

XXXII Symposium on Bioinformatics and Computer-Aided Drug Discovery
(BCADD-2026)

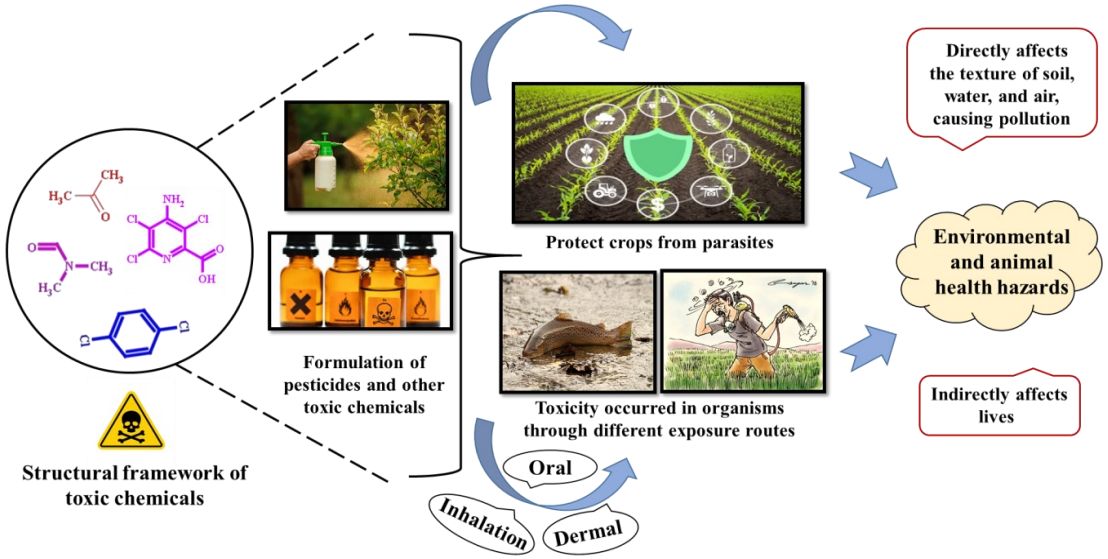
CHEMOMETRIC MODELLING FOR TOXICITY ASSESSMENT OF DIVERSE CHEMICALS IN STONEFLY: A STEP TOWARDS ENHANCED RISK EVALUATION

