MULTIVARIATE QSAR MODELING FOR IDENTIFYING STRUCTURAL DRIVERS AND PRIORITIZING TOXIC CHEMICALS AFFECTING *CRASSOSTREA VIRGINICA*

Presented By,

Prodipta Bhattacharyya
Research Scholar

Under the guidance of

Dr. Probir Kumar Ojha
Associate Professor

Drug Discovery and Development Laboratory (DDD Lab)
Department of Pharmaceutical Technology
Jadavpur University
Kolkata-700032
India

Introduction

Keystone Species & Ecosystem Engineer

- Regulates coastal ecosystems by creating complex reef habitats.
- Improves water quality through filtration (removes suspended particles and pollutants).
- Supports biodiversity by providing habitat and resources for other species.

Dietary Importance

Valued for its significant nutritional value and contribution to human health.



Crassostrea virginica
(Eastern oyster)

Critical Bio-indicator

Serves as a key indicator for monitoring environmental quality in coastal and estuarine ecosystems.

Global Status

Assessed as **G3** ("vulnerable") by **NatureServe**.

Major Threats

The decline of *C. virginica* is driven by multiple, interconnected stressors:

Overharvesting reduces population and recruitment.

Water pollution introduces harmful chemicals and pathogens.

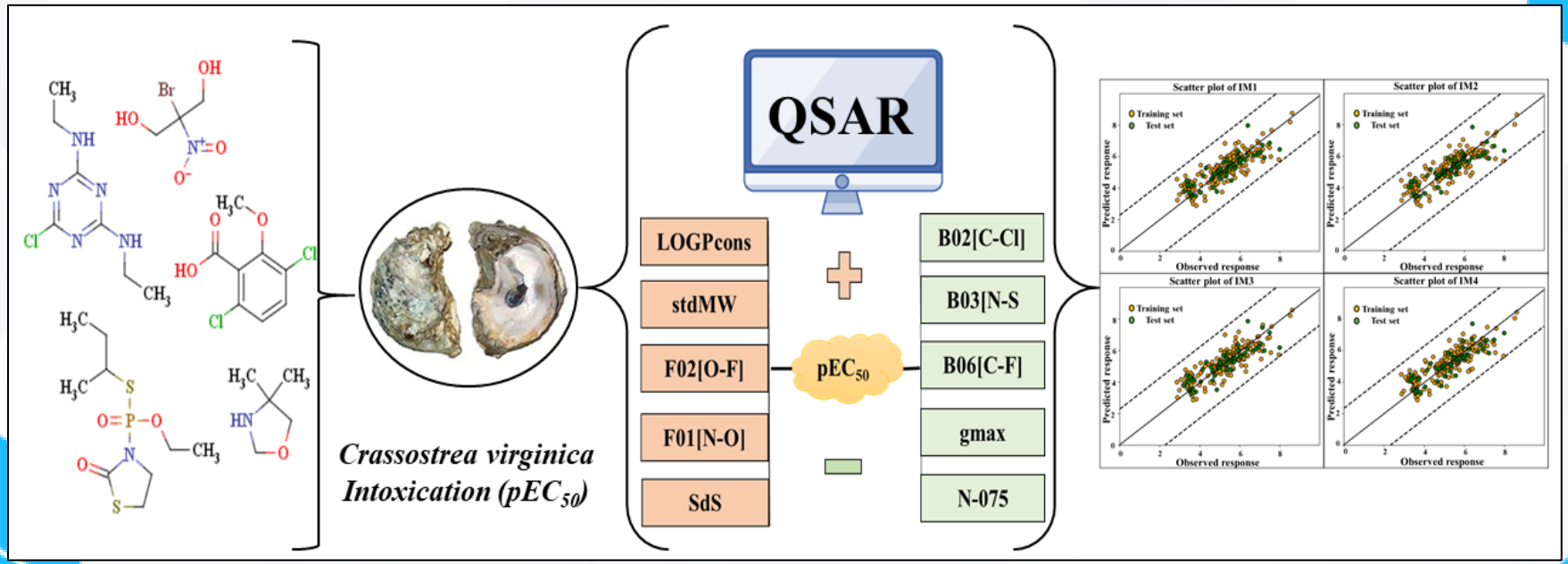
Reef habitat degradation destroys critical habitat structures.

Disease outbreaks cause mass mortality and reduce resilience.

