**ATTACHMENT 1-5. Method For Deriving Species Sensitivity Distributions For Use In Pesticide Effects Determinations for Listed Species**

1. **Introduction**

The objective of this appendix is to describe methods for deriving species sensitivity distributions (SSDs) for use in listed species effects determinations for chlorpyrifos, diazinon and malathion. A SSD is a model of the variation in sensitivity of different species exposed to a stressor (*i.e.,* either chlorpyrifos, diazinon, or malathion). SSDs are generated by fitting a statistical or empirical distribution function to the proportion of species affected as a function of stressor concentration or dose. Threshold concentrations are estimated using the 5th percentile (HC05; HC = hazardous concentration) of SSDs constructed using mortality data from acute exposures. For direct effects, the HC05 is used to estimate the threshold representing a chance of one in a million of mortality to an individual (or the 1/10,000th percentile of the fitted dose-response at the HC05). For indirect effects, a threshold representing a 10% reduction in individuals (*e.g.,* prey of a listed species) is estimated.

The methods described in this paper include four steps for deriving SSDs and their HC05 values: 1) standardization of data for inclusion in an SSD; 2) fitting an SSD; 3) estimating the HC05; and 4) estimating threshold concentrations for direct and indirect effects. This appendix includes the methods used to derive SSDs chlorpyrifos, diazinon, and methomyl. To illustrate these methods, data pertaining to malathion are used[[1]](#footnote-1). These methods may be refined after these analyses are completed based on lessons learned. The last section of this appendix describes the SSD toolbox, which was developed in order to efficiently derive SSDs and thresholds using a consistent methodology.

1. **Standardization of data for inclusion in SSD**

Because each SSD depicts relative sensitivities of different species exposed to the same stressor, it is necessary to standardize the data as much as possible to eliminate variables that would confound the relative sensitivities of species. Such variables can include study exposure durations, age class of organisms tested, saltwater versus freshwater media, and other study design factors. For chlorpyrifos, diazinon and malathion, taxon-specific SSDs are derived when possible. For these chemicals, sufficient data are available to derive SSDs for birds, aquatic invertebrates and fish (plus aquatic-phase amphibians). For each chemical data set, analyses were conducted to determine whether aquatic taxa should be separated into freshwater and saltwater species.

Insufficient data are available to derive SSDs for the following taxa: mammals, amphibians (alone), reptiles and plants. Although many terrestrial invertebrate species have been tested using these three chemicals, SSDs were not derived due to prevalence of target species in the database. Terrestrial invertebrate SSDs potentially include resistant populations that may skew the distribution to be less conservative and less representative of sensitive, non-resistant species.

The following criteria were applied when considering which registrant-submitted and open literature data that were used to derive a taxon-specific SSD:

1. The chemical exposure for each study is technical grade active ingredient (TGAI[[2]](#footnote-2))[[3]](#footnote-3) of the parent.
2. For birds:
   1. Insufficient concentration based endpoints (*i.e*., LC50s) are available from different species to derive dietary-based SSDs. Therefore, median lethal doses (LD50s) are used to derive SSDs.
   2. For dose-based exposures,
      1. The endpoint is the median lethal dose (LD50) from an acute oral toxicity study.
      2. The duration is consistent with standard toxicity studies (*i.e.,* single acute dose followed by 7-14 d observation period).
      3. The endpoint is expressed in units of mg a.i./kg-bw and is normalized to represent birds with a body weight of 100 g. LD50 values are normalized using the following equation:

The LD50 value on the right side of this equation is the endpoint reported from the study (units expressed as mg a.i./kg-bw).

TW represents the body weight (in g) of the species tested. Generally, acute oral tests involve adult animals. If body weight data are not available in the study report, the literature can be cited for species specific body weights. Default body weights for the bobwhite quail and mallard duck are 178 and 1580 g, respectively. Body weights for additional bird species can be found in Dunning 1984.

The Mineau scaling scaling factor (x) is used to adjust bird body weights. When chemical specific values are available in Mineau 1996, they should be used (Chlorpyrifos = 1.1573, Diazinon = 0.6284). If not, the default value of 1.15 should be used (*i.e.,* for malathion).

1. For fish and aquatic-phase amphibians:
   1. The endpoint is the median lethal concentration (LC50).
   2. Endpoints will be expressed in micrograms active ingredient per liter (µg a.i./L).
   3. The duration is consistent with standard toxicity studies (*i.e.,* 96h).
2. For aquatic invertebrates:
   1. The endpoint is the median lethal concentration (LC50) or effects concentration (EC50) for immobility.
   2. Endpoints are expressed in micrograms active ingredient per liter (µg a.i./L).
   3. The duration is consistent with standard toxicity studies (*i.e.,* 48h for cladocerans and 96h for all other species).

As part of this analysis, the teams considered whether or not it is relevant to combine toxicity data from freshwater (FW) and saltwater (SW) organisms. This involves deriving two separate distributions and comparing the confidence intervals associated with their HC05 values to determine if salinity influences species sensitivity. If the confidence intervals overlap the joint distribution, the distributions may be combined and a threshold for fish in general developed. If the confidence intervals do not overlap, the SSD may be kept separate with separate thresholds for FW and SW species developed. As discussed below (section 3.1.2), much of this approach relies upon best professional judgement. Several additional methods for assessing whether and how data should be divided or combined for fitting an SSD were investigated, including examination of the slopes of the SSDs (USEPA 2005) and formal model selection methods using Akaike’s Information Criterion (AIC).

In cases where multiple endpoints are available for the same species, the geometric mean of the toxicity values was estimated and that single mean value was used in the SSD. Where data from a single species indicate a notable difference in sensitivity of different life stages (juvenile vs. adult), the life stage that is closest to the standard test guideline was used. The intent of this approach is to avoid introducing the variable of age that may alter a species’ sensitivity relative to other tested species.

No attempt was made to identify and remove outliers from the datasets used to derive SSDs. This is potentially an area of future development for the SSD methodology described here. As discussed in **ATTACHMENT 1-8**, all data were reviewed when included in the ECOTOX database. A more detailed review was carried out by EPA for studies that generated LD50, LC50 or EC50 values that could potentially have an influence on the tails of the SSD. Specifically, values surrounding the 5th, 50th and 95th percentiles were reviewed in detail as well as the most sensitive values available.

1. **Fitting the SSD**

Fitting an SSD requires several important decisions, including 1) which distribution to fit, 2) how the fitted distribution will be evaluated, 3) what data to use, and 4) whether or not there is sufficient data to estimate an HC05 using an SSD. Most of these decisions are data driven, meaning that the analysis are based on evaluation of how well the distributions fit the data. As noted above, the dataset may be revised based on detailed reviews of influential data points. Some decisions, such as which distribution assumptions and fitting methods, are independent of the dataset and are applied to all chemical and taxa data sets. This appendix addresses each of these questions using examples from the malathion aquatic invertebrate dataset.

Five distributions were considered: log-normal, log-logistic, log-triangular, log-gumbel and Burr. The common log transformation was applied to data prior to fitting the first four distributions. Three methods for fitting were considered: maximum likelihood, moment estimation, and graphical methods. For the Burr distribution, only maximum likelihood was considered (because this is the only method that is available for that distribution). Detailed descriptions of these distributions and fitting methods are provided in USEPA 2011 Malathion data for aquatic invertebrates were standardized according to the guidelines established in section 2, resulting in the datasets described in **Table A 1-5.1**.

**Table A 1-5.1. Distribution of results available for Malathion tests on aquatic invertebrates**

|  |  |  |
| --- | --- | --- |
| Media | Test results | Species1 |
| All Invertebrates | 176 | 72 |
| Freshwater Invertebrates | 146 | 60 |
| Saltwater Invertebrates | 29 | 12 |

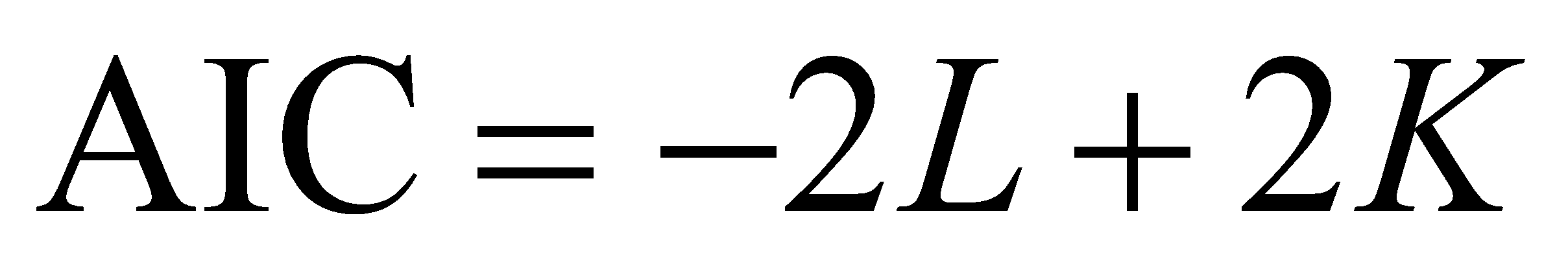
1Daggerblade grass shrimp (*Palaemonetes pugio*) was tested in both fresh and saltwater and medium was not reported for sowbug (*Alitropus typus*).

