

# QSARINS (QSAR-INSubria)

*Software for QSAR MLR model development and validation*

*From ideas of [Prof. Paola Gramatica](#), applied in “QSAR Research Unit in Environmental Chemistry and Ecotoxicology” at University of Insubria, Italy*

The image is a promotional graphic for the QSARINS software. It features a dark background with a sunset or sunrise scene. On the left, there is a circular logo of the University of Insubria, which contains a stylized green and white symbol. To the right of the logo, the word "QSARINS" is written in large, bold, cyan letters. Below the title, the text "QSAR INSubria, version 2.2.4 - 2019" is displayed in a smaller cyan font. Further down, the text "Software for QSAR MLR Model Development and Validation" and "Includes the QSARINS-Chem module" are written in cyan. The authors' names are listed in cyan: "By Gramatica P. Cassani S., Chirico N. (software developer), Kovarich S., Papa E., Sangion A." The affiliation "QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Theoretical and Applied Sciences (DiSTA), University of Insubria, Varese, Italy" is also in cyan. At the bottom, the website "http://www.qsar.it" and contact email "paola.gramatica@uninsubria.it" are provided in yellow and cyan respectively.

The approach to validated development and validation of QSAR MLR models in QSARINS is presented in the following papers:

- Gramatica, P., Principles of QSAR modeling. Comments and suggestions from personal experience, *International Journal of QSPR*, **2020**, 5(3) ,61-97 **Open Access** DOI: 10.4018/IJQSPR.20200701.0a1
- Gramatica, P., Chirico, N., Papa, E., Kovarich, S., Cassani, S. QSARINS: A New Software for the Development, Analysis, and Validation of QSAR MLR Models. *Journal of Computational Chemistry, Software news and updates*, **2013**, 34, 2121-2132, DOI: 10.1002/jcc.23361.

- Gramatica, P., Cassani, S., Chirico, N. QSARINS-Chem: Insubria Datasets and New QSAR/QSPR Models for Environmental Pollutants in QSARINS. *Journal of Computational Chemistry, Software news and updates*, **2014**, 35, 1036–1044. DOI: 10.1002/jcc.23576

The researchers interested in QSARINS (whether for QSAR models' development/validation or for application of QSARINS-Chem models or for use of the stored Insubria datasets or for analysis and validation of personal models) should read these papers to better know QSARINS.

These papers should also be kindly cited in papers where QSARINS is used.

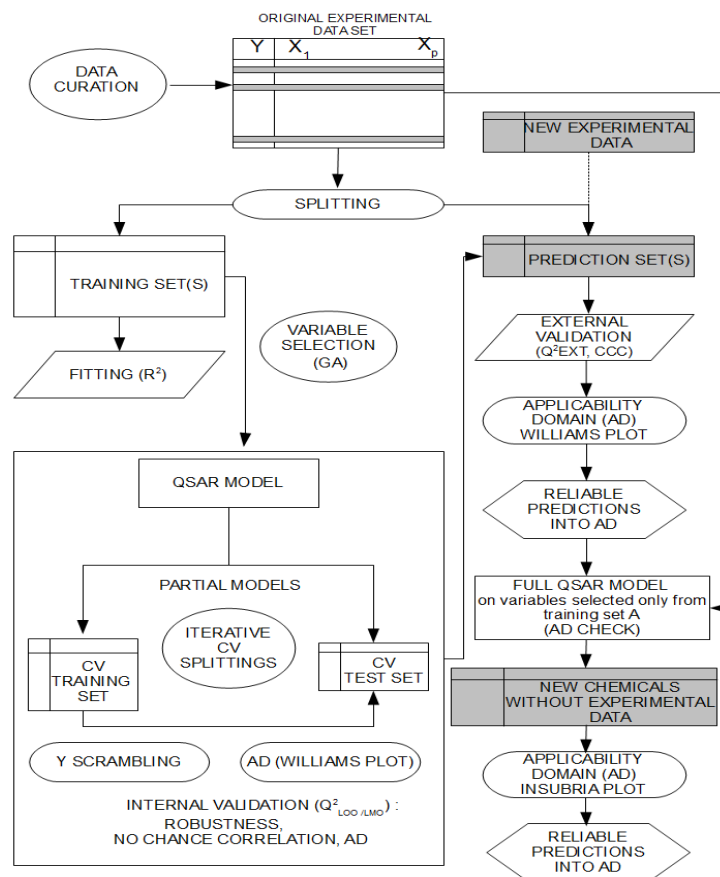
**QSARINS can be obtained by sending a request to Prof. Paola Gramatica ([paola.gramatica@uninsubria.it](mailto:paola.gramatica@uninsubria.it)), illustrating the personal experience in QSAR modeling.**

