**Application of Web-ICE to assess risks of national pesticide registrations to federally listed (threatened and endangered) species**

The National Academy of Science (NAS) recently recommended exploration of predictive tools, such as Interspecies Correlation Estimation (ICE), to estimate acute toxicity values for listed species and to support development of Species Sensitivity Distributions (SSDs). EPA explored the ability of the Web-based Interspecies Correlation Estimation (Web-ICE) tool to predict acute toxicity values for federally-listed endangered and threatened (listed) aquatic species, based on toxicity data for test species that are typically available due to FIFRA data requirements: the rainbow trout (*Oncorhynchus mykiss*), bluegill sunfish (*Lepomis macrochirus*), and fathead minnow (*Pimephales promelas*). Three organophosphates, chlorpyrifos, diazinon and malathion, were selected as test chemicals for this evaluation because they were identified as pilot chemicals for implementation of the NAS recommendations.

This work involved two separate analyses:

1. Direct estimation of toxicity values to taxa (species, genus, family) representing listed species. The endangered species module of Web-ICE uses toxicity data from one or more surrogate species to predict multiple toxicity values to represent listed fish species using all available species, genus or family level models.
2. Use of estimated values to increase the size of a toxicity database to allow for generation of a species sensitivity distribution (SSD). An SSD is a cumulative distribution function of the variation in sensitivity of different species exposed to a stressor (*i.e.,* a registered pesticide). From those SSDs, which are based on acute exposures resulting in mortality, the 5th percentile (HC05; HC = hazardous concentration) is estimated and used to calculate the threshold concentration representing a chance of one in a million of mortality to an individual. As part of the NAS recommendations, EPA is evaluating different methods for deriving SSDs for use in Steps 1 and 2 of the listed species effects determinations for the pilot chemicals, malathion, diazinon, and chlorpyrifos.  The SSD module of Web-ICE uses toxicity data from one or more surrogate species to simultaneously predict toxicity values to all species with existing Web-ICE models. This species-level toxicity data is then used to develop a cumulative probability distribution of toxicity data and generate a prescribed hazard concentration (*e.g.,* HC05). As part of this evaluation, EPA considered the appropriateness of using tools such as Web-ICE to develop SSDs, in cases where empirical datasets for particular taxa are limited, by comparing empirically derived SSDs for chemicals with large datasets (*e.g.,* chlorpyrifos, diazinon, and malathion) to SSDs derived from predicted datasets.

In addition, the Web-ICE “Rules of Thumb” from Raimondo *et al*. (2010) were re-evaluated. These Rules of Thumb are intended to be guidelines regarding the application of best professional judgment to help identify robust ICE models and predictions of acute toxicity. The new guidelines for model selection, presented here and elsewhere (Willming *et al*. in review) reduces the reliance on best professional judgement in model selection.

**Rules of Thumb**

The original Rules of Thumb were derived from trends of all ICE models and were intended to be used as approximations based on results that contained considerable variability. While they were not intended to be used as definitive limits, several questions have recently arisen as to what definitive limits could be applied that would help standardize the way ICE models are evaluated and predictions accepted for use. These rules were based on results from a leave-one-out cross-validation of all ICE models, in which each pair of acute toxicity values for surrogate and predicted taxa were systematically removed from the original model, while the remaining data were used to estimate the toxicity of the removed predicted taxa (Raimondo *et al.* 2010). The percent cross-validation success rate was calculated as the percentage of predicted values that were within 5-fold of their measured value. Throughout the recent applications of Web-ICE, however, we have demonstrated that the cross-validation approach under-represents the robustness of models with small sample sizes (N< 10). As a result, the original Rules were re-evaluated to identify new guidelines that identify definitive upper and lower bounds of model parameters that, when all are applied, result in a robust data set of predicted toxicity values.