* 1. *Deciding which distribution best fits a dataset*

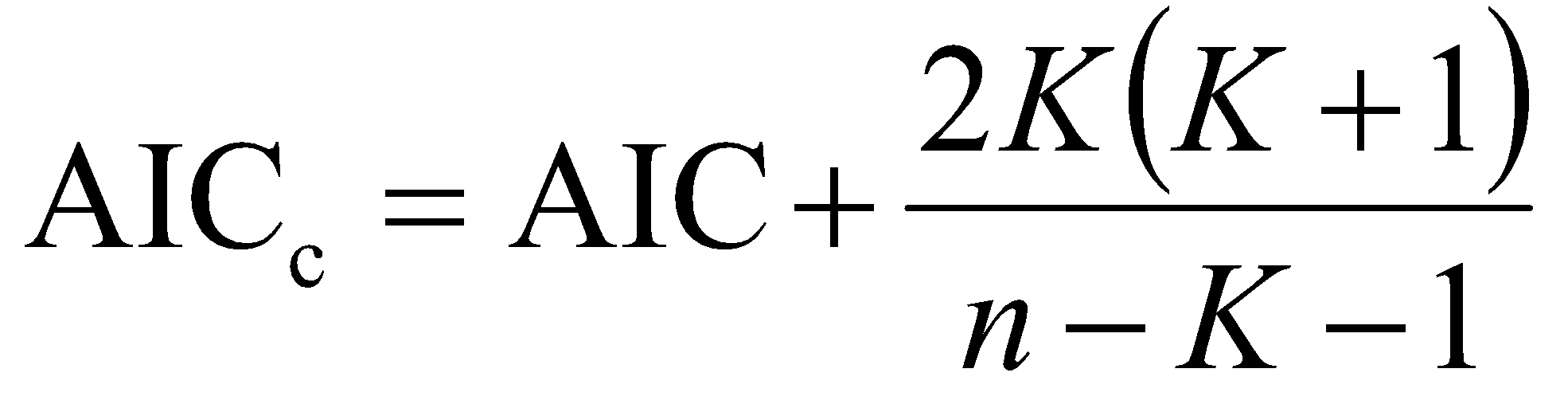
Deciding which distribution fits best is intricately linked to the decision about what data to use in fitting a distribution (*e.g.,* whether to combine freshwater and saltwater results, etc.). While deciding on the dataset may seem like a prerequisite to fitting a distribution, deciding whether data subsets should be combined or kept separate employs comparative measures of fit, visual confirmation of optimal fits, and assessment of lack of fit when visual evaluation suggests that an optimal fit models the data poorly. The following sections step through this process, starting with the evaluation of distribution fits, before returning to the question of data subsets.

3.1.1. Maximum Likelihood and Akaike’s Information Criterion (AIC)

When maximum likelihood is used, fit can be compared among distributions using Akaike’s Information Criterion (AIC). Akaike’s information criterion is a metric derived from the fitted log-likelihood function summed over all data points, with an adjustment for the number of parameters that must be estimated. Specifically,



In the above equation, *L* is the value of the log-likelihood function evaluated against the data at the maximum likelihood estimates for the parameters of the distribution. *K* is the number of estimated parameters (two for all distributions considered here, except the Burr, which has three estimated parameters). The lower the value of AIC, the better the fit; thus when multiple distributions are compared, AIC can help distinguish among competing fits. However, AIC may be biased at small sample sizes, thus we will use a sample-size corrected metric (AICc, where *n* = number of geometric mean toxicity values available):



Comparison among distributions using AICc depends only upon the differences among AICc values, not on the absolute magnitude of the AICc values themselves, which can be either negative or positive. Thus it is conventional when using AICc to report ∆AICc values, which are the difference between AICc for a given distribution and the minimum AIC over all distributions. In general, distributions (models) with a ∆AICc value greater than about 4 or 5 compared to the lowest AICc will not be very competitive (though this is by no means a strict criterion). Distributions can also be assigned weights based on their ∆AICc values, which can help resolve the confidence one should place in a given distribution as being the best for a given dataset. These weights can also be used to derive weighted HC05 estimates (described below under section 3.4). The model selection methods described here are presented in much greater detail in standard texts (*e.g.,* see Burnham and Anderson 2002 and Anderson 2008).

**Tables A 1-5.2 – A 1-5.4** show the performance of each of the five distributions, as judged by AICc, for malathion invertebrate data fit using maximum likelihood. For these analyses, the triangular distribution provided the best fit for pooled and freshwater invertebrates, whereas the gumbel distribution provided the best fit for saltwater invertebrates. In the latter two cases, separation between the best distribution and the next best distribution was almost four AICc units, indicating relatively strong preference for the best distribution over the next best competitor.

**Table A 1-5.2**. **Comparison of distributions for pooled invertebrate toxicity data for Malathion**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| distribution | AICc | ∆AICc | Weight | HC05 |
| triangular | 1095.00 | 0 | 0.69 | 0.40 |
| gumbel | 1097.40 | 2.42 | 0.20 | 0.67 |
| burr | 1099.60 | 4.66 | 0.07 | 0.66 |
| normal | 1100.60 | 5.66 | 0.04 | 0.26 |
| logistic | 1105.90 | 10.91 | 0.00 | 0.13 |

**Table A 1-5.3**. **Comparison of distributions for freshwater invertebrate toxicity data for Malathion**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| distribution | AICc | ∆AICc | Weight | HC05 |
| triangular | 932.4 | 0 | 0.78 | 0.39 |
| gumbel | 936.3 | 3.9 | 0.11 | 0.56 |
| normal | 937.5 | 5.1 | 0.06 | 0.23 |
| burr | 938.6 | 6.2 | 0.04 | 0.55 |
| logistic | 942.3 | 9.9 | 0.01 | 0.12 |

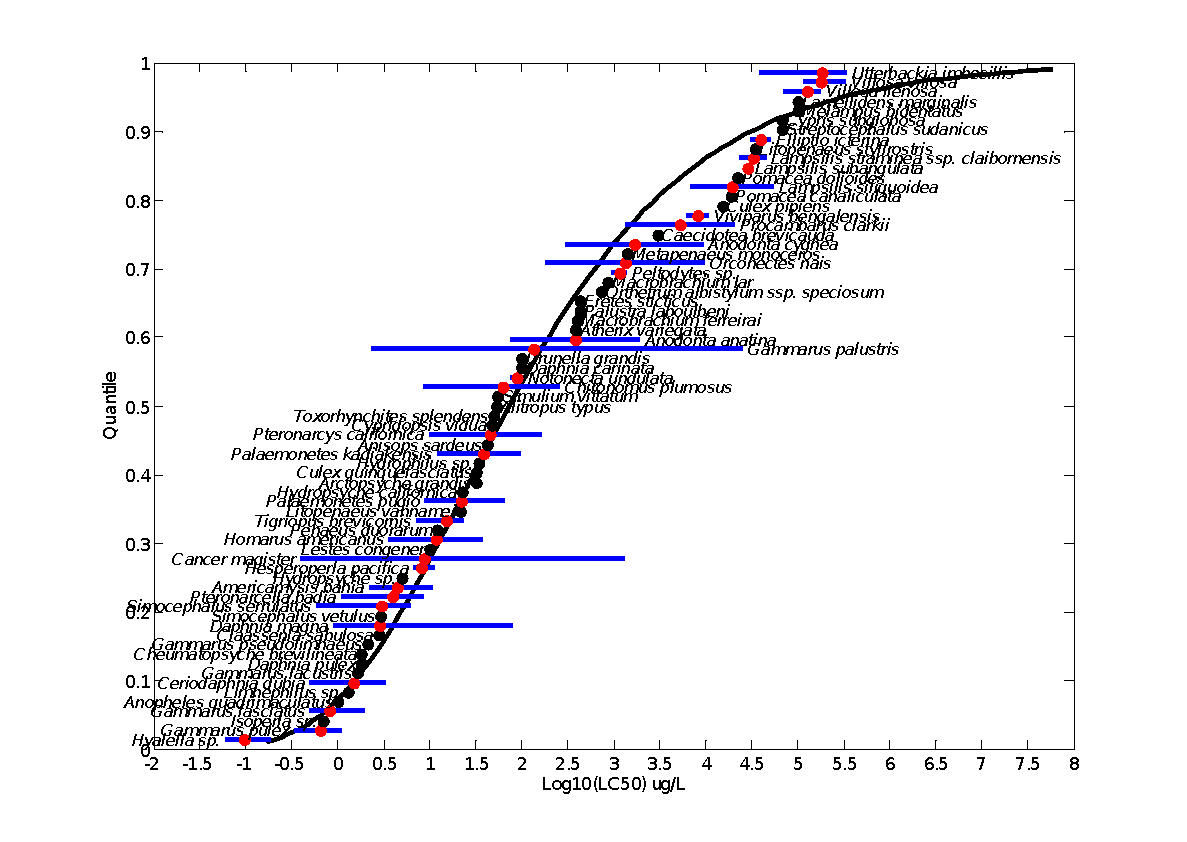
**Table A 1-5.4. Comparison of distributions for saltwater invertebrate toxicity data for Malathion**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| distribution | AICc | ∆AICc | Weight | HC05 |
| gumbel | 162.19 | 0 | 0.66 | 2.55 |
| triangular | 166.13 | 3.94 | 0.09 | 0.59 |
| burr | 166.13 | 3.94 | 0.09 | 2.56 |
| normal | 166.43 | 4.24 | 0.08 | 0.46 |
| logistic | 166.49 | 4.30 | 0.08 | 0.26 |