***D. Mahapatra*, P.K. Ojha**

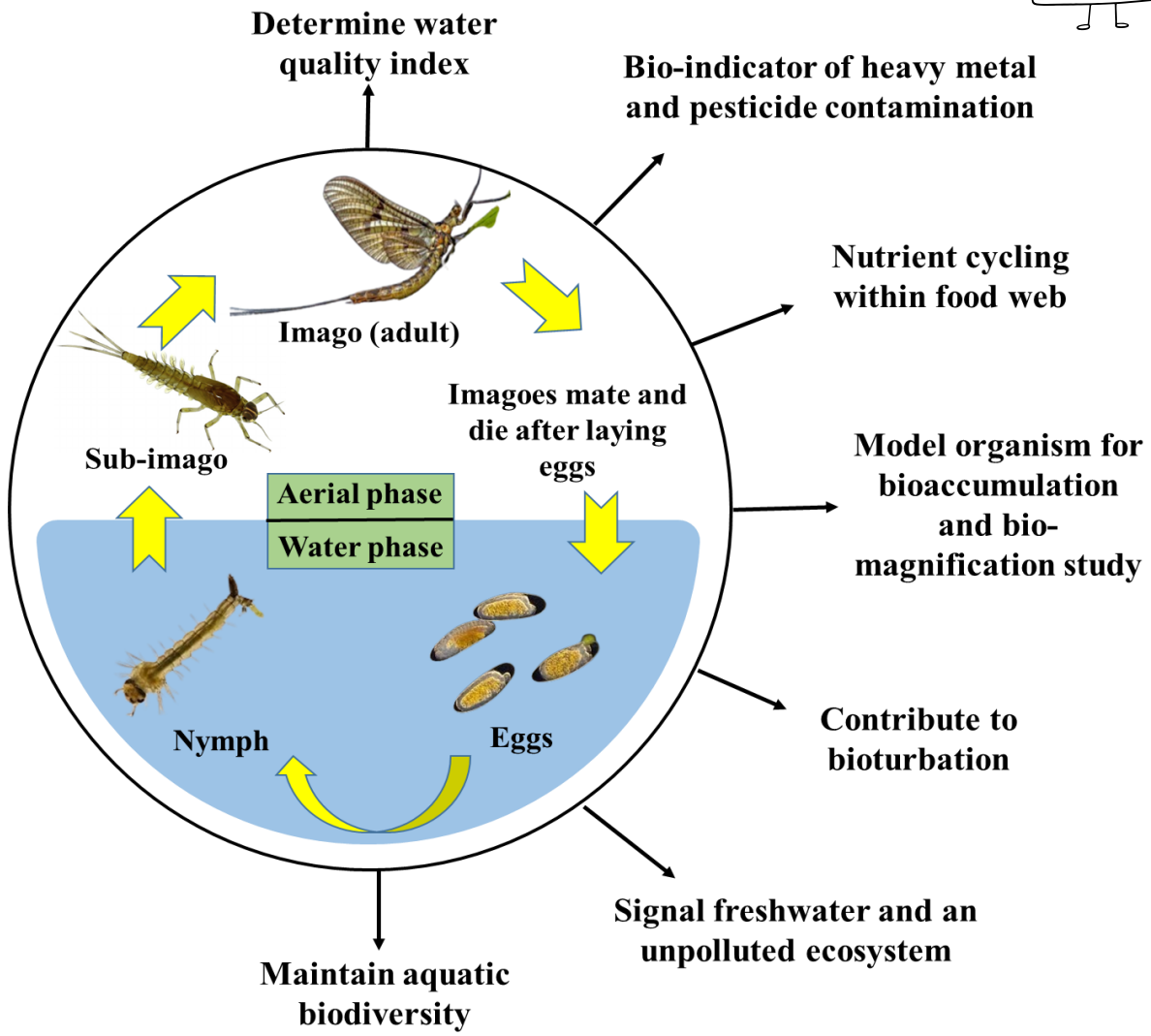
Drug Discovery and Development Laboratory, Department of Pharmaceutical Technology, Jadavp ur University, Kolkata 700032, India

1. Introduction

- **Agrochemicals** play a vital role in agriculture, but their widespread use now threatens animal habitats and the environment. Significant accumulation of these harmful organic chemicals damages **aquatic** organisms through soil runoff, acid rain, and leaching, and raises concerns about their hazardous effects on daily life.
- Here, we estimate the acute toxicity ($\mathbf{LC_{50}}$) of diverse chemicals in the **Stonefly (*Pteronarcys californica)*** species to evaluate the ecotoxicity associated with those toxicants by developing a chemometric **model (QSAR)**.

