429 **XLogP** + 2.356 m**inHother** + 15.379 **AVP-7**	**67**	**-**	**0.81**	**0.79**	**0.79**	**0.04**	**0.89cv**	**-**	**0.60**	**0.63cv**	**95%**

XlogP calculated by PaDEL-Descriptor 2.21[35] in QSARINS[25,26]

The reported parameters show that the models are stable and robust, not given by chance correlation, with a good external predictivity when applied to the three independent prediction sets. The predictive ability is similar to the calculation performance, as evident from the very similar RMSE values for training and prediction sets The final full model, recalibrated using all the 67 experimental data, once the external predictivity of the three split models were evaluated and verified, ensures a structural coverage (AD) of the whole set of 534 PCPs of 95%. In Figure 3 the graph of experimental vs. predicted values and the Williams Plot related to the Random split model are reported. The equations and the remaining plots, related to other splitting schemes and full model, are reported in ESI (Figures ES11-ES13).

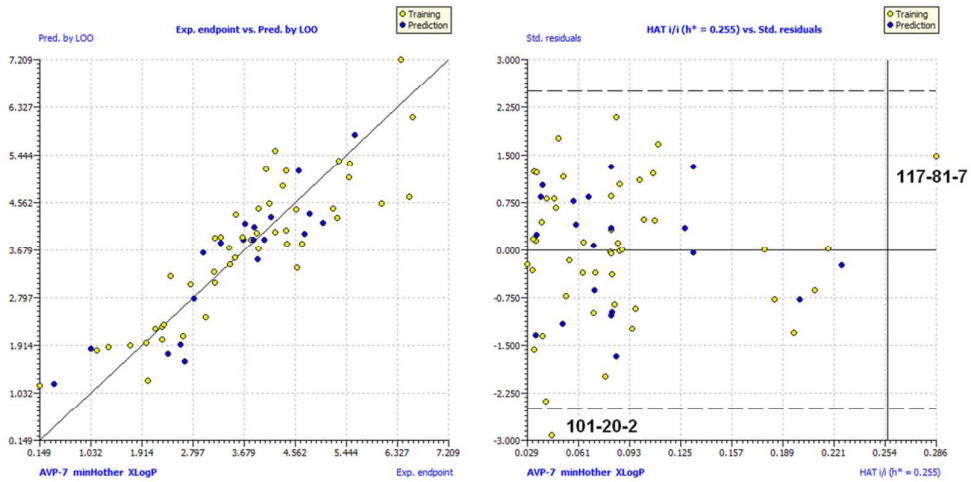


Fig. 3*. Graph of experimental vs. predicted values (on the left) and Williams Plot (on the right) for the P. promelas toxicity model (splitting by random)*

As can be seen in the Williams Plot in Figure 3, the chemical with CAS 117-81-7 (DEHP, a phthalate) is the unique structural outlier, influential for the selection of modelling variables. Furthermore, chemical with CAS 101-20-2 (triclocarban, an antibacterial agent) is the only outlier for response, being underestimated by our model. These two compounds are still outliers in the full model, as can be seen in the pertinent Williams Plot (see Figure ES11 of the Electronic Supplementary Information). The most important descriptor in the model equation is XlogP, a LogP descriptor calculated in PaDEL-Descriptor ver. 2.21.[35]

As expected, because the most lipophilic compounds are also the most bioaccumulative and potentially toxic into the organism, this descriptor has a positive sign in the equation, increasing the predicted acute toxicity of PCPs in *P. promelas*. Even though XlogP is sufficient for reaching a relatively satisfactory modelling of the modelled toxicity in fish ($Q2_{LOO}$ of 0.68), two additional descriptors were selected by the GA-VSS to increase the model performance: minHother and AVP-7, both with a positive correlation with the acute toxicity in *P.promelas*. MinHother encodes for the minimum e-state of H on aaCH (aromatic CH),

dCH2 (=CH2) or dsCH (=CH-), in our dataset this descriptor mainly discriminates the non-aromatic PCPs (minHother always 0) from aromatic PCPs (where the values of this descriptor is different from zero); the least important is AVP-7, an average valence path of order 7.[47]

Comparison with ECOSAR

A comparison with the predicted values obtained by applying ECOSAR 1.11, [24]the widely used QSAR tool for the prediction of aquatic toxicity was also performed on the here studied chemicals, and the relative RMSE values were calculated. Firstly, we compared the predictions on the entire training sets for the studied endpoints and the results are displayed in Table 4.

Then, in order to exclude any bias in this calculation for chemicals that were included in our training sets, we performed the comparison only on the prediction sets (Table ES2), generated by the three different splitting (random, ordered by response and structural similarity), and thus only on chemicals never seen during the development of our models. In Table 4 and Table ES2, the comparison A was made using the Baseline Toxicity prediction in ECOSAR (always available for all the chemicals); the comparison B was made using the "worst case" generated by ECOSAR, i.e. the lower prediction available for the studied chemical. In fact, ECOSAR can recognize one or more chemical classes within the molecular structure and apply different equations, giving thus different predictions for the studied chemicals.

Table 4: *Comparison of results from our models and ECOSAR models, with the calculated RMSE of predictions.*

Endpoint	N	RMSE UI[a]	RMSE A[b]	RMSE B[c]
P.subcapitata pEC50	20	0.40	2.28	2.51
D.magna pEC50	72	0.52	1.36	1.20
P.promelas pLC50 (with LogP)	67	0.60	0.93	0.65
P.promelas pLC50 (LogP free)	67	0.55	0.93	0.65

[a]: RMSE related to the predictions derived from QSAR models (full) presented in this work; [b]: RMSE related to the predictions derived from ECOSAR Baseline Toxicity equation; [c]: RMSE related to the predictions derived from the ECOSAR worst case scenario, i.e. the lower prediction available for the studied compound.