Environmental Hazards

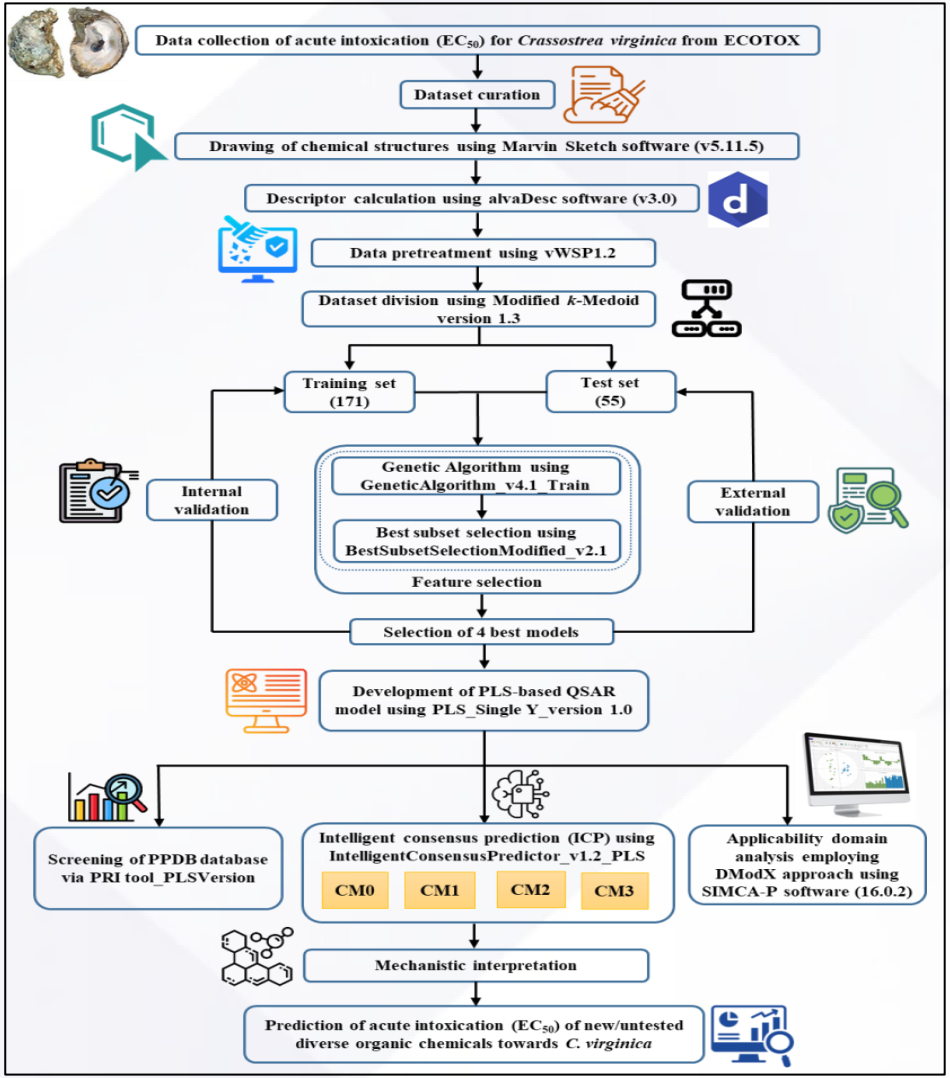
Environmental hazards pose a severe threat to aquatic species, including *C. virginica*, affecting their survival, health and long-term population stability.

Introduction



Materials and Methods

Schematic representation of the study



Results and Discussions

Model IM1

pEC_{50}= 2.97689+0.10008×$stdMW$−0.59863×$B02[C-Cl]$−0.64454×$B03[N-S]$−0.34629×$B06[C-F]$ +0.40894×$F02[O-F]$+0.57642×$LOGPcons$

Model IM2

pEC_{50}= 2.85315+0.10639×$stdMW$+0.08689×SdS−0.51979×$B02[C-Cl]$−0.66287×$B03[N-S]$ +0.34069×$F02[O-F]$+0.55643×$LOGPcons$

Model IM3

pEC_{50}=3.61182+0.09273×$stdMW$−0.26179×N−075−0.05460×$gmax$−0.48491×$B02[C-Cl]$+0.35962×$F02[O-F]$+0.58032×$LOGPcons$

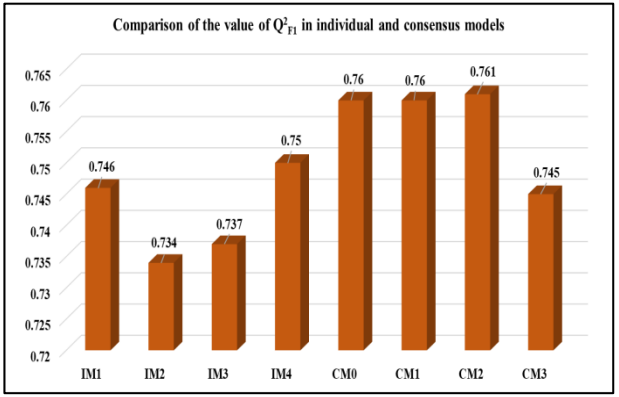
Model IM4

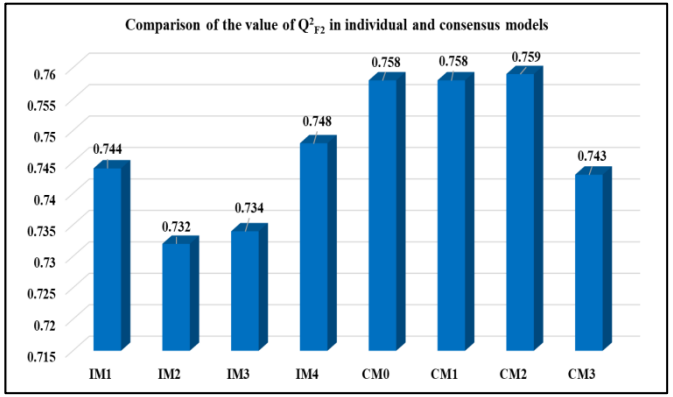
pEC_{50}= 3.00566+0.09152×$stdMW$−0.50906×$B02[C-Cl]$−0.79496×$B03[N-S]$+0.18918×$F01[N-O]$ +0.34412×$F02[O-F]$+0.54762×$LOGPcons$

