Acute toxicity prediction to threatened and endangered species using Interspecies Correlation Estimation (ICE) models

Morgan M. Willming\*†, Crystal R. Lilavois‡, Mace G. Barron‡, Sandy Raimondo‡

†Oak Ridge Institute for Science and Education, U.S. Environmental Protection Agency, Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL USA 32561

‡U.S. Environmental Protection Agency, National Health and Environmental Effects Laboratory, Gulf Ecology Division, 1 Sabine Island Drive, Gulf Breeze, FL USA 32561

\*Corresponding author: Willming.Morgan@epa.gov; Phone: 850-934-9297; Fax: 851-934-2406

**Abstract**

 Evaluating contaminant sensitivity of threatened and endangered (listed) species and protectiveness of chemical regulations often depends on toxicity data for commonly tested surrogate species. The U.S. EPA’s internet application Web-ICE is a suite of Interspecies Correlation Estimation (ICE) models that can extrapolate species sensitivity to listed taxa. ICE models are least squares regressions of the sensitivity of a surrogate species and a predicted taxon (species, genus, or family) constructed from acute toxicity values measured in each pair of taxa. Web-ICE was updated with additional toxicity records, optimized model selection guidance, and new models with the potential to predict to over 250 listed species. A case study was used to assess protectiveness of genus and family models derived from either geometric mean or minimum taxa toxicity values for U.S. federally listed species and priority chemicals from the Sacramento California and Ohio River Valleys. Genus and family models developed from the most sensitive value for each chemical were generally protective of the most sensitive species within predicted taxa, including listed species, and were more protective than geometric means models. ICE models provide robust toxicity predictions and can generate protective toxicity estimates for assessing contaminant risk to listed species.

**Introduction**

The U.S. Endangered Species Act (ESA) requires EPA to determine risks of pesticides and other chemicals to U.S. federally endangered and threatened (listed) species to ensure that chemical registration and water quality criteria are protective of listed taxa. A significant challenge in this process is determining the sensitivity of the diversity of listed species to chemicals that have only been evaluated using common test species or have limited data available. Historically, toxicity data for the majority of listed species or closely related representatives has been unavailable because of a lack of standardized culture and test methods, and limited organism availability. However, previous research suggests that listed species are not consistently more sensitive than more commonly tested species [1](#_ENREF_1), [2](#_ENREF_2). In the absence of species-specific toxicity data, conservative approaches such as applying generic safety factors to toxicity values of surrogate species have been used to develop hazard levels assumed to be protective of listed species. Recently, the U.S. National Research Council [3](#_ENREF_3) recommended the use of Interspecies Correlation Estimation (ICE) models in pesticide risk assessments as an alternative to generic safety factors.

ICE models estimate acute toxicity to aquatic or terrestrial organisms using the known toxicity data of a chemical for a surrogate species [4](#_ENREF_4), [5](#_ENREF_5). The models are log-linear least squares regressions of the sensitivity of a predicted taxon (species, genus, or family) and a surrogate species and are constructed from existing acute toxicity values determined in each pair of taxa across a number of chemicals. ICE models for aquatic species contain acute toxicity values (median lethal or median effect concentrations; LC50/EC50) for a minimum of any three different chemicals and have been demonstrated to be robust, accurate estimators of toxicity when the surrogate and predicted taxa are within the same taxonomic Order [4](#_ENREF_4). Species, genus, and family level ICE models are publically available on the U.S. EPA internet application Web-ICE (www3.epa.gov/webice).

ICE model predictions are intended to supplement toxicity databases where species of concern or species diversity have not been or cannot be adequately tested[6](#_ENREF_6). Previous studies evaluating the robustness and application of ICE model predictions have focused primarily on species level models and their use in either direct toxicity estimation or in developing species sensitivity distributions (SSDs)[4-8](#_ENREF_4). Less work has explored the protectiveness of ICE predictions in risk assessments for listed species or the use of genus and family models. Few species-specific ICE models can be developed for listed species due to limited existing toxicity data, therefore genus or family level models may be required to predict toxicity to the higher taxonomic level. Genus and family models have historically been developed using geometric means of toxicity values from multiple species, which are useful in generating toxicity estimates in cases such as water quality criteria development. However, the protectiveness of these predictions across the range of species sensitivity within the predicted taxon is uncertain, especially for listed or sensitive species.