Re-evaluation of model selection criteria are described in detail in Willming *et al.* (in review). In brief, model selection criteria were 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. Each ICE model is identified by a taxonomic distance (TD) that describes 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 related to model robustness. We identified rebuilt (during model cross-validation) 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. Since parameters are autocorrelated they were randomly adjusted over 100 iterations. 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. A comparison of the Rules is presented in **Table 1**.

**Table 1. Comparison of Rules of Thumb.**

|  |  |
| --- | --- |
| **Old “Rule of Thumb”**  **(Raimondo *et al.* 2010)** | **New ICE Guidelines (Willming *et al*. in review)** |
|  | **Suggested Limits** |
| 1. Relatively low mean square error (MSE; < 0.22) | 1. Mean Square Error < 0.951 |
| 2. High R2 value (> 0.6) | 2. High R2 value (> 0.6) |
| 3. Narrow confidence intervals | 3. Slope > 0.6 |
| 4. Low p-values (< 0.01) | 4. Narrow confidence intervals (upper 95th within 5-fold of predicted) |
| **Weight of Evidence** |
| 5. Close taxonomic distance (< 3) | 1. Close taxonomic distance (< 4) |
| 6. High cross-validation success rate (> 85%) | 2. Evaluate cross-validation success rate relative to N |
| 7. High degrees of freedom  (df > 8, N > 10) | 3. Evaluate degrees of freedom relative to confidence intervals |

1 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.

**Direct estimation of toxicity values representing listed species**

EPA evaluated the ability of Web-ICE to predict toxicity values for each listed species of fish using data for standard test species (*i.e.,* rainbow trout, bluegill sunfish, and fathead minnow). The ability of the model to do so was explored by first attempting to estimate species specific toxicity values. If a species level prediction was not available, genus level or family level toxicity values were evaluated based on the availability of models.

An initial analysis of the species to species models indicates that, using the standard test species, Web-ICE is able to predict direct toxicity of a chemical to 17 listed species (**Table 2**). Listed species not represented in **Table 2** would rely on the availability of genus and family level models for toxicity predictions.

**Table 2: Threatened and Endangered Species and Their Surrogates**

|  |  |  |  |
| --- | --- | --- | --- |
| **Threatened and Endangered Species** | **Bluegill** | **Fathead Minnow** | **Rainbow Trout** |
| **Apache trout (*Oncorhynchus gilae*)** |  | X | X |
| **Atlantic salmon (*Salmo salar*)** | X | X | X |
| **Bonytail chub (*Gila elegans*)** |  | X |  |
| **Cape fear shiner (*Notropis mekistocholas*)** |  | X | X |
| **Chinook salmon (*Oncorhynchus tshawytscha*)** | X |  | X |
| **Coho salmon (*Oncorhynchus kisutch*)** | X |  | X |
| **Colorado squawfish (*Ptychocheilus lucius*)** |  | X |  |
| **Cutthroat trout (*Oncorhynchus clarkii*)** | X | X | X |
| **Fountain darter (*Etheostoma fonticola*)** |  | X | X |
| **Gila topminnow (*Poeciliopsis occidentalis*)** |  |  | X |
| **Leon springs pupfish (*Cyprinodon bovinus*)** |  | X | X |
| **Rainbow trout (*Oncorhynchus mykiss*)** | X | X |  |
| **Razorback sucker (*Xyrauchen texanus*)** |  | X | X |
| **Shortnose sturgeon (*Acipenser brevirostrum*)** |  | X | X |
| **Spotfin chub (*Erimonax monachus*)** |  | X | X |
| **Threespine stickleback (*Gasterosteus aculeatus*)** | X | X | X |

As a case study, Web-ICE was evaluated for its ability to predict toxicity to eight listed fish species, including the Alabama Sturgeon, Slender Chub, and five species in the order Salmonidae (Apache Trout, Atlantic Salmon, Chinook Salmon, Coho Salmon, and Cutthroat Trout). Toxicity estimates were selected based on the Rules of Thumb described previously.