### **General information about QSARINS**

**QSARINS (QSAR-INSubria)** is a software for the development and validation of Multiple Linear Regression (MLR) models by Ordinary Least Squares (OLS) and Genetic Algorithm (GA) for variable selection, based on QSAR experience of Prof. Paola Gramatica since 1995 and developed by Nicola Chirico (2008-2012). It is implemented according to the statistical/chemometric approach for QSAR models' predictivity (Tropsha et al. 2003; Gramatica 2007, 2009, 2012, 2013, 2014, 2016, 2020; Chirico and Gramatica 2011, 2012; Gramatica et al. 2012; Gramatica and Sangion 2016), which is applied by the QSAR group of the University of Insubria in all the modeling researches.

In the module **QSARINS-Chem** the 3D- chemical structures of many Insubria datasets are stored, in addition to several QSAR models, based on PaDEL Descriptors (Yap 2011), developed by the Insubria QSAR group and applicable also to new chemicals, directly in QSARINS, where the PaDEL Descriptors can be calculated.

The following scheme (from Gramatica et al. 2012) summarizes the procedure for a validated QSAR modeling, applicable in QSARINS:



**QSARINS** provides a user-friendly interface that includes tools for:

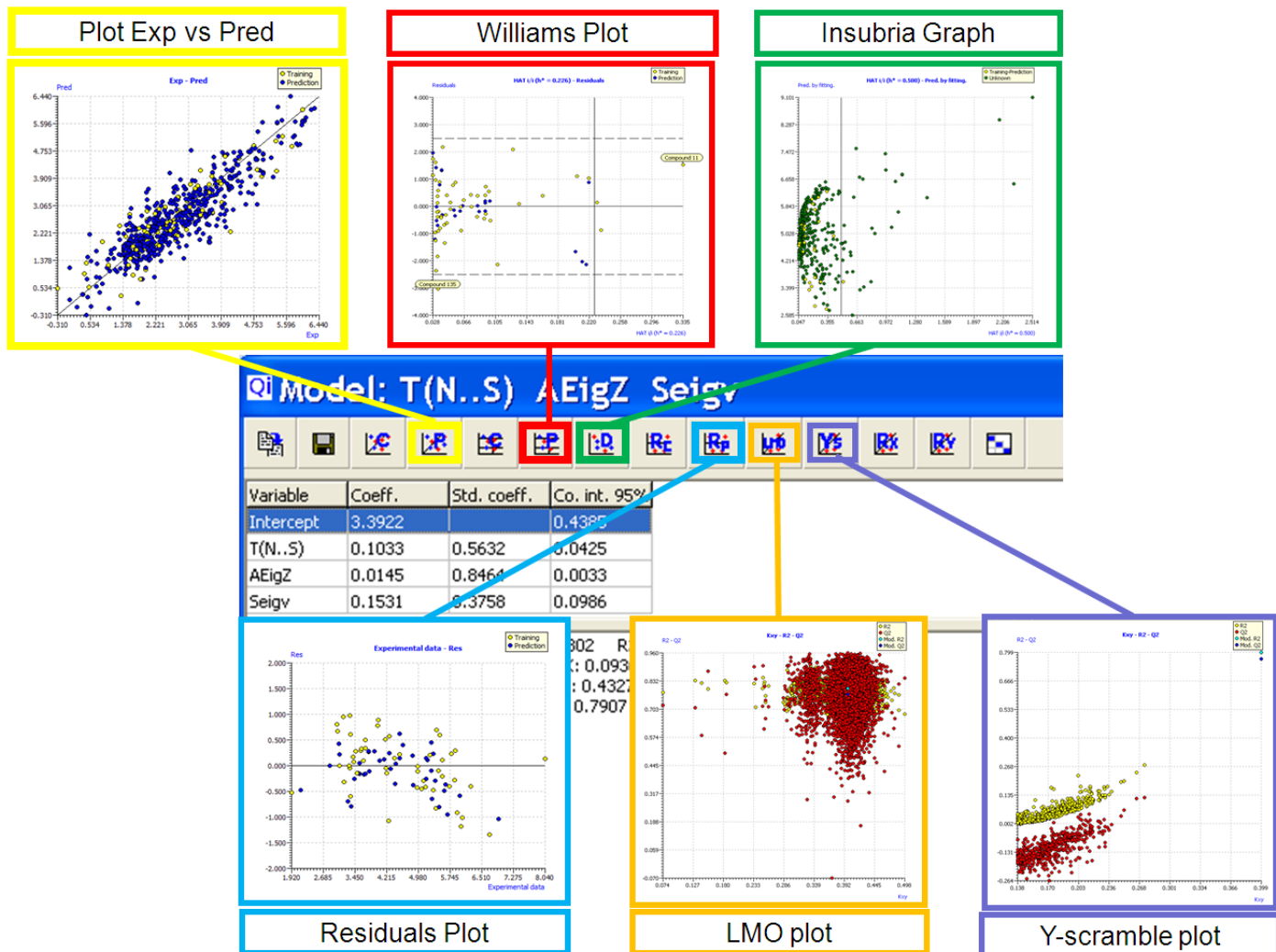
- Data normalization;
- Dataset analysis (by Principal Component Analysis (PCA), etc.);
- Splitting of the data sets (by Random, Ordered response or Structure by PCA of molecular descriptors);
- OLS Model development by All Subset Selection;
- Genetic Algorithm (GA) for Variable Selection (with tracing of the models while being developed);
- Analysis of all individual variables, as well as those more frequently selected in the models;
- Internal validation, including the following criteria:  $Q^2_{LOO}$ ,  $Q^2_{LMO}$ , Concordance Correlation Coefficient (CCC), RMSE, MAE, Y-Scrambling, etc...;
- External validation, including the following criteria:  $Q^2_{F1}$ ,  $Q^2_{F2}$ ,  $Q^2_{F3}$ , Golbraikh and Tropsha parameters,  $r^2_m$  metrics,  $CCC_{EXT}$ , RMSE, MAE;
- Structural applicability domain by the leverage from the diagonal values of the Hat matrix