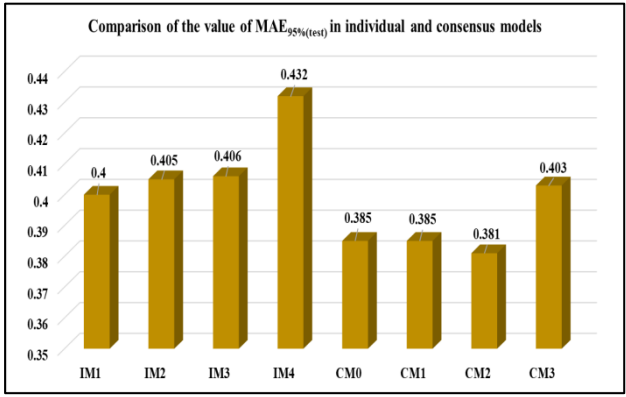
Table 1. Internal validation metrics

Individual models	DES/ LV	R^2	Q^2_{LOO}
IM1	6/3	0.714	0.690
IM2	6/3	0.714	0.690
IM3	6/3	0.712	0.689
IM4	6/2	0.700	0.673

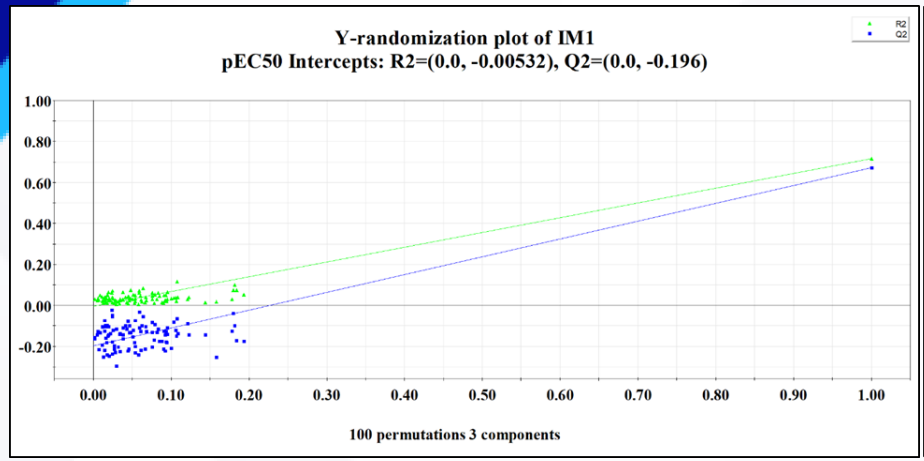
Comparison of external validation metrics of individual and consensus models