Alabama Sturgeon (*Scaphirhynchus suttkusi*)

For all three chemicals, the toxicity of the Alabama Sturgeon could be estimated using the family-level model for Acipenseridae. For both malathion and diazinon, toxicity could only be predicted using data for the rainbow trout, where chlorpyrifos could be predicted using both rainbow trout and fathead minnow data. The model for the fathead minnow had the lowest MSE, highest R2, slope closest to one, and narrowest confidence intervals for chlorpyrifos. Summary results for each of the models are presented in **Table 3**.

**Table 3. Web-ICE Model Predictions for the Alabama Sturgeon1**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Chemical** | **Species** | **Predicted Taxa** | **Model Level** | **Surrogate** | **Estimated Toxicity** | **Lower 95% Confidence Intervals** | **Upper 95% Confidence Intervals** |
| ***Malathion*** | ***Alabama sturgeon*** | ***Acipenseridae*** | ***family*** | ***Rainbow trout*** | ***49.13*** | ***21*** | ***114.94*** |
| ***Diazinon*** | ***Alabama sturgeon*** | ***Acipenseridae*** | ***family*** | ***Rainbow trout*** | ***140.87*** | ***59.43*** | ***333.92*** |
| ***Chlorpyrifos*** | ***Alabama sturgeon*** | ***Acipenseridae*** | ***family*** | ***Fathead minnow*** | ***55.46*** | ***14.66*** | ***209.8*** |
| Chlorpyrifos | Alabama sturgeon | Acipenseridae | family | Rainbow trout | 4.31 | 0.828 | 22.51 |

1Models in bold italics were selected as the most appropriate model based on the rules of thumb.

Slender Chub (*Erimystax cahni*)

For all three chemicals, the toxicity of the slender chub could be predicted by all three surrogates using the family level model Cyprinidae. The fathead minnow was the surrogate species with the closest taxonomic distance to the slender chub, and consistent with previous findings, its model for this taxa had the lowest MSE, highest R2, slope closest to one, and narrowest confidence intervals for all three chemicals. Summary results for each of the models are presented in **Table 4**.

**Table 4. Web-ICE Model Predictions for the Slender Chub1**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Chemical** | **Species** | **Predicted Taxa** | **Model Level** | **Surrogate** | **Estimated Toxicity** | **Lower 95% Confidence Intervals** | **Upper 95% Confidence Intervals** |
| Malathion | Slender Chub | Cyprinidae | family | Bluegill | 164.24 | 112.77 | 239.21 |
| ***Malathion*** | ***Slender Chub*** | ***Cyprinidae*** | ***family*** | ***Fathead minnow*** | ***9320.96*** | ***7244.3*** | ***11992.91*** |
| Malathion | Slender Chub | Cyprinidae | family | Rainbow trout | 227.52 | 164.43 | 314.83 |
| Diazinon | Slender Chub | Cyprinidae | family | Bluegill | 508.65 | 371.49 | 696.45 |
| ***Diazinon*** | ***Slender Chub*** | ***Cyprinidae*** | ***family*** | ***Fathead minnow*** | ***4641.18*** | ***3644.51*** | ***5910.39*** |
| Diazinon | Slender Chub | Cyprinidae | family | Rainbow trout | 518.58 | 389.39 | 690.63 |
| Chlorpyrifos | Slender Chub | Cyprinidae | family | Bluegill | 11.59 | 6.44 | 20.88 |
| ***Chlorpyrifos*** | ***Slender Chub*** | ***Cyprinidae*** | ***family*** | ***Fathead minnow*** | ***164.6*** | ***125.82*** | ***215.32*** |
| Chlorpyrifos | Slender Chub | Cyprinidae | family | Rainbow trout | 33.97 | 21.7 | 53.17 |

1Models in bold italics were selected as the most appropriate model based on the rules of thumb.