The present study evaluates the prediction accuracy of ICE models for listed species and compares the protectiveness of genus and family level models developed from either geometric means or minimum toxicity values. Models used in this evaluation were developed from an acute toxicity database that was substantially expanded from the previous version of Web-ICE (v3.2, release April 2013) and contains 314 species and 1501 chemicals, including new toxicity records for listed species. Of particular emphasis was the expansion of models for prediction to freshwater unionid mussels[9](#_ENREF_9). We compare the development of genus and family models using the most sensitive toxicity value for each chemical within the predicted taxon to models built with geometric means, as a potential approach to developing models that are protective of the most sensitive species within a taxon. We revised model selection guidelines to reduce reliance on professional judgement and increase reproducibility of their application. Lastly, we demonstrate the application of models and guidelines using a case study of 8 listed species exposed to a diversity of pesticides encompassing a range of aquatic toxicity modes of action (MOA).

**Methods**

*Database development*

Standardization of toxicity data for inclusion in ICE models followed the approach and selection criteria described in Raimondo et al. (2010). This included confirming chemical and species identity; compiling acute toxicity values as 48 hour EC50/LC50 for specific invertebrate taxa (e.g., daphnids, fairy shrimp) or 96 hour EC50/LC50 for fish, amphibians and other aquatic invertebrates (e.g., insects); and determining whether each data record met standardization criteria for life stage, test conditions, and water quality parameters (e.g., temperature, dissolved oxygen, salinity)[4](#_ENREF_4). The database included records from the previous Web-ICE database appended with new records from ECOTOX ([www.epa.gov/ecotox/](http://www.epa.gov/ecotox/); downloaded September 2014), and recently collected primary data on freshwater mussels[9](#_ENREF_9) and fairy shrimp (unpublished data). Data were standardized for life stage by using only juveniles for fish and decapods, immature aquatic life stages for amphibians and insects, and juveniles and spat for molluscs. All life stages were included for other taxa groups except egg or embryo stages. Specific aspects of the mussel toxicity dataset are detailed in Raimondo et al.[9](#_ENREF_9).

All chemicals in the ICE database were curated using the distributed structure-searchable toxicity database (DSSTox; www.epa.gov/ncct/dsstox/). A single name and the confirmed chemical abstract services registry number from the source material were checked against DSSTox to validate their consistency. Names that were not contained within the DSSTox list of synonyms for a particular chemical were manually checked to validate the agreement between the chemical identifiers and confirm the chemical-data linkage. The ICE database toxicity value was specified as the compound tested, except for some metal salts (recorded as the element or hardness normalized element), pentachlorophenol (pH normalized) and ammonia (pH and temperature normalized). Only records for chemicals with an active ingredient purity of ≥ 90% were accepted. Open ended (e.g., LC50 > 100 mg/L) or unconfirmed toxicity values were not used.

*ICE Model Development*

Models were developed as least squares log-linear regressions by pairing all possible surrogate species with all possible predicted taxa (i.e., a species, genus, or family) by common chemical. Each model required toxicity values from at least three different chemicals available for both the surrogate species and predicted taxon. For species level models, the geometric mean of toxicity values was used in both the surrogate and predicted species where multiple records occurred for the same species and chemical. Two sets of genus and family level models were developed: geometric mean and minimum toxicity models. Geometric mean models were developed using the geometric mean of the mean predicted species toxicity value for all species within that genus or family for each chemical. Minimum toxicity models were developed using the minimum value of toxicity data for all species within each predicted genus or family for each chemical. For both sets of genus and family level models, the predicted taxon toxicity value was paired with the geometric mean of the respective chemical for the surrogate species.

ICE models with four or more chemicals were validated using leave-one-out cross-validation[4](#_ENREF_4). In this process, each data point within a model (e.g., chemical pair of acute values for surrogate and predicted taxon) was systematically removed from the original model and a new model was rebuilt with the remaining data. The new model was used to estimate the toxicity value of the removed predicted species from the removed surrogate species toxicity value. For each removed data point, the N-fold difference was calculated as the greater value of the estimated/actual or actual/estimated for non-transformed values. For each model, the number of removed data points predicted within 5-fold of the actual value was determined and a cross-validation success rate was determined as the percentage of removed values predicted within 5-fold of the actual value [4](#_ENREF_4). The 5-fold range represents standard inter-laboratory variation of a toxicity value tested for the same species and chemical as demonstrated in previous studies[4](#_ENREF_4), [5](#_ENREF_5), [9-11](#_ENREF_9). A separate set of MOA-specific models was developed using only chemicals from a single MOA for each model. MOA-specific models have been shown to be more robust than models comprised of chemicals from multiple MOAs[4](#_ENREF_4).