As can be seen from Table 4, the models presented in this work show a lower RMSE, in every case (RMSE-UI), more evident mainly for algae and *Daphnia*. Similar results can be also generally observed in the comparison within the prediction sets (Comparison B in Table ES2). Finally, we selected the most represented class recognized within the studied chemicals (i.e. Esters), performing a direct comparison (comparison C) only on compounds that show a prediction derived from the Esters equation in ECOSAR (Table ES3). Also here, the RMSE of the QSAR models presented in this paper are the lowest, confirming the previous results and the good predictive performances of our models. This comparison can demonstrate that local QSAR models, specific for PCPs, are more able to give reliable predictions, also for chemicals not included in training set, than more general models as those included in ECOSAR.

Application of the developed models to 534 PCPs

Once the external predictive capability of our models was verified to be better for PCPs than the widely used ECOSAR tool, we

applied our developed QSAR models for algae, *Daphnia* and fish acute toxicity to the whole set of the studied 534 PCPs, in order to fill the big data gaps on these ecotoxicity end-points and also to rank the studied PCPs according to their cumulative ecotoxicity, based on the integration of experimental and predicted data. This ranking can highlight potentially more toxic compounds, prioritizing them for subsequent experimental tests.

We carried out a careful AD assessment of each previous QSAR models, using the three different approaches explained in Methods, in order to focus the further steps only on the interpolated ("In AD") or borderline ("Bor AD") chemicals, which ensure a better reliability of the predictions. The AD assessment was performed according to the three methods proposed in the Materials and Methods section. The results are summarized in the following Table (Table 5):

Table 5: *AD assessment for the developed models applied to the 534 PCPs*

Model	N	% 534 PCPs into AD			% 534 PCPs into AD ("in+bor")
		Leverage	Range Desc.	PCA	
P.subcapitata full	20	97 (n=519)	91 (n=486)	89 (n=477)	98 (n=523)
D.magna full	72	96 (n=514)	95 (n=506)	96 (n=515)	98 (n=521)
P.promelas full (logP)	67	93 (n=496)	91 (n=486)	93 (n=497)	95 (n=505)
P.promelas full (logPfree)	67	87 (n=464)	83 (n=443)	93 (n=496)	93 (n=499)

Even though the algae model is based on the smallest data set (n=20), its combined AD (In+Bor) is the largest in the whole set of models, with a coverage of 98%, as the *Daphnia* model. Therefore, we are confident on the satisfactory generalizability of this "small" model and on the reliability of the QSAR predictions for PCPs without experimental data. The fish models are only slightly lower in covering the chemical structure of the entire set of PCPs ingredients, with a percentage of chemicals inside or borderline the combined AD of 95% and 93% for model with LogP and LogPfree model respectively. In general, the leverage approach and PCA seem the most conservative methods, keeping into the structural AD the largest percentages of chemicals. We report here the magnified Insubria graphs[39] for the analysis of AD of four proposed models on chemicals without experimental data (Figure 4), while in SI the original Insubria graphs and the PCAs (Figures ES14-ES17) related to the study of AD are shown.