(including Williams plot for chemicals with experimental data, y axis: standardized residuals and Insubria graph for chemicals without data, y axis: predicted values);

- Principal Components (PCs) Regression;
- Multi-Criteria Decision Making (MCDM) for selection of the best models;
- Consensus modeling (weighted and not, different tools for selection of models to be average);
- Check and validation of imported single models;
- Calculation of molecular descriptors and fingerprints with the open source software PaDEL-Descriptor (Yap 2011) (version currently included: 2.21).
- Application of several QSAR/QSPR models for environmental pollutants, developed by PaDEL-Descriptor software and stored in the QSARINS-Chem module, with the corresponding QMRF (QSAR Model Reporting Format). In particular: the PBT Index model (Papa and Gramatica 2010), re-developed (Gramatica et al. 2013, 2015) using PaDEL-Descriptor (Yap 2011);
- Database of compounds used to develop Insubria QSAR models: chemicals can be explored in different ways (CAS, SMILES, names) and visualized in 3D;
- Ranking of chemicals, based on PCA and MCDM

All outcomes of QSARINS are reported in the software with corresponding **plots** for quick graphical interpretation.

Example of some available **plots**:



## SCREENSHOTS

**DATA SETUP** - Selection of variables and response. Three different splitting (Random, by ordered response, by structure based on PCA of molecular descriptors). PCA of the selected descriptors and generation of PC scores as variables for Principal Components Regression.

**Data Setup**

Variables setup

Select var. Select resp. Clear all

Hold var. Release resp. Set splitting

Release var. Normalize var. View resp. d.

No.	Variable	Status
1	Log Koc Exp	Response
2	Splitting	Splitting
3	MW	Selected
4	TopoPSA	Selected
5	VAdjMat	Selected
6	nH	Selected
7	nC	Selected
8	nN	Selected
9	nO	Selected
10	nS	Selected
11	nP	Selected
12	nF	Selected
13	nCl	Selected
14	nBr	Selected
15	nI	Selected
16	nX	Selected
17	nAromBond	Selected
18	WPATH	Selected
19	WPOL	Selected
20	XLogP	Selected
21	Zagreb	Selected
22	MAXDN2	Selected

Objects setup

Training Prediction Create splitting

Excluded Reset objects Delete splitting

Filter var.