Salmonids

For all chemicals, there were five Salmonid species (excluding rainbow trout) that could be directly predicted from the three surrogates: Apache trout (*Oncorhynchus gilae*), Atlantic salmon (*Salmo salar*), Chinook salmon (*Oncorhynchus tshawytscha*), Coho salmon (*Oncorhynchus kisutch*), and Cutthroat trout (*Oncorhynchus clarkii*). In general, the rainbow trout model was selected as the best predictor of toxicity at the species, genus and family level for the five Salmonid species because it has the tightest confidence intervals, smallest MSE, highest R2, and is the most taxonomically related. At the genus level, all three surrogate species could predict to the genera *Oncorhynchus*, *Salmo*, and *Salvelinus*, which represent 11, 2, and 1 endangered species, respectively, including those represented by the species-level models above. All three species could also predict to the family Salmonidae, which represented the same 14 species represented by the species and genus-level models. A summary of the species, genus and family level models for each chemical are presented in **Tables 5 - 7**.

For chlorpyrifos, the family level prediction using rainbow trout was 11.51 ug/L. The lowest value predicted using other taxonomic levels was 4.18 ug/L (Apache trout predicted using rainbow trout). The range of toxicity values predicted for salmonids using all three taxonomic level models was 4.18-13.87 ug/L. Given the close range of all of these values, 4.18 ug/L could be selected as the conservative estimate for chlorpyrifos.

Similarly, for diazinon, the range of toxicity values predicted by rainbow trout using all three taxonomic level models was 141.19-272.83 ug/L, with the most sensitive for Apache trout. The family-level model, which represents all of these species, was in the middle of this range at 215.53. As with chlorpyrifos, 141.19 ug/L may be recommended for represent endangered salmonids.

For malathion, predicted values ranged from 48.74-110.87 ug/L with the family level prediction at 88.91 ug/L and Apache trout as the most sensitive. The value of 48.74 may be used as a conservative estimate for salmonids.

**Table 5. Chlorpyrifos Web-ICE Model Predictions for the Order Salmonidae1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Predicted Taxa** | **Model Level** | **Surrogate** | **Estimated Toxicity** | **Lower 95% Confidence Intervals** | **Upper 95% Confidence Intervals** |
| Apache trout (Oncorhynchus gilae) | species | Fathead minnow | 56.76 | 19.13 | 168.35 |
| ***Apache trout (Oncorhynchus gilae)*** | ***species*** | ***Rainbow trout*** | ***4.18*** | ***2.77*** | ***6.3*** |
| Atlantic salmon (Salmo salar) | species | Bluegill | 0.786 | 0.274 | 2.25 |
| Atlantic salmon (Salmo salar) | species | Fathead minnow | 58.5 | 9.84 | 347.73 |
| ***Atlantic salmon (Salmo salar)*** | ***species*** | ***Rainbow trout*** | ***7.64*** | ***3.71*** | ***15.72*** |
| ***Chinook salmon (Oncorhynchus tshawytscha)*** | ***species*** | ***Rainbow trout*** | ***13.87*** | ***6.68*** | ***28.77*** |
| Coho salmon (Oncorhynchus kisutch) | species | Bluegill | 2.65 | 0.912 | 7.73 |
| Coho salmon (Oncorhynchus kisutch) | species | Fathead minnow | 81.4 | 25.07 | 264.32 |
| ***Coho salmon (Oncorhynchus kisutch)*** | ***species*** | ***Rainbow trout*** | ***9.98*** | ***7.75*** | ***12.86*** |
| Cutthroat trout (Oncorhynchus clarkii) | species | Bluegill | 9.99 | 3.44 | 28.97 |
| Cutthroat trout (Oncorhynchus clarkii) | species | Fathead minnow | 86.8 | 39.79 | 189.37 |
| ***Cutthroat trout (Oncorhynchus clarkii)*** | ***species*** | ***Rainbow trout*** | ***10.44*** | ***6.85*** | ***15.92*** |
| Oncorhynchus | genus | Bluegill | 3.5 | 2.71 | 4.51 |
| Oncorhynchus | genus | Fathead minnow | 113.6 | 78.65 | 164.08 |
| ***Oncorhynchus*** | ***genus*** | ***Rainbow trout*** | ***10.99*** | ***8.29*** | ***14.59*** |
| Salmo | genus | Bluegill | 1.19 | 0.597 | 2.38 |
| Salmo | genus | Fathead minnow | 49.23 | 10.26 | 236.23 |
| ***Salmo*** | ***genus*** | ***Rainbow trout*** | ***8.68*** | ***5.94*** | ***12.67*** |
| Salvelinus | genus | Bluegill | 3.67 | 1.41 | 9.54 |
| Salvelinus | genus | Fathead minnow | 68.42 | 30.88 | 151.61 |
| ***Salvelinus*** | ***genus*** | ***Rainbow trout*** | ***12.12*** | ***7.57*** | ***19.42*** |
| Salmonidae | family | Bluegill | 3.48 | 2.71 | 4.47 |
| Salmonidae | family | Fathead minnow | 108.57 | 77.88 | 151.36 |
| ***Salmonidae*** | ***family*** | ***Rainbow trout*** | ***11.51*** | ***8.91*** | ***14.86*** |