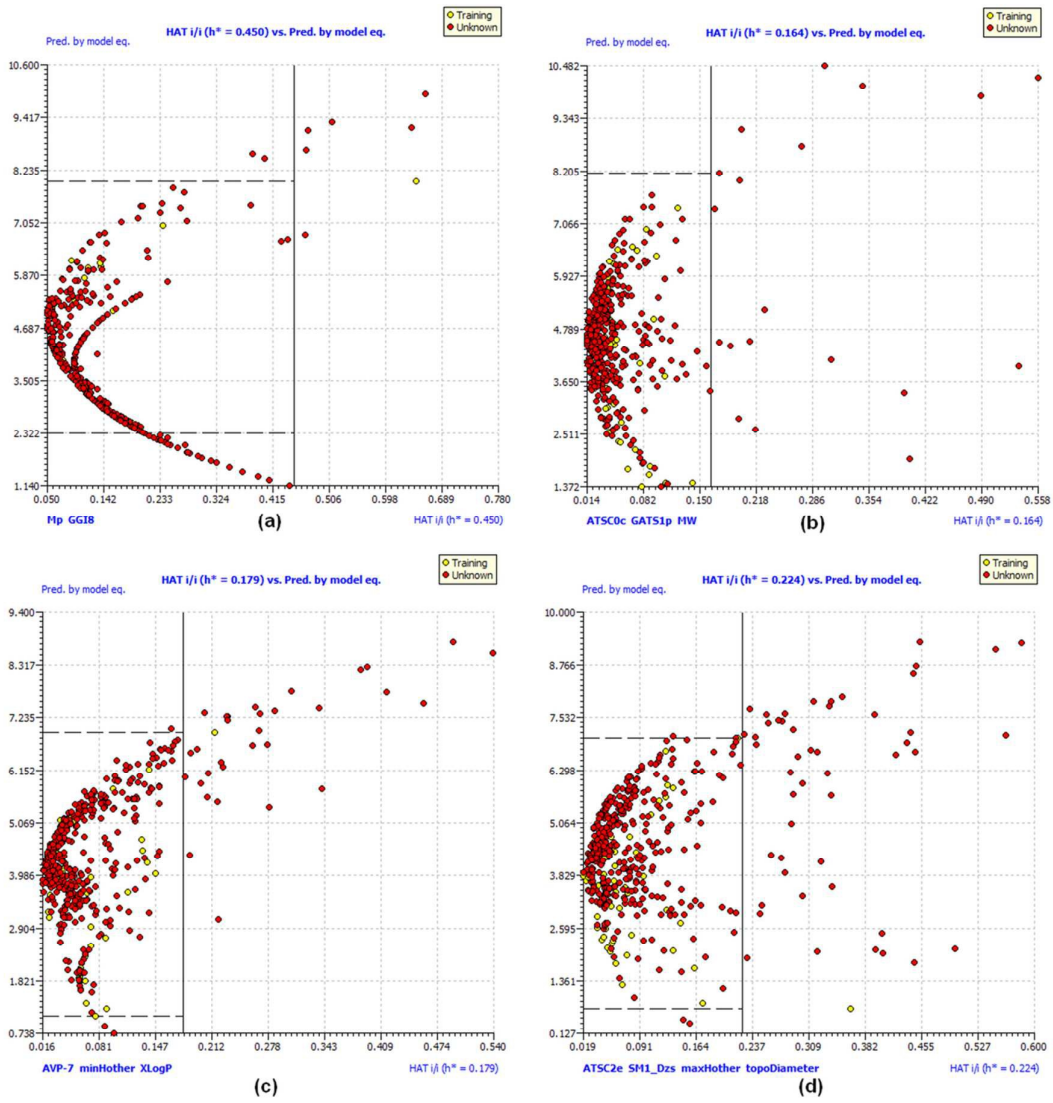


Fig. 4*. Insubria graphs, related to models for PCPs acute toxicity in algae (a), Daphnia (b), fish with logP (c) and fish LogPfree (d); in the left part of the graphs the chemicals with an h value less than the h* cut off value. The dashed lines represent the minimum and maximum value for the training set response.*

From the graphs in Figure 4, it is evident that the large majority of the studied PCPs belong to the AD of all the models, in fact they are located at the left of the h* cut off values in the structural and response zone of the training set chemicals. The predictions for these compounds can be considered more reliable, being interpolated by the model. The prediction for the chemicals that are located at the right of the leverage cut off values could be also correct, but they are extrapolated, thus they should be treated with grater care as less reliable. Overall, all the developed models show a good percentage of interpolated or borderline chemicals, leading us to consider our predictions satisfactory and reliable, also when applied to a wide set of compounds without any experimental evidence.

Trend of Aquatic Toxicity of PCPs and Priority list

Once the wide coverage of our models for the entire set of 534 PCPs was verified by the previous AD studies (Figure 4), all the

ecotoxicity data (experimental plus predictions) were finally analysed by Principal Component Analysis (PCA) in order to rank the compounds according to their cumulative aquatic toxicity and to define a trend of this cumulative toxicity for the class of PCPs, in the studied scenario, here represented by the three key organisms of the aquatic trophic level. The ultimate goal is to prioritize the compounds potentially more toxic. Chemicals outside each model ADs were excluded, in order to focus our screening exercise only on interpolated or borderline predictions ("In AD" and "Bor AD"). In this way, we excluded 50 PCPs, remaining with a final set of 484 chemicals. When experimental values were available within these 484 PCPs, they were used. Additionally, the predictions from the two *P. promelas* models were averaged in a consensus model for fish toxicity. The results are plotted in the PCA score plot (Figure 5), which explains 94% of the total variance.

In the first Principal Component (PC1 scores on x axis, explained variance about 79%), which loadings are all in the same direction, an evident trend of cumulative aquatic toxicity can be observed, from the left (the less toxic chemicals) to the right (the more toxic chemicals). Therefore, the PC1 can be defined as an Aquatic Toxicity Index (ATI).

The second principal component (PC2 scores on y axis, explained variance about 15%) discriminates the different toxicities, separating the algae from the two animal species: therefore the chemicals in the upper right zone are those more toxic on algae, while those in the down right zone are more toxic for *Daphnia* and fish.

To highlight the overall most toxic compounds in this ranking, for their inclusion in a priority list, and to define cut off values, we applied the Multi-Criteria Decision Making (MCDM) method, included in QSARINS. We selected the 40 most toxic PCPs, weighting the three different toxicity inputs (algae, *Daphnia* and fish consensus) by the same weight (weight 0.333 in every species). Therefore, in the PCA score plot of Figure 5, the predicted PCPs are differently labelled basing on the obtained MCDM score: the 40 PCPs that have obtained the highest MCMD ranking score (i.e. the more toxic in this simplified aquatic scenario) are labelled in black, while those with medium MCMD scores, corresponding to intermediate toxicity are reported in dark grey and those low MCDM scores, lower toxicity with light grey.

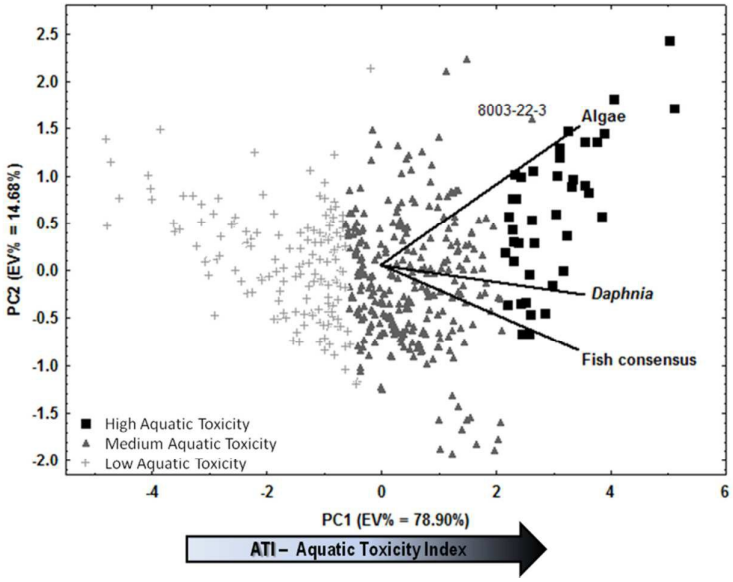


Fig. 5*. PCA score plot of the three studied toxicities and identification of an aquatic toxicity trend (Aquatic Toxicity Index (ATI)) of PCPs, in the studied simplified aquatic scenario.*