*Model Selection Guidelines*

Model selection guidelines were developed to reduce the amount of professional judgement needed to select the best available model when multiple ICE models (i.e., multiple surrogate species) were available for predicting to one taxon. Selection guidance was determined from the combination of model attributes (mean square error; MSE, R2, and slope) that optimized the percent of cross-validated data points predicted within 5-fold of the measured value. By using models rebuilt in the cross-validation, we maintained independence between the predicted and measured values. Each model was assigned a taxonomic distance (TD) that identified the relatedness of the surrogate and predicted taxon (within genus = 1, within family = 2, etc). An iterative approach was used in which MSE, R2, and slope were randomly selected because these three parameters were autocorrelated and related to model robustness. We identified rebuilt models that contained the parameters within the randomly selected limits, and determined the percent of predicted data points from these models that were within 5-fold of the actual value. Parameters were randomly adjusted one at a time in increments of 0.05. This process continued until model parameters converged on an optimized percent of accurately predicted values. Optimization was achieved at the highest MSE, lowest R2 and lowest slope that corresponded to the highest percent accuracy and the point at which no additional data points were added. Each iteration was required to have a minimum of 50 data points was required to inform the optimization process. Prior to this process, confidence intervals were calculated for all cross-validated data points. Only predicted values with an upper 95th confidence limit less than 5-fold greater than the predicted value were used because large confidence intervals are often indicative of values that are outside the model training set. This was performed separately for each taxonomic level, and separately within taxonomic levels for models with N > 10 and N < 10. The optimized values for each model attribute were then applied as the model selection guidelines. This analysis was only performed on species level models because the objective of genus and family models was protectiveness of multiple species within the taxon rather than prediction accuracy of a single point.

*Evaluation of Web-ICE for Listed Species*

A case study approach was used to evaluate the optimized selection guidelines and to assess the protectiveness of genus and family models applied to listed species. The 13 chemicals chosen for the case study represented priority chemicals (primarily pesticides) encompassing a broad range of MOAs (Table 1). The 8 selected species represented a diverse taxonomic range of listed populations, including fish, amphibians, and molluscs from either the Sacramento California[12](#_ENREF_12) or Ohio River Valleys with toxicity data and models available in the ICE database (Table 1). First, all ICE models available for predicting to the genera and families of the listed case study species were identified using both sets of models (i.e., minimum and geometric mean models). Toxicity values for all possible surrogate species were selected from data available in the Web-ICE database to ensure data quality and standardization. The geometric mean of the surrogate species toxicity values was used as the model input for each chemical and surrogate species when multiple toxicity records were available. Toxicity estimates for each chemical and 95% confidence intervals (CI) were calculated for each of the available models. When predictions from multiple surrogates were available for a given taxon and chemical, we applied optimized model selection guidance to identify the best model (Fig. 1). Model selection was first based on the taxonomic distance of the surrogate and predicted taxon, because previous analyses indicated higher prediction accuracy for models from closely related taxa. Models were then chosen with a low MSE (≤ 0.95), and a narrow 95% CI. If models had similar MSE, the model with the greater N was selected.

Protectiveness of the toxicity estimate and the lower 95% CI predicted from the best model for each taxon and chemical was determined by comparing each ICE model prediction to the measured minimum toxicity value for that taxon available in the ICE database. For example, the model prediction to the family Salmonidae for malathion was compared to the minimum value (lowest) Salmonidae LC50 for malathion available in the ICE database. If the model prediction was greater than the minimum toxicity value in the database, then the minimum value was compared to the lower 95% CI of the predicted value. The number of model predictions or lower CI values that were protective of the most sensitive value within each taxon was compared for the genus and family level and within each prediction category. Each model prediction was then categorized as either protective, the lower 95% CI protective, or not protective for genus and family models developed from both minimum toxicity values and geometric means.