1Models in bold italics were selected as the most appropriate model based on the rules of thumb.

**Table 6. Diazinon Web-ICE Model Predictions for the Order Salmonidae1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Predicted Taxa** | **Model Level** | **Surrogate** | **Estimated Toxicity** | **Lower 95% Confidence Intervals** | **Upper 95% Confidence Intervals** |
| ***Apache trout (Oncorhynchus gilae)*** | ***species*** | ***Rainbow trout*** | ***141.19*** | ***103.99*** | ***191.7*** |
| Atlantic salmon (Salmo salar) | species | Bluegill | 117.82 | 70.35 | 197.32 |
| ***Atlantic salmon (Salmo salar)*** | ***species*** | ***Rainbow trout*** | ***168.78*** | ***101.18*** | ***281.54*** |
| ***Chinook salmon (Oncorhynchus tshawytscha)*** | ***species*** | ***Rainbow trout*** | ***272.83*** | ***163.16*** | ***456.2*** |
| Coho salmon (Oncorhynchus kisutch) | species | Bluegill | 218.55 | 125.2 | 381.48 |
| Coho salmon (Oncorhynchus kisutch) | species | Fathead minnow | 1025.27 | 278.12 | 3779.61 |
| ***Coho salmon (Oncorhynchus kisutch)*** | ***species*** | ***Rainbow trout*** | ***267.61*** | ***223.57*** | ***320.32*** |
| Cutthroat trout (Oncorhynchus clarkii) | species | Bluegill | 271.28 | 161.9 | 454.57 |
| Cutthroat trout (Oncorhynchus clarkii) | species | Fathead minnow | 1504.85 | 683.29 | 3314.22 |
| ***Cutthroat trout (Oncorhynchus clarkii)*** | ***species*** | ***Rainbow trout*** | ***211.04*** | ***166.64*** | ***267.28*** |
| Oncorhynchus | genus | Bluegill | 227.03 | 198.9 | 259.14 |
| Oncorhynchus | genus | Fathead minnow | 2373.39 | 1745.76 | 3226.65 |
| ***Oncorhynchus*** | ***genus*** | ***Rainbow trout*** | ***224*** | ***187.07*** | ***268.22*** |
| Salmo | genus | Bluegill | 142.31 | 99.67 | 203.18 |
| Salmo | genus | Fathead minnow | 766.03 | 192.44 | 3049.18 |
| ***Salmo*** | ***genus*** | ***Rainbow trout*** | ***186.23*** | ***141.86*** | ***244.49*** |
| Salvelinus | genus | Bluegill | 157.7 | 101.38 | 245.29 |
| Salvelinus | genus | Fathead minnow | 954.21 | 512.07 | 1778.1 |
| ***Salvelinus*** | ***genus*** | ***Rainbow trout*** | ***198.59*** | ***147.87*** | ***266.7*** |
| Salmonidae | family | Bluegill | 224.95 | 197.34 | 256.42 |
| Salmonidae | family | Fathead minnow | 2324.33 | 1762.01 | 3066.11 |
| ***Salmonidae*** | ***family*** | ***Rainbow trout*** | ***215.53*** | ***183.27*** | ***253.48*** |

1Models in bold italics were selected as the most appropriate model based on the rules of thumb.