No.	Name	Status
1	000050-00-0	Prediction
2	000050-29-3	Training
3	000050-32-8	Prediction
4	000051-66-1	Prediction
5	000052-68-6	Prediction
6	000053-70-3	Prediction
7	000054-11-5	Training
8	000055-21-0	Prediction
9	000055-38-9	Training
10	000056-23-5	Prediction
11	000056-38-2	Prediction
12	000056-49-5	Prediction
13	000056-53-1	Prediction
14	000056-55-3	Prediction
15	000057-13-6	Prediction
16	000057-55-6	Prediction
17	000057-97-6	Prediction
18	000058-89-9	Prediction
19	000058-90-2	Prediction
20	000060-09-3	Prediction
21	000060-11-7	Prediction
22	000060-12-8	Prediction

Splitting setup

Name: Split - SOM

Random percentage: [ ]

Ordered by:  Response  Structure

Second mol.: Training

Training comp.: [ ]

Prediction comp.: [ ]

Swap

Var. PCA Des. from PCA

Correlation Single model

LOF smooth: 1.0

Selected variables:

- MW
- TopoPSA
- VAdjMat
- nH
- nC
- nN
- nO

Missing values:

Cancel OK

Training: 93 Prediction: 550 Variables: 661

Excluded: 0 Missing: 0 Unknown: 0

**SINGLE MODEL** – Model’s parameters related to internal and external validation, predicted values, HAT values, and standardized residuals are calculated.

In addition, it is possible to view the PCA and the correlation matrix of the modeling descriptors. If available, the QMRF can be exported. It is present for the large majority of the models available in QSARINS-Chem.

**IMPORTANT INFORMATION:** Any user can upload personal MLR models and use QSARINS to manage them for storing, visualization, validation, ranking etc..

Model: VP-0 nAromBond MAXDP nHBacc

Variable	Coeff.	Std. coeff.	Std. err.	(+/-) Co. int. 95%	p-value
Intercept	0.8728		0.0640	0.1256	0.0000
VP-0	0.2596	0.7481	0.0076	0.0150	0.0000
nAromBond	0.0755	0.3344	0.0045	0.0088	0.0000
MAXDP	-0.1854	-0.2777	0.0156	0.0307	0.0000

(Fitting criteria)  
R2: 0.7942      R2adj: 0.7929      R2-R2adj: 0.0013      LOF:  
Kxx: 0.3831      Delta K: 0.1057      RMSE tr: 0.5428      MAE tr: 0.4287  
RSS tr: 189.4573      CCC tr: 0.8853      s: 0.5449      F: 615.5260

(Internal validation criteria)  
Q2loo: 0.7907      R2-Q2loo: 0.0035      RMSE cv: 0.5474      MAE cv: 0.4323

