Predicting Freshwater Ecotoxicological Concentrations for Chemicals in Household Products Using Linear Regression and Random Forest

Suppiger, M.; Fantke, P.; Zhao, B.; Colacino, J.; Huang, L.; Jolliet, O.

Published in:
SOT Annual Meeting and ToxExpo 2022

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

4566 Toxicity of Copper Nanoparticle against a Representative Agricultural Crop: Tomato

S. Shikha[1], T. Ameh[1], C. Collom[1], E. Braswell[2], and C. Sayes[1]. [1]*Baylor University, Waco, TX; and* [2]*US Department of Agriculture, Edinburg, TX.*

Development of cities and towns around the world include now, more than ever, the very close proximity of plants/crops grown, and animals raised for food in and around cities near residential areas. Production, delivery, processing, and marketing of products is part of urban agriculture and has become an integral part of economic, social, and ecological systems. Chemicals, pesticides, and metalloids from anthropogenic sources have been used in agriculture for centuries to increase productivity, protection of crop losses, vector-borne disease control, and fortification. Now, with micronutrient applications, such as the use of copper nanoparticles in agricultural settings near human habitats, there is an increased need to detect, identify, and measure the impact of pesticides and metalloids in all four environmental compartments, i.e., air, water, soil, and biota. Although copper is an essential element, due to the immense use in its nanoparticulate form, environmental health concerns about copper nanoparticle impacts on crops and eventually to human health due to its consumption are warranted. Terrestrial plant species may be contaminated with copper nanoparticles via roots. In this study, foliar exposure, and soil enrichment of copper nanoparticles to *Solanum Lycopersicum* (tomato) were evaluated. The effects of the copper exposure were evaluated based on root length, shoot length, fresh weight and dry biomass, chlorophyll and carotenoid content, and fruit production. The effects of copper translocation and accumulation in plant parts were evaluated *via* inductively coupled plasma-mass spectrometry and transmission electron microscopy. Copper nanoparticle transfer rates were estimated from the total copper concentrations measured within leaves and roots. Most copper accumulation was found in leaves as compared to roots. Some accumulation was also found in fruit. These results will help to understand the fate and physiological effects of copper nanoparticles in fruit-bearing crops.

4567 Predicting Freshwater Ecotoxicological Concentrations for Chemicals in Household Products Using Linear Regression and Random Forest

M. Suppiger[1], P. Fantke[2], B. Zhao[1], J. Colacino[1], L. Huang[1], and O. Jolliet[1]. [1]*University of Michigan, Ann Arbor, MI; and* [2]*Danmarks Tekniske Universitet, Kongens Lyngby, Denmark.*

Empirical ecotoxicological data is limited for many chemicals and species. Thus, predictive models can be utilized to estimate toxic concentrations when experimental data is lacking. Linear regression and random forest models were developed to predict chronic freshwater hazardous concentrations (HC20) of the species-level chronic NOAEL for chemicals in the US EPA Chemical and Product Database (CPDAT). When acute HC50 data was available, chronic HC20 were predicted using simple linear regression (N=4,887; R^2=0.89). The univariate linear model performed similarly to a random forest model with the same input (R^2=0.88). To generate HC20 predictions for chemicals lacking acute HC50 values, chemical toxicity parameters from the OPERA and TEST models were entered into a random forest model. All 29 toxicity and chemical property inputs were ranked by variable importance (increase in mean squared error) and added to the model sequentially, until model performance was optimized. Only six inputs generated from the TEST consensus QSAR prediction method and four OPERA chemical properties were needed to maximize test R^2. Specifically, oral rat toxicity (LD50) predictions, 96-hour fish toxicity (EC50) predictions, 48-hour Daphnia toxicity (LD50) predictions, algae ecotoxicity predictions (IC50), bioconcentration (both TEST and OPERA predictions), water solubility (from TEST only), mass, and density were used in the random forest model (N=1,828; R^2=0.59). This random forest model significantly outperformed a multi-linear regression with the same inputs (R^2=0.394). Finally, when TEST model predictions were missing for a chemical, chemical properties from only the OPERA model were entered in a separate random forest model (N=4,251; R^2=0.373). Only nine out of the twelve OPERA chemical properties were needed to achieve maximum model performance (water solubility, bioconcentration, octanol-air partition coefficient, boiling point, Henry's law coefficient, octanol-water partition coefficient, vapor pressure, fish biotransformation half-life, and biodegradation). Once the models were built and trained, HC20 values were predicted for 700,000 chemicals, with the highest performing model being used in priority per chemical. HC20 predictions can be combined with chemical fate, transport, and exposure data to quantify freshwater ecosystem impacts. Quantifying ecotoxicological damages is a necessary step towards successful ecosystem conservation efforts and thus, the applicability of predictive modeling should continue to be a focus of ecotoxicological research.