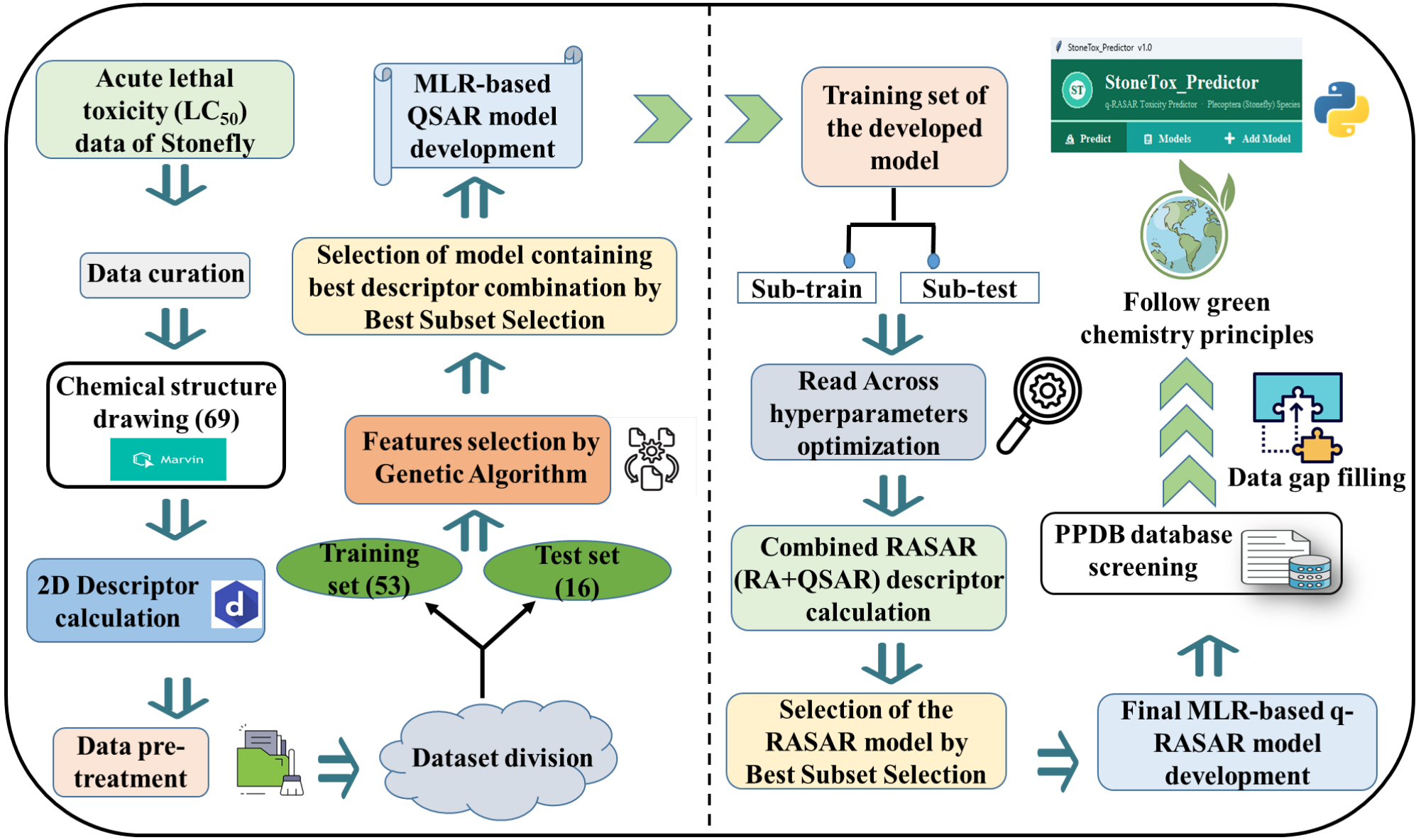


Toxicity of agrochemicals



Ecological importance of Stonefly : **highly pollution-sensitive** macro-invertebrate representing the **highest category** of clean-water indicators (**class 1 biotic index value**)

2. Methods and materials



A schematic workflow of the present study

3. Results and discussion

MLR-based QSAR model equation:

pLC_{50} = 6.779 -0.648* B08[O-Cl] -1.365* nRCOOH -11.263* X3A +0.462*MaxdsCH +0.336* MaxssO +18.178* Eta_D_AlphaA

MLR-based q-RASAR model equation:

pLC_{50} = 5.999 -7.975* X3A +1.303* MaxdsCH +0.209* MaxssO +16.350* Eta_D_AlphaA -1.462* CVsim(GK) +1.113* Pos.Avg.Sim