Std. res. thresh. 2.5

ID	Name	Status	Exp. endpoint	Pred. by model eq.	Pred.Mod.Eq.Res.	Pred. LOO	Pred. LOO Res.	HAT i/j (h*=0.0233)	Std.Pred.Mod.Eq. Res.	Std.Pred.LOO Res.
426	002008-41-5	Training	2.9000	2.3284	-0.5716	2.3242	-0.5758	0.0074	-1.0527	-1.0606
427	002008-58-4	Training	0.5300	1.9465	1.4165	1.9513	1.4213	0.0034	2.6037	2.6125
428	002032-65-7	Training	2.3200	2.7576	0.4376	2.7587	0.4387	0.0025	0.8041	0.8061
429	002050-68-2	Training	4.3000	4.1187	-0.1813	4.1174	-0.1826	0.0070	-0.3340	-0.3363
430	002051-60-7	Training	3.5200	3.8475	0.3275	3.8497	0.3297	0.0068	0.6030	0.6072
431	002051-61-8	Training	4.4200	3.8359	-0.5841	3.8320	-0.5880	0.0067	-1.0755	-1.0827
432	002077-99-8	Training	3.6100	2.8200	-0.7900	2.8155	-0.7945	0.0057	-1.4538	-1.4621
433	002104-64-5	Training	3.1200	4.2119	1.0919	4.2187	1.0987	0.0062	2.0099	2.0224
434	002122-70-5	Training	2.4800	2.8219	0.3419	2.8235	0.3435	0.0047	0.6288	0.6318
435	002136-79-0	Training	3.5100	2.2798	-1.2302	2.2754	-1.2346	0.0036	-2.2616	-2.2697
436	002150-93-8	Training	2.3400	2.2223	-0.1177	2.2220	-0.1180	0.0023	-0.2163	-0.2168
437	002163-68-0	Training	2.9500	1.7934	-1.1566	1.7833	-1.1667	0.0086	-2.1317	-2.1502
438	002164-17-2	Training	2.0000	2.1054	0.1054	2.1057	0.1057	0.0028	0.1938	0.1943
439	002212-67-1	Training	1.9400	1.7733	-0.1667	1.7723	-0.1677	0.0058	-0.3068	-0.3086
440	002234-13-1	Training	5.8900	5.3179	-0.5721	5.3092	-0.5808	0.0150	-1.0578	-1.0739
441	002303-16-4	Training	3.2800	2.5263	-0.7537	2.5195	-0.7605	0.0089	-1.3893	-1.4017
442	002303-17-5	Training	3.3500	2.7775	-0.5725	2.7714	-0.5786	0.0106	-1.0562	-1.0675
443	002307-68-8	Training	2.7600	2.7642	0.0042	2.7642	0.0042	0.0041	0.0077	0.0077
444	002310-17-0	Training	2.9800	3.3075	0.3275	3.3098	0.3298	0.0069	0.6031	0.6073
445	002312-35-8	Training	3.6000	3.6252	0.0252	3.6253	0.0253	0.0075	0.0463	0.0467
446	002327-02-8	Training	2.5300	1.9204	-0.6096	1.9186	-0.6114	0.0030	-1.1203	-1.1237
447	002385-85-5	Training	6.0000	5.6873	-0.3127	5.6717	-0.3283	0.0474	-0.5880	-0.6172
448	002425-10-7	Training	1.7100	2.2361	0.5261	2.2374	0.5274	0.0024	0.9666	0.9689



## QSARINS-Chem Module in QSARINS:

QSARINS-Chem is the module where 3012 chemicals, collected from the literature, curated and modeled by the Insubria group, are available with their 3D structure and experimental responses. In addition, 45 QSAR models of environmental end-points for organic pollutants, based on free software for molecular descriptors (Yap 2011), are available. These models, supported by their QMRF, can be applied for any new chemical without experimental data or also not yet synthesized, verifying the structural applicability domain by the Insubria graph.