**Results**

*Database and model development*

 The expanded Web-ICE database (v. 3.3) contained a substantial increase in the number of toxicity records, species, chemicals, and models, including those predicting to listed species (Table 2). The developed ICE models were able to predict toxicity values for 25 listed species and 20 genera and 20 families containing U.S. federally listed species (Table 3). Overall, species model prediction accuracy and cross validation success were consistent with previous results for a smaller dataset [4](#_ENREF_4). Models developed with predicted and surrogate species within the same family predicted within 5- and 10-fold of the actual value for 92 and 98% of data points, respectively (Table 4). All model parameters are available on the EPA Web-ICE application webpage (www3.epa.gov/webice). MOA-specific models were developed for acetylcholinesterase inhibition, electron transport inhibition, iono/osmoregulatory/circulatory impairment, narcosis, neurotoxicity, and reactivity, and had similar performance as previous MOA-specific models [4](#_ENREF_4).

 *Model selection & prediction guidance*

Optimization analysis of model parameters based on model MSE, R2, and slope indicated that models with MSE < 0.95, R2 > 0.6, and slope > 0.6 will have the highest prediction accuracy (Table 5). For models with a taxonomic distance of 6 (same Kingdom) that have N > 10, models with MSE < 0.55 will have improved prediction accuracy because of the increased variation in toxicity data associated with less closely related taxa. Distributions of each model parameter by taxonomic distance are available in the Supporting Information (Fig. S1).

Figure 1 summarizes the procedures for model selection based on the optimized model parameters when models for multiple surrogates are available. Web-ICE users should first select models with the closest taxonomic distance and then apply the optimized model parameters to identify the best model. Optimization analysis also indicated selection of models with confidence intervals within 5-fold of the predicted value, although using 3-fold slightly improves prediction accuracy, but not substantially.

*Case study analysis of model protectiveness*

The case study of 13 chemicals and 8 listed species resulted in 599 potential minimum and 613 geometric mean ICE model predictions. Predictions were dependent on the availability of surrogate toxicity values and ICE models for predicted taxa. Applying model selection guidance resulted in 39 accepted ICE predictions for the listed species for each set of models (Supporting Information SI 002; Table S1). For both minimum and geometric mean models, 22 of the 39 predicted values were from family models and 17 were genus models. The Supporting Information (SI 002) provides all available minimum models predicting to these taxa with the best selected model identified. When model predictions were available for both a family and a genus within that same family (e.g., Oncorhynchus and Salmonidae), we selected the family model if both models had similar parameters (Table S2). Because family models are comprised of greater taxa diversity, estimates derived from these models can be considered to be protective of more species. In other cases genus models may generate more conservative estimates, but model selection should ultimately be guided by the goal of the particular analysis.

The predicted value for genus level minimum models was protective of the most sensitive species within the genus in 80% of cases and for 100% of instances when using the lower 95% CI (Fig. 2). For family level minimum models, predictions were protective in 77% of cases and the lower 95% CI was protective in 86% of instances (Fig. 2). Of the family level models where the toxicity of the most sensitive species was less than the lower 95% CI of the prediction (14%, *N* = 3 models), two of these models were at a taxonomic distance of greater than 5 because no models with a closer taxonomic distance were available. For models developed with geometric means, model predictions were protective of the most sensitive species in 29 and 32% cases for genus and family models, respectively (Fig. 2). The lower 95% CI was protective in 53% of genus models and 50% of family models built with geometric means.

**Discussion**

Species extrapolation has remained a significant challenge in ecotoxicology for over 40 years. Traditionally, safety factors have been used to account for uncertainty in extrapolating between species, with larger values used for protected taxa such as endangered species[13](#_ENREF_13). While potentially conservative, safety factors fail to account for taxon-specific variability in sensitivity that may occur with chemicals of different modes of action. For example, molluscs are consistently more sensitive to metals than other invertebrates[11](#_ENREF_11), [14](#_ENREF_14). ICE models between surrogates and listed taxa should provide risk assessors a protective evaluation of relative sensitivity. Because of the unquantifiable uncertainty associated with generic safety factors, the NRC recommended ICE models as an alternative in risk assessments of listed species[3](#_ENREF_3). We demonstrate that genus and family ICE models that are developed using the minimum toxicity value for species within the predicted taxon and which meet the revised selection guidelines are consistently protective of the range of species within that taxon.