3.1.2. Visually *assessing the fit of a distribution*

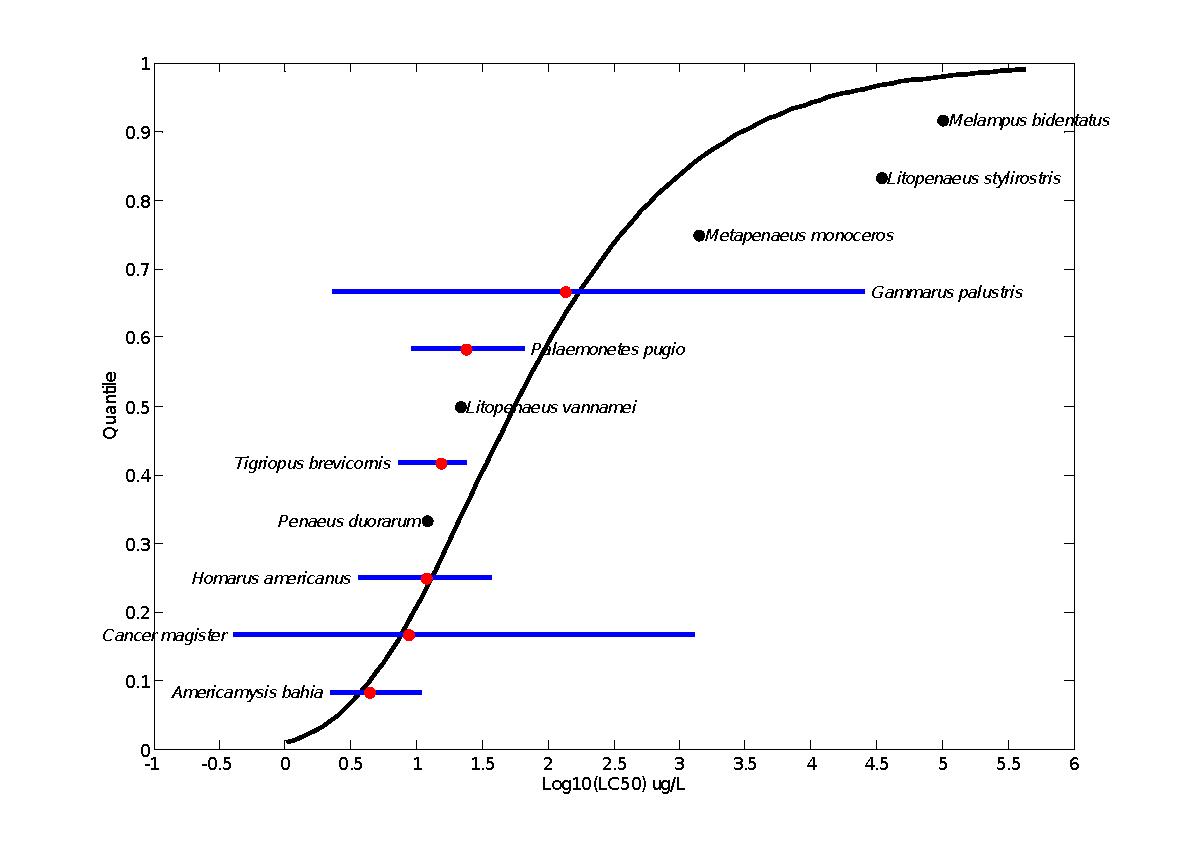
**Figures A 1-5.1 – A 1-5.3** show the best-fit distributions fit to the malathion pooled, freshwater, and saltwater invertebrate data, respectively. No plot shows strong visual evidence of lack-of-fit, though the gumbel distribution fit to the saltwater results does present some cause for concern at the upper quantiles (**Figure A 1-5.3**).

There are many different methods available in the literature for plotting points in an empirical cumulative distribution function (Cunnane 1978). The Weibull (1939) formula is used in these figures due to its satisfaction of the Gumbel (1947) postulates and its easy interpretation. The Weibull formula uses pi = ri/(n+1), where pi = plotting point for data point i, ri = rank of data point i and n = number of data points (geometric mean toxicity values).



**Figure A 1-5.1. Log-triangular distribution fit to malathion pooled invertebrate data.** Black points indicate single toxicity values. Red points indicate average of multiple toxicity values for a single species. Blue line indicates full range of toxicity values for a given species.

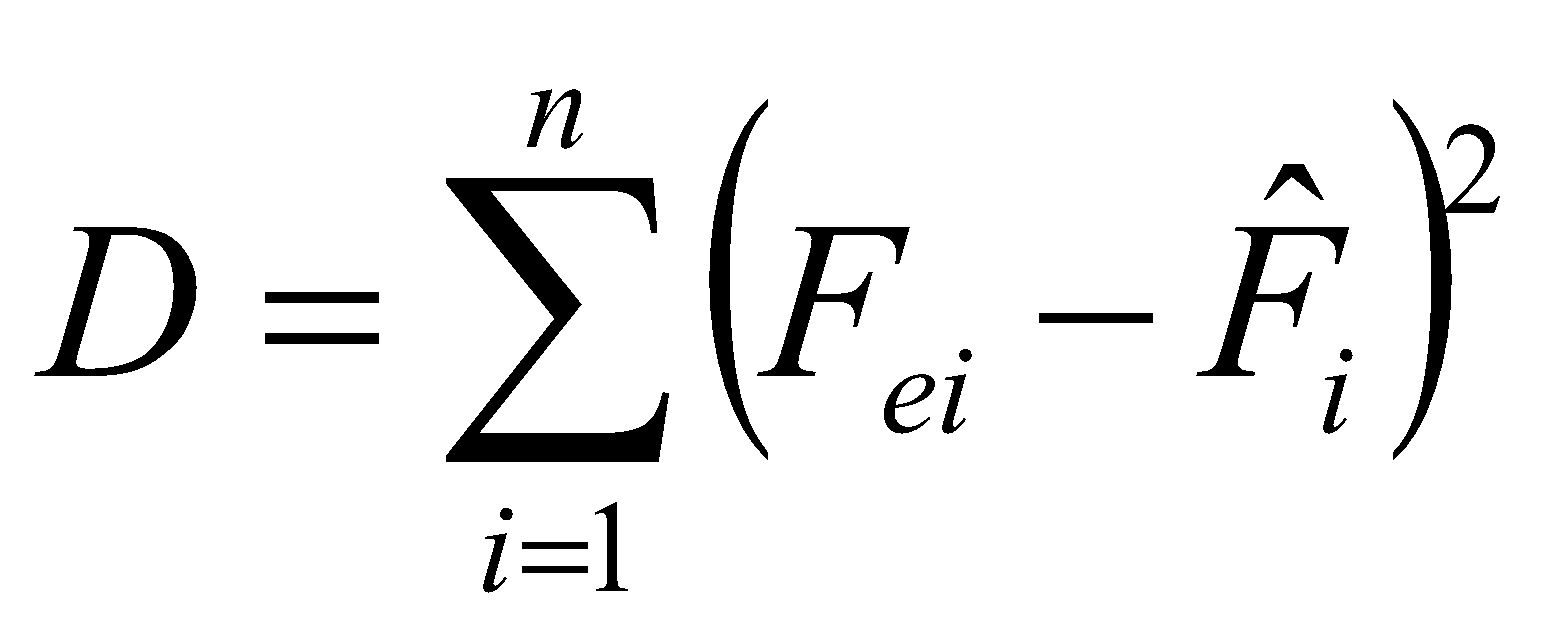
**Figure A 1-5.2. Log-triangular distribution fit to malathion freshwater invertebrate data.** Black points indicate single toxicity values. Red points indicate average of multiple toxicity values for a single species. Blue line indicates full range of toxicity values for a given species.