As can be observed in Figure 5, in the right part the most toxic PCPs according to both PCA and MCDM are located, with the only exception of chemical with CAS 8003-22-3 (D&C Yellow 11, a hair dye), which was predicted with a relative low toxicity in *P.*

promelas and, for this reason, was not included by the MCDM calculations within the 40 most toxic chemicals. The prioritized 40 chemicals are summarized in the following Table (Table 6), where the PCPs which were prioritized as potential PBTs (Persistent Bioaccumulative and Toxic) by our previous works[19,48] are additionally labeled by asterisk. The names, SMILES and relative structures of these 40 prioritized PCPs are reported in Table ES4. This is the first priority list proposed in our work, including the chemicals with more reliable predictions or experimental evidences (when available). These prioritized chemicals, which are highlighted by QSAR models and the subsequent PCA/MCDM analysis as potentially hazardous for the aquatic compartment, due to their toxicity on the three trophic levels, should be considered as compounds of higher concern. Careful experimental confirmation should be suggested on these 40 prioritized PCPs among the big studied data set of about 500 compounds. All the prioritization approaches are useful in reducing costs, time and animal tests.

Table 6: *40 most toxic chemicals in the studied aquatic scenario, derived from the PCA/MCDM ranking. The few available experimental data are in bold, other data are predicted by our QSAR models, verifying the AD.*

Rank	CAS	*pLC50 P.promelas cons.*	*pEC50 D.magna*	*PEC50 P.subcapitata*	MCDM score	Use
1	000101-20-2	**6.53**	**7.61**	9.21	0.981	Antimicrob.
2	005089-22-5*	6.08	7.28	9.95	0.967	UV-Filter
3	125304-04-3	7.23	6.00	7.58	0.963	UV-Filter
4	000084-62-8	**6.60**	6.18	7.50	0.958	Phthalate
5	052829-07-9	6.26	6.44	7.39	0.953	UV-Filter
6	000117-81-7	**6.39**	**6.39**	**6.59**	0.950	Phthalate
7	003380-34-5	**6.06**	**6.20**	**7.93**	0.948	Antimicrob
8	027987-25-3	6.08	6.29	7.30	0.946	Phthalate
9	302776-68-7	6.03	6.40	7.16	0.946	UV-Filter
10	000085-69-8	6.39	**6.68**	**6.02**	0.943	Phthalate
11	169198-72-5	5.68	6.96	8.51	0.942	UV-Filter
12	075673-16-4	6.78	6.06	5.99	0.940	Phthalate
13	025973-55-1*	5.65	7.06	7.75	0.939	UV-Filter
14	003864-99-1*	5.55	7.28	7.84	0.932	UV-Filter
15	041451-28-9	6.86	6.07	5.56	0.930	Phthalate
16	003147-75-9*	5.70	6.71	6.44	0.927	UV-Filter
17	070356-09-1	5.59	6.20	7.44	0.917	UV-Filter
18	000103-53-7	6.34	5.88	5.76	0.912	Fragrance
19	063250-25-9	5.65	6.20	6.19	0.910	UV-Filter
20	000146-50-9	6.33	5.68	6.27	0.909	Phthalate
21	000122-69-0	6.60	6.05	5.31	0.906	Fragrance
22	001843-05-6	6.37	6.07	5.34	0.904	UV-Filter
23	036437-37-3*	5.38	6.71	6.84	0.896	UV-Filter
24	000084-61-7	5.71	5.88	5.80	0.896	Phthalate
25	005320-75-2	5.96	5.70	5.97	0.895	Fragrance
26	000078-37-5	6.35	5.98	5.29	0.894	Fragrance
27	000103-41-3	6.02	5.71	5.69	0.893	Fragrance
28	000120-24-1	5.77	5.65	6.02	0.890	Fragrance
29	131812-52-7	5.64	5.88	5.53	0.883	Fragrance
30	000084-75-3	6.55	**6.27**	4.98	0.881	Phthalate
31	003846-71-7*	5.27	6.71	7.08	0.881	UV-Filter

Rank	CAS	*pLC50 P.promelas cons.*	*pEC50 D.magna*	*PEC50 P.subcapitata*	MCDM score	Use
32	000085-68-7	**5.60**	**5.61**	**6.42**	0.877	Phthalate
33	000122-68-9	6.60	6.04	4.96	0.872	Fragrance
34	005466-77-3	6.07	**6.00**	5.02	0.871	UV-Filter
35	003896-11-5*	5.15	6.80	7.45	0.868	UV-Filter
36	000575-61-1	5.41	5.74	6.28	0.867	UV-Filter
37	019224-26-1	5.64	5.43	6.80	0.862	Fragrance
38	010402-33-2	5.72	5.46	6.17	0.861	Fragrance
39	000103-64-0	5.19	6.24	6.68	0.861	Fragrance
40	036861-47-9	5.33	**5.66**	6.63	0.856	UV-Filter

In the priority list of Table 6 of the most toxic PCPs in the studied aquatic scenario, there are 17 UV filters, 11 flavor and fragrance ingredients, 10 phthalates and 2 antimicrobial agents. These results supports our observations in a previous study on PCPs,[19] where UV filters were predicted as the most environmentally hazardous subclass of PCPs, being, some of them, potential PBTs.