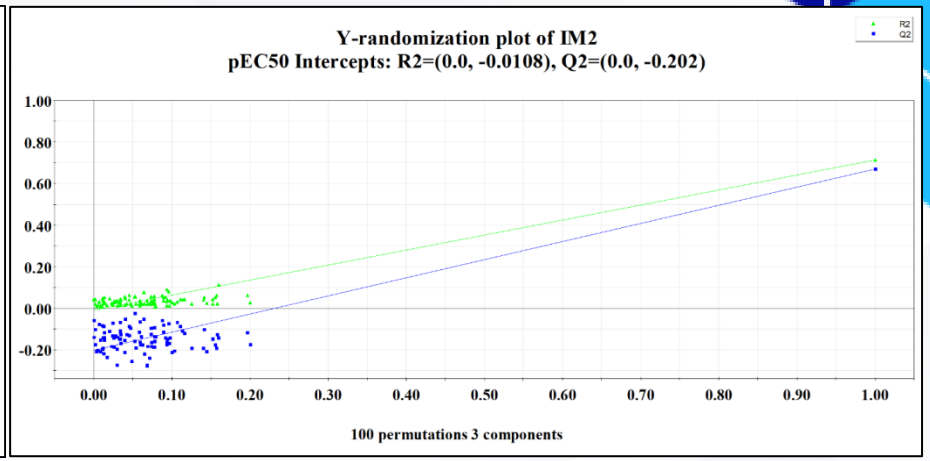


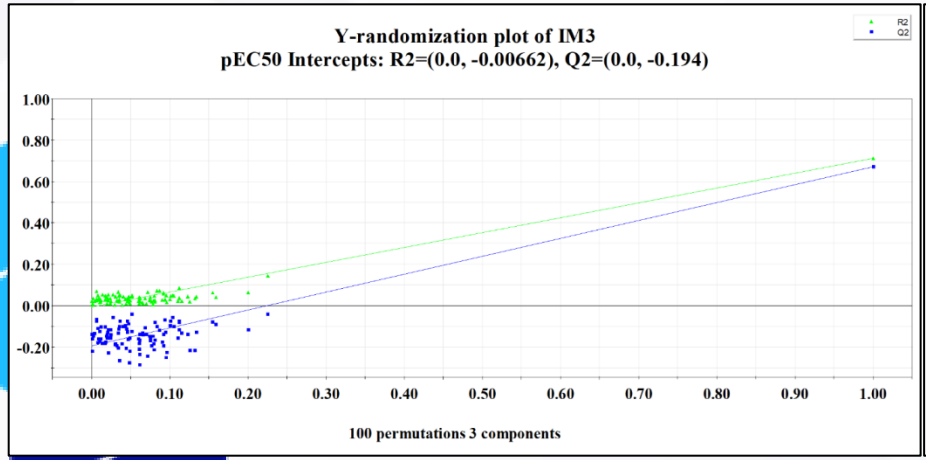


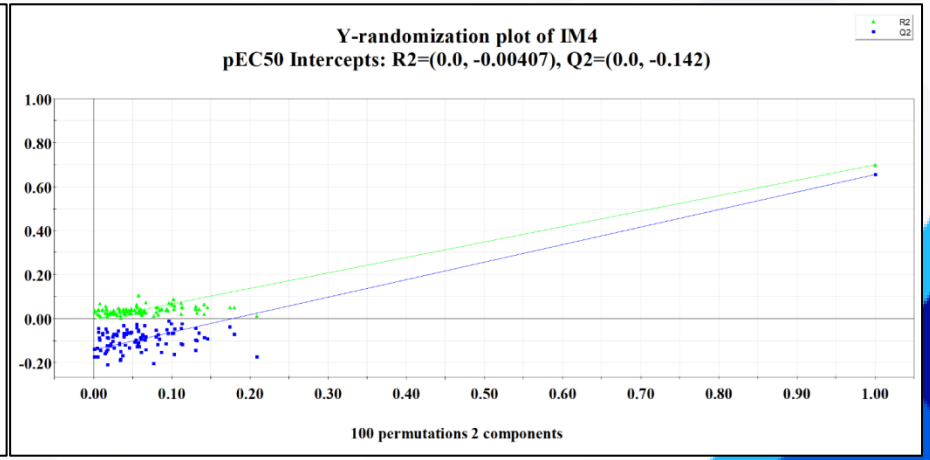


Y-randomization plots of the PLS-based QSAR models

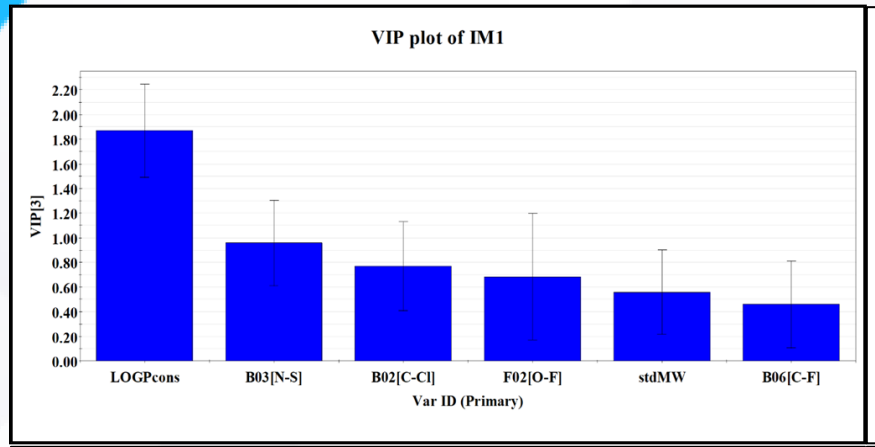


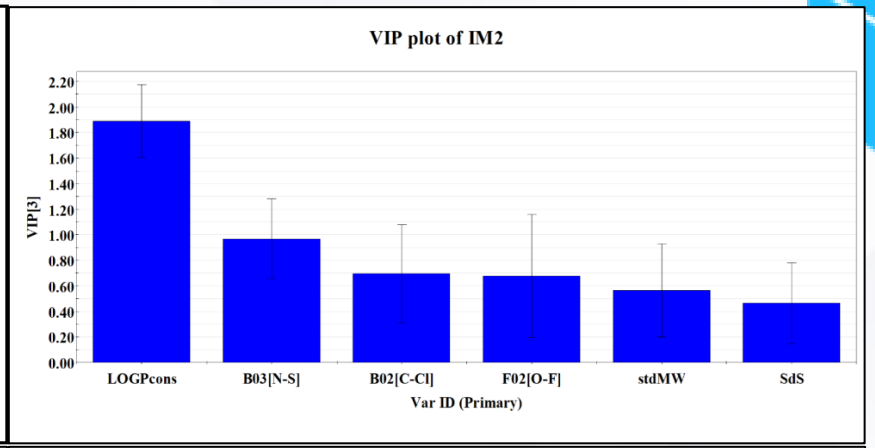