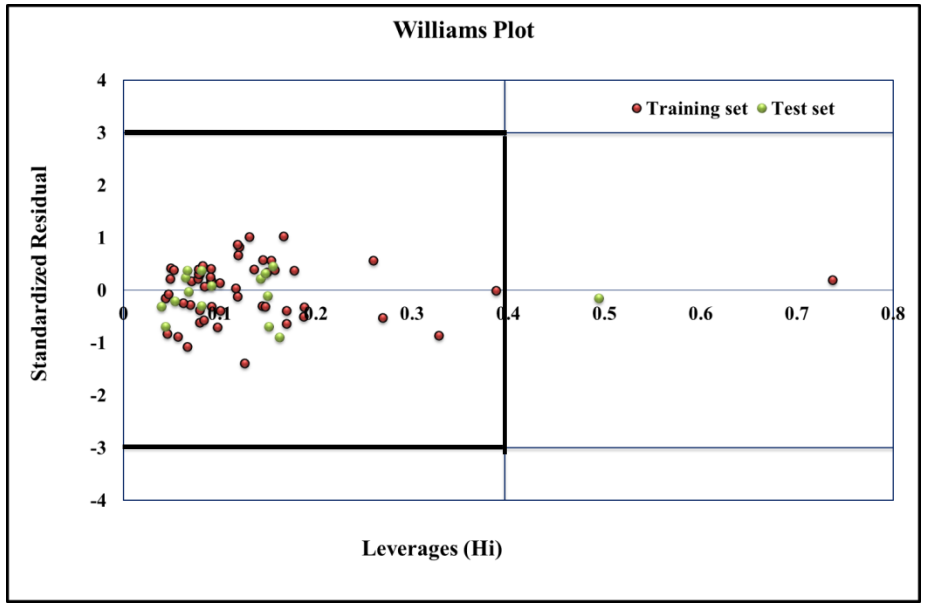
Results of Golbraikh and Tropsha's criteria

Sl. No.	Parameters	q-RASAR model	Remarks	Threshold value
1.	Q^2_{LOO} (train)	0.704	Passed	$Q^2_{LOO}>0.5$
2.	R^2(test)	0.741	Passed	R^2(test)>0.6
3.	$[(r^2-r_0^2)/ r^2]$	0.002	Passed	$[(r^2-r_0^2)/ r^2]<0.1$
4.	k	0.981	Passed	$0.85 < k < 1.15$
5.	$\|r_0^2-r'^2_0\|$	0.119	Passed	$\|r_0^2-r'^2_0\|<0.3$

Statistical validation metrics of the MLR-based QSAR and q-RASAR models

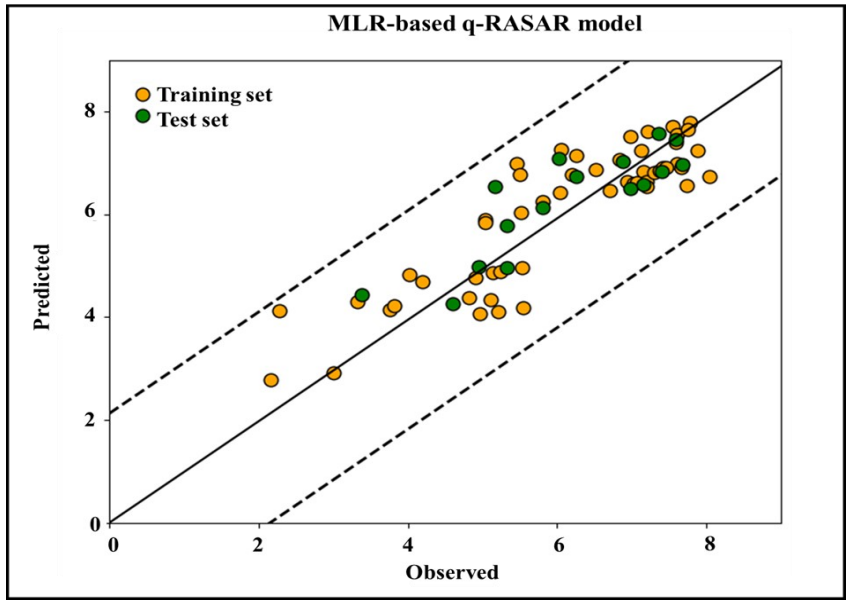
Validation metrics	QSAR model	q-RASAR model
N_{train}/N_{test}	53/16	**53/16**
No. of descriptors	6	**6**
R^2	0.754	**0.775**
$Q^2_{(LOO)}$	0.684	**0.705**
Q^2_{F1}	0.626	**0.732**
Q^2_{F2}	0.625	**0.731**
$RMSE_c$	0.744	**0.630**
$RMSE_p$	0.744	**0.630**
cRp^2	0.702	**0.716**
MAE	0.563	**0.465**