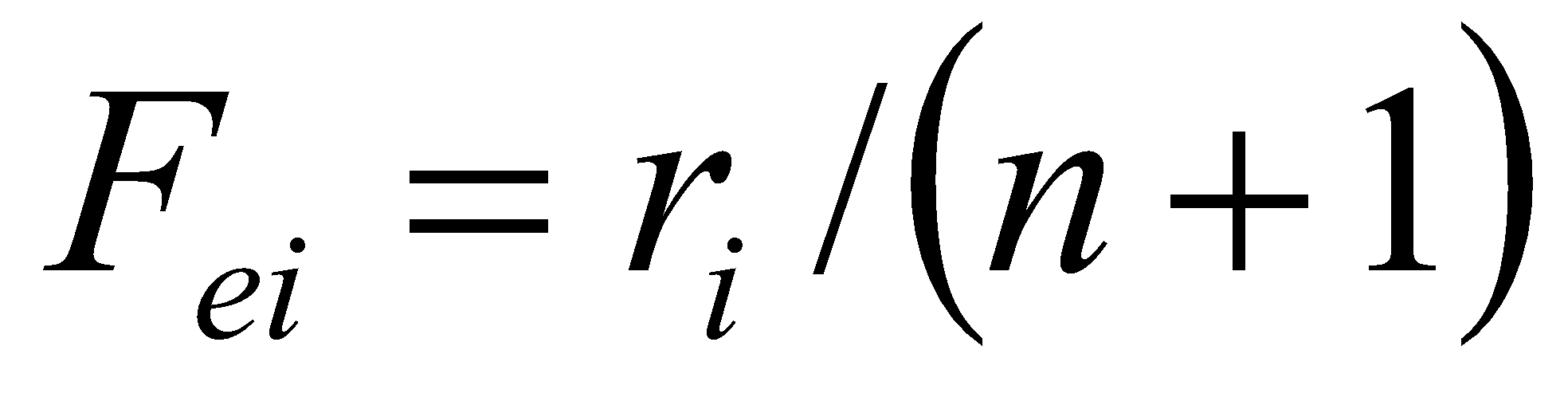
**Figure A 1-5.3. Log-gumbel distribution fit to malathion saltwater invertebrate data.** Black points indicate single toxicity values. Red points indicate average of multiple toxicity values for a single species. Blue line indicates full range of toxicity values for a given species.

3.1.3. Goodness of fit

**Figure A 1-5.3** highlights a potential weakness in relying solely on AICc to judge the fit of a distribution because AICc is useful only as a comparative metric. It is not an absolute measure of goodness-of-fit. In other words, AICc cannot tell us whether deviations from fit, such as those observed in the three points at the upper quantiles of **Figure A 1-5.3**, are significantly greater than would be expected by chance alone. For this we need a formal goodness-of-fit measure. For the fits below, a parametric bootstrap method for testing goodness-of-fit was used that compares estimates of the discrepancy between the empirical cumulative distribution function and the estimated cumulative distribution function for a candidate distribution. Specifically:



In the equation above, *Fei* is the empirical cumulative distribution function (ECDF) calculated at the *i*th data point,  is the cumulative distribution function for the fitted distribution, also calculated at the *i*th data point and *n* is the number of data points (geometric mean toxicity values). The ECDF is calculated by sorting the toxicity values in ascending order and calculating their ranks (*ri*). Then:



Significance of the discrepancy statistic (*D*) can be assessed by generating random samples with the same number of data points (*n* = geometric mean toxicity values) from the candidate distribution and calculating *D* for each iteration. If the value of *D* for the empirical data is greater than the value obtained for 95% of the random samples, then this is good evidence for lack-of-fit between the candidate distribution and the empirical data.

For all cases above (**Figures A 1-5.1 – A 1-5.3**) the best-fit candidate distributions showed little evidence for lack-of-fit using the parametric bootstrap test. However, it should be noted that power to detect lack-of-fit is low, particularly for the saltwater data, for which there were only 12 geometric mean toxicity values available.

It will generally be the case that power to detect lack-of-fit will be low at the sample sizes likely to be available for most pesticides. Thus it is important that analysts use best judgment to decide whether or not a fitted distribution is adequate. In particular, visual inspection of the plotted cumulative distribution function against the data is imperative! For example, the visual evidence for lack of fit by the gumbel distribution fit to the saltwater data for malathion shows some lack of fit at the upper quantiles. However, because the most important inference (direct and indirect effect thresholds) is made from the HC05, it may be acceptable to use the fitted gumbel in this case because the fit around the HC05 appears to be quite good. Note, however, that this is not a recommendation, simply an example of the kind of judgments that the analyst will need to make in fitting SSDs to limited datasets.

* 1. *Deciding whether to combine or separate subsets of data*

This decision will be one of the most difficult faced by the analyst in fitting SSDs to limited datasets. There is no recommended minimum sample size, because the variability in the data will also influence the fit and the confidence we should place in the HC05 estimate. The suitability of a data set for deriving an SSD is based on the fit of the distributions to the data and on the best professional judgment of the analyst. Typically, the decision to split a single dataset into two separate datasets will result in one large and one small dataset and the smaller dataset will provide relatively uncertain estimates of the HC05. Thus the analyst will be faced with the decision of whether or not to use the HC05 for a proper subset of species which may have greater uncertainty than the estimate resulting from pooling the datasets. For example, in the previous section, separate SSDs were fit to freshwater versus saltwater results for aquatic invertebrates. For freshwater invertebrates there were 60 geometric mean toxicity values available, certainly sufficient for fitting an SSD, but for saltwater invertebrates there were only 12 geometric means available and some evidence for lack-of-fit in the log-gumbel distribution, which, according to AICc provided the best fit to the saltwater data. Thus the analyst should consider whether the two datasets merit separate SSDs or whether a single invertebrate SSD should be fit. This section provides guidance on how to make these kinds of decisions.

Several methods can be used to decide whether to lump or split data for fitting SSDs. These include, 1) hierarchical modeling to specify distribution parameters as functions of class membership (*e.g.,* freshwater versus saltwater in the example below), 2) comparison of the slopes of the SSDs (USEPA 2005), and 3) visual comparison of SSDs and confidence limits plotted on common axes. Of these, the third will likely prove to be the most general purpose method.

AIC (and AICc) can also be useful for exploring the effects of experimental conditions, such as salinity, on toxicity and resulting SSDs. When SSDs are fit using maximum likelihood, the parameters of the distribution (e.g., μ and σ for the normal distribution) can be specified to be functions of measured covariates (in the example here the covariate would be FW vs. SW). In this case the SSD is fit like a generalized linear model, and the model parameters are the design parameters associated with the covariates used for modeling. For these models, the estimated parameters are θ1, θ2, θ3, and θ4, (as needed) and AICc is used to distinguish among models that incorporate different hypotheses about the influence of salinity on location and shape of the SSD.

**Table A 1-5.5. Covariate models for testing effects of salinity on toxicity of malathion**

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Description & Hypothesis | location effects1,2 | scale effects1,2 |
| null | salinity has no effect on toxicity | μ = θ1 | σ = θ2 |
| location | salinity affects mean, but not variance of distribution | μ = θ1 + θ2\*FW | σ = θ3 |
| scale | salinity affects variance, but not mean of distribution | μ = θ1 | σ = θ2 + θ3\*FW |
| full | salinity affects both mean and variance of distribution | μ = θ1 + θ2\*FW | σ = θ3 + θ4\*FW |

1FW is a binary dummy variable that takes value 1 for FW taxa and 0 for SW taxa.

2equations presented for log-normal SSD. For log-logistic SSD replace μ with α and replace σ with β. Analogous substitutions can be made for triangular and gumbel distributions.

Because FW is a simple binary variable, the models in **Table A 1-5.5** can also be interpreted as describing whether or not separate location or scale parameters are needed for SW taxa versus FW taxa. For example, the null model hypothesizes that the FW SSD and SW SSD have the same mean and same variance. The location model hypothesizes that the FW SSD has a different mean, but the same variance as the SW SSD (analogous to comparisons in USEPA 2005 that look for parallel slopes in FW versus SW SSDs). In contrast, the scale model hypothesizes that the FW SSD has the same mean, but different variance than the SW SSD. Finally, the full model states that the FW and SW SSDs have both different means and different variances (i.e., one needs to fit two entirely separate SSDs). It should be noted that the methodology suggested in **Table A 1-5.5** could handle more complex modeling situations, including, for example, cases in which a continuous covariate (*e.g*., hardness, temperature, or body weight) were thought to influence toxicity.