### List of QSAR models in QSARINS-Chem Module:

1. Brominated Flame Retardants (BFR) Log Koa (Papa et al. 2009; Gramatica et al. 2014)<sup>§</sup>
2. BFR MP (Papa et al. 2009; Gramatica et al. 2014)<sup>§</sup>
3. BFR VP (Papa et al. 2009; Gramatica et al. 2014)<sup>§</sup>
4. Benzo-Triazole (BTAZ) *D. magna* tox (Cassani et al. 2013; Gramatica et al. 2014)<sup>§</sup>
5. BTAZ Log Kow (Bhatarai and Gramatica 2011a; Gramatica et al. 2014)<sup>§</sup>
6. BTAZ MP (Bhatarai and Gramatica 2011a; Gramatica et al. 2014)<sup>§</sup>
7. BTAZ *O. mykiss* tox (Cassani et al. 2013; Gramatica et al. 2014)<sup>§</sup>
8. BTAZ *P. subcapitata* tox (Cassani et al. 2013; Gramatica et al. 2014)<sup>§</sup>
9. BTAZ Sw (Bhatarai and Gramatica 2011a; Gramatica et al. 2014)<sup>§</sup>
10. BTAZ VP (Bhatarai and Gramatica 2011a; Gramatica et al. 2014)<sup>§</sup>
11. Endocrine Disruptor Chemicals (EDC) Estrogen Receptor Binding (Li and Gramatica 2010; Gramatica et al. 2014)<sup>§</sup>
12. Esters *D. magna* tox (Papa et al. 2005a; Gramatica et al. 2014)<sup>§</sup>
13. Esters *P. promelas* tox (Papa et al. 2005a; Gramatica et al. 2014)<sup>§</sup>
14. Fish Biotrans. logHLn M1\_day (Papa et al. 2014)\*
15. Fish Biotrans. logHLn M2\_day (Papa et al. 2014)\*
16. Fish Biotrans. logHLn M3\_day (Papa et al. 2014)\*
17. Global Half-Life Index (GHLI) (Gramatica and Papa 2007; Gramatica et al. 2014)<sup>§\*</sup>
18. Human Biotrans. logHLB1\_h (Papa et al. 2018)\*
19. Human Biotrans. logHLB2\_h (Papa et al. 2018)\*
20. Human Biotrans. logHLB3\_h (Papa et al. 2018)\*
21. Human Biotrans. logHLB4\_h (Papa et al. 2018)\*
22. Human Biotrans. logHLT\_h (Papa et al. 2018)\*
23. Log Koc of Pesticides (Gramatica et al. 2007a, 2014)<sup>§\*</sup>
24. Nitro PAH Mutagenicity (Gramatica et al. 2007b, 2014)<sup>§</sup>
25. *Pimephales promelas* tox (Papa et al. 2005b; Gramatica et al. 2014)<sup>§\*</sup>
26. Persistence Bioaccumulation Toxicity (PBT) Index (Papa and Gramatica 2010; Gramatica et al. 2013)<sup>§\*</sup>
27. Personal Care Products (PCP) Aquatic Toxicity Index-ATI (Gramatica et al. 2016)\*
28. PCP *D.magna* acute tox (Gramatica et al. 2016)\*
29. PCP *P.promelas* acute tox LogP-based (Gramatica et al. 2016)\*
30. PCP *P.promelas* acute tox (Gramatica et al. 2016)\*



31. PCP *P.subcapitata* acute tox (Gramatica et al. 2016)\*
32. PerFluorinated Chemicals (PFC) Mouse Inhalation tox (Bhatarai and Gramatica 2010; Gramatica et al. 2014)<sup>§</sup>
33. PFC Rat Inhalation tox (Bhatarai and Gramatica 2010; Gramatica et al. 2014)<sup>§</sup>
34. PFC Rat Oral tox (Bhatarai and Gramatica 2011b; Gramatica et al. 2014)<sup>§</sup>
35. PFC Sw (Bhatarai and Gramatica 2011c; Gramatica et al. 2014)<sup>§</sup>
36. PFC VP (Bhatarai and Gramatica 2011c; Gramatica et al. 2014)<sup>§</sup>
37. Pharm. Aquatic Toxicity Index-ATI (Sangion and Gramatica 2016a)\*
38. Pharm. *D.magna* acute tox (Sangion and Gramatica 2016a)\*
39. Pharm. *O.mykiss* acute tox (Sangion and Gramatica 2016a)\*
40. Pharm. *P.promelas* acute tox (Sangion and Gramatica 2016a)\*
41. Pharm. *P.subcapitata* acute tox (Sangion and Gramatica 2016a)\*
42. PPCP intertox *D.magna-O.mykiss* (Sangion and Gramatica 2016b)
43. PPCP intertox *D.magna-P.promelas* (Sangion and Gramatica 2016b)
44. PPCP intertox *O.mykiss-P.promelas* (Sangion and Gramatica 2016b)
45. PPCP intertox *P.promelas-O.mykiss* (Sangion and Gramatica 2016b)

<sup>§</sup>Models originally developed using proprietary software (DRAGON v 5.5 or former versions) and updated using PaDEL Descriptor software (models are described in Gramatica et al. 2014).

\*Models reported also in QSARINS-Chem Standalone version (free downloadable from [www.qsar.it](http://www.qsar.it))

### **Additional information**

QSARINS can be used for every modeling work involving Multiple Linear Regression (MLR) calculations, based on Genetic Algorithm for variable selection and Ordinary Least Squares (OLS) as modeling method. However, any personal MLR model, even if developed by other software, can be analyzed by the plots available in QSARINS as well as validated by several statistical validation parameters (both for internal and external validation: see above list).