* The q-RASAR model has improved quality, predictivity and lower error than the corresponding QSAR model

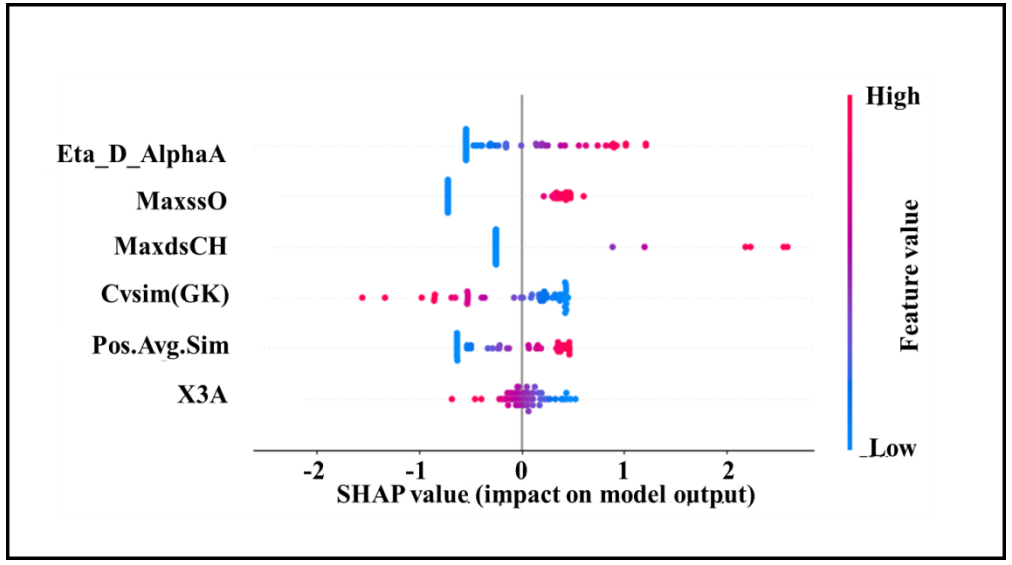


1. Williams plot of the developed model

Graphical representations

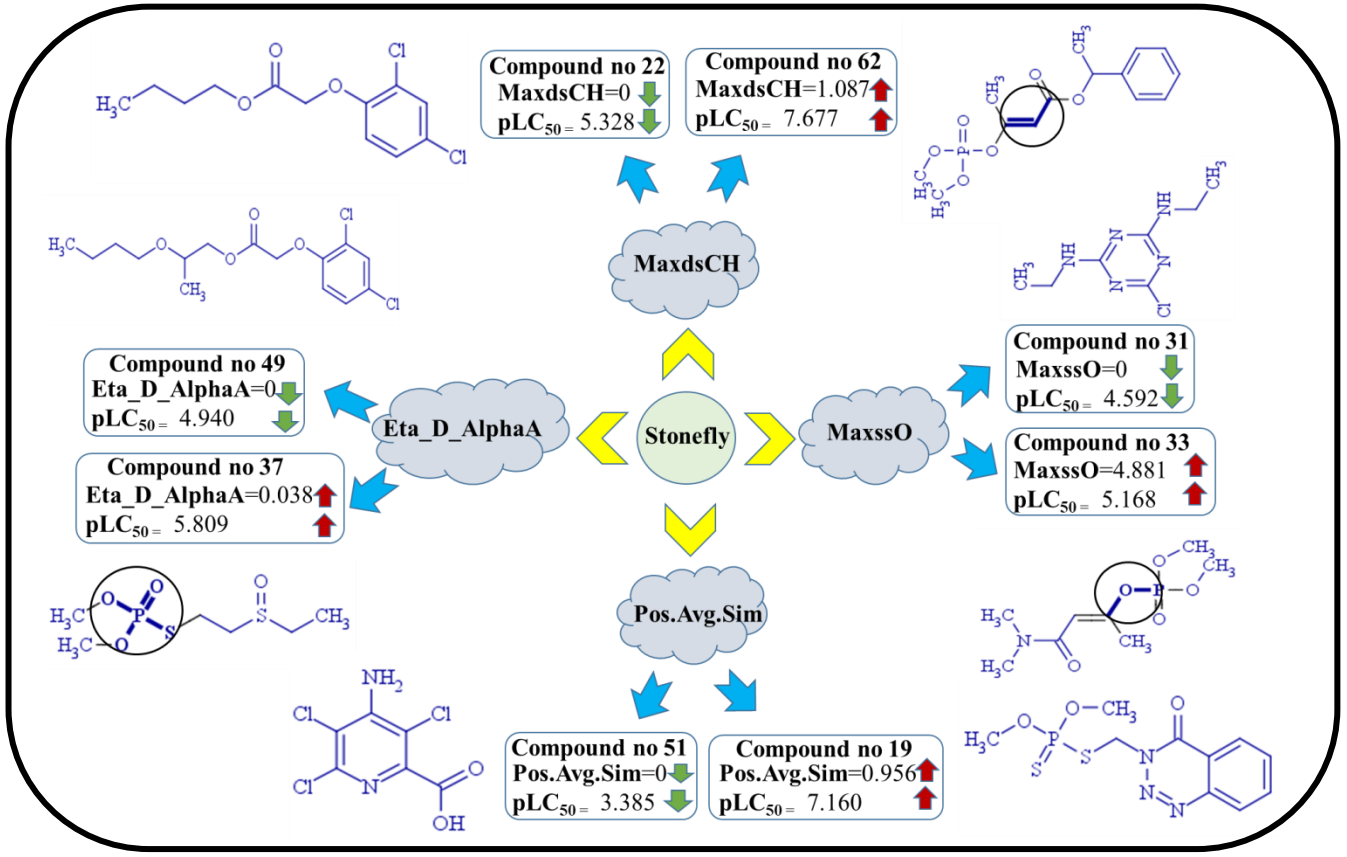


2. Scatter plot of the generated model



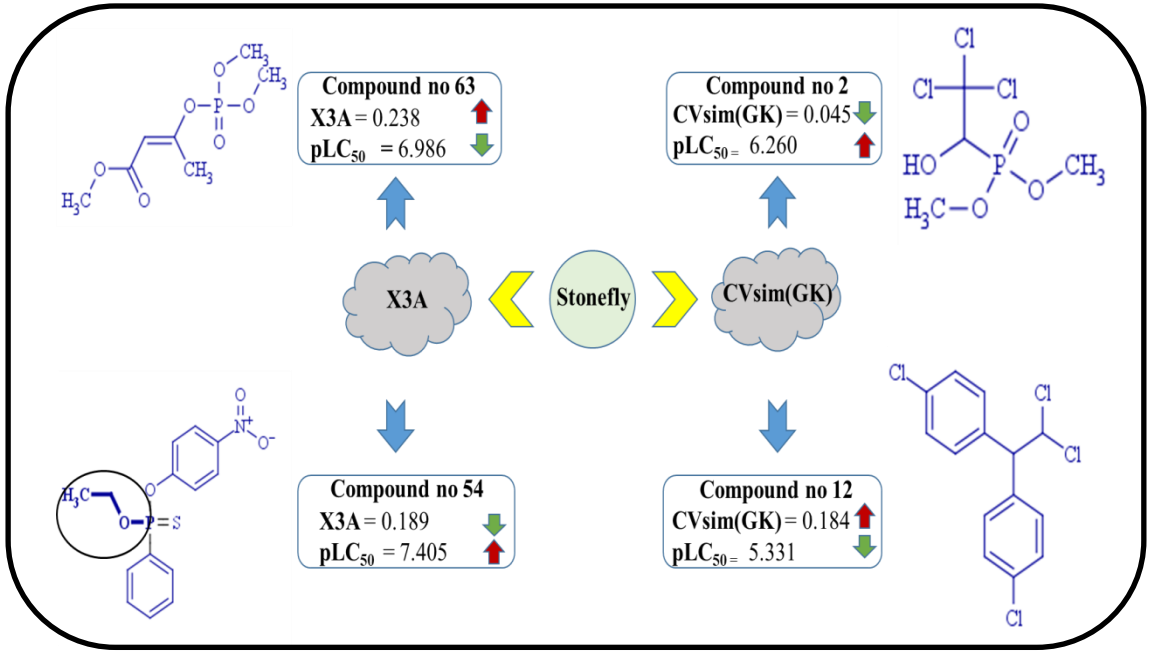
3. SHAP plot of the developed model

Mechanistic introspection



Positively contributing modelled descriptors

Negatively contributing modelled descriptors



Screening of an external set (PPDB)

Top 3 toxic screened pesticides from Pesticide Properties DataBase (PPDB).

Sl. No.	Chemicals along with their pLC_{50} value	Literature survey for validation	Real-world data source
Top 3 toxic screened compounds			
1.	Iodobonil (9.014)	Cancerous	https://comptox.epa.gov/dashboard/chemical/literature-pprt