4568 Improving the Efficiency of Literature Identification for the ECOTOXicology Knowledge Base Using Deep Learning

B. Howard[1], C. Norman[1], A. Tandon[1], R. Shah[1], J. Olker[2], and D. Hoff[2]. [1]*Sciome LLC, Research Triangle Park, NC; and* [2]*US EPA, Duluth, MN.*

The ECOTOXicology Knowledgebase (ECOTOX) is a comprehensive, publicly available resource providing single chemical environmental toxicity data on aquatic life, terrestrial plants and wildlife. The database is updated quarterly, and to identify relevant references and extract pertinent data, the ECOTOX data curation pipeline employs a methodical process similar to initial stages of systematic review. This labor-intensive workflow requires curators to regularly evaluate tens of thousands of candidate references, the majority of which are then rejected as not relevant. After the careful review of hundreds of thousands of articles, the ECOTOX database currently (as of November 2021) contains data for 12,386 chemicals and 13,621 species extracted from 52,551 references. The availability of this extensive dataset of historical screening decisions provided us with the opportunity to develop state-of-the-art neural network classifiers to partially automate title and abstract screening and to categorize (e.g. human health, fate, chemical methods) rejected references. We used this data to develop two deep learning models which were then integrated into a modified version of the SWIFT-Active Screener software, a collaborative web-based reference screening platform. The first model is a neural language-model classifier that predicts the relevance of candidate references. When used to augment the standard SWIFT-Active Screener document prioritization model, this method provides a mean improvement of 6.5% Work Saved over random Sampling (WSS) compared to the standard approach. The second model uses a separate deep learning network to conduct multi-class classification of excluded documents to predict the reason for exclusion. This model achieves F-scores in the 65-75% range for the most frequent classes and has been integrated into Active Screener to provide intelligent "default choices" for capturing exclusion reason. Using extensive simulations, we demonstrate that this modified version of Active Screener results in more than a 50% reduction, on average, in time spent screening ECOTOX references, with larger savings for the datasets having the most articles. Models have now been integrated into the literature curation pipeline at EPA and our focus is now to further refining them according to ongoing feedback obtained during the screening of new datasets.

4569 AOP Development: Aromatase Inhibition and Androgen Receptor Agonism Lead to Male Biased Sex Ratio via Impacts on Gonad Differentiation

K. J. Santana Rodriguez, D. L. Villeneuve, G. T. Ankley, and D. H. Miller. *US EPA, Duluth, MN.*

In teleost fish, the balance between steroidal estrogens and androgens is essential during sexual differentiation and this balance is in turn dependent on the availability and activity of steroid synthesizing enzymes such as aromatase, the enzyme responsible for the conversion of androgens to estrogens. Given the important role that steroid hormone signaling plays in sexual differentiation, it is not surprising that exposure to some types of endocrine active substances during critical periods of development can result in skewed phenotypic sex ratios. The present study utilized the adverse outcome pathway (AOP) framework to organize available evidence linking aromatase inhibition and androgen receptor agonism to male-biased sex ratios in fish. A literature search was conducted using 'aromatase inhibition', 'male biased sex ratios' and 'androgen sex reversal' in combination with 'fish' as the main keywords. Publications describing exposures or experimental treatments of teleost fish during the critical period of sexual differentiation and including both gonadal histology and an indicator of final population sex ratio were included as relevant. Additionally, a novel mathematical model linking male biased sex ratio to projected alterations in population trajectories was developed. Collectively, there was moderate to high scientific support for the causal linkages between aromatase inhibition or androgen receptor agonism and male biased sex ratios in fish. These two AOPs, focused on impacts during the period of sexual differentiation (i.e., early life stages), complement a broader network of AOPs documenting potential hazards that endocrine active chemicals pose to fish, thereby further supporting predictive approaches to hazard and risk assessment. *The contents of this abstract neither constitute, nor necessarily reflect, official US EPA policy.*

4570 Proteomics Reveal That Lactate Dehydrogenase Has Prognostic Value in Acetaminophen-Induced Acute Liver Failure Patients

J. H. Vazquez[1], S. Kennon-McGill[1], S. D. Byrum[1], S. G. Mackintosh[1], D. K. Williams[1], W. M. Lee[2], J. A. Dranoff[1], and M. R. McGill[1]. [1]*University of Arkansas for Medical Sciences, Little Rock, AR; and* [2]*University of Texas Southwestern Medical Center, Dallas, TX.*

Acute liver failure (ALF) is a rare yet serious condition with a high rate of mortality (~30%). Unfortunately, there are currently no biomarkers that have sufficient prognostic value to determine if a patient is unlikely to survive without a liver transplant. We obtained early and later serum samples (study days 1s and 3, respectively) from patients with Acetaminophen (APAP)-induced ALF in the Acute Liver Failure Study Group biorepository, divided into transplant-free survivor (n=28)