For an example of the hierarchical modeling approach consider the question of whether to pool or separate freshwater and saltwater results. For this example, the gumbel distribution will be used, which has two parameters, *µ* and *β*, which are location and scale parameters, respectively. By expressing *µ* and/or *β* as functions of an indicator variable distinguishing between freshwater and saltwater results, we can fit four different regression models, defined as follows:

1. null – freshwater and saltwater SSDs share the same parameters
2. location – freshwater and saltwater SSDs differ in their location parameter (*µ*) but share the scale parameter (*β*)
3. scale – freshwater and saltwater SSDs differ in their scale parameter (*β*) but share the location parameter (*µ*)
4. full – freshwater and saltwater SSDs differ in both location and scale parameters

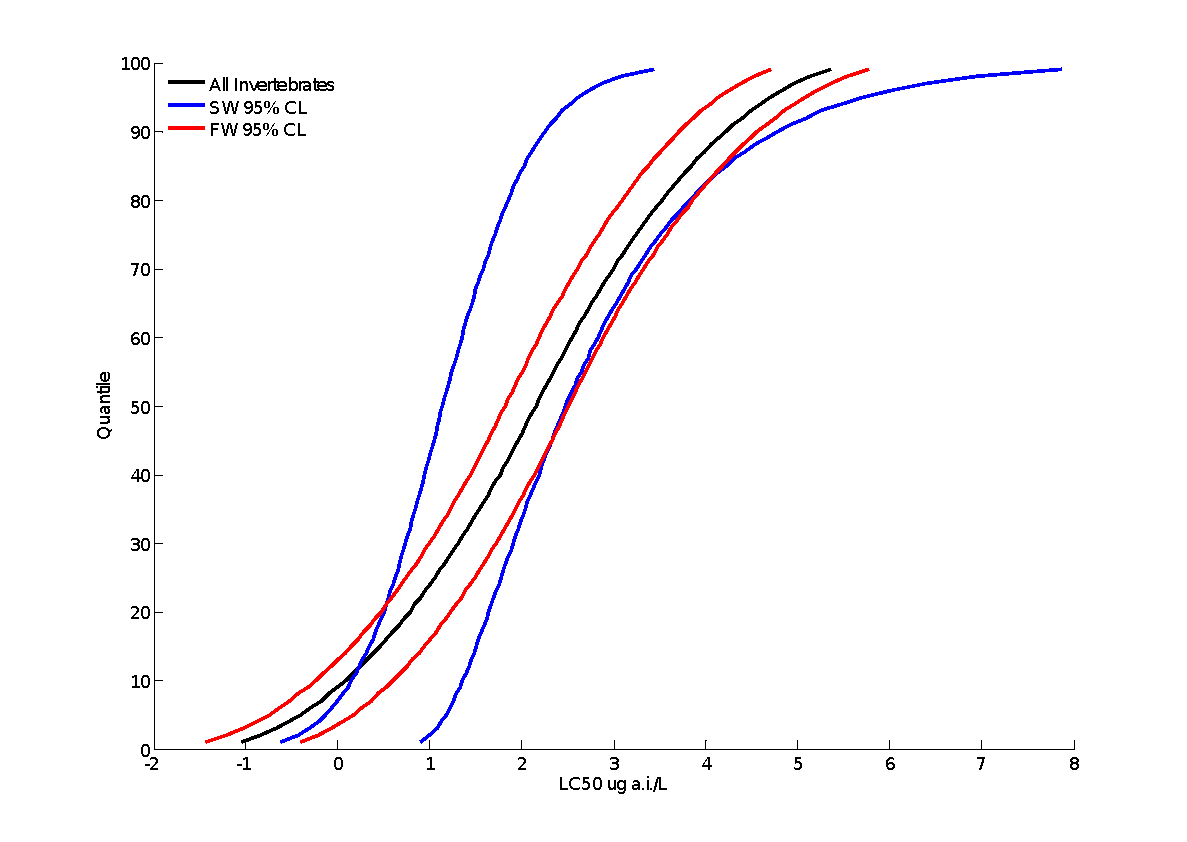
**Table A 1-5.6** shows the results of the above four models fit to all malathion invertebrate data. These results provide limited support for modeling freshwater and saltwater data separately. The best performing model (null) lends support for combining the data into a single SSD. However, the second best model (scale) suggests that the two datasets differ in their scale parameter and this model is almost as competitive as the null model. Thus the question of separating the two datasets is left unresolved by the hierarchical modeling approach.

**Table A 1-5.6. Hierarchical models fit to malathion invertebrate data**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | parameters | AIC | dAIC | Wt |
| null | 2 | 1085.1 | 0.00 | 0.41 |
| scale | 3 | 1085.6 | 0.51 | 0.32 |
| location | 3 | 1087.0 | 1.95 | 0.16 |
| full | 4 | 1087.8 | 2.70 | 0.11 |

Further exploration of the hierarchical modeling approach would require repeating the exercise, but with the triangular distribution, which was preferred for the freshwater results. However this highlights an important shortcoming to the hierarchical modeling approach, which is that the approach only works under a common distribution. When different distributions are optimal for the different data subsets, the approach cannot be used. This constraint also applies to the suggestion to compare the slopes of SSDs, because slopes can only be considered parallel if the same distribution is fit to both data subsets. These observations again reinforce the need for the analyst to use his or her own best judgment in deciding whether or not to pool or separate data subsets*.*

Plotting the fitted SSDs on similar axes can greatly assist in deciding whether to pool or separate data subsets and this method does not require that the same distribution be fit to the two subsets. For example, the log-triangular and log-gumbel fits for the separate datasets are plotted in Figure 4 against the pooled SSD for all invertebrates. **Figure A 1-5.4** shows that the pooled SSD is almost entirely within the confidence limits for the respective SSDs fit to the divided datasets. This lends support to pooling the two SSDs. However, note that in this case the pooled SSD crosses the 95% confidence limit for the saltwater SSD in the region of the HC05. Thus, an argument could be made for using the separate SSDs, because the HC05 is the primary inferential endpoint from the SSD. A further consideration would be the fact that use of the HC05 from the pooled SSD would be more conservative than use of the HC05 estimated from the saltwater SSD, perhaps tipping the balance in favor of the pooled SSD. As above, the role of critical thinking and best judgment on the part of the analyst cannot be overstated in guiding these decisions.



Log

**Figure A 1-5.4. Best SSD for pooled aquatic invertebrate test results (triangular, black), plotted against the confidence limits for the best freshwater invertebrate SSD (triangular, red), and best saltwater invertebrate SSD (gumbel, blue).**

1. **Estimating the HC05**

HC05 values are estimated from the fitted SSDs. Variance associated with each HC05 can also be estimated in order to express the variability in the distribution of the available toxicity data as well as the uncertainty associated with the estimated value.

**Tables A 1-5.7 – A 1-5.9** show the estimated HC05s generated from the five different distributions fit to the malathion invertebrate datasets. Several important patterns can be gleaned from the tables. First, lack-of-fit was detected only for the Burr distribution, though graphical methods of fitting showed a trend towards lack of fit for all distributions fit to the saltwater data. P is the best measure of fit. It is the P-value from the parametric bootstrap goodness of fit test. The confidence limits around the HC05 (LCx , UCx) and around the estimated percentile (LCp, UCp) are also valuable measures of fit. Second, the coefficients of variation (CVs) for the freshwater data were close to 1 (and less than 1 for the preferred distribution fit using maximum likelihood), but the CVs were generally large for the saltwater results, suggesting considerable uncertainty in the HC05 estimates for saltwater results. While the CV shows the sampling error of the HC05 relative to the estimated value of the HC05, another relative measure shows the confidence interval of the HC05 relative to the spread of the distribution. This is denoted in the columns LCp and UCp, which are the projections of the confidence limits of the HC05 (LCx and UCx) onto the cumulative distribution function of the fitted distribution. Thus, for example, for the malathion freshwater invertebrate results, the upper confidence limit of the HC05 for the triangular distribution fit using maximum likelihood extends to the 11th percentile of the fitted distribution. In contrast, the upper confidence limit for the gumbel distribution fit to the saltwater test results extends to the 26th percentile, highlighting the uncertainty surrounding that dataset due to its very limited sample size.

**Table A 1-5.7. Range of HC05 values for Malathion SSDs fit to all invertebrates.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| distribution | method | HC05 | SE | CV | LCx | UCx | LCp | UCp | P |
| normal | ML | 0.26 | 0.28 | 1.08 | 0.00 | 1.08 | 0.00 | 0.10 | 0.76 |
| MO | 0.25 | 0.27 | 1.06 | 0.06 | 1.02 | 0.02 | 0.10 | 0.75 |
| GR | 0.19 | 0.16 | 0.86 | 0.03 | 0.61 | 0.02 | 0.09 | 0.67 |
| logistic | ML | 0.13 | 0.18 | 1.32 | 0.03 | 0.66 | 0.03 | 0.10 | 0.63 |
| MO | 0.27 | 0.32 | 1.19 | 0.05 | 1.22 | 0.02 | 0.10 | 0.88 |
| GR | 0.17 | 0.16 | 0.93 | 0.02 | 0.58 | 0.02 | 0.08 | 0.76 |
| triangular | ML | 0.40 | 0.31 | 0.77 | 0.21 | 1.33 | 0.03 | 0.10 | 0.82 |
| MO | 0.22 | 0.21 | 0.95 | 0.07 | 0.85 | 0.02 | 0.10 | 0.61 |
| GR | 0.19 | 0.16 | 0.85 | 0.05 | 0.62 | 0.01 | 0.10 | 0.55 |
| gumbel | ML | 0.67 | 0.37 | 0.56 | 0.22 | 1.68 | 0.01 | 0.11 | 0.79 |
| MO | 0.93 | 0.58 | 0.63 | 0.32 | 2.52 | 0.01 | 0.12 | 0.92 |
| GR | 0.69 | 0.36 | 0.53 | 0.15 | 1.54 | 0.01 | 0.10 | 0.80 |
| burr | ML | 0.66 | 0.37 | 0.55 | 0.29 | 1.68 | 0.02 | 0.11 | 0.10 |