2.	Deltamethrin (8.107)	It is moderately toxic	https://npic.orst.edu/factsheets/DeltaGen.html
3.	Tralomethrin (8.066)	Tralomethrin exposure caused cardiovascular dysplasia and dysfunction	https://pubmed.ncbi.nlm.nih.gov/37471628/

Least 3 toxic screened pesticides from Pesticide Properties DataBase (PPDB).

Sl. No.	Chemicals along with their pLC_{50} value	Literature survey for validation	Real-world data source
Least 3 toxic screened compounds			
1.	Malic acid (2.443)	Relatively nontoxic in acute toxicity	https://pubmed.ncbi.nlm.nih.gov/11358110/#:~:text=Malic%20Acid%20was%20relatively%20n
2.	Hexadecanoic acid (2.679)	Not a hazardous substance	https://www.acs.org/molecule-of-the-week/archive/p/palmitic-acid
3.	Glufosinate (2.851)	Slightly toxic	https://www.mda.state.mn.us/glufosinateherbicide#:~:text=Glufosinate%20is%20slightly%20toxic

‘StoneTox_Predictor-v1.0’

StoneTox_Predictor v1.0

StoneTox_Predictor

q-RASAR Toxicity Predictor · Plecoptera (Stonefly) Species

Predict | Models | + Add Model

UPLOAD TEST DATA

Input File (CSV or XLSX)

No file selected

Browse File

OUTPUT FORMAT

- (•) Excel (.xlsx) — colour coded
- ( ) CSV (.csv)

CLASSIFICATION STANDARD

- ( ) UN GHS Only
- ( ) US EPA Only
- (•) Both (side by side)

TOXICITY SCALE

UN GHS (↑ pLC50 = ↑ Toxicity)

- Cat 1 — pLC50 ≥ 4.0 (LC50 ≤1 mg/L)
- Cat 2 — pLC50 3.0–4.0 (LC50 1–10 mg/L)
- Cat 3 — pLC50 2.0–3.0 (LC50 10–100 mg/L)
- Not Classified pLC50 < 2.0

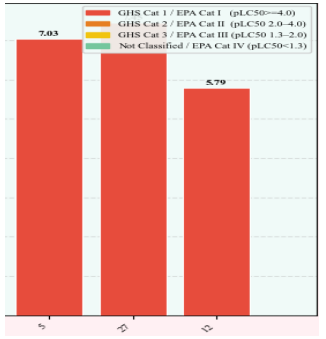
US EPA (↑ pLC50 = ↑ Toxicity)

- Cat I — pLC50 ≥ 4.0 (LC50 <1 mg/L)
- Cat II — pLC50 2.0–4.0 (LC50 1–100 mg/L)
- Cat III — pLC50 1.3–2.0 (LC50 100–500 mg/L)
- Cat IV — pLC50 < 1.3 (LC50 >500 mg/L)