**Table 7. Malathion Web-ICE Model Predictions for the Order Salmonidae1**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Predicted Taxa** | **Model Level** | **Surrogate** | **Estimated Toxicity** | **Lower 95% Confidence Intervals** | **Upper 95% Confidence Intervals** |
| ***Apache trout (Oncorhynchus gilae)*** | ***species*** | ***Rainbow trout*** | ***48.74*** | ***37.05*** | ***64.11*** |
| Atlantic salmon (Salmo salar) | species | Bluegill | 26.34 | 13.92 | 49.83 |
| Atlantic salmon (Salmo salar) | species | Fathead minnow | 5574.87 | 719.84 | 43174.63 |
| ***Atlantic salmon (Salmo salar)*** | ***species*** | ***Rainbow trout*** | ***66.24*** | ***38.41*** | ***114.24*** |
| ***Chinook salmon (Oncorhynchus tshawytscha)*** | ***species*** | ***Rainbow trout*** | ***110.87*** | ***63.78*** | ***192.75*** |
| Coho salmon (Oncorhynchus kisutch) | species | Bluegill | 58.47 | 31.37 | 108.97 |
| Coho salmon (Oncorhynchus kisutch) | species | Fathead minnow | 1740.12 | 401.53 | 7541.11 |
| ***Coho salmon (Oncorhynchus kisutch)*** | ***species*** | ***Rainbow trout*** | ***99.05*** | ***82.37*** | ***119.1*** |
| Cutthroat trout (Oncorhynchus clarkii) | species | Bluegill | 101.11 | 53.96 | 189.47 |
| Cutthroat trout (Oncorhynchus clarkii) | species | Fathead minnow | 2730.31 | 1115.7 | 6681.5 |
| ***Cutthroat trout (Oncorhynchus clarkii)*** | ***species*** | ***Rainbow trout*** | ***85.07*** | ***64.46*** | ***112.28*** |
| Oncorhynchus | genus | Bluegill | 65.23 | 55.59 | 76.55 |
| Oncorhynchus | genus | Fathead minnow | 4477.2 | 3224.13 | 6217.27 |
| ***Oncorhynchus*** | ***genus*** | ***Rainbow trout*** | ***90.07*** | ***73.84*** | ***109.88*** |
| Salmo | genus | Bluegill | 34.07 | 22.51 | 51.57 |
| Salmo | genus | Fathead minnow | 1358.81 | 287.83 | 6414.69 |
| ***Salmo*** | ***genus*** | ***Rainbow trout*** | ***73.72*** | ***55.53*** | ***97.86*** |
| Salvelinus | genus | Bluegill | 51.23 | 29.39 | 89.28 |
| Salvelinus | genus | Fathead minnow | 1654.33 | 825.25 | 3316.33 |
| ***Salvelinus*** | ***genus*** | ***Rainbow trout*** | ***85.3*** | ***61.63*** | ***118.06*** |
| Salmonidae | family | Bluegill | 64.7 | 55.22 | 75.8 |
| Salmonidae | family | Fathead minnow | 4407.05 | 3276.34 | 5927.97 |
| ***Salmonidae*** | ***family*** | ***Rainbow trout*** | ***88.91*** | ***74.35*** | ***106.31*** |

1Models in bold italics were selected as the most appropriate model based on the rules of thumb.

The protectiveness of using genus and family level models to represent listed species within those higher taxa was recently explored in Willming et al (in review, **SUPPLEMENTAL INFORMATION 1**). In this study, the range of sensitivity of species within genera and families was evaluated to determine how well models developed from geometric means of these higher taxonomic levels, such as those used here, protected the most sensitive species within the taxon. Genus and family level ICE models were developed from either geometric mean or minimum taxa toxicity values and used to predict toxicity to federally listed species 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. The Web-ICE application (www3.epa.gov/webice) is being updated to include genus and family level models developed from minimal data to provide more protective estimates of sensitivity for listed species.