Listed species are not necessarily more sensitive to toxicants than their non-listed congeners [11](#_ENREF_11). While contaminants pose a serious risk to vulnerable populations, many listed species are imperiled due to a combination of factors such as habitat degradation and loss and exotic species introductions. Listed species exposed to contaminants are at risk for larger population-level impacts than non-listed species because of smaller effective population size and the combination of multiple stressors rather than a higher inherent sensitivity. However, because of the increased vulnerability of listed populations, ensuring their chemical sensitivity is adequately represented is critical in the risk assessment of chemicals that may be applied within their ranges. ICE models provide a more accurate value for chemical sensitivity of a diversity of compounds and contain reduced, quantifiable uncertainty compared to generic safety factors.

The present study substantially expands the potential of ICE models to predict to listed species as well as provides guidelines for model selection and prediction evaluation that reduces reliance on professional judgement. Increased taxa and chemical diversity and the redevelopment of genus and family models improve model protectiveness for sensitive species and increased the number of available models for prediction to listed species. Based on analysis of model prediction accuracy and optimization of model selection guidance, ICE models continue to be a useful, robust tool for toxicity prediction to guide risk assessment of listed species and species with limited toxicity data.

A principle focus for updating Web-ICE was expansion of species diversity with a particular emphasis on increasing the number of models predicting to listed species, particularly freshwater mussels in the family Unionidae[9](#_ENREF_9). The previously available models contained only limited data for unionids, with only two ICE models that predicted toxicity to the family Unionidae. A focused effort in collaboration with other federal agencies was initiated to expand the freshwater mussel toxicity database that included developing mussel specific data standardization procedures, evaluating life stage sensitivity and optimal surrogates, and maximizing the number of validated models available for toxicity estimation in a diversity of mussel species[9](#_ENREF_9). The results of this research and development has been the substantial expansion and improvement of Web-ICE for listed species and other taxa with limited toxicity data.

Model validation and prediction success results from the present study were similar to previously developed models[4](#_ENREF_4), resulting in a larger model set with similar performance, indicating that underlying sensitivity relationships are preserved across data. Models with close taxonomic distance of the surrogate and predicted species had the highest prediction accuracy, comparable to results in Raimondo et al.[4](#_ENREF_4). The relationship of taxonomic distance and prediction accuracy is likely driven by similarities in toxicokinetics for more closely related taxa. Optimization of model selection guidance continues to recommend selecting a closely related surrogate species as a first step in model selection. In cases where models for surrogates within the same Order are not available, minimum genus and family models may still provide conservative toxicity estimates protective of the most sensitive species within the predicted taxon. Analysis of model prediction accuracy indicates that ICE models continue to be robust predictors of toxicity especially for closely related taxa.

We recommend genus and family models developed from the most sensitive species toxicity value within the predicted taxon, rather than the geometric mean of multiple species toxicity values, for conservative estimates to predict to listed species. These models provide toxicity estimates that are protective of a group of species that may have a wide range of sensitivities. Our case study of priority chemicals and listed species in the Sacramento and Ohio River Valleys indicated that when applying model selection guidelines, minimum models are considerably more protective of the most sensitive species within the predicted taxon compared to models based on geometric means. For the majority of the best selected models, the minimum model prediction or the lower CI of the prediction were protective of the most sensitive species in the predicted taxon whereas the surrogate species value was not. However, genus and family geometric mean models are still appropriate for generating mean toxicity estimates when conservative estimates of protectiveness may not be required such as in supplementing data in water quality criteria development.

ICE models have been previously demonstrated as robust tools for generating toxicity predictions for use in wildlife risk assessment and augmenting species diversity in SSDs[6](#_ENREF_6), [7](#_ENREF_7). The present study has improved and expanded their application for risk assessment of aquatic threatened and endangered species and demonstrated application of the user guidance to evaluate surrogate species and acute toxicity predictions. Genus and family models developed from minimum toxicity values are available in the endangered species module of Web-ICE v3.3, where models developed from geometric means are available on the genus and family modules of the website. The new user guidance should be applied to select models with high prediction accuracy, and minimum genus and family models provide toxicity estimates that are generally protective of the most sensitive species within the predicted taxon and can be used for toxicity prediction of listed species.

**Supporting Information**

SI 001: Model parameter output used for development of optimized model selection guidance; summary of available models used in the case study; summary of selected models for each listed species and chemical in the case study.