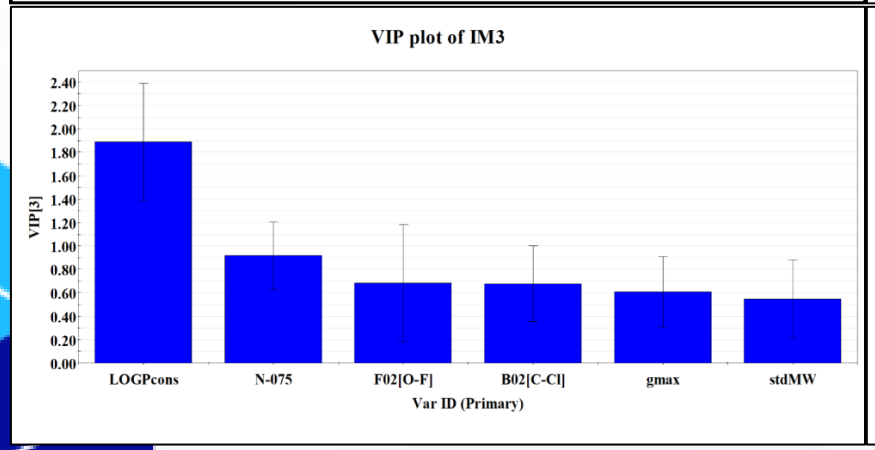


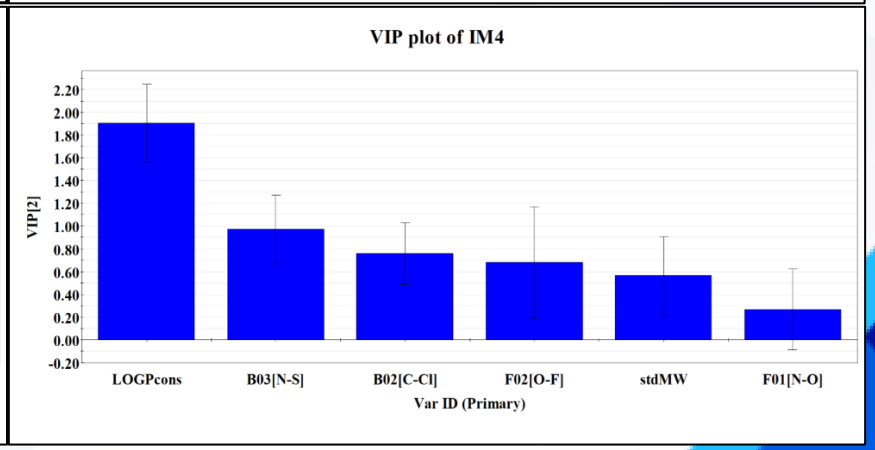


Variable importance projection (VIP) plots of the PLS-based QSAR models

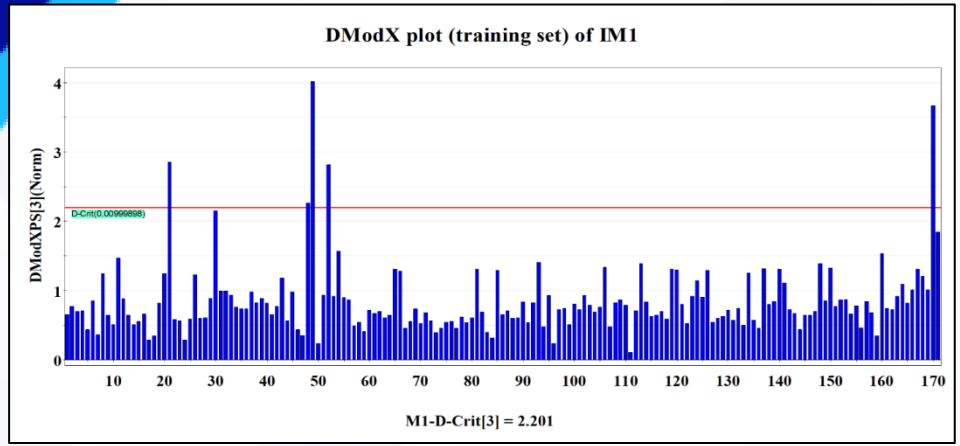


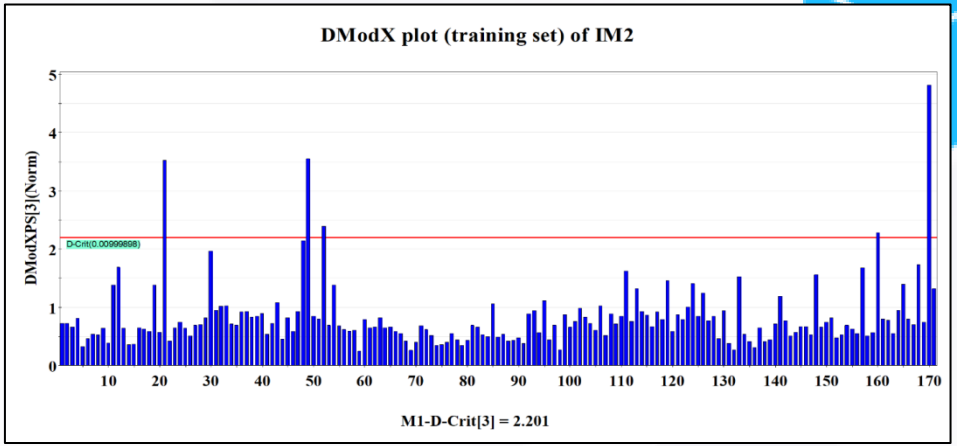


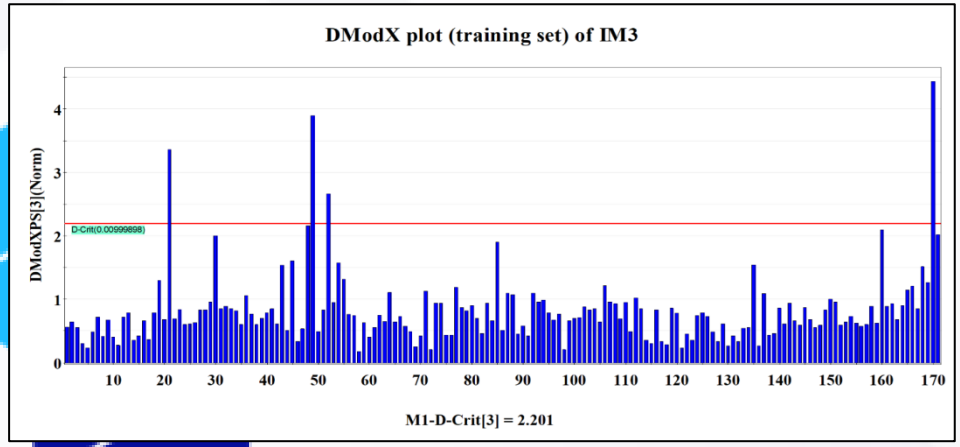