Below, we report some evidences, found in literature or in web, which can confirm our results on the potential toxicity of some PCPs in our priority list.

Triclocarban (101-20-2, ranked as the most toxic PCP ingredient in Table 6) is also classified as PT (persistent and toxic) by the ECHA Authority,[49] which, with regard to toxicity, claims: "*the toxicity values of fish, invertebrates and algae are LOEL = 0.01mg/L, EC50 = 0.0011 mg/L and LOEL = 0.01 mg/L, respectively. All these values fulfill the criteria for T classification (≤ 1 mg/l for acute and ≤ 0,01 mg/l for Chronic category)*", and it was predicted as a potential PBT by the PBT Index.[19,50] Also triclosan (3380-34-5), a well known toxic chemical,[5,51] appears as correctly ranked within the most toxic chemicals by our PCA-MCDM approach, and, similarly to triclocarban, was predicted as a potential PBT by our cumulative PBT Index model.[19,50] Di(2-ethylhexyl) phthalate (DEHP, CAS 117-81-7), ranked within the most hazardous and toxic chemicals by PCA and MCDM, is one of the most environmentally detected phthalates among the studied PCPs, even though its use in cosmetics and personal care products is being phased out in the European Union, and is also listed on Health Canada's Cosmetic Ingredient Hotlist that prohibits its use in cosmetics and PCPs.[51] DEHP has also effects on germ cell development and shows anti-androgenic effect.[52]Butylbenzyl phthalate (BBP, CAS 85-68-7), like DEHP, is already banned in US in children toys and its high acute toxicity in aquatic biota is a well-recognized concern; additionally it was included in the SVHC list by ECHA Authority and identified as a CMR substance.[53]

Apparently, the limited availability of homogeneous experimental data for all the three studied toxicity end points does not allow to validate the results of our ranking, but some among the prioritized PCPs were never included in the training sets of our QSAR models, therefore they can be considered a real validation of the obtained results. The four chemicals mentioned above, plus 2-Ethylhexyl butyl phthalate (85-69-8), Octinoxate (5466-77-3) and Enzacamene (36861-47-9), have experimental data that confirm their high toxicity (experimental pL50 > 5 mol/l) and are correctly listed in our priority list, after the combined approach of QSAR modelling and PCA-MCDM.

An additional validation of our screening and priority setting study can be done by verifying that the majority of the most toxic PCPs are UV-filters with a benzotriazole ring and phthalates. Both of these classes of chemicals have been demonstrated to reveal significant effects in acute and chronic toxicity tests.[54–58]

QSAR Model of Aquatic Toxicity Index (ATI)

The final step of our whole study was to model, by structural molecular descriptors, the trend of cumulative acute aquatic toxicity measured by the algae, *Daphnia* and fish toxicities and summarized in the aquatic toxicity index (ATI), i.e. the PC1 of the PCA score plot (reported in Figure 5), which represents the overall aquatic toxicity in the studied scenario.

The best combination of descriptors, obtained on the three independent training sets and selected from the three resulting population of models and the related statistical parameters are reported in the following Table (Table 7). The full model, redeveloped on 484 PCPs is also reported with equation and statistical parameters:

Table 7: *statistical parameters, related to the internal and external validation, of the proposed model for ATI of PCPs The equation of the full model is also reported.*

Model	N_{TR}	N_{PRED}	R^2	Q^2_{LOO}	Q^2_{LMO}	$R^2 Y_{scr}$	CCC_{EXT}	Q^2_{EXT} Fn	RMSEtr	RMSEp
Split by random	339	145	0.93	0.93	0.93	0.009	0.97	0.93-0.94	0.40	0.38
Split by ordered response	324	160	0.94	0.93	0.93	0.009	0.97	0.93-0.94	0.40	0.39
Split by structural similarity	324	160	0.94	0.94	0.94	0.009	0.96	0.91-0.92	0.37	0.43
ATI = - 14.00 + 0.34 **XlogP** + 17.97 **Mp** + 0.02 **TIC1**	**484**	**-**	**0.93**	**0.93**	**0.93**	**0.006**	**0.97cv**	**-**	**0.39**	**0.40cv**

The statistical parameters show that also QSAR models for the ATI of PCPs are internally robust and stable, not given by chance, and externally predictive on chemicals belonging to the three independent prediction sets. The prediction performances are very similar to the reproduction ability (very similar RMSE values).

LogP, here represented by the XlogP calculated in PaDEL-Descriptor ver. 2.21,[35] is the most important descriptor in predicting the toxicity of a chemical. It has, as expected, a positive sign in the equation, thus a positive influence in increasing the overall aquatic toxicity. The remaining two descriptors, Mp and TIC1, both with a positive sign in the model equation and thus also with a positive impact on the cumulative toxicity of PCPs, encode respectively for mean atomic polarizability and total information content index (neighborhood symmetry of 1-order).[47] These three descriptors are mainly related to the complexity and the dimension of the molecule, but also to the presence of polarizable atoms, giving higher values for compounds with hydrophobic chains and electronegative atoms (mainly oxygen, nitrogen and also chlorine in few chemicals)

In Figure 6, the plot of the PC1 scores (ATI values) vs. the QSAR predicted values and the Williams Plot for the AD of the ATI model (Ordered by response split) are reported.

