PESTICIDE TOXICITY IN EARTHWORMS EXPLAINED BY INTERPRETABLE MACHINE LEARNING

Mihkel Kotli, Geven Piir, Uko Maran

Institute of Chemistry, University of Tartu, Ravila 14a, 50411 Tartu, Estonia
e-mail of presenting author: mkotli@ut.ee

Chemicals, such as pesticides, that have entered nature during the rationalization of agriculture have unwanted effects on the environment and organisms. For environmental, regulatory, and commercial purposes, it is important to comprehend the impact these chemicals have. A solution is to use machine learning (ML) models, which besides avoiding animal testing, yield results faster and cheaper than testing on live animals. Retaining soil health is the foundation of agriculture and earthworms are among the most important animals for soil health. This makes earthworms an important test and model organism, and so far prediction of acute toxicity in earthworms has been an open challenge for quantitative structure-activity relationship (QSAR) modelling.

Sophisticated ML approaches are recommended for modelling structurally diverse datasets, but they lack interpretability. A possible solution is provided by Shapley analysis, which is a game-theoretic approach to explaining how a prediction is based on inputs (molecular descriptors), allowing for interpretability in ML-QSARs.

A new dataset consisting of more than 600 pesticides collected from open access sources was compiled for this study. A Random Forest (RF) classifier was optimised using a novel approach of combined genetic algorithm variable reduction and Bayesian parameter tuning. The Shapley values were then used to decompose the RF classifier, providing insight into how structural features affect toxicity.

A reliable model was obtained for assessing the toxicity of chemical compounds on earthworms with twelve molecular descriptors and providing prediction accuracies around 80% for both training and test sets. Analysis of molecular descriptors found that intramolecular connectivity coding the number of cycles determines the lethality of compounds. Also, the molecular symmetry and steric accessibility of certain structural fragments with high van der Waals radii are additional important structural features affecting earthworm lethality, whereas very high lipophilicity is marker for nontoxicity. Insights into toxicities of thioethers and triazines were provided, among others. Glutathione depletion as a mechanism finds support for further experimental investigation.


Euroopa Liit
Euroopa
Regionaalarengu Fond


Eesti
tuleviku heaks