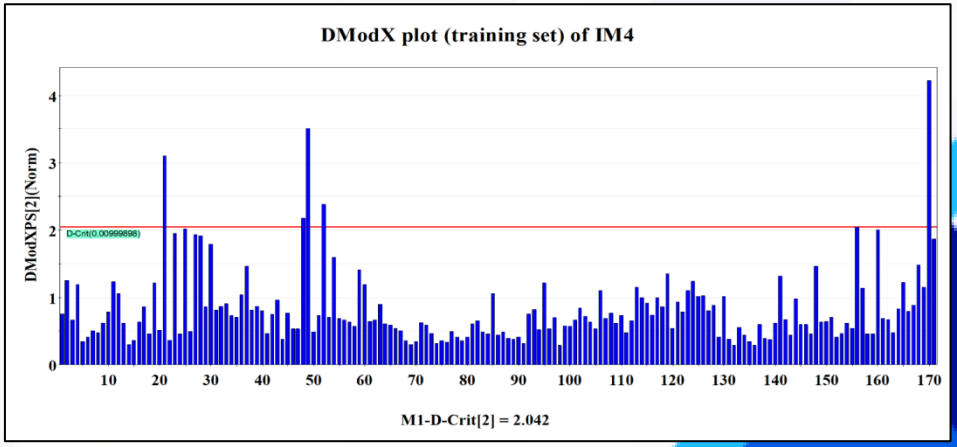


Applicability Domain (AD) analysis of the PLS-based QSAR models

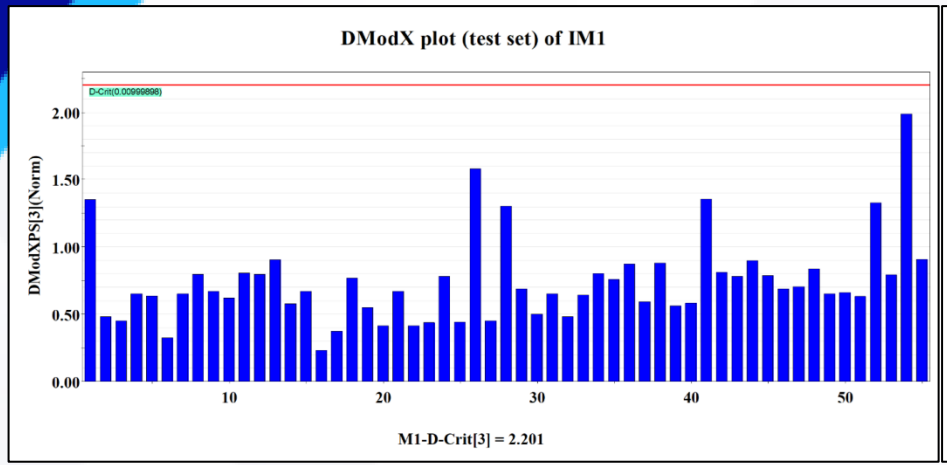


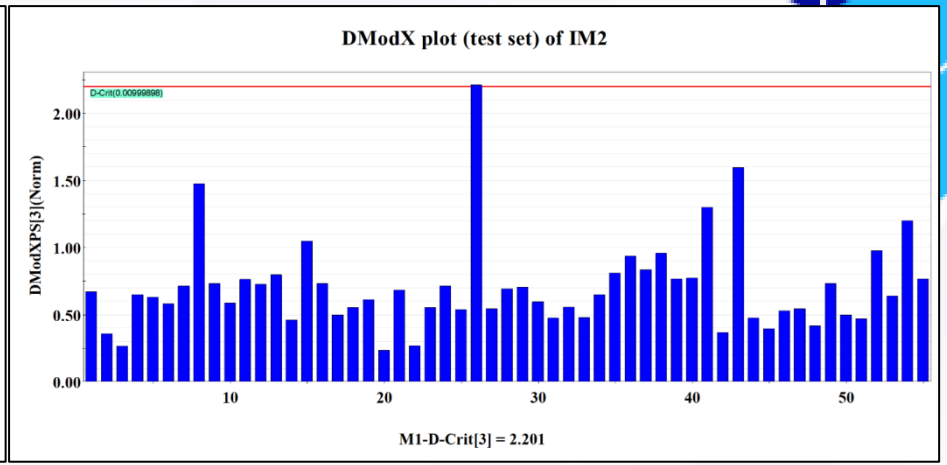


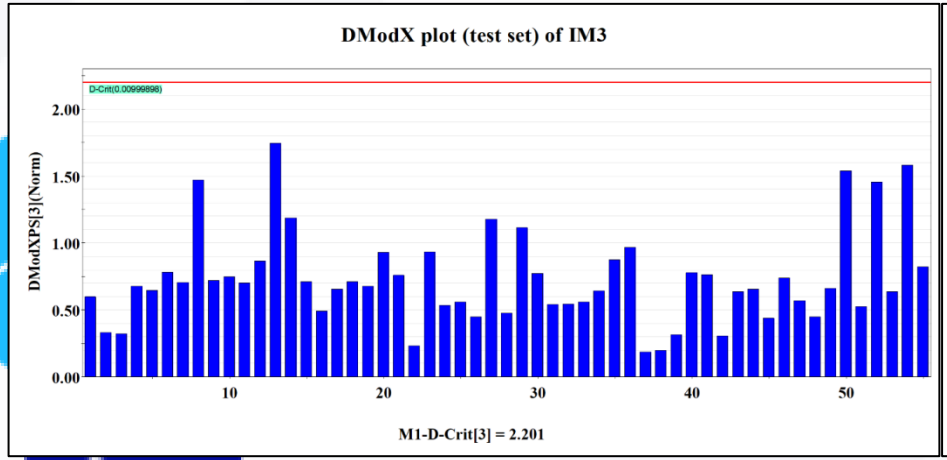


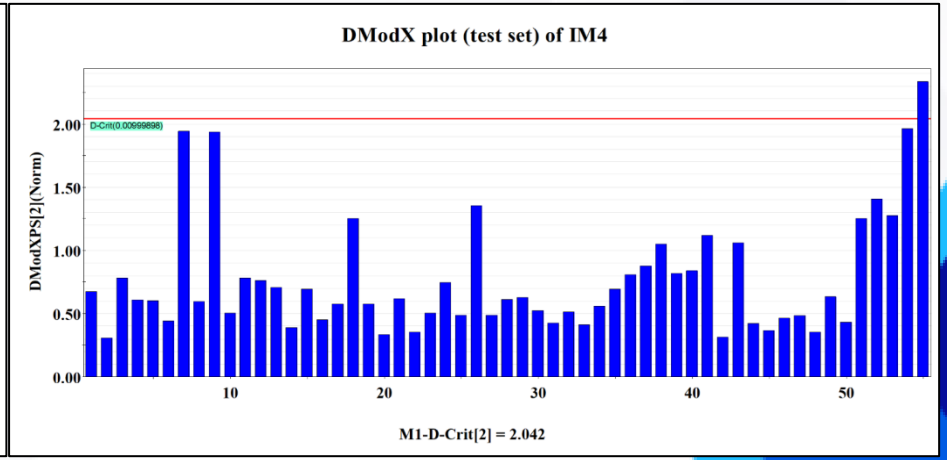