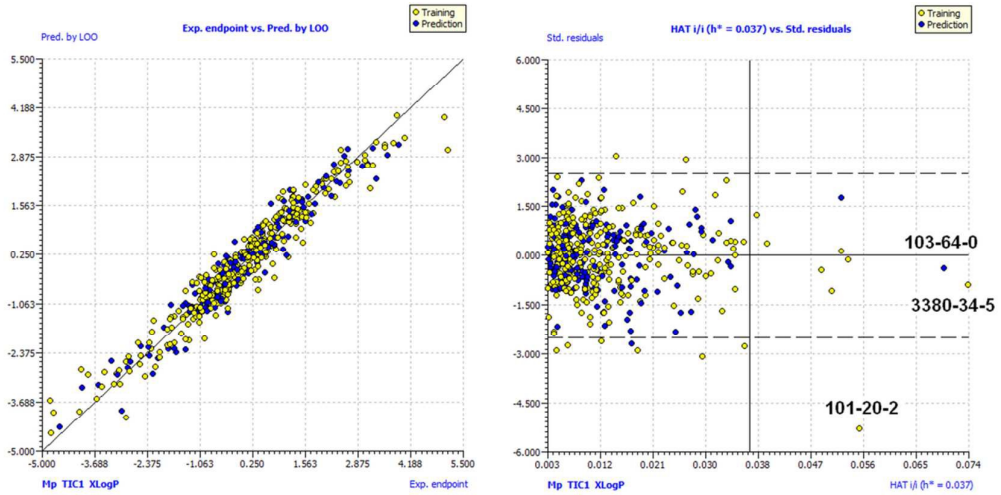


Fig. 6. *Graph of experimental vs. predicted values (on the left) and Williams Plot (on the right) for the ATI model (splitting by ordered response)*

The others graphs related to the ATI QSAR model are included in Electronic Supplementary Information material (Figures ES18-ES21).

As can be seen in the Williams Plot in Figure 6, chemical with CAS 101-20-2 (triclocarban) is a strong outlier, both for response and structure. Two others strong outliers for structure are highlighted, triclosan (CAS 3380-34-5) and bromostyrene (CAS 103-64-0). The complete list of structural and response outliers for this ATI model, is visible in Figures ES18-ES21 of the ESI, while the complete list of chemicals included in the ATI model, with the predictions and the AD belonging, are reported in Table ES5 of ESI.

Finally, once the model was internally and externally validated, the equation was applied to the 50 PCPs which were excluded from the previous PCA analysis, being outside the AD of the different developed models. Choosing an arbitrary cut-off, the chemicals predicted with an ATI greater than 2 (PC1 score of Figure 5), were considered as the most toxic for the aquatic scenario here studied. In this way, 26 additional molecules were selected for their potential hazard for their toxicity on the three key organisms of the aquatic compartment (see Table ES5).

Thus, the first priority list (i.e. the list containing the 40 most toxic compounds, predicted with more reliability, Table 6), could be incremented with the addition of these 26 PCPs, reaching a total of prioritized 66 PCPs ingredients, which are highlighted as the potentially most toxic within the studied aquatic scenario.

Looking at the chemical structures, the most hazardous PCPs are those with aromatic rings (one or more benzenes and benzotriazoles), long aliphatic chains, unsaturated esters, in general with complex structure including electronegative atoms.

It is important to note that the ATI model can be applied to any new compound, simply on its designed chemical structure, to verify before its synthesis if it could be inherently not hazardous for the aquatic environment. In fact, to predict the ATI of any new PCP, also not yet synthesized, it is sufficient to draw the hypothesized chemical structure, to derive the corresponding SMILES notation (/REF OBOILE), to compute the required three molecular descriptors and to apply the above QSAR model for ATI in QSARINS. Safer chemicals should have an ATI value as much as possible lower than 2.

It is also possible to verify if each chemical belongs or not to the structural applicability domain of the model, in order to have information on the higher reliability of the predictions for compounds interpolated into the model AD or the lower reliability for those that are extrapolated, being out of the AD. In conclusion, this model could be a useful tool for the early identification, and suggestion for the synthesis of potential safer alternatives to PCPs of highlighted concern.

Conclusions

This paper proposes new, externally validated, QSAR models specifically developed for the prediction of the acute aquatic toxicity of organic ingredients of personal care products (PCPs) in the green algae *Pseudokirchneriella subcapitata* in the crustacean *Daphnia magna* and in the fish *Pimephales promelas*, the three key organisms of the aquatic trophic level. These models, developed and validated , according to the OECD principles for QSAR acceptance in regulation, by the GA-OLS method, included in the software QSARINS, are characterized by large applicability domain, verified by different approaches, and were applied to prioritize the most toxic compounds among about 500 PCPs ingredients without experimental data. The predicted values obtained from these models are more similar to the available experimental values, if compared with those obtained by the commonly used software ECOSAR. The generated predictions allow a ranking of PCPs by PCA and MCDM according to their cumulative aquatic toxicity. An Aquatic Toxicity Index (ATI) is identified from this trend; also this index can be modelled by structural descriptors, allowing the prediction of cumulative aquatic toxicity. The application of QSAR models, PCA and MCDM approaches have yet demonstrated their utility in priority setting for characterization of potentially hazardous compounds. In particular, a total of 66 chemicals related to PCPs ingredients (mainly UV filters and phthalates) have been here selected for inclusion into a final priority list for further more definitive evaluation, focusing on them the necessary experimental tests. In this way, costs, time and sacrificed animals will be reduced. Most importantly, each QSAR model and particularly the ATI model, all applicable to a simple design of a hypothesized new compound, could allow to check *a priori* if this potential substitute could be a possible "safer alternative" of a recognized hazardous chemical, in a benign by design approach of Green Chemistry. The possibility to continuously contaminate the environment with "regrettable substitutions", which could be recognized only after they have been introduced to the market, and evidence of human health concerns have been manifested, could be highly reduced if *a priori* screenings and prioritization, by combining QSAR models and chemometric approaches, were more widely applied.[19,48,59-63]

Acknowledgements

The authors would like to thank Francois Marzetta for their collaboration. Financial support by the European Union through the project CADASTER (FP7-ENV-2007-212668) is gratefully acknowledged.