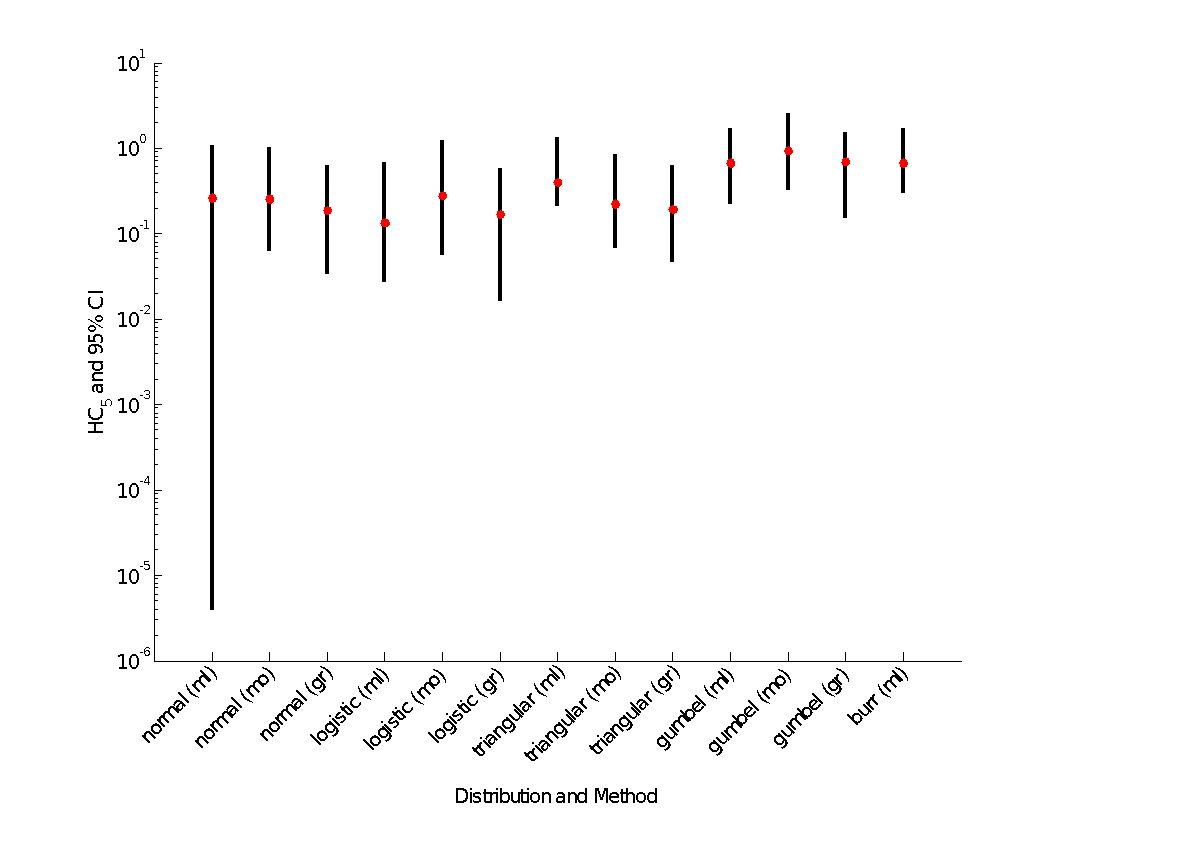
**Table A 1-5.8. Range of HC05 values for Malathion SSDs fit to freshwater invertebrates.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| distribution | method | HC05 | SE | CV | LCx | UCx | LCp | UCp | P |
| normal | ML | 0.23 | 0.31 | 1.34 | 0.00 | 1.15 | 0.00 | 0.11 | 0.90 |
| MO | 0.22 | 0.28 | 1.26 | 0.05 | 1.04 | 0.02 | 0.11 | 0.91 |
| GR | 0.16 | 0.17 | 1.08 | 0.02 | 0.65 | 0.02 | 0.10 | 0.84 |
| logistic | ML | 0.12 | 0.20 | 1.66 | 0.02 | 0.72 | 0.02 | 0.10 | 0.83 |
| MO | 0.24 | 0.37 | 1.55 | 0.04 | 1.38 | 0.02 | 0.11 | 0.96 |
| GR | 0.14 | 0.16 | 1.15 | 0.01 | 0.56 | 0.02 | 0.09 | 0.90 |
| triangular | ML | 0.39 | 0.37 | 0.96 | 0.18 | 1.49 | 0.03 | 0.11 | 0.93 |
| MO | 0.20 | 0.21 | 1.08 | 0.05 | 0.83 | 0.02 | 0.11 | 0.82 |
| GR | 0.16 | 0.16 | 0.98 | 0.03 | 0.62 | 0.01 | 0.10 | 0.77 |
| gumbel | ML | 0.56 | 0.37 | 0.66 | 0.17 | 1.54 | 0.01 | 0.11 | 0.89 |
| MO | 0.85 | 0.65 | 0.77 | 0.26 | 2.66 | 0.01 | 0.13 | 0.97 |
| GR | 0.60 | 0.38 | 0.63 | 0.11 | 1.53 | 0.01 | 0.10 | 0.91 |
| burr | ML | 0.55 | 0.38 | 0.68 | 0.22 | 1.64 | 0.02 | 0.11 | 0.06 |

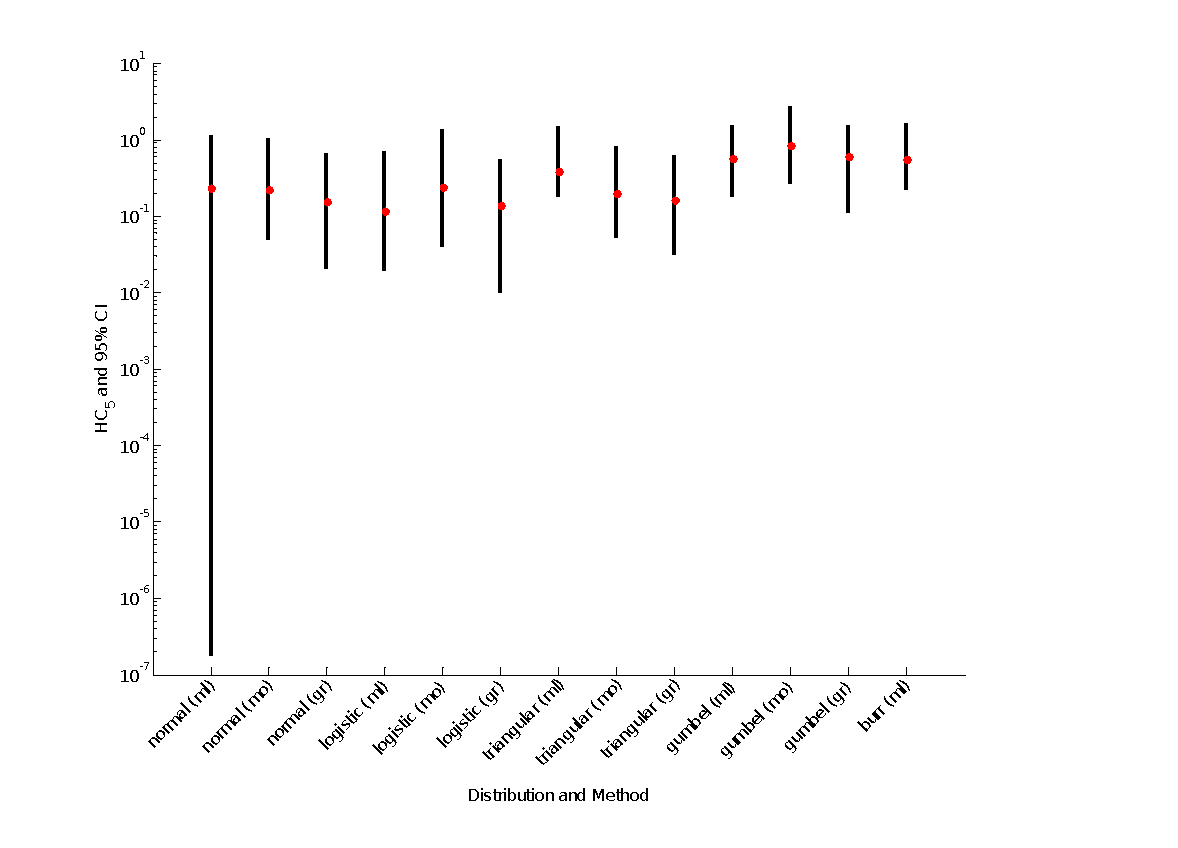
**Table A 1-5.9. Range of HC05 values for Malathion SSDs fit to saltwater invertebrates.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| distribution | method | HC05 | SE | CV | LCx | UCx | LCp | UCp | P |
| normal | ML | 0.35 | 3.64 | 10.44 | 0.01 | 8.06 | 0.00 | 0.24 | 0.12 |
| MO | 0.27 | 2.69 | 9.82 | 0.01 | 6.96 | 0.01 | 0.24 | 0.11 |
| GR | 0.10 | 1.30 | 13.08 | 0.00 | 1.67 | 0.00 | 0.17 | 0.06 |
| logistic | ML | 0.22 | 1.41 | 6.51 | 0.01 | 4.79 | 0.01 | 0.23 | 0.09 |
| MO | 0.29 | 3.89 | 13.18 | 0.01 | 8.13 | 0.01 | 0.23 | 0.13 |
| GR | 0.07 | 0.67 | 9.65 | 0.00 | 1.45 | 0.00 | 0.16 | 0.06 |
| triangular | ML | 0.61 | 5.73 | 9.43 | 0.14 | 15.64 | 0.01 | 0.26 | 0.17 |
| MO | 0.25 | 2.66 | 10.78 | 0.02 | 5.42 | 0.00 | 0.23 | 0.07 |
| GR | 0.12 | 1.67 | 13.47 | 0.00 | 2.27 | 0.00 | 0.19 | 0.05 |
| gumbel | ML | 2.01 | 2.98 | 1.48 | 0.61 | 10.70 | 0.01 | 0.26 | 0.33 |
| MO | 0.89 | 2.52 | 2.84 | 0.10 | 8.16 | 0.00 | 0.27 | 0.14 |
| GR | 0.37 | 1.04 | 2.82 | 0.01 | 2.76 | 0.00 | 0.19 | 0.05 |
| burr | ML | 2.01 | 2.78 | 1.38 | 0.42 | 10.12 | 0.00 | 0.25 | 0.02 |