Applicability Domain (AD) analysis of the PLS-based QSAR models

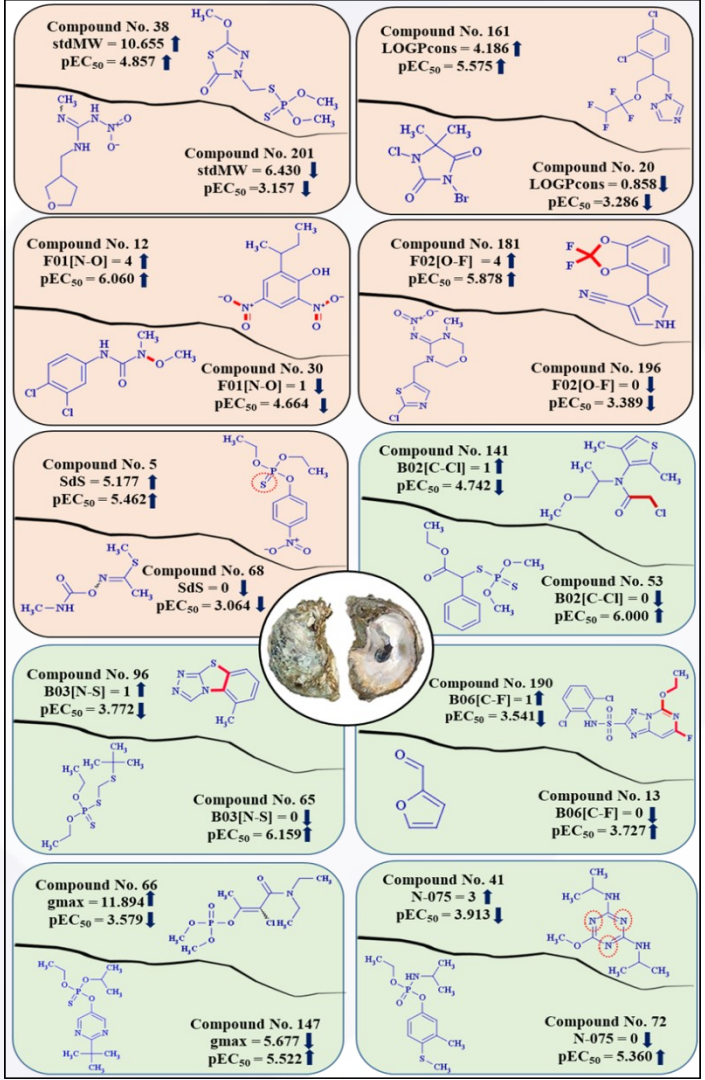








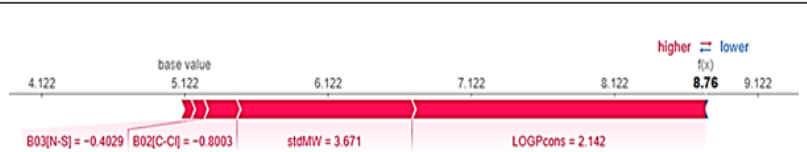
Mechanistic Interpretation of the modeled descriptors



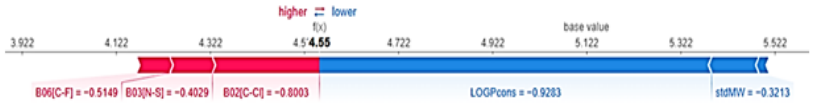
Force plots for the highest and lowest toxic compound