SI 002: Table of available minimum ICE genus and family models used in the case study of listed species.

**Acknowledgements**

This project was supported in part by an appointment to the Internship/Research Participation Program at the Office of Research and Development, U.S. Environmental Protection Agency, administered by the Oak Ridge Institute for Science and Education through in interagency agreement between the U.S. Department of Energy and EPA. The conclusions may not necessarily reflect the views of EPA and no official endorsement should be inferred.

**References**

1. Sappington, L. C.; Mayer, F. L.; Dwyer, F. J.; Buckler, D. R.; Jones, J. R.; Ellersieck, M. R., Contaminant sensitivity of threatened and endangered fishes compared to standard surrogate species. *Environmental Toxicology and Chemistry* **2001,** *20*, (12), 2869-2876.

2. Raimondo, S.; Vivian, D. N.; Delos, C.; Barron, M. G., Protectiveness of species sensitivity distribution hazard concentrations for acute toxicity used in endangered species risk assessment. *Environmental Toxicology and Chemistry* **2008,** *27*, (12), 2599-2607.

3. National Research Council *Assessing risks to endangered and threatened species from pesticides*; National Academies Press: Washington, DC, 2013; p 142.

4. Raimondo, S.; Jackson, C. R.; Barron, M. G., Influence of Taxonomic Relatedness and Chemical Mode of Action in Acute Interspecies Estimation Models for Aquatic Species. *Environmental Science & Technology* **2010,** *44*, (19), 7711-7716.

5. Raimondo, S.; Mineau, P.; Barron, M., Estimation of chemical toxicity to wildlife species using interspecies correlation models. *Environmental science & technology* **2007,** *41*, (16), 5888-5894.

6. Awkerman, J. A.; Raimondo, S.; Jackson, C. R.; Barron, M. G., Augmenting aquatic species sensitivity distributions with interspecies toxicity estimation models. *Environmental Toxicology and Chemistry* **2014,** *33*, (3), 688-695.

7. Awkerman, J. A.; Raimondo, S.; Barron, M. G., Development of Species Sensitivity Distributions for Wildlife using Interspecies Toxicity Correlation Models. *Environmental Science & Technology* **2008,** *42*, (9), 3447-3452.

8. Dyer, S. D.; Versteeg, D. J.; Belanger, S. E.; Chaney, J. G.; Raimondo, S.; Barron, M. G., Comparison of Species Sensitivity Distributions Derived from Interspecies Correlation Models to Distributions used to Derive Water Quality Criteria. *Environmental Science & Technology* **2008,** *42*, (8), 3076-3083.

9. Raimondo, S.; Lilavois, C. R.; Lee, L.; Augspurger, T.; Wang, N.; Ingersoll, C. G.; Bauer, C.; Hammer, E.; Barron, M. G., Assessing variability in chemical acute toxicity of unionid mussels: influence of intra-and inter-laboratory testing, life stage, and species. *Environmental Toxicology and Chemistry* **2016,** In print.

10. Fairbrother, A., Risk management safety factor. In *Encylopedia of Ecology*, Jorgensen, S. R.; Fath, B. D., Eds. Elsevier: Oxford, UK, 2008; Vol. 4.

11. Raimondo, S.; Vivian, D. N.; Barron, M. G., Standardizing acute toxicity data for use in ecotoxicology models: influence of test type, life stage, and concentration reporting. *Ecotoxicology* **2009,** *18*, (7), 918-928.

12. Hoogeweg, C. G.; Denton, D. L.; Breuer, R.; Williams, W. M.; TenBrook, P., Development of a Spatial-Temporal Co-occurrence Index To Evaluate Relative Pesticide Risks to Threatened and Endangered Species. In *Pesticide Regulation and the Endangered Species Act*, American Chemical Society: 2012; Vol. 1111, pp 303-323.

13. Dwyer, F. J.; Mayer, F. L.; Sappington, L. C.; Buckler, D. R.; Bridges, C. M.; Greer, I. E.; Hardesty, D. K.; Henke, C. E.; Ingersoll, C. G.; Kunz, J. L.; Whites, D. W.; Augspurger, T.; Mount, D. R.; Hattala, K.; Neuderfer, G. N., Assessing Contaminant Sensitivity of Endangered and Threatened Aquatic Species: Part I. Acute Toxicity of Five Chemicals. *Archives of Environmental Contamination and Toxicology* **2005,** *48*, (2), 143-154.