Other chemometric tools (Principal Component Analysis (PCA), Multicriteria Decision Making (MCDM)) for explorative analysis and ranking are also implemented, **therefore QSARINS is not limited to QSAR studies.**

It is also important to note that not only chemicals, but any kind of objects can be analyzed or modeled in a multivariate way.

## **Bibliography**

- Bhhatarai B, Gramatica P (2011a) Modelling physico-chemical properties of (benzo)triazoles, and screening for environmental partitioning. *Water Res* 45:1463–1471. doi: 10.1016/j.watres.2010.11.006
- Bhhatarai B, Gramatica P (2010) Per- and Polyfluoro Toxicity (LC50 Inhalation) Study in Rat and Mouse Using QSAR Modeling. *Chem Res Toxicol* 23:528–539. doi: 10.1021/tx900252h
- Bhhatarai B, Gramatica P (2011b) Oral LD50 toxicity modeling and prediction of per- and polyfluorinated chemicals on rat and mouse. *Mol Divers* 15:467–476. doi: 10.1007/s11030-010-9268-z
- Bhhatarai B, Gramatica P (2011c) Prediction of Aqueous Solubility, Vapor Pressure and Critical Micelle Concentration for Aquatic Partitioning of Perfluorinated Chemicals. *Environ Sci Technol* 45:8120–8128. doi: 10.1021/es101181g
- Cassani S, Kovarich S, Papa E, et al (2013) Daphnia and fish toxicity of (benzo)triazoles: Validated QSAR models, and interspecies quantitative activity–activity modelling. *J Hazard Mater* 258–259:50–60. doi: 10.1016/j.jhazmat.2013.04.025
- Chirico N, Gramatica P (2011) Real External Predictivity of QSAR Models: How To Evaluate It? Comparison of Different Validation Criteria and Proposal of Using the Concordance Correlation Coefficient. *J Chem Inf Model* 51:2320–2335. doi: 10.1021/ci200211n
- Chirico N, Gramatica P (2012) Real External Predictivity of QSAR Models. Part 2. New Intercomparable Thresholds for Different Validation Criteria and the Need for Scatter Plot Inspection. *J Chem Inf Model* 52:2044–2058. doi: 10.1021/ci300084j
- Gramatica P (2007) Principles of QSAR models validation: internal and external. *Qsar Comb Sci* 26:694–701. doi: 10.1002/qsar.200610151
- Gramatica P (2009) Chemometric Methods and Theoretical Molecular Descriptors in Predictive QSAR Modeling of the Environmental Behaviour of Organic Pollutants, Chapter 12 in *Recent Advances in QSAR Studies*. In: Puzyn T, Leszczynski J, Cronin MT (eds) *Recent Advances in QSAR Studies*. Springer Netherlands, pp 327–366
- Gramatica P (2012) Modelling Chemicals in the Environment. In: Livingstone D, Davis A (eds) *Drug Design Strategies Quantitative Approaches*. RSC Publishing
- Gramatica P (2013) On the Development and Validation of QSAR Models. In: Reisfeld B, Mayeno AN (eds) *Computational Toxicology*. Humana Press, pp 499–526
- Gramatica P (2014) External Evaluation of QSAR Models, in Addition to Cross-Validation: Verification of Predictive Capability on Totally New Chemicals. *Mol Inform* 33:311–314
- Gramatica P (2016) Prioritization of Chemicals Based on Chemoinformatic Analysis. In: Leszczynski J (ed) *Handbook of Computational Chemistry*. Springer Netherlands, pp 1–33.