For IM1

Compound 121

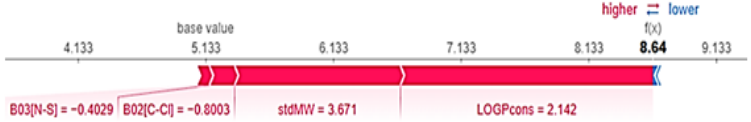


Compound 89

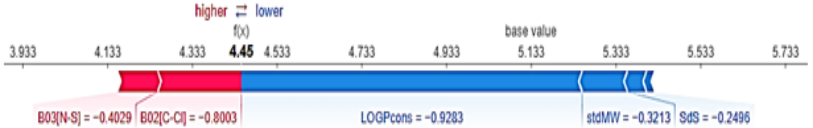


For IM2

Compound 121

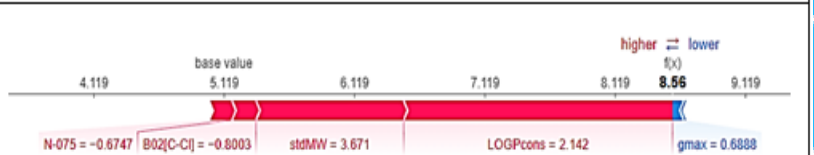


Compound 89



For IM3

Compound 121

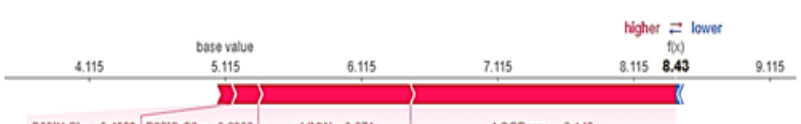


Compound 89



For IM4

Compound 121

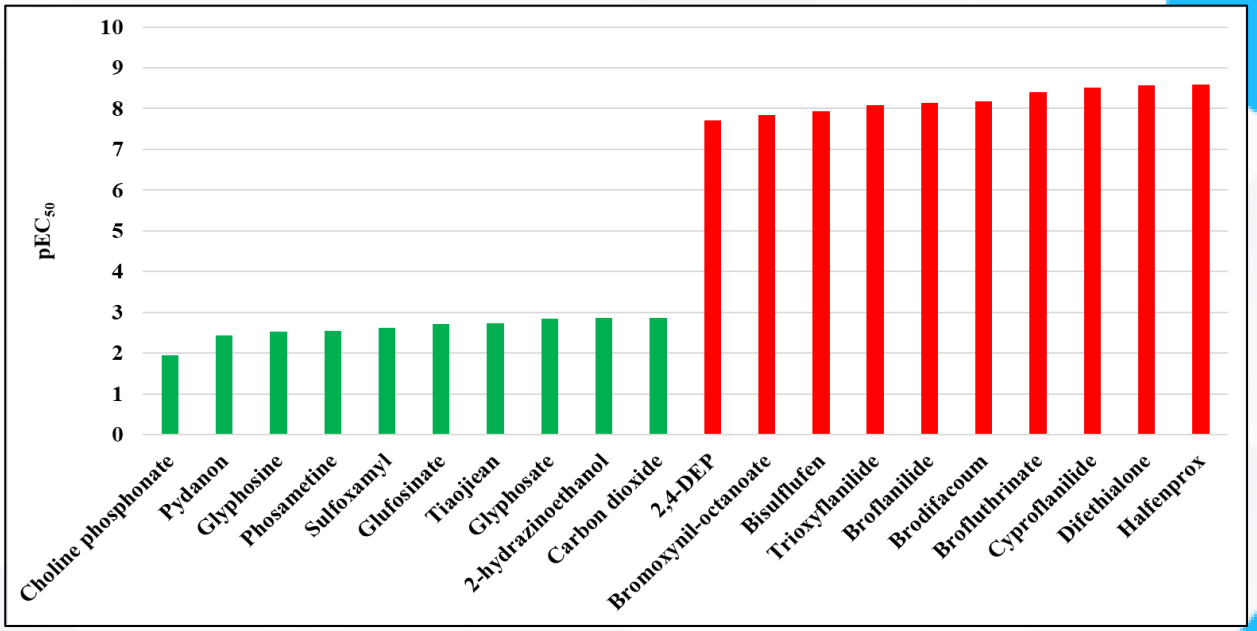


Compound 89



Screening of Pesticide Properties DataBase (PPDB)

Top and least 10 toxic screened pesticides of PPDB



Conclusion

- The present work accomplished the development of statistically sound **PLS-based QSAR models** to predict the **effect of intoxication** of diverse chemicals in ***C. virginica***.
- The **consensus model 2 (CM2)** had the **lowest prediction error ($MAE_{95\%(test)}$)** and is therefore regarded as the best model among all individual models and other consensus models.
- The **key structural findings** gained from this study will help to understand the mechanism of toxicity caused by the chemicals.
- The QSAR models were applied to **screen the external PPDB database**, enabling prioritization of the most and least toxic pesticides. This supports the application of the models in practical scenarios and support early identification of potent toxicants.
- Our study thus provides an **effective and sustainable framework** for early risk assessment in *C. virginica* to mitigate the threat associated with diverse chemicals, consequently protecting the coastal and marine ecosystems and supporting the dietary health of human beings.

Reference

- Roy, K., Ambure, P., Kar, S., Ojha, P.K., 2018a. Is it possible to improve the quality of predictions from an “intelligent” use of multiple QSAR/QSPR/QSTR models? J. Chemom. 32 (4), e2992. https://doi.org/10.1002/cem.2992.
- Roy, K., Ambure, P., Kar, S., 2018b. How precise are our quantitative structure–activity relationship derived predictions for new query chemicals? ACS Omega 3 (9), 11392–11406. https://doi.org/10.1021/acsomega.8b01647.
- Roy, K., Kar, S., Das, R.N., 2015. Statistical methods in QSAR/QSPR. In: A Primer on QSAR/QSPR Modeling: Fundamental Concepts. Springer International Publishing, Cham, pp. 37–59. https://doi.org/10.1007/978-3-319-17281-1_2.
- Wold, S., Sj¨ ostr¨ om, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58 (2), 109–130. https://doi.org/10.1016/S0169-7439 (01)00155-1.

THANK YOU