14. Wang, N.; Ingersoll, C. G.; Hardesty, D. K.; Ivey, C. D.; Kunz, J. L.; May, T. W.; Dwyer, F. J.; Roberts, A. D.; Augspurger, T.; Kane, C. M.; Neves, R. J.; Barnhart, M. C., Acute toxicity of copper, ammonia, and chlorine to glochidia and juveniles of freshwater mussels (unionidae). *Environmental Toxicology and Chemistry* **2007,** *26*, (10), 2036-2047.

Table 1. Selected priority chemicals, their mode of action (MOA) in aquatic animals, and listed species for the case study analysis.

|  |  |  |
| --- | --- | --- |
| Case study chemicals |   | Taxa containing listed species |
| Chemical  | MOA  |   | Region | Family | Species |
| Atrazine | Narcosis  |  | Sacramento Valley | Acipenseridae | *Acipenser* *medirostris* (Southern North American green sturgeon) |
| Carbaryl | AChE inhibition  |  |  | Branchinectidae | *Branchinecta lynchi* (Vernal pool fairy shrimp) |
| Chlorpyrifos | AChE inhibition |  |  | Ranidae | *Rana draytonii* (California Red-legged frog) |
| Copper | Iono/Osmoregulatory/Circulatory impairment |  |  | Salmonidae | *Oncorhynchus tshawytscha* (Chinook Salmon) |
| Cypermethrin | Neurotoxicity |  |  |  | *Oncorhynchus mykiss* (Central Valley Steelhead) |
| Diazinon | AChE inhibition |  | Ohio River Valley | Unionidae  | *Lampsilis abrupta* (Pink mucket) |
| Fipronil | Neurotoxicity |  |  |  | *Villosa fabilis* (Rayed bean mussel) |
| Glyphosate | Narcosis  |  |  |  | *Villosa trabalis* (Cumberland bean pearly mussel) |
| Imidacloprid | Neurotoxicity |  |  |  |  |
| Malathion | AChE inhibition |  |  |  |  |
| Methomyl | AChE inhibition |  |  |  |  |
| Permethrin | Neurotoxicity |  |  |  |  |
| Thiobencarb | Narcosis |   |   |   |   |

Table. 2. Web-ICE database attributes including the number of threatened and endangered (listed) species records and models available at each taxonomic level for each database version.

|  |  |  |
| --- | --- | --- |
| Web-ICE Version | Database Attributes | Number of models |
| Records | Species | Chemicals | Species | Genus | Family |
| v. 3.3 (2015) | 8203 | 314 | 1501 | 1544 | 854 | 887 |
| Listed species | 1591 | 32 | 492 | 379 | 428 | 547 |
|  |  |  |  |  |  |  |
| v. 3.2 (2013)[4](#_ENREF_4) | 5501 | 180 | 1266 | 780 | 289 | 374 |
| Listed species | 1272 | 21 | 449 | 230 | 168 | 267 |

Table 3. Threatened and endangered species model summary with the number of significant species, genus, or family ICE models available predicting to listed species (*N* surrogates) or taxa containing U.S. federally listed species (*N* species).