- Gramatica P, (2020) Principles of QSAR modeling. Comments and suggestions from personal experience, *International Journal of QSPR*, 5(3) ,61-97 Open Access  
DOI:10.4018/IJQSPR.20200701.0a1
- Gramatica P, Cassani S, Chirico N (2014) QSARINS-Chem: Insubria Datasets and New QSAR/QSPR Models for Environmental Pollutants in QSARINS. *J Comput Chem* 35:1036–1044. doi: 10.1002/jcc.23576
- Gramatica P, Cassani S, Roy PP, et al (2012) QSAR modeling is not “push a button and find a correlation”: a case study of toxicity of (benzo-)triazoles on algae. *Mol Inf* 31:817–835. doi: 10.1002/minf.201200075
- Gramatica P, Cassani S, Sangion A (2015) PBT assessment and prioritization by PBT Index and consensus modeling: Comparison of screening results from structural models. *Environ Int* 77:25–34. doi: 10.1016/j.envint.2014.12.012
- Gramatica P, Cassani S, Sangion A (2016) Aquatic ecotoxicity of personal care products: QSAR models and ranking for prioritization and safer alternatives’ design. *Green Chem* 18:4393 – 4406. doi: 10.1039/C5GC02818C
- Gramatica P, Chirico N, Papa E, et al (2013) QSARINS: A new software for the development, analysis and validation of QSAR MLR models. *J Comput Chem* 34:2121–2132. doi: 10.1002/jcc.23361
- Gramatica P, Giani E, Papa E (2007a) Statistical external validation and consensus modeling: A QSPR case study for K-oc prediction. *J Mol Graph Model* 25:755–766. doi: 10.1016/j.jmgm.2006.06.005
- Gramatica P, Papa E (2007) Screening and ranking of POPs for global half-life: QSAR approaches for prioritization based on molecular structure. *Environ Sci Technol* 41:2833–2839. doi: 10.1021/es061773b
- Gramatica P, Pilutti P, Papa E (2007b) Approaches for externally validated QSAR modelling of Nitratated Polycyclic Aromatic Hydrocarbon mutagenicity. *SAR QSAR Environ Res* 18:169–178. doi: 10.1080/10629360601054388
- Gramatica P, Sangion A (2016) A Historical Excursus on the Statistical Validation Parameters for QSAR Models: A Clarification Concerning Metrics and Terminology. *J Chem Inf Model* 56:1127–1131. doi: 10.1021/acs.jcim.6b00088
- Li J, Gramatica P (2010) The importance of molecular structures, endpoints’ values, and predictivity parameters in QSAR research: QSAR analysis of a series of estrogen receptor binders. *Mol Divers* 14:687–696. doi: 10.1007/s11030-009-9212-2
- Papa E, Battaini F, Gramatica P (2005a) Ranking of aquatic toxicity of esters modelled by QSAR. *Chemosphere* 58:559–570. doi: 10.1016/j.chemosphere.2004.08.003
- Papa E, Gramatica P (2010) QSPR as a support for the EU REACH regulation and rational design of environmentally safer chemicals: PBT identification from molecular structure. *Green Chem* 12:836–843. doi: 10.1039/B923843C

- Papa E, Kovarich S, Gramatica P (2009) Development, Validation and Inspection of the Applicability Domain of QSPR Models for Physicochemical Properties of Polybrominated Diphenyl Ethers. *QSAR Comb Sci* 28:790–796. doi: 10.1002/qsar.200860183
- Papa E, Sangion A, Arnot JA, Gramatica P (2018) Development of human biotransformation QSARs and application for PBT assessment refinement. *Food Chem Toxicol* 112:535–543. doi: screening level
- Papa E, van der Wal L, Arnot JA, Gramatica P (2014) Metabolic biotransformation half-lives in fish: QSAR modeling and consensus analysis. *Sci Total Env* 470–471:1040–1046
- Papa E, Villa F, Gramatica P (2005b) Statistically validated QSARs, based on theoretical descriptors, for modeling aquatic toxicity of organic chemicals in *Pimephales promelas* (fathead minnow). *J Chem Inf Model* 45:1256–1266. doi: 10.1021/ci050212l
- Sangion A, Gramatica P (2016a) Ecotoxicity interspecies QAAR models from *Daphnia* toxicity of pharmaceuticals and personal care products. *SAR QSAR Environ Res* 0:1–18. doi: 10.1080/1062936X.2016.1233139
- Sangion A, Gramatica P (2016b) Hazard of pharmaceuticals for aquatic environment: Prioritization by structural approaches and prediction of ecotoxicity. *Environ Int* 95:131–143. doi: 10.1016/j.envint.2016.08.008
- Tropsha A, Gramatica P, Gombar VK (2003) The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *Qsar Comb Sci* 22:69–77. doi: 10.1002/qsar.200390007
- Yap CW (2011) PaDEL-Descriptor: An open source software to calculate molecular descriptors and fingerprints. *JComput Chem* 32:1466–1474. doi: 10.1002/jcc.21707

**QSARINS is currently used by the Insubria group and several international QSAR groups: more than 800 free licenses active into 2021.**