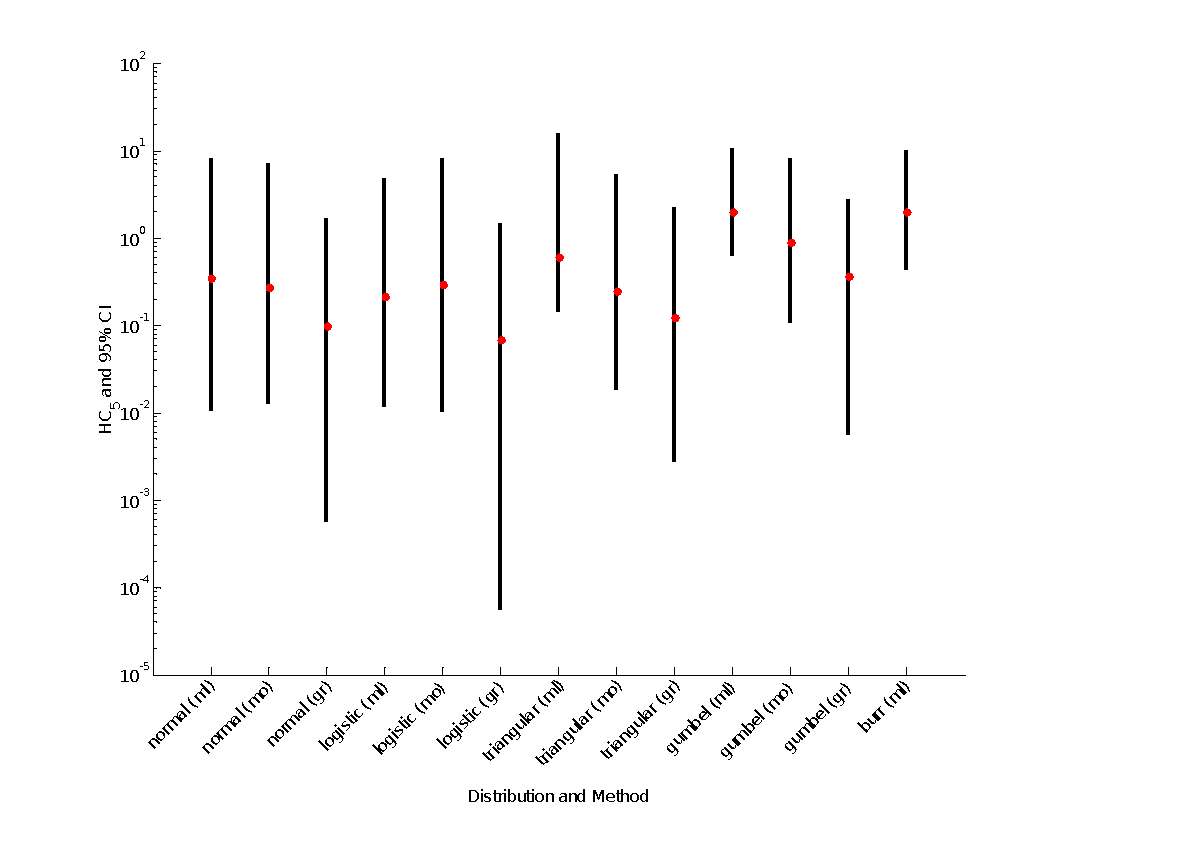
Other patterns in the HC05 can be observed by plotting the HC05 estimates, together with their 95% confidence limits, on common axes (**Figures A 1-5.5 – A 1-5.7**). First, and perhaps most obvious, is the extremely large estimate of uncertainty for the log-normal distribution fit to the freshwater results using maximum likelihood (**Figure A 1-5.5**). Another common pattern, easily observed in these results is the tendency for graphical estimates to produce lower estimates of the HC05 than either moment estimators or maximum likelihood. Finally, another commonly observed pattern is that the preferred distributions (triangular for freshwater, gumbel for saltwater), tend to have relatively narrow confidence limits in comparison to other distributions, reinforcing their selection as the best distributions. However, this pattern is not universally observed.



**Figure A 1-5.5. HC05 estimates and confidence limits for malathion pooled aquatic invertebrates**



**Figure A 1-5.6. HC05 estimates and confidence limits for malathion freshwater invertebrates**



**Figure A 1-5.7. HC05 estimates and confidence limits for malathion saltwater invertebrates**

Plots such as **Figures A 1-5.5 – A 1-5.7** may be useful for deciding between HC05 estimates when competition between distributions is tight (*e.g.,* ΔAICc values small). Another option in such conditions is to use model averaging to produce a consensus estimate of the HC05 that uses information across distributions. For the fitting methods discussed here, such methods are available only for distributions fit using maximum likelihood. In essence, this involves using the AIC weights (**Tables A 1-5.2 – A 1-5.4**) to produce a weighted average HC05 estimate with associated confidence interval, and methods are outlined in standard texts on model-averaging (e.g., Burnham and Anderson, 2002, Anderson 2008). For the two examples presented above, the model-averaged HC05 estimates and standard errors were 0.40 (0.52) μg/L for freshwater invertebrates and 1.7 (1.3) μg/L for saltwater invertebrates. The latter estimates excluded fits of the Burr distribution due to lack-of-fit. Use of model-averaged values may help protect against reliance upon a single distribution when there is a relatively large degree of uncertainty about which distribution best fits the data. Ultimately, the analyst uses best professional judgement in determining the best distribution through consideration of the available information discussed above.

1. **Estimating threshold concentrations**

The threshold concentrations can be estimated using a probit-dose-response curve for the fifth percentile species (whose sensitivity is represented by the HC05). In order to estimate the threshold concentrations, it is necessary to derive a slope estimate that is relevant to the fifth percentile species of the SSD. The species with endpoints that are near HC05 will be used to derive slope data representative of the fifth percentile species. The median estimate of the available slopes will be used as well as a range of available values. If no slope is available for a specific LC/LD/EC50, then the default central estimate of 4.5 will be used. The specific equation used to calculate the threshold are:

Log10(Threshold)=log10(HC05)+z/slope

In the above equation, z is the quantile of a standard normal distribution associated with the desired probability (10-6 for direct effects, 10-1 for indirect effects).

For direct effects, the threshold concentration will be estimated as 1/10,000th percentile of the fitted dose-response with mean LC50 equal to the estimated HC05 and slope as described immediately above. For indirect effects, the threshold concentration will be estimated as the 10% effect level based on the LC50 and slope (this is essentially the LC10). In addition to the median estimate of the thresholds, uncertainty associated with those values will also be explored using the variability in the HC05. 95% confidence limits around the thresholds are estimated by extrapolation from the 95% confidence limits around the HC05. Uncertainty around the slope of the dose-response is not currently incorporated into these uncertainty estimates, but could be included manually by users by employing different plausible slopes.

**Tables A 1-5.10 – A 1-5.12** illustrate the direct and indirect effects thresholds for pooled, freshwater, and saltwater invertebrates. These values were estimated from the HC05 values depicted in **Tables A 1-5.5** and **A 1-5.6**. No slope information was available for the species near the HC05 values for the two data sets, therefore, the default value of 4.5 was used. For freshwater invertebrates, the median estimate of the direct effects threshold is approximately 0.034 µg a.i./L (ppb), with a confidence interval that spans two orders of magnitude. For indirect effects due to impacts on invertebrate prey, the threshold is approximately 0.2 µg a.i./L (ppb), with a confidence interval spanning an order of magnitude. For saltwater invertebrates, the median estimate of the direct effects threshold is approximately 0.18 µg a.i./L, with a confidence interval that spans a little over an order of magnitude. For indirect effects due to impacts on invertebrates prey, the threshold is approximately 1.0 µg a.i./L, with a confidence interval spanning an order of magnitude.