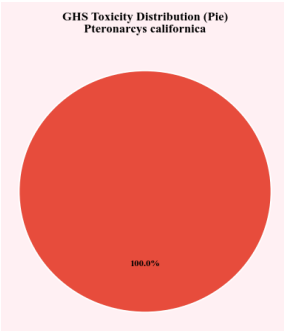
Predicted_pLC50
4.9811
4.9606
7.0723
6.8314
6.9722
4.4443
4.2509

Generate Predictions

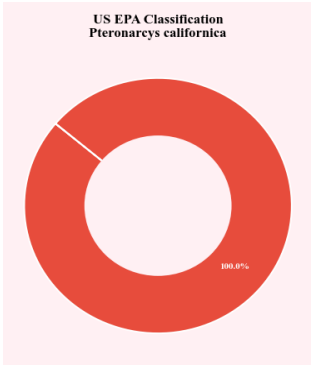
Auto-generated toxicity charts



Bar chart



Pie chart



Donut chart

Model customization

ADD CUSTOM STONEFLY MODEL

Define a new Plecoptera species model with custom descriptors and regression

MODEL INFORMATION

Model Name:

Species / Organism: Pteronarcys californica (Giant Salmonfly)

Custom name (if Other): Pteronarcys californica (Giant Salmonfly)

Endpoint: pLC50

Method: MLR q-RASAR

Latent Variables (PLS):

EQUATION — pEndpoint = Intercept + Σ(Coefficient × Descriptor)

Intercept:

Number of Descriptors:

Descriptor Name	Coefficient	
		x
		x
		x
		x
		x

+ Add Descriptor Row

Save Model | Clear Form

Conclusions

1. The developed MLR-based QSAR and q-RASAR models are **robust** and show **good statistical predictability,** following **OECD** guidelines.

2. From this study, it has been found that the presence of the **vinyl group (-CH=CH₂), ether-type Oxygen atom (R–O–R)**, and **electronegative heteroatoms (O, S, Cl, F)** plays a crucial role in regulating **aquatic toxicity**.

3. **Screening of real-world data** was performed to check the prediction accuracy of the developed model and demonstrate its feasibility, along with the ranking of toxicants according to their ecotoxicological potential and status of AD.

4. A novel **Python-based expert system** '***StoneTox_Predictor –v1.0***' was developed exclusively to calculate the pLC_{50} of any **untested** compound and give a classification of toxicants in colourful bar, pie and doughnut charts according to global standards.

5. By using these models, we gained comprehensive knowledge of how certain chemicals impact **aquatic insects** at a toxicological level. This knowledge is crucial for developing effective measures to protect the health of aquatic insects and **ecosystems.**

References

1. Feeley, H.B., Baars, J.R., Kelly-Quinn, M. and Nelson, B., 2020. Ireland Red List No. 13: Stoneflies (Plecoptera). *National Parks and Wildlife Service, Department of Culture, Heritage and the Gaeltacht, Dublin, Ireland*.

2. Das, S., Samal, A. and Ojha, P.K., 2024. Chemometrics-driven prediction and prioritization of diverse pesticides on chickens for addressing hazardous effects on public health. Journal of Hazardous Materials, 471, p.134326.

3. **Mahapatra, D**., Das, S., Samanta, P. and Ojha, P.K., 2026. QSAR meets ecology: Predictive framework for assessing pesticide toxicity against mayfly using consensus modelling. *Science of The Total Environment*, *1028*, p.181706.

4. Banerjee, A. and Roy, K., 2024. How to correctly develop q-RASAR models for predictive cheminformatics. *Expert Opinion on Drug Discovery*, *19*(9), pp.1017-1022.

5. Krno, I.J. and Holubec, M., 2009. Effects of land use on stonefly bioassessment metrics. Aquatic Insects, 31(sup1), pp.377-389.

IBMC
Institute of Biomedical Chemistry

Way2Drug
Understanding Chemical-Biological Interactions
PREDICTIVE SERVICES



XXXII Symposium on Bioinformatics and Computer-Aided Drug Discovery (BCADD-2026)