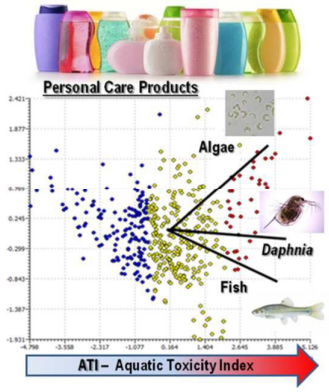
Electronic Supplementary Information

In the Electronic Supplementary Information (ESI), some tables and graphs are reported. In details, there are 5 Tables (Tables ES1-ES5) and 21 figures/graphs (Figure ES1-ES21).

Notes and References

1 M. V. Pablos, P. García-Hortigüela and C. Fernández, *Environ. Sci. Pollut. Res.*, 2015, **22**, 5417–5424.
2 A. M. Peck, *Anal. Bioanal. Chem.*, 2006, **386**, 907–939.
3 J.-W. Kim, H. Ishibashi, R. Yamauchi, N. Ichikawa, Y. Takao, M. Hirano, M. Koga and K. Arizono, *J. Toxicol. Sci.*, 2009, **34**, 227–232.
4 Q. Bu, B. Wang, J. Huang, S. Deng and G. Yu, *J. Hazard. Mater.*, 2013, **262**, 189–211.
5 S. Ortiz de García, G. P. Pinto, P. A. García-Encina and R. I. Mata, *J. Environ. Manage.*, 2013, **129**, 384–397.
6 A. Sánchez Rodríguez, M. Rodrigo Sanz and J. R. Betancort Rodríguez, *Chemosphere*, 2015, **131**, 85–90.