|  |
| --- |
| Model Level |
| Species | *N* surrogates | Genus | *N* species | *N* surrogates | Family | *N* species | *N* surrogates |
| *Acipenser brevirostrum* | 16 | Acipenser | 9 | 15 | Acipenseridae | 14 | 15 |
| *Branchinecta lynchi* | 14 | Anaxyrus | 4 | 14 | Asellidae | 1 | 38 |
| *Cyprinodon bovinus* | 13 | Branchinecta | 4 | 14 | Atherinopsidae | 1 | 11 |
| *Epioblasma capsaeformis* | 8 | Catostomus | 4 | 9 | Branchinectidae | 4 | 14 |
| *Erimonax monachus* | 15 | Cyprinodon | 7 | 21 | Bufonidae | 9 | 19 |
| *Etheostoma fonticola* | 9 | Etheostoma | 17 | 17 | Cambaridae | 3 | 4 |
| *Gasterosteus aculeatus williamsoni* | 7 | Gammarus | 4 | 35 | Catostomidae | 9 | 30 |
| *Gila elegans* | 14 | Ictalurus | 1 | 36 | Coenagrionidae | 5 | 1 |
| *Lampsilis rafinesqueana* | 4 | Lampsilis | 6 | 29 | Cyprinidae | 45 | 77 |
| *Notropis mekistocholas* | 17 | Lasmigona | 1 | 16 | Cyprinodontidae | 7 | 35 |
| *Oncorhynchus clarkii henshawi* | 37 | Lirceus | 1 | 2 | Gammaridae | 4 | 37 |
| *Oncorhynchus clarkii seleniris* | 37 | Lithobates | 2 | 22 | Ictaluridae | 8 | 38 |
| *Oncorhynchus clarkii stomias* | 37 | Menidia | 1 | 11 | Lymnaeidae | 2 | 20 |
| *Oncorhynchus gilae* | 16 | Notropis | 8 | 25 | Palaemonidae | 1 | 5 |
| *Oncorhynchus keta* | 1 | Oncorhynchus | 11 | 74 | Percidae | 23 | 30 |
| *Oncorhynchus kisutch* | 41 | Palaemonetes | 1 | 2 | Poeciliidae | 5 | 39 |
| *Oncorhynchus mykiss* | 77 | Salmo | 2 | 25 | Ranidae | 6 | 22 |
| *Oncorhynchus mykiss whitei* | 77 | Salvelinus | 1 | 34 | Salmonidae | 14 | 72 |
| *Oncorhynchus nerka* | 4 | Streptocephalus | 1 | 11 | Streptocephalidae | 1 | 11 |
| *Oncorhynchus tshawytscha* | 34 | Villosa | 3 | 16 | Unionidae | 87 | 29 |
| *Poeciliopsis occidentalis* | 3 |  |  |  |  |  |  |
| *Ptychocheilus lucius* | 12 |  |  |  |  |  |  |
| *Salmo salar* | 17 |  |  |  |  |  |  |
| *Salvelinus confluentus* | 2 |  |  |  |  |  |  |
| *Xyrauchen texanus* | 17 |  |  |  |  |  |  |

Table 4. Species model cross-validation prediction success of observations for each shared taxonomic level within each prediction category representing the fold difference in the model toxicity prediction to the measured data point.

|  |  |  |
| --- | --- | --- |
| Taxonomic level | Significant models (*N)* | Percentage within prediction category |
| 5-fold | 10-fold | 50-fold | >50-fold |
| genus | 444 | 95 | 4 | 1 | 0 |
| family | 1144 | 92 | 6 | 2 | 0 |
| order | 430 | 87 | 11 | 3 | 0 |
| class | 5734 | 77 | 10 | 10 | 3 |
| phylum | 1658 | 62 | 14 | 16 | 8 |
| kingdom | 8006 | 55 | 15 | 19 | 11 |

Table 5. Percent of cross validated data points within 5-fold of the measured toxicity value for models meeting guidelines.

|  |  |  |  |
| --- | --- | --- | --- |
| Taxonomic distance | All Models | Models *N* < 10 | Models *N* ≥10 |
| *N* | Percent | *N* | Percent | *N* | Percent |
| 1 | 417 | 94 | 79 | 95 | 338 | 94 |
| 2 | 1052 | 93 | 325 | 92 | 727 | 94 |
| 3 | 343 | 89 | 257 | 89 | 86 | 88 |
| 4 | 4376 | 80 | 560 | 80 | 3816 | 80 |
| 5 | 839 | 66 | 232 | 80 | 607 | 60 |
| 6 | 1389 | 65 | 479 | 77 | 136a | 72a |

a Mean Square Error < 0.55

TOC/Abstract Art

Figure 1

Fig. 1. ICE model selection guidance based on optimized model parameters used to identify the best model when predictions from multiple surrogates are available.

Model Selection Guidance:

1. Identify surrogates with closest taxonomic distance (highest accuracy within TD 1-3, same Order).
2. Choose model with a low MSE (≤ 0.95 for N > 10 and < 0.55 for N ≥ 10).
3. Select model with narrow 95% CI (within 5-fold of prediction).
4. If models with similar MSE’s, choose model with greater N.

Figure 2



Fig. 2. Comparison of model protectiveness for geometric mean and minimum genus and family models used with priority chemicals and taxa containing listed species in the Sacramento and Ohio River Valleys. There were 17 genus model and 22 family model predictions for both geometric mean and minimum models.