**Use of Web-ICE to derive fish SSDs for limited datasets**

To fully explore the potential utility of Web-ICE derived SSDs for chemicals with limited datasets, EPA tested five distributions (logistic, burr, normal, triangular, and gumbel) against predicted values derived using empirical data for standard test species. SSDs were developed using predicted species values derived from rainbow trout data, bluegill data, fathead minnow surrogate data as well as predicted values derived from the most robust estimate using all three surrogate species. SSDs were evaluated based on how well the predicted data fit the distribution; how the variance of the HC05 compared to the variance of the fitted SSD; how well the variance around the predicted values represent the variance from the empirical dataset; and how close the predicted parameters (HC05, HC50, HC95) were to the empirical parameters. Results of this comparison for malathion, chlorpyrifos, and diazinon are summarized below and detailed in **SUPPLEMENTAL INFORMATION 2-4**.

For malathion, predicted values from rainbow trout, bluegill, and the mixture of all three surrogate species produced estimates of the HC05 that were lower than, but comparable to the empirical data. However, predicted values from fathead minnow produced estimates of the HC05 that were almost two orders of magnitude greater than the empirical value. HC05 estimates from the predicted data appeared to be more precisely estimated, with smaller coefficients of variation and smaller estimates of sampling variance in the HC05. However, plotted SSDs for the mixed data set show severe lack of fit. For each fitted distribution, the upper confidence limit of the HC05 typically corresponded with the 10th to the 25th quantile of the distribution. Values above the 20th percentile may be indicative of substantial uncertainty in the HC05 relative to the variance of the distribution. Predicted value SSDs had smaller variances than the SSD generated from empirical data. This means that when the surrogate species is among the more sensitive species, as is the case with the rainbow trout and bluegill sunfish, the estimated HC05 from the predicted dataset is closer to the HC05 estimated from the empirical dataset. However, if the surrogate species is among the more tolerant species (*i.e.,* the fathead minnow) then the HC05 values estimated from the predicted and empirical datasets are drastically different. Based on this analysis, SSDs derived from predicted datasets should be used with some caution.

In addition, cross-tabulation of predicted values with empirical test results showed that Web-ICE predictions were significantly positively correlated with empirically measured toxicity values (correlation coefficients range from 0.72 – 0.80). However, the discrepancy between predicted and measured toxicity values increased with dissimilarity between the surrogate species used to derive the predicted values and the predicted species for all three surrogates. For species with dissimilar tolerances, the discrepancies between measured and predicted toxicity approached 2 orders of magnitude. This effect may be related to phylogeny. Thus, as previously demonstrated, predicted values get worse with taxonomic distance, as the dissimilarity in tolerance grows between two taxa.

For chlorpyrifos and diazinon, predicted values based on data for rainbow trout, bluegill sunfish, and the mixture of all three surrogates produced estimates of the HC05 that were lower than that generated from the empirical data. However, predicted values based on data for fathead minnow produced estimates of the HC05 that were an order of magnitude greater than the empirical data. In general, HC05 estimates from the predicted data appeared to be more precisely estimated, with smaller coefficients of variation and smaller estimates of sampling variance in the HC05. While there is some evidence of a lack of fit in the predicted SSDs, this was not supported by the goodness-of-fit tests. Predicted values generated using a single surrogate generally produced an SSD with less variance than that produced using the empirical data. The variance of the distribution fit using all predicted values was closer to that of the empirical test results, suggesting an SSD based on a mixture of all three surrogates would be appropriate for these two chemicals. Cross-tabulation of predicted values with empirical test results show that the values were not significantly correlated, though this may be due to lack of power. In contrast to the results for malathion, no relationship between toxicological distance and accuracy of the predicted values was found, though again this could be due to lack of power.