The value from the SSD with the best fit based on maximum likelihood will be used for thresholds. This method offers the best statistical properties (asymptotically unbiased, minimum variance). Future work will examine when other methods (graphical, moment estimators) may be better choices (*e.g.*, potentially at small sample sizes).

**Table A 1-5.10. Estimated direct and indirect effects thresholds for pooled invertebrates.** Threshold values in µg a.i./L.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Distribution | Method | Direct effects threshold  (1 in a million)\* | | | Indirect effects threshold  (10% mortality)\* | | |
| Median | Lower confidence limit | Upper confidence limit | Median | Lower confidence limit | Upper confidence limit |
| normal | ML | 0.0205 | 1.52E-08 | 0.1009 | 0.1211 | 9.00E-08 | 0.5963 |
| MO | 0.0194 | 0.0042 | 0.0914 | 0.1147 | 0.0249 | 0.54 |
| GR | 0.0137 | 0.0018 | 0.0569 | 0.0812 | 0.0104 | 0.336 |
| logistic | ML | 0.0104 | 0.0017 | 0.0632 | 0.0616 | 0.0098 | 0.3732 |
| MO | 0.0211 | 0.0034 | 0.1214 | 0.1249 | 0.0201 | 0.7171 |
| GR | 0.0122 | 0.0009 | 0.0492 | 0.0721 | 0.0051 | 0.2907 |
| triangular | ML | 0.034 | 0.0157 | 0.1308 | 0.2011 | 0.0927 | 0.7728 |
| MO | 0.0172 | 0.0045 | 0.0725 | 0.1019 | 0.0264 | 0.4287 |
| GR | 0.0144 | 0.0027 | 0.0547 | 0.0848 | 0.0162 | 0.3234 |
| gumbel | ML | 0.0494 | 0.0153 | 0.1349 | 0.2916 | 0.0903 | 0.797 |
| MO | 0.0742 | 0.0227 | 0.2339 | 0.4387 | 0.1342 | 1.382 |
| GR | 0.0527 | 0.0095 | 0.1345 | 0.3113 | 0.0562 | 0.795 |
| burr | ML | 0.0484 | 0.0189 | 0.1438 | 0.286 | 0.112 | 0.8495 |

\*Estimated using median, lower and upper confidence limits of HC05 from Table 6 and median, lower and upper bounds on slope (*i.e.,* 4.5, 2.0 and 9.0, respectively).

**Table A 1-5.11. Estimated direct and indirect effects thresholds for freshwater invertebrates.** Threshold values in µg a.i./L.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Distribution | Method | Direct effects threshold  (1 in a million)\* | | | Indirect effects threshold  (10% mortality)\* | | |
| Median | Lower confidence limit | Upper confidence limit | Median | Lower confidence limit | Upper confidence limit |
| normal | ML | 0.021 | 0.0010 | 0.07 | 0.12 | 0.05 | 0.17 |
| MO | 0.019 | 0.0009 | 0.07 | 0.11 | 0.05 | 0.16 |
| GR | 0.014 | 0.0007 | 0.05 | 0.08 | 0.04 | 0.11 |
| logistic | ML | 0.010 | 0.0005 | 0.04 | 0.06 | 0.03 | 0.09 |
| MO | 0.021 | 0.0010 | 0.07 | 0.12 | 0.06 | 0.17 |
| GR | 0.012 | 0.0006 | 0.04 | 0.07 | 0.03 | 0.10 |
| triangular | ML | 0.034 | 0.0016 | 0.11 | 0.20 | 0.09 | 0.28 |
| MO | 0.017 | 0.0008 | 0.06 | 0.10 | 0.04 | 0.14 |
| GR | 0.014 | 0.0007 | 0.05 | 0.08 | 0.04 | 0.12 |
| gumbel | ML | 0.049 | 0.0024 | 0.17 | 0.29 | 0.13 | 0.40 |
| MO | 0.074 | 0.0036 | 0.25 | 0.44 | 0.19 | 0.61 |
| GR | 0.053 | 0.0025 | 0.18 | 0.31 | 0.14 | 0.43 |
| burr | ML | 0.048 | 0.0023 | 0.16 | 0.29 | 0.13 | 0.40 |

\*Estimated using median, lower and upper confidence limits of HC05 from Table 7 and median, lower and upper bounds on slope (*i.e.,* 4.5, 2.0 and 9.0, respectively).

**Table A 1-5.12. Estimated direct and indirect effects thresholds for aquatic invertebrates.** Threshold values in mg a.i./L.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Distribution | Method | Direct effects threshold  (1 in a million)\* | | | Indirect effects threshold  (10% mortality)\* | | |
| Median | Lower confidence limit | Upper confidence limit | Median | Lower confidence limit | Upper confidence limit |
| normal | ML | 0.041 | 0.001 | 1.050 | 0.239 | 0.008 | 6.203 |
| MO | 0.031 | 0.001 | 0.853 | 0.183 | 0.007 | 5.038 |
| GR | 0.011 | 0.000 | 0.219 | 0.062 | 0.000 | 1.291 |
| logistic | ML | 0.023 | 0.001 | 0.624 | 0.136 | 0.005 | 3.684 |
| MO | 0.033 | 0.001 | 0.992 | 0.198 | 0.005 | 5.864 |
| GR | 0.007 | 0.000 | 0.176 | 0.042 | 0.000 | 1.039 |
| triangular | ML | 0.052 | 0.011 | 1.649 | 0.309 | 0.066 | 9.746 |
| MO | 0.028 | 0.002 | 0.689 | 0.165 | 0.010 | 4.069 |
| GR | 0.013 | 0.000 | 0.282 | 0.079 | 0.001 | 1.668 |
| gumbel | ML | 0.224 | 0.066 | 1.322 | 1.325 | 0.387 | 7.813 |
| MO | 0.101 | 0.011 | 1.032 | 0.598 | 0.065 | 6.096 |
| GR | 0.040 | 0.000 | 0.337 | 0.235 | 0.003 | 1.988 |
| burr | ML | 0.225 | 0.038 | 1.347 | 1.326 | 0.225 | 7.959 |

\*Estimated using median, lower and upper confidence limits of HC05 from Table 8 and the default slope for the dose-response (4.5).

1. **SSD toolbox**

Software (the SSD Toolbox) was developed for fitting species sensitivity distributions. The SSD Toolbox provides a variety of algorithms to support fitting and visualization of simple SSDs. This tool implements the methods described above and generates the figures and tables displayed in this appendix. This toolbox was coded in Matlab. It will eventually be made publicly available, along with a user’s guide.

1. **References**

Anderson, D. 2008. Model based inference in the life sciences: a primer on evidence. Springer, New York.

Burnham, K.P. and D.R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach, 2nd Ed., Springer-Verlag, New York.

Cunnane, C. 1978. Unbiased plotting positions – a review. Journal of Hydrology 37:205-222.

Dunning, J.B. 1984. Body weights of 686 species of North American birds. Western Bird Banding Assoc. Monograph No. 1.

Gumbel, E.J., 1947. Discussion on paper by B.F. Kimball, 1946. (q.v.) Trans. Am. Geophys. Union, 28:951--952.

Mineau, P., Collins, B.T., and A. Baril. 1996. On the use of scaling factors to improve interspecies extrapolation of acute toxicity in birds. Regulatory Toxicology and Pharmacology, 24: 24-29.

Weibull, W., 1939. A statistical theory of strength of materials. Ing. Vet. Ak. Handl. (Stockholm), 151.

U.S. EPA. 2005. Methods/Indicators for Determining when Metals are the Cause of Biological Impairments of Rivers and Streams: Species Sensitivity Distributions and Chronic Exposure-Response Relationships from Laboratory Data. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH. EPA/600/X-05/027.

U.S.EPA. 2011. Review of methods for characterizing effects of pesticides and other chemical stressors to aquatic organisms. December 20, 2011. A white paper developed for FIFRA SAP. Available online at: http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OPP-2011-0898-0005

Weibull, W., 1939. A statistical theory of strength of materials. Ing. Vet. Ak. Handl. (Stockholm), 151.

1. The malathion data set used in this appendix has minor differences compared to the malathion data used in the effects characterization to generate SSDs used for thresholds. These differences arose because the SSD appendix was developed before the malathion dataset was completely reviewed. The figures in this appendix are for illustration of the methods described in this appendix only. [↑](#footnote-ref-1)
2. What percentage of active ingredient is considered a TGAI is dependent on the pesticide. For most pesticides, TGAI contain greater than 90% active ingredient. However, for some pesticides the TGAI may have a much lower percentage of active ingredient. The percentage of active ingredient in the products produced as manufacturing or technical products may be used to determine what percentages are considered appropriate for the TGAI. [↑](#footnote-ref-2)
3. Note that although formulation data are not being considered for SSDs, they will still be used in the weight of evidence analysis used in making effects determinations. [↑](#footnote-ref-3)