7 M. M. P. Tsui, H. W. Leung, B. K. Y. Kwan, K.-Y. Ng, N. Yamashita, S. Taniyasu, P. K. S. Lam and M. B. Murphy, *J. Hazard. Mater.*, 2015, **292**, 180–187.
8 J. M. Brausch and G. M. Rand, *Chemosphere*, 2011, **82**, 1518–1532.
9 S. Ramos, V. Homem, A. Alves and L. Santos, *Sci. Total Environ.*, 2015, **526**, 278–311.
10 M. Parolini, S. Magni, I. Traversi, S. Villa, A. Finizio and A. Binelli, *J. Hazard. Mater.*, 2015, **285**, 1–10.
11 C. G. Daughton and T. A. Ternes, *Environ. Health Perspect.*, 1999, **107**, 907–938.
12 K. Fent, A. Zenker and M. Rapp, *Environ. Pollut.*, 2010, **158**, 1817–1824.
13 H. Nakata, R.-I. Shinohara, Y. Nakazawa, T. Isobe, A. Sudaryanto, A. Subramanian, S. Tanabe, M. P. Zakaria, G. J. Zheng, P. K. S. Lam, E. Y. Kim, B.-Y. Min, S.-U. We, P. H. Viet, T. S. Tana, M. Prudente, D. Frank, G. Lauenstein and K. Kannan, *Mar. Pollut. Bull.*, 2012, **64**, 2211–2218.
14 M. Bachelot, Z. Li, D. Munaron, P. Le Gall, C. Casellas, H. Fenet and E. Gomez, *Sci. Total Environ.*, 2012, **420**, 273–279.
15 P. Gago-Ferrero, M. S. Díaz-Cruz and D. Barceló, *Anal. Bioanal. Chem.*, 2012, **404**, 2597–2610.
16 P. Gago-Ferrero, M. B. Alonso, C. P. Bertozzi, J. Marigo, L. Barbosa, M. Cremer, E. R. Secchi, A. Azevedo, J. Lailson-Brito Jr., J. P. M. Torres, O. Malm, E. Eljarrat, M. S. Díaz-Cruz and D. Barceló, *Environ. Sci. Technol.*, 2013, **47**, 5619–5625.
17 M. Picot Groz, M. J. Martinez Bueno, D. Rosain, H. Fenet, C. Casellas, C. Pereira, V. Maria, M. J. Bebianno and E. Gomez, *Sci. Total Environ.*, 2014, **493**, 162–169.
18 K. H. Langford, M. J. Reid, E. Fjeld, S. Øxnevad and K. V. Thomas, *Environ. Int.*, 2015, **80**, 1–7.
19 S. Cassani and P. Gramatica, *Sustain. Chem. Pharm.*, 2015, **1**, 19–27.
20 C. Schmitt, M. Oetken, O. Dittberner, M. Wagner and J. Oehlmann, *Environ. Pollut.*, 2008, **152**, 322–329.
21 A. Sieratowicz, D. Kaiser, M. Behr, M. Oetken and J. Oehlmann, *J. Environ. Sci. Health Part -ToxicHazardous Subst. Environ. Eng.*, 2011, **46**, 1311–1319.
22 X. Zhang, Q. Xu, S. Man, X. Zeng, Y. Yu, Y. Pang, G. Sheng and J. Fu, *Environ. Sci. Pollut. Res.*, 2012, **20**, 311–322.
23 R. E. Dodson, M. Nishioka, L. J. Standley, L. J. Perovich, J. G. Brody and R. A. Rudel, *Environ. Health Perspect.*, 2012, **120**, 935–943.
24 US EPA, *The ECOSAR (ECOlogical Structure Activity Relationship) Class Program*, 2012.
25 P. Gramatica, N. Chirico, E. Papa, S. Cassani and S. Kovarich, *J Comput Chem*, 2013, **34**, 2121–2132.
26 P. Gramatica, S. Cassani and N. Chirico, *J.Comput.Chem.*, 2014, **35**, 1036–1044.
27 US EPA, *ECOTOX User Guide: ECOTOXicology Database System. Version 4.0. Available: http:/www.epa.gov/ecotox/ (acessed 29-01-2016)* 2016.
28 A. Tropsha, *Mol. Inform.*, 2010, **29**, 476–488.
29 OECD, *Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test*, Organisation for Economic Co-operation and Development, Paris, 2011.
30 M. Cassotti, D. Ballabio, V. Consonni, A. Mauri, I. V. Tetko and R. Todeschini, *ATLA*, 2014, 31–41.
31 OECD, *Test No. 202: Daphnia sp. Acute Immobilisation Test*, Organisation for Economic Co-operation and Development, Paris, 2004.
32 OECD, *Test No. 203: Fish, Acute Toxicity Test*, Organisation for Economic Co-operation and Development, Paris, 1992.
33 Hypercube, inc, *HyperChem(TM)*, Hypercube, Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA, 2002.
34 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminformatics*, 2011, **3**, 33.
35 C. W. Yap, *JComput Chem*, 2011, **32**, 1466–1474.
36 OECD, *Principles for the validation, for regulatory purposes, of (Quantitative) Structure-Activity Relationship Models,* Organisation for Economic Co-operation and Development, Paris, 2004.
37 P. Gramatica, *Qsar Comb. Sci.*, 2007, **26**, 694–701.
38 P. Gramatica, *Mol. Inform.*, 2014, **33**, 311–314.
39 P. Gramatica, S. Cassani, P. P. Roy, S. Kovarich, C. W. Yap and E. Papa, *Mol Inf*, 2012, **31**, 817–835.
40 L. M. Shi, H. Fang, W. D. Tong, J. Wu, R. Perkins, R. M. Blair, W. S. Branham, S. L. Dial, C. I. Moland and D. M. Sheehan, *J. Chem. Inf. Comput. Sci.*, 2001, **41**, 186–195.
41 G. Schüurmann, R.-U. Ebert, J. Chen, B. Wang and R. Kuehne, *J. Chem. Inf. Model.*, 2008, **48**, 2140–2145.
42 V. Consonni, D. Ballabio and R. Todeschini, *J. Chem. Inf. Model.*, 2009, **49**, 1669–1678.
43 L. Lin, *Biometrics*, 1992, **48**, 599–604.
44 N. Chirico and P. Gramatica, *J. Chem. Inf. Model.*, 2011, **51**, 2320–2335.
45 N. Chirico and P. Gramatica, *J. Chem. Inf. Model.*, 2012, **52**, 2044–2058.
46 OECD, *Guidance document on the validation of (Quantitative) Structure-Activity Relationships [(Q)SAR] models*, Organisation for Economic Co-operation and Development, Paris, 2007.
47 R. Todeschini and V. Consonni, *Handbook of Molecular Descriptors*, John Wiley & Sons, 2008.
48 P. Gramatica, S. Cassani and A. Sangion, *Environ. Int.*, 2015, **77**, 25–34.
49 ECHA, *Triclocarban PBT Assessment*, European Chemicals Agency, Helsinki, 2014.
50 E. Papa and P. Gramatica, *Green Chem*, 2010, **12**, 836–843.
51 ECHA, *Triclosan PBT Assessment*, European Chemicals Agency, Helsinki, 2015.
52 D. Koniecki, R. Wang, R. P. Moody and J. Zhu, *Environ. Res.*, 2011, **111**, 329–336.
53 ECHA, *SVHC Support Document for Benzyl butyl phthalate*, European Chemicals Agency, Helsinki, 2008.
54 N. Scholz, *Chemosphere*, 2003, **53**, 921–926.
55 A. Seeland, M. Oetken, A. Kiss, E. Fries and J. Oehlmann, *Environ. Sci. Pollut. Res.*, 2012, **19**, 1781–1790.
56 X. Zhao, Y. Gao and M. Qi, *Ecotoxicology*, 2014, **23**, 626–632.
57 M. D. Alotaibi, A. J. McKinley, B. M. Patterson and A. Y. Reeder, *Water. Air. Soil Pollut.*, 2015, **226**, 226.
58 H. Xu, X. Dong, Z. Zhang, M. Yang, X. Wu, H. Liu, Q. Lao and C. Li, *Fish Shellfish Immunol.*, 2015, **45**, 286–292.
59 P. Gramatica, in *Recent Advances in QSAR Studies*, eds. T. Puzyn, J. Leszczynski and M. T. Cronin, Springer Netherlands, 2009, pp. 327–366.
60 P. Gramatica, in *Drug Design Strategies Quantitative Approaches*, eds. D. Livingstone and A. Davis, RSC Publishing, 2012.
61 J. B. Zimmerman and P. T. Anastas, *Science*, 2015, **347**, 1198–1199.
62 P. Fantke, R. Weber and M. Scheringer, *Sustain. Chem. Pharm.2015, **1**, 1-8.*
63 P. Gramatica, S. Cassani and A. Sangion, *J Hazard Mater,* 2016, **306**, 237-246



New externally validated QSAR models for aquatic toxicity of PCPs are proposed and applicable in QSARINS for *a priori* chemical design of environmentally safer PCPs.